Theoretical Economics 15 (2020), 715–761 1555-7561/20200715

Optimal contracts with a risk-taking agent

Daniel Barron
Kellogg School of Management, Northwestern University

George Georgiadis
Kellogg School of Management, Northwestern University

Jeroen Swinkels
Kellogg School of Management, Northwestern University

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

Keywords. Risk-taking, contract theory, gaming.

JEL classification. D86, M2, M5.

1. Introduction

Contractual incentives motivate employees, suppliers, and partners to exert effort, but improperly designed incentives can instead encourage excessive risk-taking. These risk-taking motives are most obvious when they have dramatic consequences for society as a whole. For instance, following the 2008 financial crisis, Federal Reserve Chairman Ben Bernanke stated that "compensation practices at some banking organizations have led to misaligned incentives and excessive risk-taking, contributing to bank losses and financial instability" (The Federal Reserve 2009). Garicano and Rayo (2016) suggest that

Daniel Barron: d-barron@kellogg.northwestern.edu
George Georgiadis: g-georgiadis@kellogg.northwestern.edu
Jeroen Swinkels: j-swinkels@kellogg.northwestern.edu

We are grateful to the three anonymous referees, as well as to Simon Board, Ralph Boleslavsky, Luis Cabral, Odilon Câmara, Gabriel Carroll, Bengt Holmström, Hugo Hopenhayn, Johannes Hörner, Eddie Lazear, Elliot Lipnowski, Erik Madsen, Niko Matouschek, John Matsusaka, Meg Meyer, Konstantin Milbradt, Dimitri Papanikolaou, Andrea Prat, Lars Stole, Doron Ravid, Luis Rayo, Patrick Rey, Balazs Szentes, Dimitri Vayanos, Glen Weyl, and seminar participants at Cambridge University, New York University, Northwestern University, the University of Chicago, the University of Southern California, University College London, the University of Miami, the University of Exeter, the University of Rochester, the University of Warwick, the 2016 ASU Theory Conference, the 2016 Transatlantic Theory Conference, CRETE 2017, CSIO-IDEI 2017, and the Spring 2018 NBER Organizational Economics meeting for helpful comments and discussions.

 https://doi.org/10.3982/TE3660

poorly designed incentives led the American International Group (AIG) to expose itself to massive tail risk in exchange for the appearance of stable earnings. Rajan (2011) echoes these concerns and suggests that misaligned incentives worsened the effects of the crisis.

Even without such disastrous outcomes, agents face opportunities to game their incentives by engaging in risk-taking in many other settings. Portfolio managers can choose riskier investments, as well as exert effort, to influence their returns (Brown et al. 1996, Chevalier and Ellison 1997, de Figueiredo et al. 2015). Executives and entrepreneurs work hard to innovate, but also choose whether to pursue moonshot or incremental projects (Matta and Beamish 2008, Rahmandad et al. 2018, Vereshchagina and Hopenhayn 2009). In what we will see is a related phenomenon, salespeople can both work to sell more products and choose when those sales count toward their quotas (Oyer 1998, Larkin 2014).

In addition to the obvious social costs of excessive risk, the fact that agents can game their incentives in this way has a second cost as well: the possibility of risk-taking makes it harder for firms to motivate their agents to work hard. In this paper, we focus on this incentive cost by exploring how risk-taking constrains optimal contracts in a canonical moral hazard setting. We argue that the fact that the agent can game his incentives in this way renders convex incentives ineffective. Consequently, the principal can do no better than to offer a contract that makes the agent's utility concave in output. This simple but central result spurs us to analyze optimal *concave* contracts, with the goal of exploring how this additional concavity constraint changes the structure of incentives, profits, and productivity.

Our model considers a principal who offers an incentive contract to a potentially liquidity-constrained and risk-averse agent. If the agent accepts the contract, then he exerts costly effort that produces a noncontractible intermediate output, the distribution of which satisfies the increasing marginal likelihood ratio property. The key twist on this canonical framework is that the agent can engage in risk-taking by costlessly adding mean-preserving noise to this intermediate output, which in turn determines the contractible final output.

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

In Section 4, we consider the case of a risk-neutral agent and a weakly risk-averse principal. Absent the no-gaming constraint, the principal would like to offer a convex contract in this setting so as to concentrate high pay on high outcomes and so inexpensively motivate the agent while respecting his limited liability constraint. As a result, we show that the no-gaming constraint binds everywhere, which means that a linear (technically, affine) contract is optimal, remains so regardless of the principal's attitude

toward risk (even if she is risk-loving), and is uniquely optimal if the principal is risk-averse. In particular, relative to any strictly concave contract, we show that there is a linear contract that both better motivates the agent *and* better insures the principal.

Section 5 explores the consequences of risk-taking in the case of a risk-averse agent and a risk-neutral principal. In this setting, the no-gaming constraint implies that the agent's *utility* must be concave in output. Similar to Section 4, the optimal contract makes the agent's utility linear wherever this constraint binds. Unlike that section, however, the no-gaming constraint does not necessarily bind everywhere, so the agent's payoff under the optimal contract might have both linear and strictly concave regions.

We develop a set of necessary and sufficient conditions that characterize the unique profit-maximizing contract in this setting. We cannot directly apply the techniques of Mirrlees (1976) and Holmström (1979) because the resulting contract might violate the no-gaming constraint. Instead, we identify two perturbations of a candidate contract that respects this constraint while changing either the level or the slope of the agent's utility over appropriate *intervals* of output. Perhaps surprisingly, we prove that it suffices to consider these two perturbations so that a contract is profit-maximizing if, and only if, it cannot be improved by them.

We use this characterization to identify how the constraint that incentives be concave shapes the optimal contract. If the limited liability constraint binds and the participation constraint is slack in this setting, then the optimal contract follows a logic similar to the case with a risk-neutral agent. The principal would like to offer a contract that makes the agent's payoff convex over any output that suggests less than the desired effort. The profit-maximizing contract therefore makes the agent's utility linear over low outputs. Unlike the case with a risk-neutral agent, however, the optimal incentive scheme might make the agent's utility strictly concave following high output, since the principal finds it increasingly expensive to give the agent higher and higher utility.

If the limited liability constraint is slack, then the optimal contract is shaped by the same trade-off between incentives and insurance that arises in classic moral hazard problems. In the absence of risk-taking, the optimal contract would equate *output-by-output* the principal's marginal cost of paying the agent to the marginal benefit of relaxing his participation and incentive constraints (as in Mirrlees 1976 and Holmström 1979). Where this constraint binds, however, optimizing output-by-output would violate the no-gaming constraint. Over such regions, we show that the optimal contract is *ironed*, in the sense that it is linear in utility over an interval. *Expected* marginal benefits equal *expected* marginal costs on that interval. For instance, if the no-gaming constraint binds for low output but not for high output, then the optimal contract makes the agent's utility linear for low outputs and otherwise sets marginal benefits equal to marginal costs output-by-output. If no-gaming is slack everywhere, then the contract characterized by Mirrlees (1976) and Holmström (1979) is optimal; if it binds everywhere, then the optimal contract makes the agent's utility linear in output.

The final part of this section presents simulations of the optimal contract in a discrete approximation of the model. We show that optimal incentives are characterized by a standard convex optimization program, and we consider examples that illustrate how parameters of the model influence the optimal contract.

The unifying idea behind all of our results is that the possibility of risk-taking renders convex incentives ineffective. Section 6 extends this intuition to three other settings, all of which assume that both the principal and the agent are risk-neutral. First, we modify the agent's payoff so that he incurs a cost that is increasing in the variance of his risk-taking distribution. It turns out that this extension can be reformulated as a variant of our analysis in Section 4. We show that the unique optimal contract is strictly convex in output, but not so convex as to induce gaming, and that this contract converges to a linear contract as gaming becomes costless.

Second, we alter the timing of the model so that the agent engages in risk-taking *before* he observes intermediate output. We show that the possibility of ex ante risk-taking leads optimal incentives to be a concave function of the agent's *effort*, rather than a concave function of intermediate output. This modified no-gaming constraint binds under mild conditions, in which case a linear contract is optimal.

Finally, we exhibit a close connection between risk-taking and another type of gaming: manipulating the *timing* of output. To do so, we study a dynamic setting in which the principal offers a stationary contract that the agent can game by choosing *when* output is realized over an interval of time. For example, Oyer (1998) and Larkin (2014) document how salespeople accelerate or delay sales so as to game convex incentive schemes over a sales cycle. We show that this setting is equivalent to our risk-taking model. Thus, a linear contract is optimal, since a strictly convex contract would induce the agent to bunch sales over short time intervals and a strictly concave contract would provide subpar effort incentives.

Our analysis is inspired by Diamond (1998) and Garicano and Rayo (2016). The latter includes a model of risk-taking that is similar to ours, but it fixes an exogenous (nonconcave) contract to focus on the social costs of excessive risk. The former is a seminal exploration of optimal contracts when the agent can both exert effort and make other choices that affect the output distribution. In particular, part of Diamond (1998) argues that linear contracts are (nonuniquely) optimal in an example with risk-neutral parties, binary effort, and an agent who can choose any mean-preserving spread of output. Our Proposition 2 expands this result to settings with a risk-averse principal as well as more general effort choices and output distributions. In doing so, we identify an additional advantage of linear contracts with a risk-neutral agent: relative to any strictly concave contract, they better insure the principal and so are *uniquely* optimal if the principal is even slightly risk-averse.

The rest of our analysis departs further from Diamond (1998). Section 3 shows that the fundamental consequence of agent risk-taking is to constrain incentives to be *concave*, not necessarily linear. Linear contracts are instead a consequence of this concavity constraint binding everywhere, as it does if the agent is risk-neutral. However, as Section 5 demonstrates, the concavity constraint need not necessarily bind everywhere if the agent is risk-averse, in which case the optimal contract may make utility strictly concave in output. Our analysis shows how risk-taking affects contracts in a classic moral hazard setting. Section 6 explores how a similar logic shapes optimal contracts in several related settings.

Our model of risk-taking is embedded in a classic moral hazard problem. With a risk-neutral agent, our model builds on Innes (1990), Poblete and Spulber (2012), and other papers in which limited liability is the central contracting friction. With a risk-averse agent, we build on Mirrlees (1976) and Holmström (1979) if the limited liability constraint is slack, and on Jewitt et al. (2008) if it binds. Within the classic agency literature, our analysis is conceptually related to papers that study principal–agent relationships in which the agent both exerts effort and makes other decisions. Classic examples include Lambert (1986) on how agency problems in information-gathering can lead to inefficient investment in risky projects and Holmström and Ricart i Costa (1986) on project selection under career concerns. Malcolmson (2009) presents a general model of such settings, but differs from our analysis by assuming that decisions are contractible. Other papers consider settings in which the principal also chooses actions other than the agent's wage contract, such as an endogenous performance measure; see, for example, Halac and Prat (2016) and Georgiadis and Szentes (2019).

A growing literature studies agent risk-taking. Some papers in this literature assume that an agent chooses from a parametric class of risk-taking distributions in either static (Palomino and Prat 2003, Hellwig 2009) or dynamic (DeMarzo et al. 2013) settings. We differ by allowing our agent to choose any mean-preserving spread of output, which means that our optimal contract must deter a more flexible form of gaming. Therefore, we join other papers that study nonparametric risk-taking, again in either static (Robson 1992, Diamond 1998, Hébert 2018) or dynamic (Ray and Robson 2012, Makarov and Plantin 2015) settings. We differ from these papers by identifying *concavity* as the key constraint on the optimal incentive scheme if the agent can costlessly take on risk and then characterizing optimal incentives given this constraint.[1]

More broadly, our work is related to a longstanding literature that argues that optimal contracts must both induce effort and deter gaming. A seminal example is Holmström and Milgrom (1987), who display a dynamic environment in which linear contracts are optimal. Ederer et al. (2018) show how opacity (i.e., randomization over compensation schemes) can be used to deter gaming. Others, including Chassang (2013), Carroll (2015), and Antić (2016), depart from a Bayesian framework and prove that simple contracts perform well under min-max or other non-Bayesian preferences. In contrast, our paper considers contracts that deter gaming in a setting that lies firmly within the Bayesian tradition.

While Carroll's paper considers a max-min rather than a Bayesian solution concept, its intuition is related to ours. In that paper, Nature selects a set of actions available to the agent so as to minimize the principal's expected payoffs. As in our setting, Nature might allow the agent to take on additional risk to game a convex incentive scheme. However, Nature might also allow the agent to choose a distribution with *less* risk to game a concave incentive scheme, while we allow the agent to add risk but not reduce it. That is, we model a moral hazard problem in which output is intrinsically risky and that risk cannot be completely hedged away. This difference is most striking if the agent

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

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. One advantage of our approach is that our model results in a canonical contracting problem with an additional concavity constraint. Consequently, our technology would be straightforward to embed in Bayesian models of other applications.

2. Model

We consider a 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 $[\underline{y}, \overline{y}] \equiv \mathcal{Y} \subseteq \mathbb{R}$ be the set of contractible outputs, with $\underline{y} < 0$. The timing is as follows.

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

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

(iii) If the agent accepts, he chooses effort $a \geq 0$.

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

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

(vi) 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 $\pi(y - s(y))$ and $u(s(y)) - c(a)$, respectively.

