



Centre Interuniversitaire sur le Risque,  
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **12-20**

## **The First-Order Approach when the Cost of Effort is Money**

Marie-Cécile Fagart  
Claude Fluet

Avril/April 2012

---

Fagart: Université Paris Descartes  
Fluet: Corresponding author. UQAM and CIRPÉE  
[fluet.claude-denys@uqam.ca](mailto:fluet.claude-denys@uqam.ca)

**Abstract:**

We provide sufficient conditions for the first-order approach in the principal-agent problem when the agent's utility has the non-separable form  $u(y - c(a))$  where  $y$  is the contractual payoff and  $c(a)$  is the money cost of effort. We first consider a decision-maker facing prospects which cost  $c(a)$  with distributions of returns  $y$  that depends on  $a$ . The decision problem is shown to be concave if the primitive of the cumulative distribution of returns is a convex function, a condition we call Concavity of the Cumulative Quantile (CCQ). Next we apply CCQ to the distribution of outcomes (or their likelihood-ratio transforms) in the principal-agent problem and derive restrictions on the utility function that validate the first-order approach. We also discuss a stronger condition, log-convexity of the distribution, and show that it allows binding limited liability constraints, which CCQ does not.

**Keywords:** Principal-agent models, moral hazard, stochastic decision problem, quantile function, information systems

**JEL Classification:** D81, D82, D86

# 1 Introduction

The principal-agent problem is central to the economics of moral hazard. Yet the theoretical literature has focused on a particular version of the problem. In this version, the agent's utility has the additively separable form  $v(y, a) = u(y) - \psi(a)$  where  $u(y)$  is a utility of income function and  $\psi(a)$  is the disutility of effort. The reason is tractability and the convenience of the so-called first-order approach in characterizing the optimal contract. Standard conditions validating the approach rely on additive separability.

Many situations clearly do not fit the restriction. It is particularly awkward when the agent is a profit-maximizing firm whose actions are more appropriately viewed as involving unverifiable money costs; for instance, the cost incurred by a subcontractor operating at arm's length or that of workplace safety measures under an experience rated industrial insurance plan. The agent's utility is then  $v(y, a) = u(y - c(a))$ , where  $c(a)$  is the money cost of effort,  $y$  is the gross payoff which may depend on the contract signed with the principal, and  $y - c(a)$  is net profit. We provide sufficient conditions for the validity of the first-order approach when effort entails a money cost.<sup>1</sup>

The approach replaces the incentive compatibility constraint with the first-order condition of the agent's maximization problem. When the agent's utility is additively separable, it is well-known that such an approach is valid when the distribution of outcomes satisfies the monotone likelihood ratio property (MLR) and the convexity of the distribution function condition (CDF); see Rogerson (1985). The CDF condition is generally considered as

---

<sup>1</sup>By contrast, there is a large applied literature where the agent's utility is precisely of the form  $u(y - c(a))$ . Borrowing from Holmstrom and Milgrom (1987), this literature draws on the simplicity of the LEN model (for Linear Exponential Normal). The applications are many, e.g., procurement and subcontracting (McAfee and McMillan 1986; Kawasaki and McMillan 1987) or public-private partnerships (Martimort and Pouyet 2008). Contracts are restricted to be linear in outcome, so the validity of the first-order approach is not an issue.

quite restrictive: it is not satisfied by most of the probability distributions commonly used and is difficult to reconcile with outcomes generated by a stochastic production function exhibiting decreasing marginal returns. Jewitt (1988) proposed an alternative set of conditions avoiding CDF and relying on the curvature of the contract; the trade-off is then that restrictions need to be imposed on the agent's utility of income function. Most of the literature has borrowed from one set of conditions or the other (e.g., Sinclair-Desgagné 1994; Carlier and Dana 2005; Jewitt and al. 2008; Conlon 2009).

When utility is nonseparable, however, a risk-averse agent's preferences over income lotteries will generally depend on his action. Grossman and Hart (1983) remark that random incentive schemes may then be superior to deterministic schemes. Similar observations are made by Arnott and Stiglitz (1988). More recently, Kadan et al. (2011) investigated the existence of optimal contracts in very general environments allowing a nonseparable utility. They do not discuss the first-order approach but give conditions under which randomization is unnecessary.

Alvi (1997) provides a set of sufficient conditions in the nonseparable case  $v(y, a)$ . Given MLR and CDF conditions, he shows that the first-order approach is valid if  $-v_{yy}/v_y$  is nondecreasing in  $a$  and  $v_{ya} \leq 0$ . The first condition means that the agent's absolute risk aversion over income lotteries is nondecreasing in effort; the second is interpreted as normality of leisure (with  $a$  as work effort). However, when effort entails a money cost,  $v(y, a) = u(y - c(a))$  so that  $-v_{yy}/v_y = -u''/u'$ . Alvi's first condition then reduces to nonincreasing absolute risk aversion (NIARA) but his second condition cannot be met because  $v_{ya} = -u''c' > 0$ .

Another related paper is Abraham et al. (2011). They consider a two-period principal-agent problem with hidden borrowing and lending. In our notation, the agent's utility is  $[u(w_0 - s) - \psi(a)] + \beta u(y + s)$ . The expression in brackets is the first-period utility, where  $w_0$  is initial wealth and  $s$  is saving; the other term is the present value of the second-period utility with  $\beta$  as

discount factor. The agent chooses his effort in the first period and receives a contractual payment  $y$  contingent on the outcome observed in the second period. Because the agent simultaneously chooses  $s$  and  $a$ , both of which are unobservable, the principal faces two incentive compatibility constraints. As in our one-period problem, the agent's preferences over income lotteries then depend on his actions. Abraham et al. (2011) show that the first-order approach is valid if the utility function  $u$  satisfies NIARA and the distribution of outcomes satisfies the MLR condition and is log-convex in the agent's effort, a condition obviously stronger than CDF.

We proceed as follows. We first abstract from the intricacies of the principal-agent problem and consider risk-averse decision-makers facing random prospects indexed by  $a$ . The gross return is  $y$  and the cost of the prospects is  $c(a)$ , an increasing convex function. We seek conditions for the decision-makers' problem to be concave in the action  $a$ . This question is of interest in itself and is common in the finance and economics of risk literature. For instance, Jullien et al. (1998), Eeckhoudt and Gollier (2005) and Meyer and Meyer (2011) consider the case where  $c(a)$  refers to prevention expenditure against accidental losses. Conditions are given for the problem to be concave in a two-outcome prospect, but no equivalent conditions are proposed for the many-outcome case.

We show that the decision-maker's problem is concave if the primitive of the cumulative distribution of returns is jointly convex in  $a$  and  $y$ . The condition amounts to a decreasing marginal returns property and is satisfied by many common probability distributions under an appropriate parameterization. In financial jargon, it is equivalent to the requirement is that, at any confidence level, the expected tail return (or average Value at Risk) is concave in  $a$ . For reasons that will become clear, we refer to this condition as Concavity of the Cumulative Quantile (CCQ).

Next we apply the CCQ condition to the principal-agent problem when effort entails a money cost. The gross returns are now determined by the

incentive scheme designed by the principal as a function of the observable outcomes. The validity of the first-order approach then boils down to whether the approach yields a payment scheme with the appropriate stochastic properties. We discuss various conditions on the distributions of outcomes, all of which include standard MLR property.

The weakest condition is in terms of likelihood-ratio transforms of the observable outcome, the requirement being that the probability distribution of the transforms satisfies the CCQ condition. The interpretation is that the agent faces decreasing marginal returns to effort in the generation of “favorable information”. Again, such a condition is satisfied by many common probability distributions. The validity of the first-order approach then follows if, roughly speaking, the agent’s utility function is such that absolute prudence is less than twice absolute risk aversion and does not decrease too fast as the agent’s wealth increases. This trivially holds when absolute risk aversion is constant.

