The Form of Incentive Contracts: The Limited Liability Agency Model with Moral Hazard*

Joaquín Poblete[†], Daniel Spulber[‡]

June 2010

Abstract

The paper obtains and characterizes the optimal agency contract in a model with moral hazard and limited liability. The framework features a parametric representation of uncertainty. The analysis shows that the optimal contract can be characterized based on the properties of a critical ratio that describes incentives per unit of expected return in each state. The critical ratio equals the hazard rate of the otcome times the marginal return to effort. The properties of the critical ratio hold for a wide range of problems in economics and finance. The analysis yields a general framework for comparing moral hazard with adverse selection.

Key Words: Agency, Moral Hazard, Debt, Limited Liability.

---

*We would like to thank Sandeep Baliga, Marco Ottaviani, Yuk-fai Fong, Yogmin Chen, Tom Gresik and Alessandro Pavan for interesting comments. Poblete is grateful to the Searle Center for Law, Regulation and Economic Growth for research support. Spulber is grateful for the support from a research grant from the Ewing Marion Kauffman Foundation on Entrepreneurship.

[†]London School of Economics , Houghton Street, London WC2A 2AE, United Kingdom E-mail: j.j.poblete-lavanchy@lse.ac.uk

[‡]Elinor Hobbs Distinguished Professor of International Business, Management & Strategy, Kellogg School of Management, Northwestern University, 2001 Sheridan Road, Evanston, IL, 60208. E-mail: jems@kellogg.northwestern.edu

1 Introduction

Incentive contracts between principals and agents are fundamental for a wide range of economic and financial transactions. Principals apply agency contracts to provide incentives for employees, managers, entrepreneurs, insurance buyers, sales personnel, business representatives, independent contractors, tenant farmers, and regulated firms. When the agent's action is unobservable, the optimal agency contract is designed to maximize the joint benefits of the parties while mitigating moral hazard. We derive a novel and intuitive characterization of the optimal contract in an environment that corresponds to a wide range of applied problems in economics and finance.

We study a model where the outcome of a task performed by the agent depends on the effort of the agent and a random shock. The principal observes the outcome of the task, but cannot observe the agent's effort or the random shock. The principal is unable to sell the task to the agent because of the agent's limited liability constraint, which prevents the contract from attaining the first-best outcome. The second best contract effectively sells the agent the proceeds from the task in those states that provide more incentives per unit of expected return. The main result of the analysis characterizes the optimal contract based on the properties of a critical ratio that describes incentives per unit of expected return in each state. Intuitively the critical ratio corresponds to the rate of return of increasing the slope of the compensation for the agent in each state. The critical ratio equals the hazard rate of the outcome distribution times the marginal return to effort.

The critical ratio provides an intuitive and easily derived condition for characterizing the form of the optimal contract. When the critical ratio is increasing in the random shock, higher states are better for providing incentives. Therefore, when the critical ratio is increasing in the random shock, it follows that the optimal contract is debt. When the critical ratio is strictly decreasing in the random shock, lower states are more efficient at providing incentives and the optimal contract is a call option. When the critical ratio is linear in the random shock, all states are equally efficient and only in that case are linear contracts efficient. Applying these types of conditions in a moral hazard setting yields a consistent characterization of the optimal contract based on assumptions that can be easily verified and occur in most applications in economics and finance.

Because the critical ratio is the hazard rate of the outcome multiplied by the marginal return to effort, determining its properties is straightforward. A sufficient condition for the critical ratio to be increasing in the random shock is for both of its factors to be increasing. This property holds for a wide range of applied models in economics and finance. The hazard rate is nondecreasing for many probability distributions including the exponential family of distributions and the uniform distribution. Effort and the shock are complements when the marginal return to effort is increasing in the random shock. Complementarity is consistent with many types of models including additive and multiplicative outcome functions. Outcome functions such that the marginal return to effort is increasing in the random shock are used in models with uncertainty regarding prices, taxes, subsidies, outputs, and technology. Such outcome functions also occur in models with uncertainty regarding discount rates, depreciation rates, failure rates, and natural shocks such as weather and demographic changes.