We assume that $\pi(\cdot)$ and $u(\cdot)$ are strictly increasing and weakly concave, with $u(\cdot)$ onto, and that $c(\cdot)$ is infinitely differentiable, strictly increasing, and strictly convex. We also assume that $F(\cdot|a)$ has full support for all $a \in [0, \overline{y})$, satisfies $\mathbb{E}_{F(\cdot|a)}[x] = a$, and is infinitely differentiable with a density $f(\cdot|a)$ that is strictly monotone likelihood ratio property- (MLRP-) increasing in a, with $f_a(\cdot|a)/f(\cdot|a)$ uniformly bounded for all a.[3]

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

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

denote the set of mappings $x \mapsto G_x$. Without loss, we can treat the agent as choosing a and $G \in \mathcal{G}$ simultaneously.

Intermediate output has different interpretations in different settings. For instance, chief executive officers (CEOs) typically have advance information about whether they

[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.

[3] We assume that $\overline{y}$ is sufficiently large such that the principal never offers a contract that induces the agent to choose $a = \overline{y}$. Together with $\underline{y} < 0$ and $a \geq 0$, this also ensures that the agent can always choose a nondegenerate distribution G_x.

will hit their earnings targets in a given quarter, and they can cut maintenance or research and development expenditures if they are likely to fall short, taking on tail risk for the appearance of higher earnings (Rahmandad et al. 2018). Similarly, portfolio managers are typically compensated based on their annual returns and can adjust the riskiness of their investments over the course of the year so as to game those incentives (Chevalier and Ellison 1997).

After the agent observes x but before y is realized, we have a setting with both a hidden type and a hidden action. The principal might therefore benefit from asking the agent to report x before y is realized. By punishing differences between this report and y, the principal might be able to dissuade at least some gambling.[4] We do not allow such mechanisms in our analysis. This restriction makes sense if the principal cannot intervene between the realization of x and the outcome of gambling, as is the case if x is realized at a random time and gambling is instantaneous. We think that this is the economically correct modeling assumption in many settings. For instance, financial advisors realize their expected returns and choose their investment strategies over time, rendering it impossible to identify a single moment at which intermediate output has been realized but final output has not. The spirit of the model is that the principal cannot catalog the precise moments or ways in which an agent might engage in risk-taking.

3. Risk-taking and optimal incentives

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

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

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

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

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

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

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

[4] Allowing these types of mechanisms does not necessarily eliminate the agent's gaming incentives. The the supplementary file on the journal website, http://econtheory.org/supp3660/supplement.pdf, includes a supplemental section that studies this case and shows that gaming continues to constrain incentives. Indeed, if both parties are risk-neutral, then linear contracts are optimal.

Proposition 1 (The no-gaming constraint). *Suppose* $(a, G, v(\cdot))$ *satisfies (IC$_F$)–(LL$_F$). Then there exists a weakly concave* $\hat{v}(\cdot)$ *such that* $(a, G^D, \hat{v}(\cdot))$ *satisfies (IC$_F$)–(LL$_F$) and gives the principal a weakly higher expected payoff.*

The proof of Proposition 1 is in Appendix A. For an arbitrary incentive scheme $v(\cdot)$, define $v^c(\cdot) : \mathcal{Y} \to \mathbb{R}$ as its concave closure:

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

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

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

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

$$\text{subject to} \quad a \in \arg\max_{\tilde{a}} \big\{\mathbb{E}_{F(\cdot|\tilde{a})}\big[v(y)\big] - c(\tilde{a})\big\}, \tag{IC}$$

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

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

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

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

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

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

This result, which follows from the theorem of the maximum, is an implication of Proposition 9 in Appendix D. Existence is guaranteed by (NG); for example, without this constraint, no profit-maximizing contract would exist with risk-neutral parties.[5] If

[5]With risk-neutral parties, the principal wants to pay the agent only after an arbitrarily narrow range of the highest outputs, since those outputs are most indicative of high effort. See, e.g., Innes (1990).

at least one player is strictly risk-averse, then Jensen's inequality implies that a convex combination of two different contracts that implement a also implements a and gives the principal a strictly higher payoff, which proves uniqueness.

4. Optimal contracts for a risk-neutral agent

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

For any effort level a, define

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

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

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

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

The proofs for all results in this section can be found in Appendix A. To see the intuition for Proposition 2, consider $s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$, which both implements a^{FB} and provides full insurance to the principal. If $s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$ satisfies (IR) with equality, then it is clearly optimal.

Suppose instead that (IR) is slack for $s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$, in which case (LL) must bind. Suppose that $(a^*, s^*(\cdot))$ is optimal and $s^*(y) \neq s_{a^*}^{\mathrm{L}}(y)$ for at least some y. To prove the result, we construct a linear contract that satisfies (IC)–(NG) and performs better than $s^*(\cdot)$. Toward this goal, define $\hat{s}(\cdot)$ to be the linear contract that agrees with $s^*(\cdot)$ at $\underline{y}$ and gives the agent the same utility as $s^*(\cdot)$ if he optimally responds to that contract. As shown in Figure 1, $\hat{s}(\cdot)$ must single-cross $s^*(\cdot)$ from below, effectively moving payments from low to high outputs. Since $F(\cdot|a)$ satisfies MLRP, this shift in pay from low to high outputs motivates more effort: $\hat{s}(\cdot)$ implements some $\hat{a} \geq a^*$.

The effort $\hat{a}$ might be either larger or smaller than first-best effort, a^{FB}. If $\hat{a} \geq a^{\mathrm{FB}}$, then it must be that $\hat{s}(\cdot) \geq s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$. But then $s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$ implements a^{FB}, perfectly insures the principal, and entails smaller payments than $\hat{s}(\cdot)$. The principal therefore prefers $s_{a^{\mathrm{FB}}}^{\mathrm{L}}(\cdot)$ to $s^*(\cdot)$.

If instead $\hat{a} < a^{\mathrm{FB}}$, then the slope of $\hat{s}(\cdot)$ must be strictly less than 1, which means that the principal's wealth under $\hat{s}(\cdot)$, $y - \hat{s}(y)$, is increasing in y. Consequently, the principal prefers high outputs and so she likes that $\hat{a} \geq a^*$. Moreover, $\hat{s}(y) > s^*(y)$ exactly when output is high and so her marginal utility is low (and vice versa), which means that $\hat{s}(\cdot)$ also insures the principal better than $s^*(\cdot)$. Therefore, the principal prefers $\hat{s}(\cdot)$ to $s^*(\cdot)$. She a fortiori prefers $s_{\hat{a}}^{\mathrm{L}}(\cdot)$, which lies weakly below $\hat{s}(\cdot)$, to $s^*(\cdot)$. We conclude that any

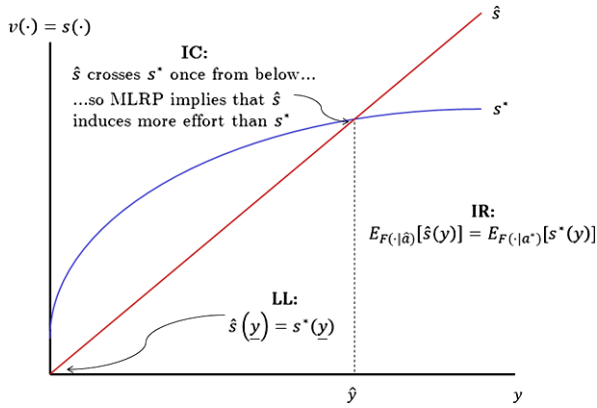

FIGURE 1. Intuition for the proof of Proposition 2.

optimal contract $s^*(\cdot)$ must coincide with the the linear contract that implements a^*, $s^*(\cdot) \equiv s^{\mathrm{L}}_{a^*}(\cdot)$.

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

For any $a > 0$, the agent's promised utility under $s^{\mathrm{L}}_{a}(\cdot)$ depends on $\underline{y}$, the *worst* possible outcome over which the agent can gamble. In particular, $s^{\mathrm{L}}_{a}(\cdot)$ starts at $\underline{y}$ and has a strictly positive slope, so that the agent's expected compensation increases without bound as $\underline{y}$ decreases. That is, as the agent's ability to take on left-tail risk becomes arbitrarily severe, motivating effort while deterring risk-taking becomes arbitrarily costly to the principal. Consequently, the optimal effort level converges to 0 as $\underline{y}$ becomes arbitrarily negative.[6]

The possibility of risk-taking unambiguously harms the principal. However, the agent might either benefit or be harmed by risk-taking. The reason is that risk-taking both increases the agent's rent for a fixed effort level and changes the optimal effort level, which changes the agent's rent. Consequently, we can find examples in which the agent earns higher rent when we impose (NG) as well as examples in which he earns strictly lower rent.

In some applications, the principal might have risk-seeking preferences over output, for instance because she also faces convex incentives. For example, Rajan (2011) argues that, anticipating the possibility of bailouts, shareholders of financial institutions might have had an incentive to encourage risk-taking prior to the 2008 financial crisis. We can model such settings by allowing $\pi(\cdot)$ to be any strictly increasing and continuous function. Proposition 1 does not directly apply in this case because the principal might

[6]If the principal is risk-neutral, then we can prove the stronger result that effort is strictly increasing in $\underline{y}$: as the agent's ability to take left-tail risks becomes more severe, the principal responds by inducing lower effort. See the supplementary file on the journal website, http://econtheory.org/supp3660/supplement.pdf.

strictly prefer the agent to at least sometimes engage in risk-taking. Nevertheless, we can modify the argument from Proposition 2 to show that a linear contract is optimal.

COROLLARY 1 (Risk-neutral agent, risk-loving principal). *Let $u(s) \equiv s$ and let $\pi(\cdot)$ be an arbitrary continuous and strictly increasing function that has concave closure $\pi^c(\cdot)$. If a^* is optimal, then $a^* \leq a^{\text{FB}}$ and $s^{\text{L}}_{a^*}(\cdot)$ is optimal.*

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

5. Optimal contracts if the agent is risk-averse

This section characterizes the unique contract that implements a given $a > 0$ at maximum profit in a setting with a risk-averse agent and a risk-neutral principal. In Section 5.1, we develop necessary and sufficient conditions that characterize the profit-maximizing contract in this setting. We explore the implications of this characterization in Section 5.2; in Section 5.3, we show how to numerically derive the optimal contract in a discrete approximation of the model.

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

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

Replacing (IC) with (IC-FOC) entails no loss under mild regularity conditions on $F(\cdot|\cdot)$. Given (NG), Proposition 5 of Chade and Swinkels (2016) shows that the agent's expected utility is concave in effort as long as expected output is concave in effort and $F_{aa}(\cdot|a)$ is never first negative and then positive. For a fixed effort $a \geq 0$, define the principal's problem

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

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

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

Define $\rho(\cdot)$ as the function that maps $1/u'(\cdot)$ into $u(\cdot)$; that is, for every z in the range of $1/u'(\cdot)$, $\rho(z) = u((u')^{-1}(1/z))$. Then $\rho^{-1}(v(y))$ equals the marginal cost to the principal of giving the agent extra utility at y.

If $\underline{u} > -\infty$, then Lemma 1 implies that a unique solution to (P) exists. If $\underline{u} = -\infty$, then one can show that a unique solution exists as long as $u'(\cdot)$ is not too convex. In particular, we can define the *concavity* of a positive function $h(\cdot)$, $\text{con}(h)$, as the largest number t such that $\frac{h^t}{t}$ is concave. If h is concave, then $\text{con}(h) \geq 1$, while if h is log concave, then $\text{con}(h) \geq 0$. For the case $\underline{u} = -\infty$, an optimal contract exists as long as $\text{con}(u') \geq -2$, which is weaker than $u'(\cdot)$ being log concave.[7] Our results in this section apply in either setting. Unless otherwise noted, proofs for this section can be found in Appendix B.

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

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

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

Let us ignore (LL) for the moment. Absent (NG), the optimal contract would set $n(y) = 0$ output-by-output and so $v(\cdot) = \rho(\lambda + \mu l(\cdot|a))$. Indeed, this incentive scheme (with the appropriate λ and μ) is the *Holmström–Mirrlees* (HM) contract characterized in Mirrlees (1976) and Holmström (1979). However, setting $n(y) = 0$ at each y might violate (NG). In the following section, we develop necessary and sufficient conditions for a profit-maximizing contract. These conditions guarantee that the contract cannot be improved by a set of perturbations that respect (NG) and affect an interval of an incentive scheme. We show that these perturbations are enough to pin down the optimal contract.

5.1 *A characterization*

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

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

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

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

Consider the following two perturbations, formally defined in Appendix B and illustrated in Figure 2. *Raise* increases the level of $v(\cdot)$ by a constant over an interval, while

[7] For a (rather complicated) proof of existence for $\underline{u} = -\infty$, available in a supplementary file on the journal website, http://econtheory.org/supp3660/supplement.pdf. This condition is satisfied for, for instance, $u(w) = w^\alpha$ for $\alpha < \frac{1}{2}$. See Prékopa (1973) and Borell (1975) for details.