Finally, we note that the conditions in Abraham et al. (2011) — that is, NIARA and log-convexity of the distribution of outcomes with respect to the agent’s action — are indeed also sufficient for our problem. We show that they are consistent with binding payment restrictions such as can arise when the agent faces a limited liability constraint, thus extending the results obtained by Jewitt and al. (2008) for the separable case. By contrast, the weaker CCQ condition does not allow a binding downward bounded payment constraint.

## 2 Conditions on the Distribution of Returns

A decision-maker faces random prospects defined by the cumulative distribution functions  $G(y | a)$  where  $y$  is the gross return and  $a \in [0, \bar{a}]$  is the action taken. To simplify notation, the cost of the prospects is  $c(a) \equiv a$  so that final wealth is  $w = y - a$ . All our results also hold for  $c(a)$  increasing

and convex.

For all  $a$ , the support of  $G$  is an interval contained in  $[\underline{y}, \bar{y}]$  where  $\underline{y}$  and  $\bar{y}$  need not be finite.  $G(\underline{y} | a)$  is twice-continuously differentiable in  $a$ ; for  $y$  in the open interval  $(\underline{y}, \bar{y})$ ,  $G(y | a)$  is twice-continuously differentiable in  $a$  and  $y$  with density  $g(y | a) \equiv G_y(y | a)$ . The utility of final wealth is  $u(w)$ , a smooth function with  $u' > 0$  and  $u'' \leq 0$ ; the domain is an interval containing  $[\underline{y} - \bar{a}, \bar{y}]$ .

The setup is consistent with an absolutely continuous probability distribution or with a mixed distribution with probability masses at  $\underline{y}$  or  $\bar{y}$ . Specifically,  $G(y | a) = 0$  for  $y < \underline{y}$  and

$$G(y | a) = G(\underline{y} | a) + \int_{\underline{y}}^y g(y | a) dy \text{ for } y \in [\underline{y}, \bar{y}), \quad G(\bar{y} | a) = 1. \quad (1)$$

Note that the support need not be invariant in  $a$ .

Many situations of interest involve mixed distributions. For example, the expenditure  $a$  refers to prevention measures bearing on accidental losses; the return is  $y = w_0 - z$  where  $w_0$  is initial wealth and  $z \geq 0$  is the loss. The distribution has a discontinuity at  $\bar{y} = w_0$  if the no-loss event has positive probability. There is a discontinuity at  $\underline{y}$  if  $y = \max(w_0 - z, \underline{y})$  and losses larger than  $w_0 - \underline{y}$  have positive probability; for instance,  $\underline{y}$  captures the protection provided by social security or the decision-maker's limited liability in an investment context. Mixed distributions can also arise in the principal-agent framework when contracts are subject to exogenous bounded payment constraints.

The individual chooses  $a$  to maximize expected utility

$$U(a) \equiv \int_{\underline{y}}^{\bar{y}} u(y - a) dG(y | a). \quad (2)$$

We seek conditions for the first-order condition  $U'(\hat{a}) = 0$  to be sufficient for a maximum at  $\hat{a}$ . What is needed is some form of decreasing marginal returns. In a stochastic environment, this may be characterized in several ways.

**Concavity of the Cumulative Quantile (CCQ).** We consider a condition expressed in terms of the primitive of the cumulative distribution of returns,

$$\tilde{G}(y, a) \equiv \int_{\underline{y}}^y G(z | a) dz.$$

The condition is  $\tilde{G}(y, a)$  jointly convex in  $a$  and  $y$ . Note that this is stronger than mere convexity in  $a$ , a condition which has been discussed in the principal-agent literature (see Section 3).

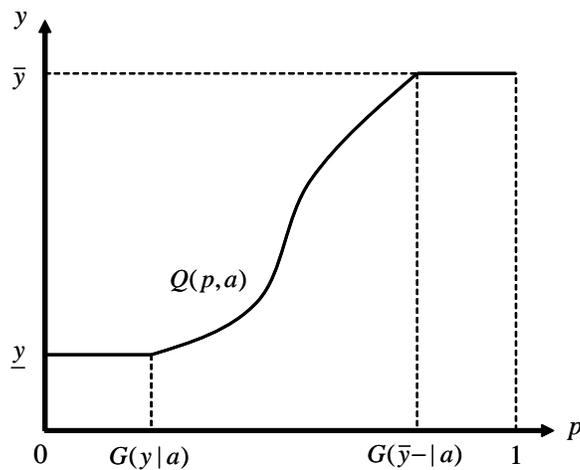


Figure 1. The Quantile Function

The interpretation is straightforward when the distribution is expressed in terms of its quantile function. This is defined by

$$Q(p, a) = \inf\{y \mid G(y | a) \geq p\}, \quad p \in (0, 1).$$

Our assumptions imply that the quantile is a continuous function, although it may be only piece-wise differentiable. Figure 1 gives an illustration in the  $(p, y)$  plane when both the lower and upper bound of the support have a probability mass. In the interior of the support, the quantile is defined by  $G(Q(p, a) | a) \equiv p$ . When the distribution is absolutely continuous, the identity holds everywhere and the curve has no “flat spots”.

**Lemma 1**  $\tilde{G}(y, a)$  convex in  $(y, a)$  is equivalent to  $\int_0^p Q(\eta, a) d\eta$  concave in  $a$  for all  $p$ .

In view of Lemma 1, we henceforth refer to the condition as Concavity of the Cumulative Quantile (CCQ). It is equivalent to

$$E(Y | a, G(Y | a) \leq p) = \int_0^p \frac{Q(\eta, a)}{p} d\eta$$

concave in  $a$ . In words, the average return below any  $p$ -percentile is concave in  $a$ . Letting  $p$  tend to unity, the overall expected return is also concave.

In terms of the quantile function, expected utility is

$$U(a) \equiv \int_0^1 u(Q(p, a) - a) dp. \quad (3)$$

If  $Q(p, a)$  is concave in  $a$ , so is the integrand in (3) and therefore the expected utility. A concave quantile function obviously implies CCQ. The latter turns out to be sufficient for the concavity of the decision-maker's problem.

**Proposition 1** *Expected utility is concave in  $a$  if the distribution of returns satisfies CCQ.*

When the distribution is absolutely continuous,  $G(y | a)$  quasiconvex in  $a$  and  $y$  is equivalent to  $Q(p, a)$  concave in  $a$  for all  $p$  (see Jewitt 1988). The return can then be interpreted as generated by a stochastic production function  $y = \varphi(a, \varepsilon)$  that is concave in  $a$  in each state of the world  $\varepsilon$ . The CCQ condition then corresponds to the weaker requirement that marginal returns to investment are on average decreasing in bad states.

Many common probability distributions satisfy CCQ given an appropriate parameterization. To illustrate, if returns have the exponential distribution with a mean that is concave in  $a$ , the quantile function is also concave in  $a$  and therefore CCQ holds. If returns are normally distributed with mean  $\mu(a)$  and standard deviation  $\sigma(a)$ , then  $y = \mu(a) + \sigma(a)\varepsilon$  where  $\varepsilon$  is the standard normal

variable. Its quantile is  $Q_\varepsilon(p)$  so that  $\int_0^1 Q_\varepsilon(p) dp = 0$ . The quantile of returns is then  $Q(p, a) = \mu(a) + \sigma(a)Q_\varepsilon(p)$ . When the standard deviation is constant,  $\mu''(a) \leq 0$  implies that the quantile of returns is concave. When the standard deviation varies with  $a$ , the quantile cannot be concave because  $Q_\varepsilon(p)$  varies between minus and plus infinity. However, because  $\int_0^p Q_\varepsilon(\eta) d\eta \leq 0$ ,

$$\int_0^p Q_{aa}(\eta, a) d\eta = \mu''(a) + \sigma''(a) \int_0^p Q_\varepsilon(\eta) d\eta \leq 0$$

if the mean is a concave function and dispersion is convex. The same argument also holds if  $\varepsilon$  is a continuous but otherwise arbitrary random variable.