The critical ratio approach introduced here for moral hazard models turns out to be analogous to familiar conditions in adverse selection models. First, requirement that the probability distribution has a non-decreasing hazard rate corresponds to a common requirement for the distribution of agent types in adverse selection problems. Second, the requirement that the random shock and the agent's effort are complementary corresponds to the familiar single-crossing or Spence-Mirrlees condition in adverse selection problems. An important benefit of our approach is that it allows us to identify fundamental connections between moral hazard and adverse selection.

This approach highlights the importance of wealth constraints, see Innes (1990) and Holmstrom and Tirole (1997). Our approach is related to that of Innes (1990) who shows that the optimal agency contract takes the form of debt. Innes requires contracts to be chosen such that the principal's benefit is monotonic in output. This feasibility condition is desirable because it guarantees that the principal does not have an incentive to sabotage the outcome. Otherwise, the principal would wish to impede the agent's efforts so as to avoid paying the reward for effort. The monotonicity requirement also is desirable in situations where agents can shirk and costlessly reduce their performance reports. Given the monotonicity condition, Innes shows that the optimal contract must resemble debt. Our results confirm Innes' conclusion about the optimality of debt-style contracts while substantially extending the analysis of optimal contracts and providing new intuition for the results. Our results remedy a shortcoming in Innes' (1990) analysis. His main assumptions, including implementability and MLRP, are difficult to satisfy and do not hold in most economics and finance analyses.

Our state space representation differs from the now standard approach of working directly with the induced probability distribution on outcomes that depends on effort. Explicit recognition of uncertainty is present in Spence and Zeckhauser (1971), Ross (1973), and Harris and Raviv (1976) who apply a first-order approach with risk-averse agents. The reduced-form approach originates with Mirrlees (1974, 1976) and Holmstrom (1979). By considering uncertainty explicitly, we can obtain a characterization of the optimal contract based on the underlying form of risk and the production function. Parametric uncertainty occurs in models with demand uncertainty, supply uncertainty, and strategic uncertainty. Parametric uncertainty can take the form of uncertain lotteries with linear probabilities and random financial valuations. Parametric uncertainty is important because it is consistent with empirical analysis in economics and finance.

There is a natural correspondence between the state space and the induced distribution representation. We discuss this correspondence in the paper and show that the critical ratio depends on the hazard rate of the induced distribution and on changes in the marginal return to effort. The critical ratio cannot be represented in terms only of the induced distribution and thus the state space representation becomes useful.

Our analysis of the problem of moral hazard with a state space representation has implications for various economic and finance applications. The optimality of debt-style contracts implies that such contracts perform better under uncertainty than other contractual forms.

Perhaps most significantly, debt-style contracts perform better than equity-style sharing rules under uncertainty. The economic model of agency has its origins in labor contracts in agrarian economics, particularly sharecropping and piece-rate labor contracts, see Otsuka et al. (1992). Economic studies often assume that the contract is a linear sharing rule, appealing to the dynamic aggregation result of Holmstrom and Milgrom (1987). Debt-style contracts clearly have empirical implications. In practice, debt-style contracts are widely used and correspond to all kinds of compensation agreements with threshold effects, such as bonuses for employees.

Our analysis showing the greater efficiency of debt in comparison with equity also has important implications for managerial finance. Debt-style contract provide a means of reducing the manager's incentives to shirk. This addresses the vast literature on corporate finance, beginning with the work of Jensen and Meckling (1976). When managers are risk neutral and have limited liability, debt-style contracts perform better that equity in providing incentives to perform. Debt-style contracts also correspond to financial assets including securities and bonds whose features resemble options, see for example Cox and Rubinstein (1985). Real options are an important tool for analyzing investment under uncertainty, see Dixit and Pindyck (1994).The simplicity of debt-style contracts and their optimality in a broad range of environments helps to explain their widespread application.

The paper is organized as follows. Section 2 presents the basic agency model. Section 3 derives the optimal contract under different conditions. Section 4 discusses the main result and compare it with existing results. Section 5 Concludes.

2 The Basic Model of Agency

Consider two risk-neutral economic actors who enter into a contract. The actor designated as the principal owns a task and the actor designated as the agent performs the task. The task can represent various projects including the following: the agent performs a service under authority delegated by the principal, the agent produces a good within a firm owned by the principal, or the agent is an entrepreneur who uses financial capital provided by the principal to establish a firm.