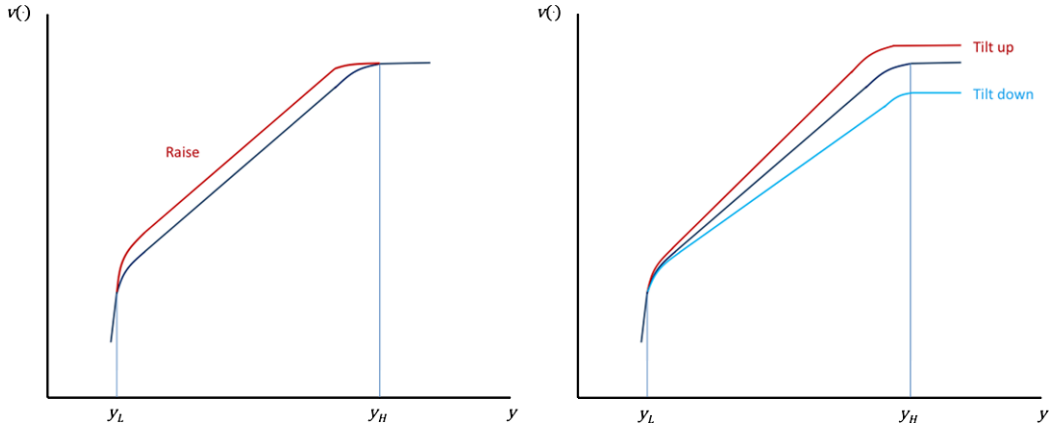

FIGURE 2. *Raise* and *tilt*. These perturbations require care around y_L and y_H to ensure that concavity is preserved. For this reason, we need both y_L and y_H to be free for *raise*. For *tilt up*, we need y_L to be free, while y_H must be free for *tilt down*.

tilt increases the slope of $v(\cdot)$ by a constant over an interval. Raising an interval typically introduces nonconcavities into $v(\cdot)$ at both endpoints of the interval. Tilting it a positive amount may introduce a nonconcavity at the lower end of the interval and tilting it a negative amount may introduce a nonconcavity at the upper end of the interval. Appendix B shows that for small perturbations, we can repair these nonconcavities on an arbitrarily small interval as long as the relevant endpoints are free.

Raise and *tilt* affect both (IR) and (IC-FOC). However, Appendix B uses the fact that $F(\cdot|a)$ satisfies MLRP to show that these two perturbations have noncollinear effects on (IR) and (IC-FOC), which means that we can construct combinations of them to affect each constraint separately. Therefore, as long as there exists at least one free point $\hat{y} < \overline{y}$ such that $v(\hat{y}) > \underline{u}$, we can use *raise* and *tilt* on $[\hat{y}, \overline{y}]$ to establish the shadow values λ and μ of relaxing (IR) and (IC-FOC). If no such point exists, then $v(\cdot)$ is linear and $v(\underline{y}) = \underline{u}$.

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

$$\int_{y_L}^{y_H} n(y) f(y|a)\, dy \geq 0. \tag{2}$$

If $v(y_L) > \underline{u}$, then we can raise $v(\cdot)$ by a negative amount on $[y_L, y_H]$, in which case (2) holds with equality.

Similarly, if y_L is free, then tilting $v(\cdot)$ on $[y_L, y_H]$ must have nonnegative expected net cost,

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

where the first term represents the fact that *tilt* increases the slope of $v(\cdot)$ from y_L to y_H and the second represents the resulting higher level of $v(\cdot)$ from y_H to $\overline{y}$. If y_H is free,

then applying negative *tilt* yields the reverse inequality:

$$\int_{y_L}^{y_H} n(y)(y-y_L)f(y|a)\,dy + (y_H - y_L)\int_{y_H}^{\overline{y}} n(y)f(y|a)\,dy \leq 0. \tag{4}$$

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

Definition 2. A contract $v(\cdot)$ is generalized Holmström–Mirrlees (GHM) if (IC-FOC) holds with equality, (IR), (LL), and (NG) are satisfied, there exist $\lambda \geq 0$ and $\mu > 0$ such that

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

and for any $y_L < y_H$,

(i) if y_L and y_H are free, then (2) holds, and holds with equality if $v(y_L) > \underline{u}$;

(ii) if y_L is free, then (3) holds;

(iii) if y_H is free, then (4) holds.

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

Proposition 3 (Risk-averse agent, risk-neutral principal). *Suppose $u(\cdot)$ is strictly concave and $\pi(y) \equiv y$. Then for any $a > 0$, $v(\cdot)$ implements a at maximum profit if and only if it is GHM.*

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

One implication of Proposition 3 is that net cost equals 0 for any output where both (LL) and (NG) are slack.

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

At any point of normal concavity y, we can find two free points that are arbitrarily close to y.[8] Proposition 3 implies that (2) holds with equality between these points;

[8] See Claim 1 in Appendix B.

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

5.2 *Implications of the no-gaming constraint*

This section builds on Proposition 3 to illustrate how risk-taking affects the trade-off between insurance and incentives that lies at the heart of this moral hazard problem. For a broad class of settings, we show that optimal incentives are linear in output where (NG) binds and otherwise equate the marginal costs and benefits of incentive pay at each output.

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

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

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

The proof of Lemma 2 can be found in Appendix D.2. The requirement that $\operatorname{con}(\rho') + \operatorname{con}(l_y) > -1$ is relatively mild. It is automatic if ρ' and l_y are log concave, but it also holds, for example, if l_y is strictly log concave and the agent's utility function is from a broad class that satisfy hyperbolic absolute risk aversion, including $u(w) = \log w$.[9]

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

PROPOSITION 4 (Slack (LL)). *Fix $a \geq 0$ and $\pi(y) \equiv y$. Let $v^*(\cdot)$ solve (P), let λ and μ be the shadow values on (IR) and (IC-FOC), respectively, and suppose that $v^*(\underline{y}) > \underline{u}$.*[10] *Suppose there exists y_I such that $\rho(\lambda + \mu l(\cdot|a))$ is convex on $[\underline{y}, y_I)$ and concave on $(y_I, \overline{y}]$. Then*

[9]Recall that hazard analysis and risk assessment utility satisfies $u(w) = \frac{1-\gamma}{\gamma}(\frac{\alpha w}{1-\gamma} + \beta)^{\gamma}$. If l_y is strictly log concave, then $\rho(\lambda + \mu l(\cdot|a))$ is first convex and then concave if either $\gamma < 0$ or $\gamma \in (\frac{1}{2}, 1)$. To see this result for $u(w) = \log w$, observe that $\frac{1}{u'(w)} = w$ in this case. By definition, $\rho(\frac{1}{u'(w)}) = u(w)$, which means that $\rho(w) = \log w$. Then $\rho'(w) = \frac{1}{w}$, which satisfies $\operatorname{con}(\rho'(w)) = -1$ because $\frac{(1/w)^{-1}}{-1} = -w$. The assumption that l_y is strictly log concave then ensures that $\operatorname{con}(l_y) > 0$ and, hence, $\operatorname{con}(\rho') + \operatorname{con}(l_y) > -1$.

[10]The supplementary file on the journal website, http://econtheory.org/supp3660/supplement.pdf, gives conditions under which an optimal contract exists even if $\underline{u} = -\infty$. Under those conditions, this existence proof also shows that (LL) is slack if $\underline{u} > -\infty$ is sufficiently negative.

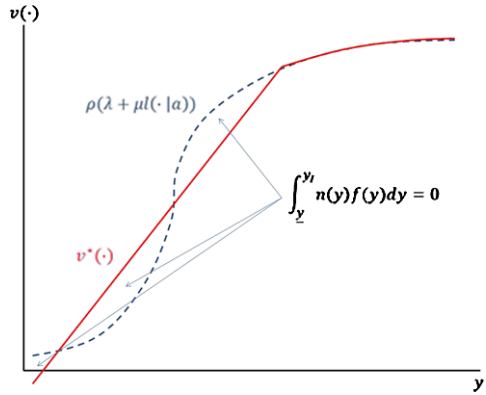

FIGURE 3. Illustration of $\rho(\lambda + \mu l(\cdot|a))$ and the profit-maximizing $v^*(\cdot)$.

$v^*(\cdot)$ satisfies (IR) and (IC-FOC) with equality, and there exist $\hat{y} \geq y_I$, $\underline{v} \in \mathbb{R}$, and $\alpha \in \mathbb{R}_+$ such that $v^*(\cdot)$ is continuous,

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

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

Under the condition that $\rho(\lambda + \mu l(\cdot|a))$ is first convex and then concave, and (LL) is slack, the profit-maximizing contract $v^*(\cdot)$ is linear in utility for low output and otherwise sets $n(y) = 0$ output-by-output. Moreover, on the linear region of $v^*(\cdot)$, expected net costs equal 0. See Figure 3 for an illustration.

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

Proposition 4 focuses on the case where (LL) is slack, but (NG) has a similar effect if (IR) is slack so that (LL) binds. In that case, the principal would like to pay the agent as little as possible for any y with $l(y|a) < 0$, since paying for low output both increases the

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

[12] In particular, $\rho(\lambda + \mu l(\cdot|a))$ is concave (convex) if both $\rho(\cdot)$ and $l(\cdot|a)$ are concave (convex). Recalling that prudence is $-u'''(\cdot)/u''(\cdot)$ and absolute risk aversion is $-u''(\cdot)/u'(\cdot)$, we can show that $\rho(\cdot)$ is convex if the ratio of prudence to absolute risk aversion exceeds 3. Note that this condition is equivalent to $\text{con}(u') < -2$. The border case is $u(w) = \sqrt{w}$, which corresponds to a linear $\rho(\cdot)$.

agent's rent and tightens (IC-FOC) (Jewitt et al. 2008). But paying the agent as little as possible for low output and rewarding high output would violate (NG), so this constraint binds following low output.

PROPOSITION 5 (Slack (IR)). *Fix $a \geq 0$ and $\pi(y) \equiv y$. Let $v^*(\cdot)$ solve (P) and suppose that (IR) is slack under $v^*(\cdot)$. Define y_0 such that $l(y_0|a) = 0$. Then $v^*(\cdot)$ is linear on $[\underline{y}, y_0]$.*

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

5.3 *Numerical examples*

In this section, we present simulations of the profit-maximizing contract for a version of the model with discrete outputs. Fix $N \in \mathbb{N}$ and define

$$\mathcal{Y} = \left\{ y \mid y = \underline{y} + i(\overline{y} - \underline{y})/N \text{ for some } i \in \{1, \ldots, N\} \right\}.$$

We constrain output to satisfy $y \in \mathcal{Y} = \{y_1, \ldots, y_N\}$, where the probability that $y = y_i$ is $p_i(a) \equiv \int_{y_{i-1}}^{y_i} f(z|a)\, dz$ and we define $y_0 = \underline{y}$.

For any $a > 0$, the profit-maximizing contract in this discrete setting solves the discrete version of (P),

$$\begin{aligned} \max_{v \in \mathbb{R}^N} \quad & \sum_{i=1}^{N} \left[y_i - u^{-1}(v_i) \right] p_i(a) && (\mathrm{P}_N) \\ \text{subject to} \quad & \sum_{i=1}^{N} v_i p_i'(a) \geq c'(a), \\ & \sum_{i=1}^{N} v_i p_i(a) - c(a) \geq u_0, \\ & v_i \geq \underline{u} \quad \text{for all } i, \\ & 2v_i \geq v_{i-1} + v_{i+1} \quad \text{for all } i \in \{2, \ldots, N-1\}, && (5) \end{aligned}$$

where v_i represents the agent's utility following output y_i. The benefit of this discrete setting is that the analog to the no-gaming constraint, (5), can be written in a way that is linear in v_i. Therefore, the contracting problem is a convex optimization program that can be solved using standard techniques. It can be shown that the solution of (P_N) and the corresponding payoffs converge to the solution of the original problem as $N \to \infty$.

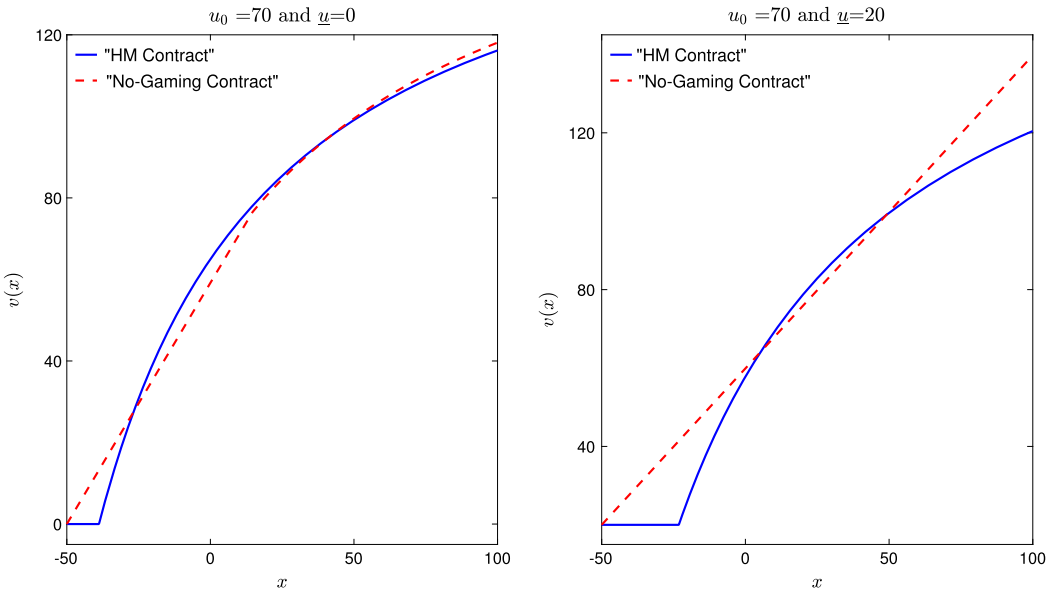

FIGURE 4. An example of the contracts that implement a given effort at maximum profit with and without the concavity constraint.