Two points need to be emphasized. First, the convexity of  $\tilde{G}$  is inconsistent with a probability mass at the lower bound of the support, except in the particular case where the probability does not vary with  $a$ .<sup>2</sup> Second, the same condition is required for the quasiconvexity of  $G$  to imply the concavity of the quantile in  $a$ . Indeed, the CCQ condition implies that  $Q(p, a)$  is concave in  $a$  for  $p$  sufficiently small. This is easily shown to be consistent with a probability mass at the lower bound only if the probability is constant.<sup>3</sup>

**Log-Convexity of the Distribution Function (LCDF).** We now discuss a more restrictive set of conditions. By contrast with CCQ, however, they are compatible with a non-constant probability mass at the lower bound of the support. The distribution of returns satisfies LCDF if  $\ln G$  is convex in  $a$ . We combine this property with the condition that the

---

<sup>2</sup>For instance,  $G(y | a)$  is always equal to zero. An example with a positive constant is the case of accidental losses when the probability of no accident does not depend on prevention expenditures, although prevention affects the severity of losses should an accident occur.

<sup>3</sup>To illustrate, let  $G(y | a) = 1 - \varphi(a) \exp(-(y/\psi(a)))$  where  $y \geq 0$  with  $\varphi(a) \in (0, 1)$  and  $\psi(a)$  twice differentiable and nondecreasing concave. Conditional on  $y > 0$ , the distribution of returns has the exponential density with mean  $\psi(a)$ . The quantile is  $Q(p, a) = 0$  if  $\varphi(a) \leq 1 - p$  and  $Q(p, a) = -\psi(a) \ln((1 - p)/\varphi(a))$  otherwise.  $Q_{aa}(p, a) \leq 0$  whenever  $\varphi(a) \neq 1 - p$ . The distribution is quasiconvex but the quantile is not concave in  $a$  unless  $\varphi(a)$  is a constant.

decision-maker's utility function satisfies NIARA; that is, absolute risk aversion  $r(w) \equiv -u''(w)/u'(w)$  is nonincreasing.

**Proposition 2** *Expected utility is concave in  $a$  if  $G$  satisfies LCDF and the utility function satisfies NIARA.*

The proof is similar to that in Abraham et al. (2011) except that we allow probability masses at the bounds of the support. LCDF is stronger than CDF (i.e.,  $G$  convex in  $a$ ), itself a restrictive condition as already noted. In the Appendix we show that, because of wealth effects, CDF is not sufficient to ensure that expected utility is concave.<sup>4</sup>

LCDF has the following interpretation in terms of decreasing marginal returns. The ratio  $g/G$  is the reverse hazard rate, the probability of the realization  $y$  conditional on the outcome being no more than  $y$ . It is easily seen that

$$-\ln(G(y | a)) = \int_y^{\bar{y}} \frac{g(\eta | a)}{G(\eta | a)} d\eta. \quad (4)$$

LCDF means that the cumulative reversed hazard rate is concave in  $a$ . When a larger expenditure increases returns in the sense of first-order stochastic dominance, the cumulative reversed hazard rate is increasing in  $a$ . LCDF imposes that it increases at a decreasing rate.

### 3 The Principal-Agent Problem

The random return facing the individual, henceforth the agent, now originates from the contract designed by the principal. The observable outcome is  $x$  distributed according to  $F(x | a)$ , a twice-continuously differentiable distribution with density  $f(x | a)$ ; the support is  $[\underline{x}, \bar{x}]$  for all  $a$ . We make the

---

<sup>4</sup>An example of a distribution satisfying LCDF is the function used in Rogerson (1985) to illustrate CDF,  $G(y | a) = (y/\bar{y})^\alpha$ ,  $y \in [0, \bar{y}]$ ,  $a > 0$ .

standard assumption that the distribution satisfies the strict MLR condition,  $f_a(x | a)/f(x | a)$  strictly increasing in  $x$  for all  $a$ . We also assume that the likelihood ratio is bounded below. This precludes near-forcing contracts as in Mirrlees (1999). When the agent faces a limited liability constraint, it also allows for the possibility that the constraint is not binding in the optimal contract.<sup>5</sup>

A contract is a piece-wise differentiable function  $y(x)$ . This describes a deterministic scheme. In a general framework permitting nonseparable utility, Kadan et al. (2011) show that randomization is unnecessary when the principal and agent have “weakly conflicting preferences” over rewards. The condition is trivially satisfied in our setup.<sup>6</sup>

The agent’s expected utility is

$$U(a, y) \equiv \int_{\underline{x}}^{\bar{x}} u(y(x) - a) f(x | a) dx. \quad (5)$$

The agent is strictly risk averse and the principal is risk neutral. We consider the cost to the principal of implementing some level of effort  $\hat{a} \in (0, \bar{a})$ . Equivalently, one could look for the contract that maximizes the agent’s expected utility subject to the principal’s profit constraint.

The agent’s participation constraint is

$$U(\hat{a}, y) \geq u_0, \quad (IR)$$

where  $u_0$  is reservation utility. His incentive compatibility constraint is

$$\hat{a} \in \max_{a \in [0, \bar{a}]} U(a, y). \quad (IC)$$

---

<sup>5</sup>We focus on the case where  $F$  is absolutely continuous in order to simplify the exposition. Our results also hold when  $F$  is discontinuous at  $\bar{x}$ . For instance, in the insurance context,  $x = -z$  where  $z$  is the accidental loss and the no-loss event has positive probability.

<sup>6</sup>In a two-outcome framework with the nonseparable utility  $v(y, a)$ , Arnott and Stiglitz (1988) show that a sufficient condition for the suboptimality of randomizing contracts is that  $v_{yya} \leq 0$ . In our setting,  $v_{yya} = -u'''$ . Arnott and Stiglitz’ condition would therefore follow from NIARA.

Payments are constrained by the bounded payment conditions

$$\underline{y} \leq y(x) \leq \bar{y} \text{ for all } x. \quad (BP)$$

The lower bound will be referred to as the agent's limited liability constraint.

The principal's cost minimization problem, denoted  $P$ , is

$$\min_y C(\hat{a}, y) \equiv \int_{\underline{x}}^{\bar{x}} y(x) f(x | \hat{a}) dx \quad \text{s.t. } IR, IC, \text{ and } BP. \quad (P)$$

**The relaxed problem.** The first-order approach replaces  $IC$  with the necessary first-order condition of the agent's problem,  $U_a(\hat{a}, y) = 0$ . In validating the approach, it has been common to look at the (doubly) relaxed problem where  $IC$  is replaced by

$$U_a(\hat{a}, y) \geq 0. \quad (ICR)$$

We assume that the relaxed problem is nondegenerate; that is, there exists piecewise differentiable contracts such that the constraints are strictly slack.

Let  $y(x)$  denote a solution to the relaxed problem, henceforth problem  $R$ . Given the strict MLR condition, a contract can equivalently be written as a function of the transformed variable  $l \equiv f_a(x | \hat{a})/f(x | \hat{a})$ . Abusing notation, we denote this representation of the contract by  $y(l)$ . Define

$$\Delta(y, \mu) \equiv \frac{1 + \mu u''(y - \hat{a})}{u'(y - \hat{a})}. \quad (6)$$

A solution  $y(l)$  to the relaxed problem satisfies

$$\Delta(y(l), \mu) = \begin{cases} \Delta(\underline{y}, \mu) & \text{if } \lambda + \mu l < \Delta(\underline{y}, \mu), \\ \lambda + \mu l & \text{if } \Delta(\underline{y}, \mu) \leq \lambda + \mu l \leq \Delta(\bar{y}, \mu), \\ \Delta(\bar{y}, \mu) & \text{if } \lambda + \mu l > \Delta(\bar{y}, \mu), \end{cases} \quad (7)$$

for some multipliers  $\lambda \geq 0$  and  $\mu \geq 0$  satisfying the complementary slackness conditions  $\lambda(U(\hat{a}, y) - u_0) = 0$  and  $\mu U_a(\hat{a}, y) = 0$ . Clearly  $\mu > 0$ , otherwise

condition (7) implies a constant payment, which is inconsistent with *ICR*. Hence,  $U_a(\hat{a}, y) = 0$ , meaning that the solution to the relaxed problem satisfies the first-order condition of the true incentive-compatibility condition *IC*.<sup>7</sup>

**Lemma 2** *Let  $y(l)$  be a solution to problem  $R$ . Then  $y(l)$  is nondecreasing.*

The proof does not rely on the Lagrangian conditions (see the Appendix).<sup>8</sup> Combined with (7), the lemma implies that the set of  $l$  values for which *BP* does not bind is an interval. If the lower bound binds, it does so only at realizations below some critical  $l_0$ ; if the upper bound binds, it does only above some critical  $l_1$ . Moreover, when *BP* does not bind,  $y(l)$  cannot be constant; by Lemma 2, it is therefore strictly increasing.