The agent provides the production technology for the project and supplies productive effort, a. The agent owns a production technology given by

$$\Pi = \Pi(\theta, a), \tag{1}$$

where Π represents the outcome and θ is a random variable. The principal observes the outcome but cannot observe either the agent's effort a or the random variable θ. The agent chooses effort before the realization of the random variable occurs. Assume that Π is twice differentiable in a and θ. The random variable θ has a density function $f(\theta)$ with support $[0, \overline{\theta}]$. We also allow for the support of $f(\theta)$ to be $[0, \infty)$.

Assumption 1 *The outcome function, $\Pi(\theta,a)$, is increasing in θ.*

This assumption allows us to define $\widehat{\theta}(a, \Pi)$ as the shock that satisfies

$$\Pi(\widehat{\theta}(a, \Pi), a) = \Pi, \tag{2}$$

for an effort level a, and realization of revenues, Π.

Assumption 2 *The outcome function, $\Pi(\theta,a)$, is increasing and concave in a.*

Assumption 2 guarantees the existence of a first best effort level $a^{FB} = \arg\max_a \Pi(\theta, a) - a$, and implies that the induced distribution satisfies first order stochastic dominance in a. To guarantee that $a^{FB} > 0$ we further assume that $lim_{a \to 0} \frac{\partial}{\partial a} \Pi(\theta, a) = \infty$.

The contract can be based only on the outcome Π because the agent's choice of effort a, and the realization of the random variable θ, are not observable to the principal. The agent has a disutility of effort given by a. The contract between the principal and the agent is fully described by a function of the realized benefit,

$$w = w(\Pi). \tag{3}$$

The agent′s net benefit is given by

$$u(w, a, \Pi) = w(\Pi) - a. \tag{4}$$

Given the contract, w, the agent chooses effort to maximize his expected net benefit,

$$U(w, a) = \int_0^{\overline{\theta}} w(\Pi(a, p))f(p)dp - a. \tag{5}$$

The agent's effort is said to satisfy the incentive compatibility constraint if

$$a \in \arg\max_a \int_0^{\overline{\theta}} w(\Pi(\theta, a))f(\theta)d\theta - a.$$

The principal's cost of providing the task to the agent is $K > 0$. For any realization of Π, the principal's net benefit is

$$v(w, \Pi) = \Pi - w(\Pi) - K. \tag{6}$$

The principal's expected net benefit given the form of the outcome function and the distribution of the random shock equals

$$V(w, a) = \int_0^{\overline{\theta}} (\Pi(\theta, a) - w(\Pi(\theta, a)))f(\theta)d\theta - K. \tag{7}$$

The optimal contract maximizes the net benefit of the agent subject to an individual rationality constraint for the principal. The principal's individual rationality constraint is $V(w, a) \geq 0$.

We define a feasible contract based on several requirements.

Definition 1 *A contract w is said to be feasible if it satisfies two conditions. (a) The principal's and agent's net benefit, $v(w, \Pi)$, and $u(w, \Pi)$ are non-decreasing in Π, b) The agent's payments are non-negative $w \geq 0$.*

The monotonicity requirement on the principal's net benefit, $v(w, \Pi)$, follows Innes (1990). As Innes explains, making the principal's net benefit non-decreasing in Π rules out situations in which the parties may subvert the contract. Otherwise, if the principal's return is decreasing in the outcome, the principal may attempt to sabotage the task to avoid making payments to the agent. Alternatively, an entrepreneur may borrow money to supplement the return of the firm and thereby increase his returns based on reported performance. The monotonicity requirement rules out forcing contracts that would lead to this type of behavior. The monotonicity requirement for the agent plays a similar role and it greatly simplifies the proofs, however it is not necessary to prove our main result. [1] The non-negativity requirement represents the fact that the agent has limited wealth, and it prevents the agent from buying the firm and achieving the first best.

Definition 2 *An effort level a is implementable if there exists a feasible contract w such that a)The effort level a is incentive compatible $a \in \arg\max U(w,a)$, and b) The principal's participation constraint holds $V(w,a) \geq 0$.*

We assume that the set of implementable effort levels is non empty. This is basically a requirement that the principal's cost, K, is not too large.

Assumption 3 *There exists an implementable effort level $a > 0$.*

Of course the model can be re-phrased in terms of the mirrlees representation. We study the mirrlees representation in more detail in section 2.