Figure 4 gives an example that fixes the effort level and varies the lower bound on the agent's utility, $\underline{u}$. In each panel, the contract that solves (P_N), the no-gaming contract, is denoted by a dashed line, while the contract without risk-taking, the HM contract, does not impose (5) and is denoted by a solid line. For all of our examples, $N = 1{,}000$.[13]

Consider the left panel of Figure 4. In this example, the no-gaming constraint binds, and so the no-gaming contract is linear in utility for an interval of outputs including the lowest one. This result echoes our observation that the optimal contract resembles an ironed version of the contract without risk-taking. Note, however, that the no-gaming constraint affects the optimal contract even for outputs where (NG) is slack. This global effect arises because the no-gaming constraint distorts the multipliers λ and μ and so changes the net cost of paying the agent following *any* output realization. In both panels, the limited liability constraint is binding and so the no-gaming contract makes the agent's utility linear following low output, as Proposition 5 suggests. Increasing $\overline{u}$ makes this liability constraint binding over a wider range of outputs and so expands the region over which the agent's utility is linear.

Figure 5 uses a similar example to illustrate how the profit-maximizing contract changes with the agent's outside option, u_0.[14] As u_0 increases, the limited liability constraint becomes "less binding" in the sense that it binds for a smaller range of outputs.

[13]These examples assume that $u(\omega) = 2\sqrt{\omega}$, $c(a) = a^2/100$, $u_0 = 70$, and $\underline{u} = 0$ ($\underline{u} = 20$) in the left (right) panel. The set of contractible outputs satisfies $\mathcal{Y} \subseteq [-50, 100]$ and the density of intermediate output is $f(y|a) = \alpha y + \beta$. The parameters α and β are pinned down by the constraints $\int f(y|a)\,dy = 1$ and $\int yf(y|a)\,dx = a$, and the principal aims to implement effort $a = 40$.

[14]Apart from u_0, this example is identical to the left panel in Figure 4.

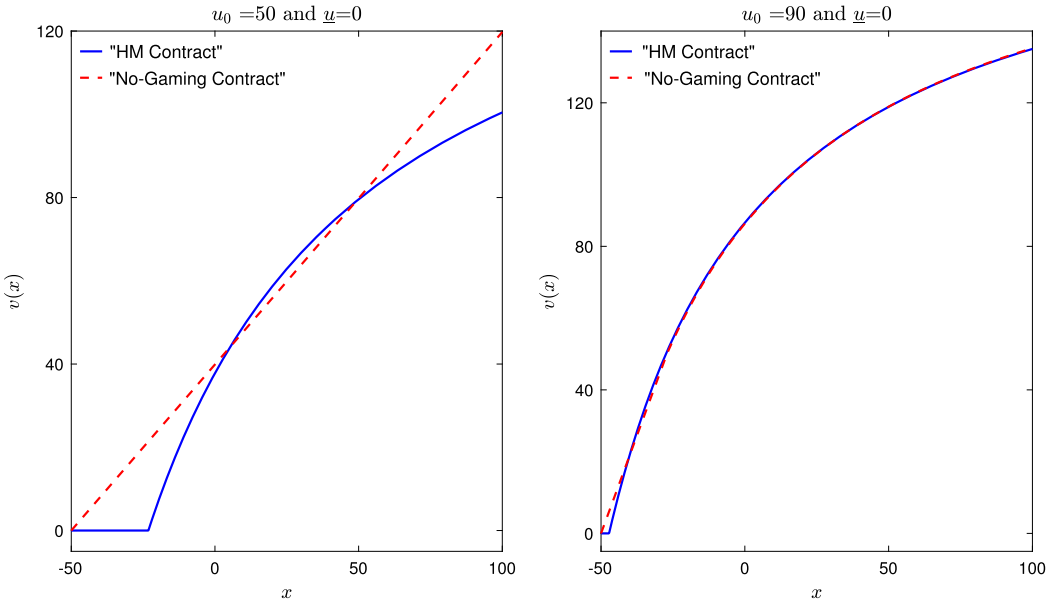

FIGURE 5. An example of the contracts that implement a given effort at maximum profit with and without the concavity constraint for different outside options, u_0.

Again consistent with Proposition 5, the no-gaming contract is linear over a smaller range of outputs as u_0 increases.

Thus far in this section, we have characterized the profit-maximizing contract for a fixed effort level. We can numerically solve for the (approximately) optimal effort level by solving (P_N) for a fine grid of efforts. Of course, the possibility of gaming unambiguously increases the total incentive cost of inducing any fixed effort level. However, imposing the no-gaming constraint has an ambiguous effect on the incentive cost of inducing increased effort. Consequently, the possibility of risk-taking can either increase or decrease the optimal effort level. Indeed, Figure 6 illustrates examples in which each of these possibilities obtains.[15]

6. Extensions and reinterpretations

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

[15]These examples are identical to the example from Figure 4 except that $u_0 = 20$, $\underline{u} = 2\sqrt{10}$, $c(a) = 0.004a^2$ in the left panel, and $c(a) = 0.002a^2$ in the right panel.

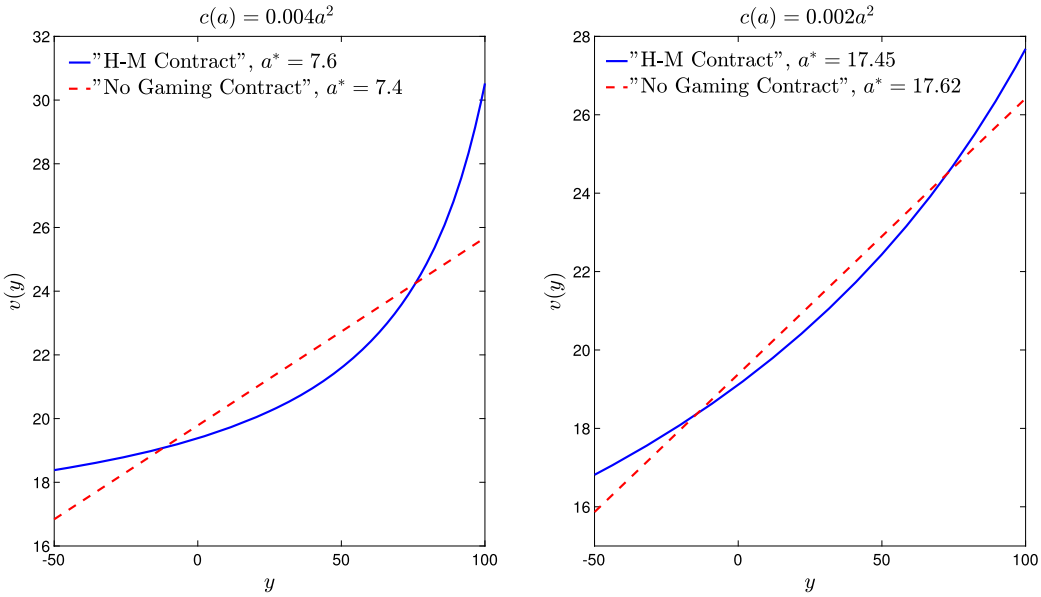

FIGURE 6. The possibility of risk-taking can result in either lower effort (the left panel) or higher effort (the right panel).

6.1 *Costly risk-taking*

In many settings, the agent might have to bear a cost to engage in risk-taking. A portfolio manager, for example, might spend time and effort to identify investments that allow for risk-taking without being detected. If larger gambles are harder to hide from investors, then the manager's cost is increasing in the dispersion of the risk-taking distribution. In this section, we adapt the arguments in Propositions 1 and 2 to a model with costly risk-taking. The resulting contracts are strictly convex, providing a rationale for such contracts in practice.

Consider the model from Section 2, and suppose that the agent must pay a private cost $\mathbb{E}_{G_x}[d(y)] - d(x)$ to implement distribution G_x following the realization of x, where $d(\cdot)$ is smooth, strictly increasing, and strictly convex, with $d(\underline{y}) = 0$. For example, this cost function equals the variance of G_x if $d(y) = y^2$. More generally, $d(\cdot)$ captures the idea that the agent must incur a higher cost to take on more dispersed risk. The principal's and agent's payoffs are $y - s(y)$ and $s(y) - c(a) - d(y) + d(x)$, respectively.[16]

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

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

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

As in Section 3, the agent chooses G_x so that his expected payoff equals $\tilde{v}^c(x)$. Since $\tilde{\pi}(\cdot)$ is strictly concave, the principal prefers to deter risk-taking by offering a contract

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

that makes the agent's payoff $\tilde{v}(\cdot)$ concave. Consequently, we can modify the proof of Proposition 2 to show that the principal's optimal contract makes $\tilde{v}(\cdot)$ linear. The optimal $s(\cdot)$ equals $\tilde{v}(\cdot) + d(\cdot)$ and is, therefore, strictly convex.

PROPOSITION 6 (Costly risk-taking). *Assume $\tilde{c}(\cdot)$ is strictly increasing and strictly convex. For optimal effort $a^* \geq 0$, define $s^*(y) = \tilde{c}'(a^*)(y - \underline{y}) + d(y) - \tilde{w}$, where $\tilde{w} = \min\{M, \tilde{c}'(a^*)(a - \underline{y}) - \tilde{c}(a^*) - u_0\}$. Then $s^*(\cdot)$ is optimal.*

This result follows a logic similar to Proposition 2, where the optimal $s^*(\cdot)$ ensures that $\tilde{v}(\cdot)$ is linear. Intuitively, $s^*(\cdot)$ is the most convex contract that deters the agent from gambling. Note that the principal earns more if risk-taking is costly, since she can offer somewhat convex incentives without inducing gaming.

6.2 *Risk-taking before intermediate output is realized*

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

Consider the following game.

Move 1. The principal offers a contract $s(y) : \mathcal{Y} \to [-M, \infty)$.

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

Move 3. The agent chooses an effort $a \geq 0$ and a distribution $G(\cdot) \in \Delta(\mathcal{Y})$ subject to the constraint $\mathbb{E}_G[x|a] = a$.[17]

Move 4. The outcome of the gamble $x \sim G(\cdot)$ is realized and final output is realized according to $y \sim F(\cdot|x)$. We assume that $F(\cdot|x)$ has full support, with $\mathbb{E}_{F(\cdot|x)}[y] = x$ and a density $f(\cdot|x)$ that satisfies strict MLRP in x.

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

By choosing $G(\cdot)$, the agent essentially randomizes his level of effort. This feature means that the contract cannot increase the agent's expected payoff following effort a without also increasing the expected payoff of exerting less effort and randomizing between $x = a$ and some lower x. The agent will therefore engage in risk-taking whenever his expected payoff as a function of *effort* is convex. One advantage of this model is that our tools extend naturally to it, a feature that is not shared by every model of ex ante risk-taking.[18]

As an example of the kind of risk-taking that fits this setting, suppose the principal is an investor and the agent is an entrepreneur who chooses among many possible

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

[18] For example, if the agent could instead choose the distribution of an additively separable noise term that affects output, then linear contracts would not necessarily be optimal.

projects. The entrepreneur can exert more effort to identify better projects, but he can also work less hard and choose a riskier project that succeeds wildly in some environments but fails miserably in others. The inherent riskiness of the project is then captured by the entrepreneur's choice of $G(\cdot)$, while $F(\cdot|x)$ represents residual uncertainty that remains even if the entrepreneur picks the "safest" project that he has identified.[19]

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

$$V_s(x) \equiv \int_{\underline{y}}^{\overline{y}} s(y) f(y|x)\, dy.$$

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

$$\begin{aligned} \max_{a, s(\cdot)} \quad & a - V_s^c(a) \\ \text{subject to} \quad & a \in \arg\max_{\tilde{a}} \left\{ V_s^c(\tilde{a}) - c(\tilde{a}) \right\}, \\ & V_s^c(a) - c(a) \geq u_0, \\ & s(\cdot) \geq -M. \end{aligned} \tag{6}$$

We prove that a linear contract solves this problem.

Proposition 7 (Ex ante risk-taking). *If $a^* \geq 0$ is optimal in the program (6), then $a^* \leq a^{\mathrm{FB}}$ and $s^{\mathrm{L}}_{a^*}(\cdot)$ is optimal.*

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

6.3 *Manipulating the timing of Output*[20]

In this section, we argue that risk-taking is very similar to another common form of gaming: manipulating *when* output is realized over time. To make this point, we consider 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. This model turns out to be equivalent to the setting in Section 4.