So far, nothing ensures that  $y(l)$  is continuous. One possibility is illustrated in Figure 2. When *BP* does not bind, condition (7) requires that  $y$  satisfies  $\Delta(y, \mu) = \lambda + \mu l$ . Lemma 2 implies that the solution is on the increasing portions of the curve represented in the figure. As shown,  $y(l)$  is discontinuous at  $l = l_c$ . Payments then belong to the intervals  $[\underline{y}, y']$  or  $(y'', \bar{y}]$ .

Contracts satisfying the constraints of problem  $P$  also satisfy those of problem  $R$ . If a solution to the relaxed problem satisfies the true incentive-compatibility condition *IC*, it is therefore also a solution to the original problem. Sufficient conditions are that the contract solving  $R$  yields a random payment satisfying the conditions set out in Section 2. The agent's expected utility is then concave and *IC* is therefore satisfied.

---

<sup>7</sup>The *IR* constraint need not be binding, possibly because of a binding limited liability constraint. Thiele and Wambach (1999) remark that, in a nonseparable setting, *IR* may be slack irrespective of bounded payment constraints.

<sup>8</sup>In the separable case,  $v(y, a) = u(a) - \psi(e)$ . The solution to the relaxed problem satisfies the condition (7) but with  $\Delta(y, \mu) \equiv 1/u'(y)$ . That  $y(l)$  is increasing then follows directly from (7). This is no longer so in our problem because  $\Delta(y, \mu)$  as defined in (6) need not be increasing in  $y$ .

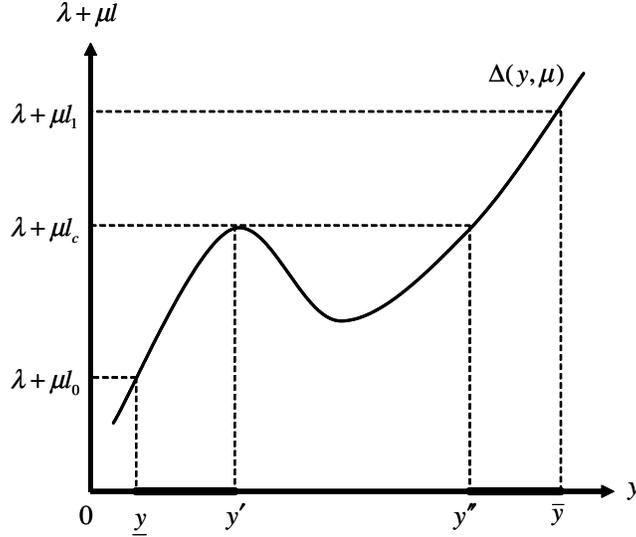


Figure 2. A Solution to the Relaxed Problem

**Validity of the first-order approach.** The first set of conditions is in the spirit of Proposition 1. As in Jewitt (1988) for the separable setting, the conditions rely on the curvature of the contract. In Jewitt’s analysis, the conditions on the distribution of outcomes include the so-called *concave* MLR condition,  $f_a(x | a)/f(x | a)$  increasing and concave in  $x$  for all  $a$ , together with one of the following two:

- (a)  $F(x | a)$  is quasiconvex in  $(x, a)$ ,
- (b)  $\int_{\underline{x}}^x F(\eta | a) d\eta$  is convex in  $a$  for all  $x$  and  $\int_{\underline{x}}^{\bar{x}} x f(x | a) dx$  is concave in  $a$ .

We will substitute the CCQ condition for (a) or (b). CCQ is weaker than (a) but implies (b). In order to apply Proposition 1 while allowing for an upward bounded payment constraint, we need the following lemma.

**Lemma 3** *If the random variable  $Z$  satisfies the CCQ condition, then the variable  $\min(\varphi(Z), k)$  also satisfies CCQ for any constant  $k$  and  $\varphi$  nondecreasing concave and twice differentiable.*

In the separable setting, the conditions in Jewitt (1988) are not sufficient when the lower bound on payments is binding. The same is true of our second set of conditions for the nonseparable case. A binding limited liability constraint implies that the distribution of payments has a probability mass at  $\underline{y}$ . Recalling section 2, this is inconsistent with CCQ. In what follows,  $p(w) \equiv -u'''(w)/u''(w)$  is absolute prudence.

**Proposition 3** *Let  $y(x)$  be a solution to problem R. Sufficient conditions for  $y(x)$  to solve problem P are:*

- (i)  $y(\underline{x}) > \underline{y}$ ;
- (ii)  $F$  satisfies the CCQ and concave MLR conditions;
- (iii)  $p(w) \leq 2r(w)$  and  $p'(w) \geq \min[0, (r(w) - p(w))(2r(w) - p(w))]$  for all  $w \in [\underline{y} - \bar{a}, \bar{y}]$ .

Condition (i) states that the lower bound on payments is not binding in the solution to the relaxed problem. Condition (iii), which is further discussed below, implies that  $\Delta(y, \mu)$  is convex in  $y$  whenever it is nondecreasing. Condition (ii) then implies that the random payment satisfies the CCQ condition, so Proposition 1 applies. See the Appendix for details. In the next subsection we discuss how both the CCQ and concave MLR conditions can be relaxed.

In condition (iii), the inequality  $p \leq 2r$  is equivalent to  $1/u'$  convex in wealth.<sup>9</sup> This is satisfied, for instance, if relative risk aversion is not less than unity and is nondecreasing. Roughly speaking, the second inequality in condition (iii) states that prudence does not decrease too fast over the domain of feasible wealth levels. It suffices that prudence be nondecreasing.

---

<sup>9</sup>This condition was used by Grossman and Hart (1983) in the separable setting to ensure that the wage is a concave function of the likelihood ratio. Jewitt (1988) imposes only that  $u(v(z))$  be concave in  $z$ , where  $v(z) \equiv (u')^{-1}(1/z)$ . This is equivalent to  $p \leq 3r$ . See Conlon (2009) for a discussion.

However, noting that  $r' = r(r - p)$ , the condition is also consistent with decreasing prudence when absolute risk aversion is itself decreasing.<sup>10</sup>

To illustrate, consider the family of utility functions with constant prudence,  $p(w) = \beta > 0$  for all  $w$ . Solving the differential equation yields  $u(w) = \alpha w - e^{-\beta w}$  where  $\alpha \geq 0$ . When  $\alpha = 0$ , this is the CARA utility function with absolute risk aversion equal to  $\beta$ , which obviously satisfies condition (iii). When  $\alpha > 0$ , risk aversion is

$$r(w) = \frac{\beta^2}{\beta + \alpha e^{\beta w}}$$

and is decreasing in wealth. The condition  $p(w) \leq 2r(w)$  then amounts to  $\alpha e^{\beta w} \leq \beta$ . For  $\alpha < \beta$ , the condition is satisfied as long as  $w \leq \ln(\beta/\alpha)/\beta$ . Thus, it suffices that the upper bound  $\bar{y}$  be less than this quantity. Another class of utility functions that is easily seen to satisfy condition (iii) consists of functions with marginal utility satisfying

$$u'(w) = (\alpha - \beta w)^\gamma, \alpha > 0, \beta > 0, \gamma \geq 1, 0 \leq w \leq \alpha/\beta.$$

For  $\gamma = 1$ , this is the quadratic utility function, hence prudence is zero.