3 The Optimal Contract.

The optimal contract between the principal and the agent maximizes the agent's expected net benefit over the set of feasible contracts,

$$\max_{w(.),a} U(w,a) \text{ subject to}$$

$$a \in \arg\max U(w,a),$$

$$v(w,\Pi) \text{ is non-decreasing in } \Pi,$$

$$V(w,a) \geq 0,\ w(\cdot) \geq 0.$$

The maximization of the agent's net benefit helps to highlight the effects of the agent's limited liability. There is no loss of generality. The results are consistent with analyses of

[1] As explained later we don't use this assumption in the proof of proposition 3.

agency that maximize the principal's net benefit. Varying the agent's endowment generates other allocations of rents between the principal and the agent without changing the characterization of the optimal contract.

To characterize the optimal contract we first define the critical ratio.

Definition *The critical ratio is*

$$\rho(\theta, a) = \frac{f(\theta)}{1 - F(\theta)} \frac{\Pi_a}{\Pi_\theta}.$$

Observe that the critical ratio is given by the product of the hazard rate of the distribution of θ times the ratio of the marginal return to effort to the marginal effect of the shock in the output. Assumptions on the hazard rate are commonly used in adverse selection models such as auctions and nonlinear pricing. A large class of distributions satisfies monotone hazard rates. Hazard rate conditions as applied in studies of reliability lends itself to empirical testing and calibration [2] .The second term Π_a/Π_θ is the marginal return to effort over the marginal return to the shock, this term is increasing in θ if there are complementarities between the effort of the agent a and the state of nature θ. Similar expressions appear frequently in the literature on adverse selection where it is referred to as the single-crossing or Spence-Mirrlees condition. In section *4.2*, equation (27) we express this ratio as a function of the induced distribution and show that it corresponds to the hazard rate of the induced distribution times the marginal return to effort.

It will be shown that the optimal agency contract depends on the behavior of the critical ratio. We will show this by describing the optimal contract in three cases. First when the critical ratio is constant in θ we show that linear incentive schemes are optimal, second when the critical ratio is decreasing in θ we show that the optimal contract is an option contract, and third we show that when the critical ratio is increasing the optimal contract is standard debt. The intuition for the relation between the critical ratio and the optimal contract is presented in the next section.

[2] See for example Hall and Van Keilegom (2005).

Linear contracts are often used in practice, taking the form of sharecropping in agriculture, piece rates in manufacturing, cost sharing in procurement, sales commissions, and other types of proportional reward sharing between principals and agents. Linear contracts are often assumed in studies of agency in economics and finance.

Proposition 1 *If the critical ratio* $(\Pi_a/\Pi_\theta)f(\theta)/(1-F(\theta))$ *is constant in* θ*, then an optimal incentive scheme exists and linear contracts are optimal.*

PROOF First observe that with a linear contract of the form

$$w(\Pi) = l\Pi, \tag{8}$$

We will show that without loss of optimality we can restrict to linear contracts. To prove this let $w()$ be an arbitrary contract that implements $\widehat{a}$, we will show that there exists a linear contract that implements the same effort and gives both agents the same expected utility. First notice that feasible contracts are continuous and weakly increasing (otherwise payments are decreasing for either the principal or the agent) and therefore differentiable a.e, let $w'(\Pi)$ be the derivative of $w(\Pi)$ with respect to Π, then at $\widehat{a}$ the contract $w(\Pi)$ must satisfy the first order condition

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)d\theta = 1. \tag{9}$$

Let $\widehat{l}$ be the linear contract that at $\widehat{a}$ satisfies

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)d\theta = \widehat{l}\int_0^{\overline{\theta}} \Pi_a f(\theta)d\theta. \tag{10}$$

It is clear that $\widehat{l}$ is between 0 and 1 since w' is bounded between 0 and 1 (otherwise payments are decreasing for either the principal or the agent). It is also clear that a is incentive compatible with the contract $\widehat{l}$ since Π is concave in a . I will show that this linear contract$\widehat{l}$ gives both the agent and the principal the same expected utility and implements the same effort as the contract $w()$. Since ρ is constant and given (10) it must be the case that

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)/\rho d\theta = \widehat{l}\int_0^{\overline{\theta}} \Pi_a f(\theta)/\rho d\theta. \tag{11}$$