[19] If $\int_{\underline{y}}^{z} F_{xx}(y|x)\, dy \geq 0$ for all $z \in \mathcal{Y}$ and x, then a riskier $G(\cdot)$ leads to a riskier distribution over final output (in each case, in the sense of second-order stochastic dominance).

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

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

Move 1. The principal offers a stationary contract $s(y) : \mathcal{Y} \to [-M, \infty)$.

Move 2. The agent accepts or rejects. If he rejects, he earns u_0 and the principal earns 0.

Move 3. The agent chooses an effort $a \geq 0$.

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

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

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

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

Crucially, the principal must offer a stationary contract $s(\cdot)$ in this model. Without this restriction, the principal could eliminate gaming incentives entirely, for instance, by paying only for cumulative output at $t = 1$. While stationarity is a significant restriction, we believe it is realistic in many settings: as documented by Oyer (1998) and Larkin (2014), contracts tend to be stationary over some period of time (such as a quarter or a year).

This problem is equivalent to one in which, rather than choosing the realized output $y_x(t)$ at each time t, the agent instead decides *what fraction* of time in $t \in [0, 1]$ to spend producing each possible output $y \in \mathcal{Y}$. In particular, define $G_x(y)$ as the fraction of time for which $y_x(t) \leq y$.[21] Then $G_x(\cdot)$ is a distribution that satisfies $\mathbb{E}_{G_x}[y] = x$, and the agent's and principal's payoffs are $\mathbb{E}_{G_x}[s(y)] - c(a)$ and $\mathbb{E}_{G_x}[y - s(y)]$, respectively. That is, intertemporal gaming plays exactly the same role as gambling in our baseline model.

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

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

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

7. Concluding remarks

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

Nevertheless, our framework provides a natural starting point to consider why contracts might not deter risk-taking. Corollary 1 suggests one reason: the principal might be risk-seeking, for instance, because her own incentives are nonconcave. A second reason is implicit in our assumption that the principal can commit to an incentive scheme. Commitment might be difficult in some settings, for instance, because output can serve as the basis for future compensation (Chevalier and Ellison 1997, Makarov and Plantin 2015). More generally, an agent's competitive context shapes the incentives they face, which in turn determine the kinds of risks they optimally pursue; see Fang and Noe (2016) for a step in this direction. Our model provides a foundation on which to study the consequences of risk-taking behavior for markets, organizations, and society.

Appendix A: Proofs for Sections 3 and 4

For notational convenience, we use the indefinite integral to indicate an integral on $[\underline{y}, \overline{y}]$ in all of the appendices. Proofs are ordered based on where the corresponding results appear in the text. Some proofs depend on later results. We point out each of these dependencies as they arise; see footnotes 22 and 24.

A.1 *Proof of Proposition 1*

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

Fix $x \in \mathcal{Y}$. Since v is upper semicontinuous, there exist $p \in [0,1]$ 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

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

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

A.2 *Proof of Lemma 1*

Existence follows from Proposition 9 in Appendix D.[22] 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 semicontinuous and concave, they must differ on an interval of positive length. But then the contract $v^*(\cdot) \equiv \frac{1}{2}(v(\cdot) + \tilde{v}(\cdot))$ satisfies (IC_F)–(LL_F) for effort a, and the principal's payoff under v^* is

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

by Jensen's inequality, where at least one of the inequalities is strict. □

A.3 *Proof of Proposition 2*

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

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

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

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

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

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

$$\int \big(\hat{s}_{a^*}(y) - \big(s^*(y) + \hat{u}\big)\big) f\big(x|a^*\big)\, dx = 0 \tag{7}$$

[22] Proposition 9 is self-contained and thus presents no circularities.

and suppose to the contrary that $\hat{u} > 0$. Then, since $\hat{s}_{a^*}(\underline{y}) < s^*(\underline{y}) + \hat{u}$, and since $\hat{s}_{a^*}(\cdot)$ is linear and $s^*(\cdot) + \hat{u}$ is concave, there exists $\tilde{y} > \underline{y}$ such that $\hat{s}_{a^*}(\cdot) - (s^*(\cdot) + \hat{u})$ is strictly negative below $\tilde{y}$ and strictly positive above $\tilde{y}$. Hence, since $f_a(\cdot|a^*)/f(\cdot|a^*)$ is strictly increasing, by Beesack's inequality,[23] (7) implies that

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

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

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

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

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

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

$$\pi'(y - t) \geq \pi'\big(y - s^{\mathrm{L}}_{\hat{a}}(y)\big) \geq \pi'\big(\hat{y} - s^{\mathrm{L}}_{\hat{a}}(\hat{y})\big),$$

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

$$\pi'(y - t) \leq \pi'\big(y - s^{\mathrm{L}}_{\hat{a}}(y)\big) \leq \pi'\big(\hat{y} - s^{\mathrm{L}}_{\hat{a}}(\hat{y})\big),$$

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

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

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

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

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

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

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

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

A.4 *Proof of Corollary 1*

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

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

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

Appendix B: Proofs for Section 5

First we prove some preliminary properties of optimal incentives schemes. If $\underline{u} > -\infty$, Lemma 1 has shown that any profit-maximizing incentive scheme $v(\cdot)$ must be unique, and supplementary file on the journal website, http://econtheory.org/supp3660/supplement.pdf, show the same for $\underline{u} = -\infty$. We prove that $v(\cdot)$ must be monotonically increasing and satisfy (IC-FOC) with equality.

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

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

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

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

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

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

In Section B.1.2, we will carefully define the perturbations *raise* and *tilt* and show that they respect concavity.

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

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

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

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

PROOF. Note that for $y > y_H$, $t_{\hat{y},y_H}(y) = y_H - \hat{y} = (y_H - \hat{y}) r_{y_H,\overline{y}}(y)$. Since v satisfies (IC), since $a > 0$, and since v is concave and weakly increasing, v must be strictly increasing near $\underline{y}$. Hence, since $y_H > \underline{y}$, $v(y_H) > \underline{u}$. We thus have $\int n(y) r_{y_H,\overline{y}}(y) f(y|a)\, dy = 0$ by Definition 2(i). Hence, by Definition 2(iii), we have

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

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

Similarly, note that if $v(y_L) > \underline{u}$, then $\int n(y) r_{y_L,\overline{y}}(y) f(y|a)\, dy = 0$ by Definition 2(i), and so by Definition 2(ii),

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

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

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

B.1 *Proof of Proposition 3*

The discussion prior to the statement of Proposition 3 proves necessity, given well defined perturbations that satisfy concavity and well defined shadow values. This section begins by formally defining the relevant perturbations, showing that they preserve concavity, and then showing how they can be used to establish shadow values for (IR) and (IC-FOC). We then turn to sufficiency.[24]

B.1.1 *Preliminaries* Definition 2 and Proposition 3 are phrased in terms of free points. But not every free point is a convenient place to define a perturbation. Instead, for any given v, let C_v be the set of points y at which there exists a supporting plane L such that $L(y') > v(y')$ for all $y' \neq y$.

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

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

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

So fix $\delta > 0$. Since $\hat{y}$ is not on the interior of a linear segment and not a kink point, there is at least one side of $\hat{y}$, without loss of generality the right side, such that $v(\cdot)$ is not linear on $(\hat{y}, \hat{y} + \delta)$. Let $S(\cdot)$ be the correspondence that, for each y, assigns the set of slopes of supporting planes at y and let $s(\cdot)$ be any selection from $S(\cdot)$. Note that since v is concave, for any $y'' > y'$, $\max\{S(y'')\} \leq \min\{S(y')\}$ and, hence, s is decreasing. Assume first that there is a point $\tilde{y} \in (\hat{y}, \hat{y} + \delta)$, where $s(\cdot)$ jumps downward, say from s'' to $s' < s''$. Then the supporting plane at $\tilde{y}$ with slope $(s' + s'')/2$ qualifies. Assume instead that $s(\cdot)$ is continuous on $(\hat{y}, \hat{y} + \delta)$. It cannot be everywhere constant, since $v(\cdot)$ is not linear on $(\hat{y}, \hat{y} + \delta)$. Hence, since $s(\cdot)$ is continuous, there is a point $\tilde{y}$ at which it is strictly decreasing, so that, specifically, $s(\tilde{y}) < s(y)$ for all $y < \tilde{y}$ and $s(\tilde{y}) > s(y)$ for all $y > \tilde{y}$. The supporting plane at $\tilde{y}$ with slope $s(\tilde{y})$ then qualifies. □

[24]The proof of sufficiency uses Lemma 6 from Appendix D.3, which establishes the existence of an optimal contract when $\underline{u} = -\infty$. Again, it is easy to verify that this lemma is self-contained and, thus, presents no circularities.

To see that why Claim 1 is helpful, assume that some part of Definition 2 is violated. For example, assume some optimal contract has a pair of free points y_L and y_H such that $\int n(y) r_{y_L,y_H} f(y)\,dy < 0$. If either y_L or y_H is a kink point, then it is also an element of C_v. If not, then we can apply Claim 1 to replace each relevant point by a sufficiently close-by element of C_v that the strict inequality is maintained. Hence, it is enough to prove Proposition 3 when each restriction to a free point is tightened to a restriction to C_v.

B.1.2 *Formal definition and properties of the perturbations* This section defines *raise* and *tilt*, being careful, in particular, to maintain concavity at the endpoints of the perturbed interval. We need to consider as many as three perturbations at once, where, given the previous discussion, we require the relevant points to be in C_v. First, we have some small amount ε_p of a perturbation p, where p could be r_{y_L,y_H} or t_{y_L,y_H} in each case with ε_p positive or negative. Second, for some $\hat{y} \in C_v$, we need to consider some amount ε_t of $t_{\hat{y},\overline{y}}$ and ε_r of $r_{\hat{y},\overline{y}}$. Intuitively, we use $t_{\hat{y},\overline{y}}$ and $r_{\hat{y},\overline{y}}$ to establish shadow values for (IC-FOC) and (IR), and then, for any particular perturbation p, we consider the three deviations together where one uses $t_{\hat{y},\overline{y}}$ and $r_{\hat{y},\overline{y}}$ to undo the effect of p on (IC-FOC) and (IR).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\begin{aligned} \hat{v}_{\varepsilon_p}(\cdot; \mathbf{0}) &= p(\cdot), \\ \hat{v}_{\varepsilon_t}(\cdot; \mathbf{0}) &= t_{\hat{y}, \overline{y}}(\cdot), \\ \hat{v}_{\varepsilon_r}(\cdot; \mathbf{0}) &= r_{\hat{y}, \overline{y}}(\cdot). \end{aligned} \tag{8}$$

B.1.3 *Shadow values* We need to establish that starting from $\boldsymbol{\varepsilon} = \mathbf{0}$, the effects of perturbation p can be undone via $t_{\hat{y}, \overline{y}}$ and $r_{\hat{y}, \overline{y}}$. To do so, let

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

The top row of Q tracks the rate at which ε_t and ε_r, respectively, affect (IC-FOC), while the bottom row tracks the rate at which ε_t and ε_r, respectively, affect (IR). Then, from (8),

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

and so

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

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

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

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

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

Let λ be the rate of change of costs as one relaxes (IR) using $t_{\hat{y},\overline{y}}$ and $r_{\hat{y},\overline{y}}$. That is, if we let

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

then

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

Similarly, if

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

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

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

Given the shadow values λ and μ, the argument in Section 5 (prior to Definition 2) completes the proof of necessity in Proposition 3. □

B.1.4 *Proof of sufficiency* We begin by proving the following useful result.

LEMMA 4. *Let $v(\cdot)$ be GHM and suppose $y \in (\underline{y}, \overline{y})$ is free. Then $n(y) \leq 0$, and $n(y) = 0$ if y is a point of normal concavity (as defined immediately before Corollary 2).*

PROOF. If y is a kink point, then Lemma 3 applied to the left of y implies that $n(y) \leq 0$. If y is a point of normal concavity, then by Lemma 1, there exist sequences of points $\{y_k^{\mathrm{L}}\}, \{y_k^H\} \in C_v$ such that $y_k^{\mathrm{L}} < y < y_k^H$ for all $k \in \mathbb{N}$ and $\lim_k y_k^{\mathrm{L}} = \lim_k y_k^H = y$. These points are free, so (2) holds with equality on each interval $[y_k^{\mathrm{L}}, y_k^H]$. Hence, in the limit, $n(y) = 0$. □

Now let v, with associated λ and μ, be GHM. Let us show that v is optimal. We argue by contradiction. Assume v is not optimal, and let v^* be a lower-cost contract satisfying (IC-FOC) and (IR)–(NG). As in the argument at the beginning of Appendix B, v^* can be taken to be increasing and to satisfy (IC-FOC) exactly, and as in the proof of Lemma 6 in Appendix D.3, $v^*(\overline{y})$ and $v^*(\underline{y})$ can be taken to be finite.

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

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

it follows that

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

and so, since every point in $\mathcal{Y} \setminus S$ 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

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

where the first equality follows by Lemma 4.