The second set of conditions is with respect to Proposition 2. As in Jewitt et al. (2008) for the separable case, the following holds under both downward and upward bounded payments constraints.

**Corollary 1** *Suppose that the agent's utility function satisfies NIARA and that  $F$  satisfies the MLR and LCDF conditions. Then the solution to  $R$  is also the solution to the original problem  $P$ .*

The proof follows from the observation that, if a random variable satisfies LCDF, so does any nondecreasing transformation. MLR together with Lemma 2 therefore imply that the distribution of payments satisfies the condition of Proposition 1.

---

<sup>10</sup>The condition  $p' \geq (r - p)(2r - p)$  is easily seen to be equivalent to  $r'' \leq 0$ .

**Information systems.** Jewitt (1988) and Conlon (2009) interpret the concave MLR condition as meaning that variations in output become “less informative” at high levels of output. Yet in many situations the outcome is a pure signal. When it satisfies the MLR condition, so does any strictly increasing transformation. Moreover, the transformed signal provides the same information. Unless the transformation is concave, properties such as CCQ and concave MLR need not be preserved. Nevertheless, if a contract conditioning on the original signal solves the principal’s problem, a suitably specified contract conditioning on the transformed signal will do the same. For validating the first-order approach, one should therefore also look for properties that are robust to arbitrary strictly increasing transformations of the signal. Such properties will be referred to as *informational* because they characterize the underlying information system.

LCDF is preserved under increasing transformations. The condition on the signal distribution in Corollary 1 is therefore informational. We look for informational properties to substitute for the non informational properties in condition (ii) of Proposition 3.

**Definition 1** *Let  $X$  be distributed according to  $F(x | a)$  with associated density  $f(x | a)$  satisfying the MLR condition.  $F$  satisfies Likelihood-CCQ if, for all  $a'$ , the distribution of  $f_a(X | a')/f(X | a')$  satisfies CCQ.*

The random variable  $f_a(X | a')/f(X | a')$  will be referred to as the likelihood ratio transform at  $a'$ . Note that its distribution depends on the agent’s effort  $a$ .

**Corollary 2** *If  $X$  satisfies the MLR and Likelihood-CCQ conditions, so does any strictly increasing transformation. Moreover, condition (ii) in Proposition 3 can be replaced by:  $F$  satisfies the MLR and Likelihood-CCQ conditions.*

The second claim follows from the fact that the agent’s wage will then satisfy CCQ when the other conditions in Proposition 3 hold. The first claim states that Likelihood-CCQ is an informational property. To see this, let  $Z \equiv \varphi(X)$  where  $\varphi$  is strictly increasing. The distribution of the transformed signal is  $F(\varphi^{-1}(z) | a)$ . Its likelihood ratio transform at  $a'$  is the random variable

$$f_a(\varphi^{-1}(Z) | a')/f(\varphi^{-1}(Z) | a').$$

Clearly this has the same distribution as the likelihood ratio transform at  $a'$  of the original signal.

The condition means that the agent faces decreasing marginal returns to effort in the generation of “favorable information”. If  $X$  satisfies both the CCQ and concave MLR conditions, then it satisfies Likelihood-CCQ. Any strictly increasing transformation will then also satisfy Likelihood-CCQ, although the transformed signal need not satisfy CCQ or the concave MLR condition. To illustrate, suppose the signal has the Weibull distribution  $F(x | a) = 1 - \exp(-x/a^\gamma)^k$ , where  $\gamma$  and  $k$  are positive parameters and  $x \geq 0$ . The signal satisfies CCQ if  $\gamma \leq 1$  and the concave MLR condition if  $k \leq 1$ . Likelihood-CCQ is the weaker condition  $\gamma k \leq 1$ .

## 4 Concluding remarks

There will obviously be cases where the first-order approach is valid even though the above sufficient conditions are not met; that is, the solution to the relaxed problem is nevertheless the solution to the true problem. Lemma 2 shows that the payment to the agent is then increasing with “more favorable” information, as in the standard model where utility is additively separable in income and effort. The principal’s cost is then increasing in the effort level that is to be implemented, an important property of the first-order approach.

The CCQ condition or the weaker Likelihood-CCQ may be appealing in the additively separable case as well. When the agent’s limited liability

condition is not binding, Likelihood-CCQ ensures the validity of the first-order approach under Jewitt's (1988) conditions with respect to the agent's utility of income function. Because Proposition 3 specifies more restrictive conditions, they are also sufficient in the mixed case where effort entails both psychological disutility and money expenditure; that is, the agent's utility function is  $u(y - c(a)) - \psi(a)$  where the money cost and the disutility are both increasing convex functions.

## Appendix

**Non sufficiency of CDF.** We give an example showing that  $U'(a) = 0$  characterizes a global minimum even though the distribution of returns satisfies CDF. The decision-maker's utility function is CARA with  $u(w) = -\exp(-rw)$ . The return is the function  $y(X)$  of the random variable  $X$  with support  $[0, 1]$  and cdf  $F(x | a) = x - \beta(x)\gamma(a)$ , where  $\beta(x) = x(1 - x)$ ,  $\gamma(a) = (1 - k)a + k \ln(1 + a)$ ,  $k \in [0, 1]$ . The cdf belongs to a family of distributions satisfying both the MLR and CDF conditions; see LiCalzi and Spaeter 2003. Because these properties are robust to arbitrary increasing transformations, the returns satisfy MLR if  $y(\cdot)$  is an increasing function. Let  $r = k = 0.1$  and  $y(X) = bX$ . We consider  $U(a)$  for  $a \in [0, 1]$ . Figure A1 depicts the graph for  $b = 6$ ; the curve is convex with a maximum at  $a = 1$  and a global minimum at  $a = 0.285$ . In Figure A2,  $b = 5.85$ ; the curve is again convex but with a maximum at  $a = 0$  and a global minimum at  $a = 0.769$ .

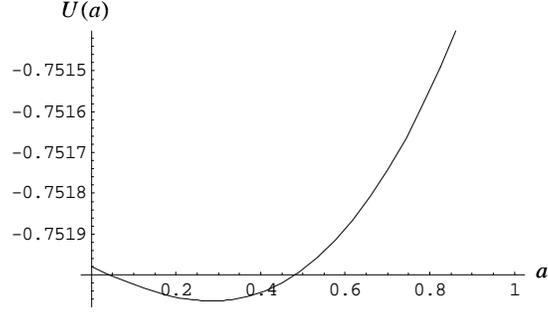


Figure A1:  $b = 6$

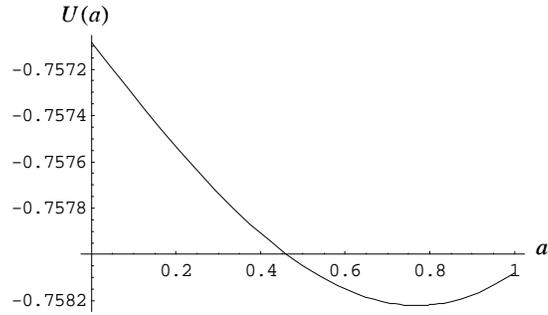


Figure A2:  $b = 5.85$

**Proof of Lemma 1.** Define

$$\tilde{Q}(p, a) \equiv \int_0^p Q(\eta, a) d\eta \text{ and } \tilde{G}(y, a) \equiv \int_y^1 G(z | a) dz.$$