Replacing ρ this can be written as

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_\theta(1 - F(\theta))d\theta = \widehat{l} \int_0^{\overline{\theta}} \Pi_\theta(1 - F(\theta))d\theta. \tag{12}$$

Which integrating by parts and replacing can be written as

$$\int_0^{\overline{\theta}} w(\Pi)f(\theta)d\theta = \widehat{l} \int_0^{\overline{\theta}} f(\theta)d\theta \tag{13}$$

The last equation shows that the linear contract $\widehat{l}$ gives the agent the same expected utility as $w()$. The expected benefit of the principal in both cases is also the same since both players are risk neutral and under both contracts the same action is being induced and thus the same expected surplus is created. So we have shown that without loss of optimality we can restrict attention to linear incentive schemes. Existence of an optimal contract follows from the fact that without loss of optimality we can restrict to linear contracts and the fact that the set of linear contracts is compact.□

In this case, however linear contracts are not the only optimal contracts. From the proof it is easy to see that any contract that implements an effort level equal to the optimal linear contract gives the agent the same expected utility and therefore it is also optimal.

An example of a production function of this kind is $\Pi(\theta, a) = \theta + Q(a)$ where $Q(a)$ is a concave function and $f(\theta)$ is exponential.

Call option style contracts are used in practice in finance when the buyer has the right but not the obligation to buy a commodity or financial asset at a particular time and at a strike price from the seller. Call options are offered in practice as incentive stock options for managers and employees.The next proposition shows that call options are optimal when ρ is decreasing.

Proposition 2 *If the critical ratio* $\rho = (\Pi_a/\Pi_\theta)f(\theta)/(1 - F(\theta))$ *is decreasing in the shock*

θ, then a unique optimal contract exists and it takes the form of a call option contract,

$$\Pi - w(\Pi) = Max\{\Pi - r, 0\} \;\; \textit{for some} \;\; r > 0.$$ [3]

Proof: See the appendix.

Any additive concave technology with any distribution with decreasing hazard rate satisfy the conditions of proposition 2.

The third proposition deals with the case when the ratio is increasing in θ. In this case however we need to make a further assumption regarding the concavity of the problem, this assumption replaces the CDF or GCDF conditions usually used in moral hazard problems. The proposition shows that the optimal contract takes the form of debt. Letting r be the face value of the debt, the agent obtains nothing if the outcome is less than or equal to the face value of the debt. Otherwise, the agent obtains the difference between the outcome and the face value of the debt. With a debt contract, the agent chooses the effort level that solves:

$$\max_a \int_0^{\overline{\theta}} \max\{\Pi(\theta, a) - r, 0\} f(\theta) d\theta - a, \tag{14}$$

The next condition is a regularity condition that rules out pathological cases when there is a debt contract.

Weak implementability condition *There exists at most one local interior maximum in a standard debt contract.*

The condition is not stated in terms of the parameters of the model, but as shown in the next examples it is easily satisfied.[4] We have been unable to find an example with an increasing

[3] Remember that $\Pi - w(\Pi)$ is payment to the principal, so as usual we look at payments from the principal's perspective.

[4] In terms of the parameters of the model, a sufficient condition for the weak implementability condition to hold is the following.

$$\int_{\widehat{\theta}(r,a)}^{\overline{\theta}} \Pi_{aaa}(\widehat{\theta}, a) f(\theta) d\theta - 2\frac{\partial \widehat{\theta}}{\partial a} \Pi_{aa}(\widehat{\theta}, a) f(\widehat{\theta}) + \frac{\partial^2 \widehat{\theta}}{\partial a^2} \Pi_{aa}(\widehat{\theta}, a) f(\widehat{\theta}) < 0 \; \forall a, r$$

This is equivalent to a concavity condition on Π.

critical ratio that fails to satisfy the weak implementability condition.