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

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

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

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

Making these two substitutions thus yields

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

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

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

or, equivalently,

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

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

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

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

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

Note that y_H and $\overline{y}$ are free. Note also that as in the proof of Lemma 3, $v(y_H) > \underline{u}$. It follows from Definition 2(i) that since $\tilde{\delta}$ is constant on $[y_H, \overline{y}]$,

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

and, hence,

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

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

$$\tilde{\delta}(y) = \delta(y_0)r_{y_0,\overline{y}}(y) + \sum_{k=1}^{2^{\hat{K}}-1}(s_k - s_{k+1})t_{y_0,y_k}(y) + s_{2^{\hat{K}}}t_{y_0,y_{2^{\hat{K}}}}(y). \tag{9}$$

To see (9), note first that for $y < y_0 = y_L$, both sides of the equation are 0. At y_0, each side is $\delta(y_0)$, since $r_{y_0,\overline{y}}(y_0) = 1$ and since $t_{y_0,y_k}(y_0) = 0$ for all k. Thus, since both sides are continuous and piecewise linear on $[y_0, \overline{y}]$, it is enough that the two sides have that same derivative where defined. So fix $\hat{k} \in \{1, \dots, 2^{\hat{K}}\}$ and let $y \in (y_{\hat{k}-1}, y_{\hat{k}})$. Note that for $k < \hat{k}$, $t'_{y_0,y_k}(y) = 0$, and for $k \ge \hat{k}$, $t'_{y_0,y_k}(y) = 1$. Hence, the derivative of the right-hand side is

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

as desired, and so, noting that $\tilde{\delta}'(y) = 0$ for $y > y_K = y_H$, we have established (9).

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

(i) $\delta(y_0)\int n(y)r_{y_0,\overline{y}}(y)f(y|a)\,dy < 0$;

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

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

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

B.2 *Proof of Corollary 2*

This result follows immediately from Proposition 3 and Lemma 4.

B.3 *Proof of Proposition 4*

Suppose that there exists some $y_I \in [\underline{y}, \overline{y}]$ such that $\rho(\lambda + \mu l(\cdot|a))$ is convex on $[\underline{y}, y_I]$ and concave on $[y_I, \overline{y}]$, let $v^*(\cdot)$ implement $a \geq 0$ at maximum profit, and suppose $v^*(\underline{y}) > \underline{u}$. Since $v^*(\cdot)$ is increasing, (LL) must be slack.

First, we show that $v^*(\cdot)$ has no more than one linear segment. Since $v^*(\cdot)$ implements a at maximum profit, it is GHM by Proposition 3. Consequently, if $v^*(\cdot)$ has more than one linear segment, then Lemma 3 implies that $n(\cdot)$ must be positive, then negative, then positive over each segment. Hence, $v^*(\cdot) - \rho(\lambda + \mu l(\cdot|a))$ must be negative, then positive, then negative over each linear segment. But then $\rho(\lambda + \mu l(\cdot|a))$ must have two disjoint nonconcave regions, which is ruled out by assumption.

If $y_I = \underline{y}$, then $v^*(\cdot)$ cannot have any linear segments, since on any such segment, $v^*(\cdot) - \rho(\lambda + \mu l(\cdot|a))$ would be positive, then negative, then positive. But then any interior free point must be a point of normal concavity, and so Corollary 2 implies that $v^*(\cdot) = \rho(\lambda + \mu l(\cdot|a))$ over $(\underline{y}, \overline{y})$.

If $y_I > \underline{y}$, then $v^*(\cdot)$ must have a linear segment because it cannot coincide with $\rho(\lambda + \mu l(\cdot|a))$ everywhere. We claim that this linear segment must be $[\underline{y}, \hat{y}]$ for some $\hat{y} \geq y_I$. If the linear segment starts at some $\tilde{y} > \underline{y}$, then every $y \in (\underline{y}, \tilde{y})$ must be a point of normal concavity. But then $v^*(\cdot) = \rho(\lambda + \mu l(\cdot|a))$ on $(\underline{y}, \tilde{y})$, which violates (NG) because $\rho(\lambda + \mu l(\cdot|a))$ is convex on that region by assumption. Similarly, if $\hat{y} < y_I$, then every $y \in (\hat{y}, y_I)$ must be a point of normal concavity, which again violates (NG). So $v^*(\cdot)$ has a single linear segment $[\underline{y}, \hat{y}]$, where $\hat{y} \geq y_I$. Since $v^*(\cdot)$ is GHM and $v^*(\underline{y}) > \underline{u}$, (2) holds with equality on this linear segment and so $\int_{\underline{y}}^{\hat{y}} n(y) f(y)\, dy = 0$.

Finally, any $y \in (\hat{y}, \overline{y})$ is again a point of normal concavity, and so $v^*(\cdot) = \rho(\lambda + \mu l(\cdot|a))$ at all such points. This proves the result. □

B.4 *Proof of Proposition 5*

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

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

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

$\tilde{v}(\cdot)$ satisfies (IR). Furthermore,

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

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

Appendix C: Proofs for Section 6

C.1 *Proof of Proposition 6*

Given the definition of $\tilde{v}(\cdot)$, $\tilde{c}$, and $\tilde{\pi}$, the optimal a and $\tilde{v}(\cdot)$ solve

$$\max_{a, G\in\mathcal{G}, \tilde{v}(\cdot)} \quad \mathbb{E}_{F(\cdot|a)}\big[\mathbb{E}_{G_x}\big[\tilde{\pi}(y) - \tilde{v}(y)\big]\big] \tag{10}$$
$$\text{subject to} \quad a, G \in \arg\max_{\tilde{a}, \tilde{G}\in\mathcal{G}} \big\{\mathbb{E}_{F(\cdot|\tilde{a})}\big[\mathbb{E}_{\tilde{G}_x}\big[\tilde{v}(y)\big]\big] - \tilde{c}(\tilde{a})\big\},$$
$$\mathbb{E}_{F(\cdot|a)}\big[\mathbb{E}_{G_x}\big[\tilde{v}(y)\big]\big] - \tilde{c}(a) \geq u_0,$$
$$\tilde{v}(y) \geq -M - d(y) \quad \forall y \in \mathcal{Y}.$$

As in Proposition 1, following any intermediate output x, the agent optimally chooses G_x so that $\mathbb{E}_{G_x}[\tilde{v}(x)] = \tilde{v}^c(x)$, where $\tilde{v}^c(\cdot)$ is the concave closure of $\tilde{v}(\cdot)$. Therefore, the principal's payoff following x equals $\mathbb{E}_{G_x}[\tilde{\pi}(y) - \tilde{v}(y)] \leq \tilde{\pi}(x) - \tilde{v}^c(x)$. Since $\tilde{\pi}(\cdot)$ is strictly concave, this inequality holds with equality only if G_x is degenerate. Consequently, we can restrict attention to contracts in which $\tilde{v}(\cdot)$ is concave, and, hence, for every x, the agent will optimally choose $G_x(y) = \mathbb{I}_{\{y\geq x\}}$.

Relax the limited liability constraint so that it must be satisfied only at $y = \underline{y}$. Then (10) can be written as

$$\max_{a, \tilde{v}(\cdot)} \quad \mathbb{E}_{F(\cdot|a)}\big[\tilde{\pi}(y) - \tilde{v}(y)\big]$$
$$\text{subject to} \quad a \in \arg\max_{\tilde{a}} \big\{\mathbb{E}_{F(\cdot|\tilde{a})}\big[\tilde{v}(y)\big] - \tilde{c}(\tilde{a})\big\},$$
$$\mathbb{E}_{F(\cdot|a)}\big[\tilde{v}(y)\big] - \tilde{c}(a) \geq u_0,$$
$$\tilde{v}(\underline{y}) \geq -M,$$
$$\tilde{v}(\cdot) \text{ is concave.}$$

Fix any effort $a \geq 0$ and any concave incentive scheme $\tilde{v}(\cdot)$ that implements a. As in the proof of Proposition 2, let $\tilde{v}^{\mathrm{L}}(\cdot)$ be the unique linear incentive scheme that satisfies $\tilde{v}^{\mathrm{L}}(\underline{y}) = \tilde{v}(\underline{y})$ and $\mathbb{E}_{F(\cdot|a)}[\tilde{v}^{\mathrm{L}}(y)] = \mathbb{E}_{F(\cdot|a)}[\tilde{v}(y)]$. Then $\tilde{v}^{\mathrm{L}}(\cdot) - \tilde{v}(\cdot)$ single-crosses 0 from

below and, hence, Beesack's inequality implies

$$\int \big(\tilde{v}^{\mathrm{L}}(y) - \tilde{v}(y)\big)\frac{f_a(y|a)}{f(y|a)}f(y|a)\,dx \geq 0$$

with strict inequality if $\tilde{v}^{\mathrm{L}}(y) \neq \tilde{v}(y)$ for some y. Consequently, $\tilde{v}^{\mathrm{L}}(\cdot)$ implements some $\tilde{a} \geq a$, with $\tilde{a} > a$ if $\tilde{v}^{\mathrm{L}}(y) \neq \tilde{v}(y)$ for some y.

Define $\tilde{v}^*(y) = \tilde{c}'(a)(y - \underline{y}) - \tilde{w}$, where $\tilde{w} = \min\{M, \tilde{c}'(a)(a - \underline{y}) - \tilde{c}(a) - u_0\}$, and suppose that $\tilde{v}^*(\underline{y}) = -M$. Then $\tilde{v}^*(y) \leq \tilde{v}^{\mathrm{L}}(y)$ for all $y \geq \underline{y}$ and strictly so if $\tilde{a} > a$. Therefore, $\tilde{v}^*(\cdot)$ uniquely implements $a \geq 0$ at maximum profit in the relaxed problem. But $\tilde{v}^*(y) \geq -M \geq -M - d(y)$ for all $y \in \mathcal{Y}$, so $\tilde{v}^*(\cdot)$ satisfies the limited liability constraint and, hence, implements a in the original problem.

Next suppose that $\tilde{v}^*(\underline{y}) > -M$. Then by construction, $\mathbb{E}_{F(\cdot|a)}[\tilde{v}^*(y)] = u_0 + \tilde{c}(a) \leq \mathbb{E}_{F(\cdot|a)}[\tilde{v}(y)]$, which implies that $\mathbb{E}_{F(\cdot|a)}[\tilde{\pi}(y) - \tilde{v}^*(y)] \geq \mathbb{E}_{F(\cdot|a)}[\tilde{\pi}(y) - \tilde{v}(y)]$; i.e., $\tilde{v}^*(\cdot)$ implements a at maximum profit.

Finally, note that the preceding holds for any $a \geq 0$, proving that $\tilde{v}^*(\cdot)$ or, equivalently, $s^*(y) = \tilde{c}'(a)(y - \underline{y}) - d(y) - \tilde{w}$, is optimal. □

C.2 *Proof of Proposition 7*

Since $s(\cdot) \geq -M$, $V_s(x) = \int s(y)f(y|x)\,dy \geq -M$ and so $V_s^c(\cdot) \geq -M$. Consider relaxing (6) so that the principal can choose any $V_s(\cdot)$ that is concave and satisfies $V_s(\cdot) \geq -M$. In this relaxed problem, the principal solves

$$\begin{aligned}
\max_{a, V_s(\cdot)} \quad & a - V_s(a) \\
\text{subject to} \quad & a \in \arg\max_{\tilde{a}} \big\{V_s(\tilde{a}) - c(\tilde{a})\big\}, \\
& V_s(a) - c(a) \geq u_0, \\
& V_s(y) \geq -M \quad \text{for all } y \in \mathcal{Y}, \\
& V_s(\cdot) \text{ is weakly concave.}
\end{aligned}$$

This problem is identical to (Obj)–(NG) with a degenerate distribution over intermediate output.

Suppose $(a^*, V_s(\cdot))$ is optimal in this relaxed program. Note that $s^{\mathrm{L}}_{a^*}(\cdot)$ is feasible in this relaxed problem, so $V_s(a^*) \leq s^{\mathrm{L}}_{a^*}(a^*)$. Suppose $s^{\mathrm{L}}_{a^*}(\cdot)$ is not optimal, so $V_s(a^*) < s^{\mathrm{L}}_{a^*}(a^*)$. Then $s^{\mathrm{L}}_{a^*}(a^*) - c(a^*) > u_0$ and so $s^{\mathrm{L}}_{a^*}(\underline{y}) = -M$. Define $s^{\mathrm{L}}(\cdot)$ as the linear function that intersects $V_s(\cdot)$ at $\underline{y}$ and a^*, so

$$s^{\mathrm{L}}(y) = V_s(\underline{y}) + \frac{V_s(a^*) - V_s(\underline{y})}{a^* - \underline{y}}(y - \underline{y}).$$

Since $V_s(a^*)$ is concave, $s^{\mathrm{L}}(y) \leq V_s(y)$ for all $y \in [\underline{y}, a^*]$.