Because the quantile is the convex conjugate of the cumulative distribution, for all  $p \in (0, 1)$  we have

$$\tilde{Q}(p, a) = pQ(p, a) - \tilde{G}(Q(p, a), a) \tag{8}$$

$$\geq py - \tilde{G}(y, a) \text{ for all } y \in [y, \bar{y}]. \tag{9}$$

We first show that the convexity of  $\tilde{G}(y, a)$  implies the concavity of  $\tilde{Q}(p, a)$  in  $a$ . Let  $a = \lambda a' + (1 - \lambda)a''$  and  $y = \lambda y' + (1 - \lambda)y''$  where  $\lambda \in (0, 1)$ . From (8) and (9),

$$\tilde{G}(y, a) \leq \lambda \tilde{G}(y', a') + (1 - \lambda) \tilde{G}(y'', a'')$$

implies

$$\tilde{Q}(p, a) \geq \lambda(py' - \tilde{G}(y', a')) + (1 - \lambda)(py'' - \tilde{G}(y'', a'')).$$

Therefore, setting  $y' = Q(p, a')$  and  $y'' = Q(p, a'')$  yields

$$\tilde{Q}(p, a) \geq \lambda\tilde{Q}(p, a') + (1 - \lambda)\tilde{Q}(p, a'').$$

Next we show that  $\tilde{Q}(p, a)$  concave in  $a$  implies the convexity of  $\tilde{G}(y, a)$ . From (8) and (9),

$$\begin{aligned} \lambda\tilde{G}(y', a') + (1 - \lambda)\tilde{G}(y'', a'') &\geq py - (\lambda\tilde{Q}(p, a') + (1 - \lambda)\tilde{Q}(p, a'')) \\ &\geq py - \tilde{Q}(p, a), \end{aligned}$$

where the second inequality follows from the concavity of  $\tilde{Q}(p, a)$  in  $a$ . Setting  $p = G(y | a)$ , the right-hand side in the string of inequalities equals

$$G(y | a)y - \tilde{Q}(G(y | a), a) = \tilde{G}(y, a),$$

implying that  $\tilde{G}(y, a)$  is convex. QED

**Proof of Proposition 1.** We first prove a preliminary result.

*Step 1.* For  $p \in [0, 1]$  and  $a \in [0, \bar{a}]$ , let  $w(p, a)$  be a continuous function, nondecreasing in  $p$  and piecewise differentiable with respect to  $p$  and  $a$ . Define

$$\tilde{w}(p, a) \equiv \int_0^p w(\eta, a) d\eta \text{ and } U(p, a) \equiv \int_0^p u(w(\eta, a)) d\eta.$$

We show that  $U(p, a)$  is concave in  $a$  if  $\tilde{w}(p, a)$  is concave in  $a$ . Observe that  $\tilde{w}(p, a)$  and  $U(p, a)$  are differentiable with respect to  $a$ . Recall that  $u$  is increasing concave and twice-continuously differentiable. The concavity implies that, for any  $a$  and  $a^*$ ,

$$u(w(\eta, a)) - u(w(\eta, a^*)) \geq u'(w(\eta, a))(w(\eta, a) - w(\eta, a^*))$$

and therefore

$$U(p, a) - U(p, a^*) \geq \int_0^p u'(w(\eta, a))(w(\eta, a) - w(\eta, a^*)) d\eta. \quad (10)$$

Denote by  $w_{p+}(p, a)$  the right-hand derivative of  $w(p, a)$  with respect to  $p$ . Integrating by parts, the right-hand side of (10) equals

$$u'(w(p, a)) (\tilde{w}(p, a) - \tilde{w}(p, a^*)) - \int_0^p u''(w(\eta, a)) w_{p+}(\eta, a) (\tilde{w}(\eta, a) - \tilde{w}(\eta, a^*)) d\eta.$$

Substituting in (10) and noting that the concavity of  $\tilde{w}(p, a)$  implies

$$\tilde{w}(\eta, a) - \tilde{w}(\eta, a^*) \geq \tilde{w}_a(\eta, a)(a - a^*),$$

we obtain

$$U(p, a) - U(p, a^*) \geq (a - a^*) \left( u'(w(p, a)) \tilde{w}_a(p, a) - \int_0^p u''(w(\eta, a)) w_{p+}(\eta, a) \tilde{w}_a(\eta, a) d\eta \right).$$

Integrating by parts the expression in the parentheses,

$$U(p, a) - U(p, a^*) \geq (a - a^*) \left( \int_0^p u'(w(\eta, a)) w_a(\eta, a) d\eta \right) = (a - a^*) U_a(p, a).$$

Hence  $U(p, a)$  is concave in  $a$  for all  $p$ .

*Step 2.* We now turn to the claim in the proposition. Write the expected utility using the quantile formulation (3). By Lemma 1, CCQ implies that  $\int_0^p (Q(\eta, a) - a) d\eta$  is concave in  $a$ . Letting  $w(p, a) \equiv Q(p, a) - a$  and applying Step 1 then completes the proof. QED

**Proof of Proposition 2.** We start with a preliminary result. Let  $H(y, a)$  be a positive continuous function, differentiable and increasing in  $y \in (y_0, y_1) \subset (y, \bar{y})$  and log-convex in  $a$ . Define

$$U(a, y_0, y_1, H) \equiv u(y_0 - a)H(y_0, a) + u(y_1 - a)(1 - H(y_1, a)) + \int_{y_0}^{y_1} u(y - a)H_y(y, a) dy.$$

Integrating by parts,

$$U(a, y_0, y_1, H) = u(y_1 - a) - \int_{y_0}^{y_1} u'(y - a)H(y, a) dy. \quad (11)$$

When  $r(w)$  is nonincreasing,  $u'(y - a)$  is log-convex in  $a$ . It follows that  $u'(y - a)H(y, a)$  is also log-convex in  $a$ , hence the integral in (11) is convex, implying that  $U(a, y_0, y_1, H)$  is concave.

Now consider the expected utility in (2) and assume NIARA. We discuss the cases where the support of  $G$  is either unbounded at both ends or bounded at both ends; the argument for the mixed cases is similar. When  $\underline{y} = -\infty$  and  $\bar{y} = +\infty$ ,

$$U(a) = \lim_{\substack{y_1 \rightarrow +\infty \\ y_0 \rightarrow -\infty}} U(a, y_0, y_1, G).$$

Therefore, if  $G$  is log-convex in  $a$ ,  $U(a)$  is concave. For the case of a bounded support, let  $y_0 = \underline{y}$  and  $y_1 = \bar{y}$  with

$$H(y, a) = G(\underline{y} | a) + \int_{\underline{y}}^y g(z | a) dz, \quad y \in [\underline{y}, \bar{y}].$$

Hence  $H(y, a) = G(y | a)$  for  $y < \bar{y}$  and  $H(\bar{y}, a) \leq G(\bar{y} | a)$ , where the inequality is strict if there is a jump at the upper bound. It follows that  $U(a) = U(a, \underline{y}, \bar{y}, H)$ . If  $G$  is log-convex in  $a$ , so is  $H$  and therefore  $U(a)$  is concave. QED

**Proof of Lemma 2.** If  $y(x)$  solves problem  $R$ , the agent's expected utility is

$$U(\hat{a}, y) = \int_{\underline{x}}^{\bar{x}} u(y(x) - \hat{a})f(x | \hat{a}) dx \geq u_0$$

and marginal expected utility satisfies

$$U_a(\hat{a}, y) = \int_{\underline{x}}^{\bar{x}} u(y(x) - \hat{a})f_a(x | \hat{a}) dx - \int_{\underline{x}}^{\bar{x}} u'(y(x) - \hat{a})f(x | \hat{a}) dx = 0.$$

The cost to the principal is

$$C(\hat{a}, y) \equiv \int_{\underline{x}}^{\bar{x}} y(x)f(x | \hat{a}) dx.$$

Let  $Q_X(p, a)$  be the quantile of the signal. Define

$$l(p) \equiv f_a(Q_X(p, \hat{a}) \mid a) / f(Q_X(p, \hat{a}) \mid \hat{a}).$$

Strict MLR implies that  $l(p)$  is strictly increasing. Let  $w(p) \equiv y(l(p))$ . The above expressions can then be rewritten as

$$\begin{aligned} U(\hat{a}, w) &= \int_0^1 u(w(p) - \hat{a}) dp, \\ U_a(\hat{a}, w) &= \int_0^1 u(w(p) - \hat{a}) l(p) dp - \int_0^1 u'(w(p) - \hat{a}) dp. \quad (12) \\ C(\hat{a}, w) &\equiv \int_0^1 w(p) dp. \end{aligned}$$

If  $y(l)$  is not nondecreasing, the same is true of  $w(p)$ . Hence, there exists two disjoint intervals  $[p', p' + \delta] \subset [0, 1]$  and  $[p'', p'' + \delta] \subset [0, 1]$ ,  $p' + \delta < p''$ , such that  $w(p) > w(\tilde{p})$  for all  $p \in [p', p' + \delta]$  and  $\tilde{p} \in [p'', p'' + \delta]$ . Consider the contract

$$\hat{w}(p) = \begin{cases} w(p) & \text{if } p \notin [p', p' + \delta] \cup [p'', p'' + \delta], \\ w(p'' - p' + p) & \text{if } p \in [p', p' + \delta], \\ w(p' - p'' + p) & \text{if } p \in [p'', p'' + \delta]. \end{cases}$$

For any function  $\varphi$ ,

$$\int_0^1 \varphi(\hat{w}(p)) dp = \int_0^1 \varphi(w(p)) dp.$$

Indeed

$$\begin{aligned} \int_{p'}^{p'+\delta} \varphi(\hat{w}(p)) dp &= \int_0^\delta \varphi(\hat{w}(p' + \eta)) d\eta \\ &= \int_0^\delta \varphi(w(p'' + \eta)) d\eta = \int_{p''}^{p''+\delta} \varphi(w(p)) dp, \\ \int_{p''}^{p''+\delta} \varphi(\hat{w}(p)) dp &= \int_0^\delta \varphi(\hat{w}(p'' + \eta)) d\eta \\ &= \int_0^\delta \varphi(w(p' + \eta)) d\eta = \int_{p'}^{p'+\delta} \varphi(w(p)) dp. \end{aligned}$$

Expected utility and the principal's cost are therefore unchanged under  $\widehat{w}(p)$ . The same is true of the second term on the right-hand side of (12). However,  $U_a(\widehat{a}, \widehat{w}) > U_a(\widehat{a}, w)$  because

$$\int_0^1 u(\widehat{w}(p) - \widehat{a})l(p) dp > \int_0^1 u(w(p) - \widehat{a})l(p) dp,$$

To see this,  $u(p) = u(w(p) - \widehat{a})$ ,  $\widehat{u}(p) = u(\widehat{w}(p) - \widehat{a})$  and observe that

$$\begin{aligned} \int_{p'}^{p'+\delta} \widehat{u}(p)l(p) dp &= \int_0^\delta \widehat{u}(p' + \eta)l(p' + \eta) d\eta = \int_0^\delta u(p'' + \eta)l(p' + \eta) d\eta, \\ \int_{p''}^{p''+\delta} \widehat{u}(p)l(p) dp &= \int_0^\delta \widehat{u}(p'' + \eta)l(p'' + \eta) d\eta = \int_0^\delta u(p' + \eta)l(p'' + \eta) d\eta. \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^1 \widehat{u}(p)l(p) dp - \int_0^1 u(p)l(p) dp \\ = \int_0^\delta (u(p'' + \eta) - u(p' + \eta)) \left( l(p' + \eta) - l(p'' + \eta) \right) d\eta. \end{aligned}$$

On the right-hand side, both factors in the integrand are negative, so the expression is positive.

Finally,  $U_a(\widehat{a}, \widehat{w}) > U_a(\widehat{a}, w)$  implies that  $w(p)$  cannot be a solution to problem  $R$ . Let  $\varepsilon(p) \equiv \widehat{w}(p) - C(\widehat{a}, w)$  so that  $\int_0^1 \varepsilon(p) dp = 0$ . Consider the contracts

$$\widetilde{w}(p) \equiv C(\widehat{a}, w) + \beta\varepsilon(p), \beta \in [0, 1].$$

All these contracts yield the same costs to the principal when the agent's effort is  $\widehat{a}$ . Denote expected utility and marginal expected utility by  $U(\widehat{a}, \beta)$  and  $U_a(\widehat{a}, \beta)$  respectively. For all  $\beta < 1$ ,  $U(\widehat{a}, \beta) > U(\widehat{a}, 1) = u_0$  because  $\beta$  is a mean preserving decrease in risk. From the previous argument  $U_a(\widehat{a}, 1) > 0$ . Therefore, by continuity, for some  $\beta < 1$  both the  $IR$  and  $ICR$  constraints are slack. For  $\eta$  positive but small enough, they will remain slack under the contract

$$\widetilde{w}(p) \equiv C(\widehat{a}, w) + \beta\varepsilon(p) - \eta$$

which costs less than  $C(\hat{a}, w)$ . QED

**Proof of Lemma 3.** Denote by  $Q_Z(p, a)$  the quantile function of  $Z$ . The quantile of  $\varphi(Z)$  is then  $Q_\varphi(p, a) \equiv \varphi(Q_Z(p, a))$ . First, observe that the distribution of  $\varphi(Z)$  satisfies CCQ when  $Z$  does. To see this, use the argument in Step 1 of Proposition 2 with  $Q_Z(p, a) \equiv w(p, a)$  and  $\varphi \equiv u$ . Next we that the distribution of  $\min(\varphi(Z), k)$  also satisfies CCQ. Its quantile function is  $Q_m(p, a) \equiv \min(\varphi(Q_Z(p, a)), k)$ . Obviously

$$\tilde{Q}_m(p, a) \equiv \int_0^p Q_m(p, a) d\eta \leq \tilde{Q}_\varphi(p, a) \equiv \int_0^p Q_\varphi(p, a) d\eta.$$

It suffices to show that the concavity of  $\tilde{Q}_\varphi(p, a)$  in  $a$  implies that of  $\tilde{Q}_m(p, a)$ . Let  $\hat{p}(a)$  be the the solution to  $Q_\varphi(\hat{p}(a), a) = k$ . When  $p \leq \hat{p}(a)$ ,  $\lambda \in [0, 1]$  and  $a = \lambda a' + (1 - \lambda)a''$  imply

$$\begin{aligned} \tilde{Q}_m(p, a) &= \tilde{Q}_\varphi(p, a) \\ &\geq \lambda \tilde{Q}_\varphi(p, a') + (1 - \lambda) \tilde{Q}_\varphi(p, a'') \\ &\geq \lambda \tilde{Q}_m(p, a') + (1 - \lambda) \tilde{Q}_m(p, a''). \end{aligned} \tag{13}$$

For  $p > \hat{p}(a)$ ,

$$\tilde{Q}_m(p, a) = \tilde{Q}_\varphi(\hat{p}(a), a) + (p - \hat{p}(a))k.$$

Observe that

$$(p - \hat{p}(a))k \geq \lambda \int_{\hat{p}(a)}^p Q_m(\eta, a') d\eta + (1 - \lambda) \int_{\hat{p}(a)}^p Q_m(\eta, a'') d\eta.$$

Combining with (13) yields

$$\begin{aligned} \tilde{Q}_m(p, a) &= \tilde{Q}_\varphi(\hat{p}(a), a) + (p - \hat{p}(a))k \\ &\geq \lambda \tilde{Q}_m(\hat{p}(a), a') + (1 - \lambda) \tilde{Q}_m(\hat{p}(a), a'') \\ &\quad + \lambda \int_{\hat{p}(a)}^p Q_m(\eta, a') d\eta + (1 - \lambda) \int_{\hat{p}(a)}^p Q_m(\eta, a'') d\eta \\ &\geq \lambda \tilde{Q}_m(p, a') + (1 - \lambda) \tilde{Q}_m(p, a''), \end{aligned}$$

which completes the proof. QED

**Proof of Proposition 3.** Write the contract solving  $R$  as  $y(l)$ ,  $l \in [\underline{l}, \bar{l}]$  where  $\underline{l} = f_a(\underline{x} | \hat{a})/f(\underline{x} | \hat{a})$ ,  $\bar{l} = f_a(\bar{x} | \hat{a})/f(\bar{x} | \hat{a})$ .