For the agent to be willing to choose a^* under $V_s(\cdot)$, it must be that $\partial^- V_s(a^*) \geq c'(a^*)$, where $\partial^- V_s(y)$ is the left derivative of $V_s(\cdot)$ at y. Since V_s is concave,

$$\frac{V_s(a^*) - V_s(\underline{y})}{a^* - \underline{y}} \geq \partial^- V_s(y) \geq c'(a^*).$$

Since $V_s(\underline{y}) \geq M$, we conclude that $s^{\mathrm{L}}(y) \geq s^{\mathrm{L}}_{a^*}(y)$ for all $y \in \mathcal{Y}$. But then $V_s(a^*) = s^{\mathrm{L}}(a^*) \geq s^{\mathrm{L}}_{a^*}(a^*)$, which gives a contradiction. So $(a^*, s^{\mathrm{L}}_{a^*}(\cdot))$ is also optimal. Note that for any $a^* > a^{\mathrm{FB}}$, $(a^*, s^{\mathrm{L}}_{a^*}(\cdot))$ is strictly dominated by $(a^{\mathrm{FB}}, s^{\mathrm{L}}_{a^{\mathrm{FB}}}(\cdot))$, which generates higher total surplus and gives a (weakly) lower payment to the agent. So $a^* \leq a^{\mathrm{FB}}$ and $s^{\mathrm{L}}_{a^*}(\cdot)$ is optimal in this relaxed problem.

Finally, note that for any $a \geq 0$ and $x \in \mathcal{Y}$, $V_{s^{\mathrm{L}}_a}(x) = \mathbb{E}_{F(\cdot|x)}[s^{\mathrm{L}}_a(y)] = s^{\mathrm{L}}_a(x)$ because $\mathbb{E}_{F(\cdot|x)}[y] = x$. But then $V^c_{s^{\mathrm{L}}_a}(a) = V_{s^{\mathrm{L}}_a}(a) = s^{\mathrm{L}}_a(a)$, and so the optimal linear $V_s(\cdot)$ in the relaxed problem can be implemented in the full problem by $s^{\mathrm{L}}_{a^*}(\cdot)$. □

C.3 *Proof of Proposition 8*

It suffices to prove that for any total output x,

$$\max_{y_x:[0,1]\to[\underline{y},\overline{y}]} \left\{ \int_0^1 s\big(y_x(t)\big)\,dt \text{ subject to } \int_0^1 y_x(t)\,dt = x \right\} = s^c(x).$$

Consider the following y_x: if $s(x) = s^c(x)$, then $y_x(t) = x$ for all t. If $s(x) < s^c(x)$, then there exist w, z, and $\alpha \in [0,1]$ such that $\alpha w + (1-\alpha)z = x$ and $\alpha s(w) + (1-\alpha)s(z) = s^c(x)$. For $t \leq \alpha$, $y_x(t) = w$, with $y_x(t) = z$ for $t > \alpha$. This function y_x guarantees that the agent earns $s^c(x)$.

Now $s(y_x(t)) \leq s^c(y_x(t))$ for all $y_x(t)$. Since s^c is weakly concave and $\int_0^1 y_x(t)\,dt = x$, we conclude that $\int_0^1 s(y_x(t))\,dt \leq \int_0^1 s^c(y_x(t))\,dt \leq \int_0^1 s^c(x)\,dt = s^c(x)$. So the agent earns (and the principal pays) $s^c(x)$ following intermediate output x, which proves the claim. □

Appendix D: Additional results

The first part of this section proves existence and some properties of the optimal contract for the case of a finite limited liability constraint. The second part gives sufficient conditions on ρ and l for Proposition 4. The final part proves a result about how the optimal contract varies in $\underline{u}$ that we use in Appendix B.

D.1 *Proof of existence, uniqueness, and continuity for $\underline{u}$ finite*

Proposition 9. *Let U and Π be the set of increasing concave utility functions for the agent and the principal satisfying our assumptions, and let V be the set of concave (but not necessarily increasing) functions from $[\underline{y}, \overline{y}]$ to $\mathbb{R}$, where each of U, Π, and V has the topology of almost everywhere pointwise convergence. Fix a. Then (i) for each $z = (M, u_0, \pi, u)$, there exists an optimal contract v that implements a given z and (ii) at any point z where*

at least one of π or u is strictly concave, the optimal contract implementing a is unique and continuous in z.

PROOF. The proof relies on Berge's theorem. Fix a. For any given $z = (M, u_0, u, \pi)$, let $v^{\mathrm{L}}(\cdot|z)$ be given by $v^{\mathrm{L}}(y|z) = c'(a)(y - \underline{y}) + \beta$, where $\beta = \min(u(-M), c(a) + u_0 - c'(a)(a - \underline{y}))$, be the maximum-profit linear (in utils) contract that implements a. In particular, $v^{\mathrm{L}}(\cdot|z)$ satisfies (IC) since, under our assumptions, the agent's utility from income given $v^{\mathrm{L}}(\cdot|z)$ is linear in effort while $-c(\cdot)$ is concave and so the first-order condition implies (IC)

Let $B : \mathbb{R} \times \mathbb{R} \times \Pi \times U \to\!\!\to V$ be the correspondence that for each $M \in \mathbb{R}$, $u_0 \in \mathbb{R}$, $\pi \in \Pi$, and $u \in U$ gives the set of contracts v such that

$$\begin{aligned} E_{F(\cdot|a)}\big[\pi\big(y - u^{-1}\big(v(y)\big)\big)\big] &\geq E_{F(\cdot|a)}\big[\pi\big(y - u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big)\big] - 1, \qquad (11)\\ a &\in \arg\max_{\tilde{a}} \big\{E_{F(\cdot|\tilde{a})}\big[v(y) - c(\tilde{a})\big]\big\},\\ E_{F(\cdot|a)}\big[v(y) - c(a)\big] &\geq u_0,\\ v(\underline{y}) &\geq u(-M),\\ v &\in V, \end{aligned}$$

where the second through fifth constraints are simply the translations of (IC)–(NG) when z is a parameter, and the first constraint restricts attention to contracts that come within 1 util for the principal of $v^{\mathrm{L}}(\cdot|z)$. Since $v^{\mathrm{L}}(\cdot|z) \in B(z)$, this constraint is innocuous, and it also follows that B is non-empty-valued.

For any given $v \in V$, define $v_{\max} = \max_{y \in [\underline{y}, \overline{y}]} v(y)$. We begin by proving the following statement.

(*) For each compact subset $Z \subseteq \mathbb{R} \times \mathbb{R} \times \Pi \times U$, there is $\overline{u}$ such that $v_{\max} \leq \overline{u}$ for all $z \in Z$ and $v \in B(z)$.

To see (*), begin by noting that $v^{\mathrm{L}}(\cdot|\cdot)$ is continuous on the compact set $[\underline{y}, \overline{y}] \times Z$, and so $-\infty < m \equiv \min_{y \in [\underline{y}, \overline{y}] \times Z} \{\pi(y - u^{-1}(v^{\mathrm{L}}(y|z)))\}$. Using that Z is compact, let $u^* < \infty$ satisfy that for all $z \in Z$, $\pi(\overline{y} - u^{-1}(u^*)) \leq m - 2$, so that any time the principal gives the agent utility u^* or above, the principal is at least 2 utils worse off than under $v^{\mathrm{L}}(\cdot|z)$.

Fix $z \in Z$ and $v \in B(z)$. Choose $y_{\max}$ so that $v(y_{\max}) = v_{\max}$. Let $u_{\min} = \min_{z \in Z} u(-M)$, and define $\hat{v}$ as the function that equals $u_{\min}$ at $\underline{y}$ and $\overline{y}$, equals $v_{\max}$ at $y_{\max}$, and is linear to the left and right of $y_{\max}$. That is, $\hat{v}(y_{\max}) = y_{\max,}$ and

$$\hat{v}(y) = \left\{ \begin{array}{ll} u_{\min} + \dfrac{v_{\max} - u_{\min}}{y_{\max} - \underline{y}}(y - \underline{y}), & y \in [\underline{y}, y_{\max}) \\ u_{\min} + \dfrac{v_{\max} - u_{\min}}{\overline{y} - y_{\max}}(\overline{y} - y), & y \in (y_{\max}, \overline{y}] \end{array} \right\}.$$

Note that

$$E_{F(\cdot|a)}\big[\pi\big(y - u^{-1}\big(\hat{v}(y)\big)\big)\big] \geq E_{F(\cdot|a)}\big[\pi\big(y - u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big)\big] - 1, \qquad (12)$$

using that the concave function v is everywhere at or above $\hat{v}$ and the first constraint in (11).

We show that (12) implies a uniform bound on $v_{\max}$. Intuitively, when $v_{\max}$ is large, the piecewise linear function $\hat{v}(y)$ is above u^* for nearly all of $[\underline{y}, \overline{y}]$, implying losses compared to $v^{\mathrm{L}}(\cdot|z)$ that contradict (12).

A uniform bound on $v_{\max}$ is, of course, trivial for v such that $v_{\max} \leq u^*$. So assume $v_{\max} > u^*$. Let $y_L \in [\underline{y}, y_{\max})$ solve $\hat{v}(y_L) = u^*$, where if $y_{\max} = \underline{y}$, we let $y_L = \underline{y}$ and, similarly, define $y_H \in (y_{\max}, \overline{y}]$ by $\hat{v}(y_H) = u^*$, where if $y_{\max} = \overline{y}$, $y_H = \overline{y}$.

Since $\hat{v}(\cdot)$ is concave, $\hat{v}(y) \geq u^*$ for all $y \in [y_L, y_H]$ and, hence,

$$\pi\big(y - u^{-1}\big(\hat{v}(y)\big)\big) - \pi\big(y - u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big) \leq -2,$$

while for any y,

$$\pi\big(y - u^{-1}\big(\hat{v}(y)\big)\big) - \pi\big(y - u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big) \leq b,$$

where $b \equiv \pi(\overline{y} + \max_{z \in Z} M) - m$. So from (12) we must have

$$\big(F(y_H|a) - F(y_L|a)\big)(-2) + \big(1 - \big(F(y_H|a) - F(y_L|a)\big)\big)b \geq -1$$

or, equivalently,

$$F(y_H|a) - F(y_L|a) \leq \frac{1+b}{2+b}, \tag{13}$$

where the right-hand side is strictly less than 1 because $\infty > b > 0$. But if $y_L \neq \underline{y}$, then

$$\begin{aligned} y_L &= \underline{y} + \frac{u^* - u_{\min}}{v_{\max} - u_{\min}}(y_{\max} - \underline{y}) \\ &\leq \underline{y} + \frac{u^* - u_{\min}}{v_{\max} - u_{\min}}(\overline{y} - \underline{y}), \end{aligned}$$

and so as $v_{\max} \to \infty$, $y_L \to \underline{y}$. Similarly, if $y_H \neq \overline{y}$, then

$$\begin{aligned} y_H &= \overline{y} - \frac{u^* - u_{\min}}{v_{\max} - u_{\min}}(\overline{y} - y_{\max}) \\ &\geq \overline{y} - \frac{u^* - u_{\min}}{v_{\max} - u_{\min}}(\overline{y} - \underline{y}), \end{aligned}$$

and so as $v_{\max} \to \infty$, $y_H \to \overline{y}$. But then by (13), $v_{\max}$ is bounded, establishing (*).

From (*) and the dominated convergence theorem, each expectation in (11) is continuous in z, and, hence, noting that each of (IC) and (NG) can be expressed as a collection of weak inequalities, $B(\cdot)$ is upper hemicontinuous.

Let us next show that $B(\cdot)$ is lower hemicontinuous. To see this, fix z, let $v \in B(z)$, and let $z_k \to z$. For $\varepsilon \in (0,1)$ and $\delta > 0$, define $\tilde{v}(\cdot|\varepsilon, \delta)$ by $\tilde{v}(y|\varepsilon, \delta) = (1-\varepsilon)v(y) + \varepsilon v^{\mathrm{L}}(y|M - \delta, u_0 + \delta, u, \pi)$.

By Jensen's inequality, for each y,

$$y - u^{-1}\big((1-\varepsilon)v(y) + \varepsilon v^{\mathrm{L}}(y|z)\big) \geq (1-\varepsilon)\big(y - u^{-1}\big(v(y)\big)\big) + \varepsilon\big(y - u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big),$$

since $-u^{-1}$ is concave. Hence, since π is increasing and concave,

$$\begin{aligned}\pi\big(y-u^{-1}\big((1-\varepsilon)v(y)+\varepsilon v^{\mathrm{L}}(y|z)\big)\big) &\geq \pi\big((1-\varepsilon)\big(y-u^{-1}\big(v(y)\big)\big)+\varepsilon\big(y-u^{-1}v^{\mathrm{L}}(y|z)\big)\big)\\ &\geq (1-\varepsilon)\pi\big(y-u^{-1}\big(v(y)\big)\big)+\varepsilon\pi\big(y-u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big),\end{aligned}$$

and so the same is true in expectation. Since the first constraint in (11) holds weakly for $v(\cdot)$ and strictly for $v^{\mathrm{L}}(\cdot|z)$, we have that for each $\varepsilon\in(0,1)$,

$$E_{F(\cdot|a)}\big[\pi\big(y-u^{-1}\big((1-\varepsilon)v(y)+\varepsilon v^{\mathrm{L}}(y|z)\big)\big)\big]>E_{F(\cdot|a)}\big[\pi\big(y-u^{-1}\big(v^{\mathrm{L}}(y|z)\big)\big)\big]-1.$$

It follows from continuity that for each $n\in\{1,2,\ldots\}$ there exists $\frac{1}{n}>\delta_n>0$ sufficiently small that the first constraint in (11) is slack for $v_n\equiv\tilde{v}(\cdot|\frac{1}{n},\delta_n)$.