*Step 1.* Suppose that  $y(l)$  is continuous and the  $BP$  constraints are not binding for  $l \in [l_0, l_1] \subset [\underline{l}, \bar{l}]$ . By Lemma 2 and because  $\Delta(y(l), \mu) = \lambda + \mu l$ , it follows that  $y(l)$  is strictly increasing in that interval. Hence,  $\Delta_y(y, \mu) \geq 0$  for  $y \in [y(l_0), y(l_1)]$  with strict inequality almost everywhere. We show that condition (iii) then implies  $\Delta_{yy}(y, \mu) \geq 0$  for  $y \in (y(l_0), y(l_1))$ . For  $y$  in this interval,

$$\Delta_y(y, \mu) = \frac{\mu u''' u' - u''(1 + \mu u'')}{(u')^2} \geq 0. \quad (14)$$

Then

$$\begin{aligned} \Delta_{yy}(y, \mu) &= \frac{\mu u'''' u' - u'''(1 + \mu u'')}{(u')^2} - 2u'' \frac{\mu u''' u' - u''(1 + \mu u'')}{(u')^3} \\ &= \frac{\mu u'''' u' - u'''(1 + \mu u'')}{(u')^2} + 2 \left( -\frac{u''}{u'} \right) \Delta_y(y, \mu). \end{aligned} \quad (15)$$

Note that  $p' = (u'''/u'')^2 - (u''''/u'')$ . Hence  $u'''' \geq (u''')^2/u''$  is equivalent to  $p' \geq 0$ . Under this condition,

$$\begin{aligned} \Delta_{yy}(y, \mu) &\geq \frac{\mu \left( \frac{(u''')^2}{u''} \right) u' - u'''(1 + \mu u'')}{(u')^2} + 2 \left( -\frac{u''}{u'} \right) \Delta_y(y, \mu) \\ &= (2r - p) \Delta_y(y, \mu). \end{aligned}$$

Hence,  $p \leq 2r$  and  $p' \geq 0$  are sufficient for  $\Delta_{yy} \geq 0$  whenever  $\Delta_y \geq 0$ . Noting that  $r' = r(r - p)$ , differentiating (14) once more is easily seen to yield

$$\Delta_{yy}(y, \mu) = \frac{r(2r - p)}{u'} - \mu r [(r - p)(2r - p) - p'].$$

Hence,  $p \leq 2r$  and  $p' \geq r'(2 - p/r)$  are also sufficient for  $\Delta_{yy} \geq 0$ .

*Step 2.* Condition (i) implies that  $y(l)$  is an interior solution solving  $\Delta(y, \mu) = \lambda + \mu l$  for all  $l \in [\underline{l}, l_1]$  for some  $l_1 \in (\underline{l}, \bar{l}]$ . Without loss of generality,

we can discard the possibility that  $y(l)$  is discontinuous at  $\underline{l}$  (otherwise it is an isolated point). Thus,  $y(l)$  is continuous in a right neighborhood of  $\underline{l}$  and Step 1 therefore applies; that is, for  $y(l)$  in a right neighborhood of  $\underline{l}$ ,  $y(l)$  is strictly increasing,  $\Delta_y(y(l), \mu) \geq 0$  with strict inequality almost everywhere and  $\Delta_{yy}(y(l), \mu) \geq 0$ . But this then implies that the same is true for all  $y > y(l)$ . Hence the contract satisfies

$$y(l) = \min(\Delta^{-1}(\lambda + \mu l), \bar{y}),$$

where  $\Delta^{-1}$  is the inverse of  $\Delta(y, \mu)$  with respect to its first argument. Because  $\Delta^{-1}$  is increasing concave and applying Lemma 3, the distribution function of the payment then satisfies CCQ if the distribution of  $f_a(X | \hat{a})/f(X | \hat{a})$  does. The latter, again using Lemma 3, is ensured by condition (ii). QED

## References

- Abraham, A., Koehne, S. and N. Pavoni (2011), “On the First-Order Approach in Principal-Agent Models with Hidden Borrowing and Lending.” *Journal of Economic Theory* 146, 1331-1361.
- Alvi, E. (1997), “First-Order Approach to Principal-Agent Problems: A Generalization.” *Geneva Papers on Risk and Insurance Theory* 22, 59-65.
- Arnott, R. and J.E. Stiglitz (1988), “Randomization with Asymmetric Information.” *RAND Journal of Economics* 19, 344-362.
- Carrier, G. and R.A Dana (2005), “Existence and Monotonicity of Solutions to Moral Hazard Problems.” *Journal of Mathematical Economics* 41, 826-843.
- Conlon, J.R. (2009), “Two New Conditions Supporting the First-Order Approach To Multisignal Principal-Agent Problems.” *Econometrica* 77, 249-278.

- Eeckhoudt, L. and Ch. Gollier (2005), “The Impact of Prudence on Optimal Prevention.” *Economic Theory* 26, 989-994.
- Grossman, S. and O. Hart (1983), “An Analysis of the Principal-Agent Problem.” *Econometrica* 51, 7–45.
- Holmstrom, B. and P. Milgrom (1987), “Aggregation and Linearity in the Provision of Intertemporal Incentives.” *Econometrica* 55, 303-328.
- Jewitt, I. (1988), “Justifying the First-Order Approach to Principal-Agent Problems.” *Econometrica* 51, 1177-1190.
- Jewitt, I., O. Kadan, J.M. Swinkels (2008), “Moral Hazard with Bounded Payments.” *Journal of Economic Theory* 143, 59-82.
- Jullien, B., R. Salanié and F. Salanié (1998), “Should More Risk-Averse Agents Exert More Effort?” *Geneva Papers on Risk and Insurance Theory* 24, 19-28.
- Kadan, O., Ph.J. Reny and J.M. Swinkels (2011), “Existence of Optimal Contracts in the Moral Hazard Problem.” MFI working paper no 2011-02.
- Kawasaki, S. and J. McMillan (1987), “The Design of Contracts: Evidence from Japanese Subcontracting.” *Journal of the Japanese and International Economies* 1, 327-349.
- LiCalzi, M. and S. Spaeter (2003), “Distributions for the First-Order Approach to Principal-Agent Problems.” *Economic Theory* 21, 167-173.
- McAfee, P. and J. McMillan (1986), “Bidding for Contracts: A Principal-Agent Analysis.” *RAND Journal of Economics* 17, 326-328.
- Martimort, D. and J. Pouyet (2008), “To Build or Not to Build: Normative and Positive Theories of Public-Private Partnerships.” *International Journal of Industrial Organization* 26, 393-411.

- Meyer, D.J. and J. Meyer (2011), “A Diamond-Stiglitz Approach to the Demand for Self-Protection.” *Journal of Risk and Uncertainty* 42, 45-60.
- Mirrlees, J. (1999) “The Theory of Moral Hazard and Unobservable Behavior: Part I.” *Review of Economic Studies* 66, 3-21.
- Rogerson, W.P. (1985), “The First-Order Approach to Principal-Agent Problems.” *Econometrica* 53, 1375–67.
- Sinclair-Desgagné, B. (1994), “The First-Order Approach to Multidimensional Principal-Agent Problems.” *Econometrica* 62, 459-465.
- Thiele, H. and A. Wambach (1999), “Wealth Effects in the Principal Agent Model.” *Journal of Economic Theory* 89, 247-260.