It is immediate that the third and fourth constraints in (11) hold strictly at v_n, while the second and fifth constraints (which do not depend on z) continue to hold, since v_n is a convex combination of concave contracts satisfying (IC). Hence, for each n, there is K_n such that for all $k\geq K_n$, $v_n\in B(z_k)$. Let $k_n=\max\{n,\max_{n'\leq n}K_n\}$. Then, for each n, $v_{k_n}\in B(z_{k_n})$, and since $k_n\to\infty$ and $\delta_n\to 0$, $v_{k_n}\to v$. Hence, $B(\cdot)$ is lower hemicontinuous, and thus continuous.

Fix z and let $\{v_k\}$ be a sequence in $B(z)$. Since each v_k is concave and thus has variation at most $2(\overline{u}-u(-M))$, it follows from Helly's selection theorem that $\{v_k\}$ has a convergent subsequence. Thus, B is compact-valued and, from Berge's theorem, the set of maximizers of $E_{F(\cdot|a)}[\pi(y-u^{-1}(v(y)))]$ on $B(\cdot)$ is non-empty-valued and upper hemicontinuous.

Finally, consider any z where at least one of π and u is strictly concave. Then if $v_1,v_2\in B(z)$, it is direct that $(v_1+v_2)/2\in B(z)$ is strictly more profitable than either v_1 or v_2. Thus, the maximum is unique and, hence, continuous in z. □

D.2 *Mild sufficient conditions for Proposition 4*

This appendix gives sufficient conditions under which $\rho(\lambda+\mu l(\cdot|a))$ is first convex and then concave. We show that this case obtains if $\mathrm{con}(\rho')+\mathrm{con}(l_y)>-1$, where for an interval $X\subseteq\mathbb{R}$ and analytic function $h:X\to\mathbb{R}_+$, $\mathrm{con}(h)=\inf_X\{1-(hh'')/(h')^2\}$. For any analytic function q with domain a subset of the reals, let $q^{(k)}$ be the kth derivative of q.

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

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

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

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

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

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

Using this lemma, we can prove the following claim, from which our sufficient condition is immediate.

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

Proof. Let

$$\theta(\cdot) = \big(g\big(h(\cdot)\big)\big)'' = g''\big(h'\big)^2 + g'h''. \tag{14}$$

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

To see this, note that

$$\theta' = g'''\big(h'\big)^3 + 3g''h'h'' + g'h'''.$$

Consider any point $\hat{y}$ at which $\theta = 0$. Consider first the case that $g''(\hat{y})h''(\hat{y}) \neq 0$. Then, since $g' > 0$, it follows by (14) that $g''(\hat{y})$ and $h''(\hat{y})$ have opposite sign. Hence, $g''(\hat{y})h''(\hat{y})h'(\hat{y}) < 0$ and so, evaluated at $\hat{y}$,

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

where in the second line we substitute for $(h')^2$ in the first term using (14) and that $\theta(\hat{y}) = 0$, and similarly for g' in the third term. Hence, θ is negative on an interval to the right of $\hat{y}$.

Assume instead that $g''(\hat{y})h''(\hat{y}) = 0$, where, since $\theta(\hat{y}) = 0$, it follows that $g''(\hat{y}) = h''(\hat{y}) = 0$. Thus, since $\operatorname{con}(g') > -\infty$, it follows from Lemma 5 applied to $q = g'$ that the

first nonzero derivative of g' is strictly negative and similarly for h'. But then the first nonzero derivative of θ will be of the form $g^{(k)}(h')^k + g'h^{(k)}$, with $k \geq 3$, and at least one term strictly negative, and so, taking a Taylor expansion, θ is strictly negative on an interval to the right of $\hat{y}$ and we are done. □

D.3 *Stability of optimal contract as $\underline{u}$ decreases*

This appendix shows that if $v_{\underline{u}}(\cdot)$ is an optimal contract for some limited liability constraint $\underline{u}$ and $v_{\underline{u}}(\underline{y}) > \underline{u}$, then $v_{\underline{u}}(\cdot)$ remains optimal in the problem with any less binding limited liability constraint $\underline{u}'$, including $\underline{u}' = -\infty$.

LEMMA 6. *Assume that for some $\underline{u} > -\infty$, $v_{\underline{u}}(\underline{y}) > \underline{u}$. Let $\underline{u}' < \underline{u}$. Then $v_{\underline{u}'} = v_{\underline{u}}$.*

PROOF. Assume $v_{\underline{u}}$ has $v_{\underline{u}}(\underline{y}) > \underline{u}$, but that when the limited liability constraint is some $\underline{u}' < \underline{u}$, there exists a superior concave contract $\hat{v}$ that implements a. We show that this leads to a contradiction.

Assume first that $\hat{v}(\underline{y}) > -\infty$ (as is automatic if $\underline{u}'$ is finite). Then, for small enough ε, the contract $(1-\varepsilon)v_{\underline{u}}(\cdot) + \varepsilon\hat{v}(\cdot)$ is both strictly cheaper than $v_{\underline{u}}$ (since u is strictly concave) and implements a subject to limited liability constraint $\underline{u}$, yielding the desired contradiction.

Assume instead that $\hat{v}(\underline{y}) = -\infty$. Begin by picking any point $x' > \underline{y}$ where $x' \in C_{\hat{v}}$ (since $\hat{v}(\underline{y}) = -\infty$, such points exist) and construct $\tilde{v}$ by applying a sufficiently small positive amount of $t_{x',\overline{y}}$ such that $\tilde{v}$ remains strictly cheaper than $v_{\underline{u}}$. Since this adds a positive increasing function to $\hat{v}$, both (IC-FOC) and (IR) are strictly slack at $\tilde{v}$.

For each $y \in [\underline{y}, \overline{y}]$, let $h_y(\cdot)$ be a supporting plane to $\tilde{v}$ at y. Let the concave contract $v_y(\cdot)$ be given by $v_y(x) = \tilde{v}(x)$ for $x > y$ and by $v_y(x) = h_y(x)$ for $x \leq y$. For each x, $v_y(x)$ is weakly decreasing in y, with $\lim_{y\to\underline{y}} v_y(x) = \tilde{v}(x)$. Thus, by the monotone convergence theorem, as $y \to \underline{y}$, $\int v_y(x)f(x|a)\,dx \to \int \tilde{v}(x)f(x|a)\,dx$, $\int v_y(x)f_a(x|a)\,dx \to \int \tilde{v}(x)f_a(x|a)\,dx$, and $\int u^{-1}(v_y(x))f(x|a)\,dx \to \int u^{-1}(\tilde{v}(x))f(x|a)\,dx$. Hence, for y close enough to $\underline{y}$, v_y implements a and is cheaper than $v_{\underline{u}}$. For any such y, $v_y(\underline{y})$ is finite and we are back to the previous case. □

REFERENCES

Antić, Nemanja (2016), "Contracting with unknown technologies." Unpublished paper, Princeton University. [719]

Beesack, Paul R. (1957), "A note on an integral inequality." *Proceedings of the Americal Mathematical Society*, 8, 875–879. [740]

Borell, Christer (1975), "Convex set functions in d-space." *Periodica Mathematica Hungarica*, 6, 111–136. [726]

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni Michaely (2005), "Payout policy in the 21st century." *Journal of Financial Economics*, 77, 483–527. [737]

Brown, Keith C., W. V. Harlow, and Laura T. Starks (1996), "Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry." *Journal of Finance*, 51, 85–110. [716]

Carroll, Gabriel (2015), "Robustness and linear contracts." *American Economic Review*, 105, 536–563. [719]

Chade, Hector and Jeroen Swinkels (2016), "The no-upward-crossing condition and the moral hazard problem." Unpublished paper, Northwestern University. [725]

Chassang, Sylvain (2013), "Calibrated incentive contracts." *Econometrica*, 81, 1935–1971. [719]

Chevalier, Judith and Glenn Ellison (1997), "Risk taking by mutual funds as a response to incentives." *Journal of Political Economy*, 105, 1167–1200. [716, 721, 738]

de Figueiredo, Rui, Evan Rawley, and Orie Shelef (2015), "Bad bets: Excessive risk-taking, convex incentives, and performance." Unpublished paper. [716]

DeMarzo, Peter M., Dmitry Livdan, and Alexei Tchistyi (2013), "Risking other people's money: Gambling, limited liability, and optimal incentives." Unpublished paper. [719]

Diamond, Peter (1998), "Managerial incentives: On the near linearity of optimal compensation." *Journal of Political Economy*, 106, 931–957. [718, 719]

Ederer, Florian, Richard Holden, and Margaret Meyer (2018), "Gaming and strategic opacity in incentive provision." *RAND Journal of Economics*, 49, 819–854. [719]

Fang, Dawei and Thomas H. Noe (2016), "Skewing the odds: Strategic risk taking in contests." Unpublished paper. [738]

Garicano, Luis and Luis Rayo (2016), "Why organizations fail: Models and cases." *Journal of Economic Literature*, 54, 137–192. [715, 718]

Georgiadis, George and Balazs Szentes (2019), "Optimal monitoring design." Unpublished paper, Northwestern University. [719]

Halac, Marina and Andrea Prat (2016), "Managerial attention and worker performance." *American Economic Review*, 106, 3104–3132. [719]

Hébert, Benjamin (2018), "Moral hazard and the optimality of debt." *Review of Economic Studies*, 85, 2214–2252. [719]

Hellwig, Martin (2009), "A reconsideration of the Jensen–Meckling model of outside finance." *Journal of Financial Intermediation*, 18, 495–525. [719]

Holmström, Bengt (1979), "Moral hazard and observability." *Bell Journal of Economics*, 10, 74–91. [717, 719, 726]

Holmström, Bengt and Paul R. Milgrom (1987), "Aggregation and linearity in the provision of intertemporal incentives." *Econometrica*, 55, 303–328. [719]

Holmström, Bengt and Joan Ricart i Costa (1986), "Managerial incentives and capital management." *Quarterly Journal of Economics*, 101, 835–860. [719]

Innes, Robert D. (1990), "Limited liability and incentive contracting with ex-ante action choices." *Journal of Economic Theory*, 52, 45–67. [719, 722]

Jensen, Michael C. and William H. Meckling (1976), "Theory of the firm: Managerial behavior, agency costs and ownership structure." *Journal of Financial Economics*, 4, 305–360. [716]

Jewitt, Ian, Ohad Kadan, and Jeroen M. Swinkels (2008), "Moral hazard with bounded payments." *Journal of Economic Theory*, 143, 59–82. [719, 731]

Lambert, Richard A. (1986), "Executive effort and selection of risky projects." *RAND Journal of Economics*, 17, 77–88. [719]

Larkin, Ian (2014), "The cost of high-powered incentives: Employee gaming in enterprise software sales." *Journal of Labor Economics*, 32, 199–227. [716, 718, 737]

Makarov, Igor and Guillaume Plantin (2015), "Rewarding trading skills without inducing gambling." *Journal of Finance*, 70, 925–962. [719, 738]

Malcolmson, James M. (2009), "Principal and expert agent." *B.E. Journal of Theoretical Economics*, 9(1). [719]

Matta, Elie and Paul W. Beamish (2008), "The accentuated CEO career horizon problem: Evidence from international acquisitions." *Strategic Management Journal*, 29, 683–700. [716]

Mirrlees, James A. (1976), "The optimal structure of incentives and authority within an organization." *Bell Journal of Economics*, 7, 105–131. [717, 719, 726]

Oyer, Paul (1998), "Fiscal year ends and nonlinear incentive contracts: The effect on business seasonality." *Quarterly Journal of Economics*, 113, 149–185. [716, 718, 737]

Palomino, Frédéric and Andrea Prat (2003), "Risk taking and optimal contracts for money managers." *RAND Journal of Economics*, 34, 113–137. [719]

Poblete, Joaquín and Daniel Spulber (2012), "The form of incentive contracts: Agency with moral hazard, risk neutrality, and limited liability." *RAND Journal of Economics*, 43, 215–234. [719]

Prékopa, András (1973), "On logarithmic concave measures and functions." *Acta Scientiarum Mathematicarum*, 34, 335–343. [726]

Rahmandad, Hazhir, Rebecca Henderson, and Nelson P. Repenning (2018), "Making the numbers? "Short termism" and the puzzle of only occasional disaster." *Management Science*, 64, 1328–1347. [716, 721]

Rajan, Raghuram G. (2011), *Fault Lines: How Hidden Fractures Still Threaten the World Economy*. Princeton University Press, Princeton, New Jersey. [716, 724]

Ray, Debraj and Arthur Robson (2012), "Status, intertemporal choice, and risk-taking." *Econometrica*, 80, 1505–1531. [719]

Robson, Arthur (1992), "Status, the distribution of wealth, private and social attitudes to risk." *Econometrica*, 60, 837–857. [719]

The Federal Reserve (2009), "Federal Reserve press release." Available at https://www.federalreserve.gov/newsevents/pressreleases/bcreg20091022a.htm. [715]

Vereshchagina, Galina and Hugo A. Hopenhayn (2009), "Risk taking by entrepreneurs." *American Economic Review*, 99, 1808–1830. [716]

Co-editor Thomas Mariotti handled this manuscript.

Manuscript received 26 February, 2019; final version accepted 26 September, 2019; available online 30 September, 2019.