To illustrate how easy it is to satisfy this condition, consider the additively-separable model,

$$\Pi(\theta, a) = \theta + Q(a). \tag{15}$$

In this case the agent gets paid as long as $\theta > r - Q(a)$ and therefore the agent's problem can be expressed as

$$\max_a \int_{r-Q(a)}^{\overline{\theta}} [\theta + Q(a) - r] f(\theta) d\theta - a \tag{16}$$

Any local interior maximum satisfy the first order condition,

$$Q_a \int_{r-Q(a)}^{\overline{\theta}} f(\theta) d\theta = 1 \tag{17}$$

The second order condition for a maximum is:

$$Q_{aa} \int_{r-Q(a)}^{\overline{\theta}} f(\theta) d\theta + Q_a^2 f\left(r - Q(a)\right) \leq 0 \tag{18}$$

The second order condition is satisfied as long as:

$$\frac{Q_{aa}}{Q_a^2} + \frac{f\left(r - Q(a)\right)}{1 - F(r - Q(a))} \leq 0 \tag{19}$$

If (19) is decreasing; then the weak implementability condition is satisfied. The reason is that if two local interior maximums exist then there exists a local minimum in between two local maximums; but that is incompatible with the second derivative being decreasing in a.

The second term in (19) is the hazard rate, it is decreasing in a if the distribution has a monotone hazard rate. If the function Q is sufficiently concave, the first term in (19) will also be decreasing, and the weak implementability condition will hold. For example if $Q(a) = Ln(a + 1)$ the condition is satisfied for any distribution of θ with monotone hazard rate. Recall that these conditions are sufficient, many additive examples that do not satisfy these properties still satisfy the weak implementability condition. To further illustrate the generality of the assumptions, we show the multiplicatively separable case in the appendix.

We now present our main result. The optimal contract takes the form of a debt obligation. Debt-style contracts are not only widely used in finance but also are commonly applied in the form of incentives with bonuses and other rewards based on thresholds.

Proposition 3 *If the critical ratio is increasing in θ, and the weak implementability conditions hold, then a unique optimal contract exists and it takes the form of standard debt,*

$$\Pi - w(\Pi) = \min\{\Pi, r\} \text{ for some } r \geq 0.$$

Proof: See the appendix.

For generality, in the proof of proposition 3 we don't assume that the agent's net benefit needs to be non-decreasing in output. An example that satisfies all the assumption in proposition 3 is $\Pi(\theta, a) = \theta\sqrt{a}$ with $\theta \sim \exp(\lambda)$.

Propositions 1,2 and 3 are important because they specify the form of the optimal contract in an agency model with moral hazard and limited liability. The underlying assumptions governing uncertainty and the production function are readily verified and fit with a wide range of standard models in economics and finance. It should be emphasized that the contract obtained in Proposition 3 applies to any type of incentive contract with moral hazard. This includes situations that do not involve financial obligations, including regulation, procurement, and incentives for effort. The face value of the debt specified by the contract is a cut-off level such that the agent receives no payment when the outcome is below the cut-off and the agent receives the difference between the outcome and the cut-off value otherwise.

4 Discussion and Extension.

4.1 Intuition and the Critical Ratio

This section provides an intuitive explanation for the critical ratio and how it determines the optimal contract. The general idea is that contracts have two roles, to provide the agent with incentives and to compensate the principal for the task. The optimal contract is the one that combines these roles most effectively. To explain the intuition in more detail, assume that the solution to the problem is continuous and differentiable. Moreover suppose that it is the case that the first order condition approach is valid, and that any effort level is implementable with some debt contract. Then, given a contract $w(\Pi)$, the effort of the agent is determined by the first order condition

$$\int_0^{\overline{\theta}} w_{\Pi}(\Pi(\theta, a))\Pi_a(\theta, a)f(\theta)d\theta = 1, \tag{20}$$

where $w_\Pi(\Pi)$ is used for $\partial w(\Pi)/\partial\Pi$ and Π_a for $\partial\Pi(\theta, a)/\partial a$. The derivative of the left-hand side of equation (20) with respect to the slope of the compensation (w_Π) is $\Pi_a f(\theta)$. The higher $\Pi_a f(\theta)$ is, the less we need to increase the slope of the payoff w_Π to induce a given effort level. The term $\Pi_a f(\theta)$ represents how powerful is a given state in providing incentives. Intuitively, the greater are the likelihood of a state and the marginal return to effort, the more efficient is the state in providing incentives.

The second role of the contract is to provide compensation. The benefit of the contract to the agent is given by,

$$\int_0^{\overline{\theta}} w(\Pi(\theta, a))f(\theta)d\theta - a.$$

We can rewrite the benefit as a function of w_Π using integration by parts

$$u(w, a) = \int_0^{\overline{\theta}} w_{\Pi}(\Pi(\theta, a))(1 - F(\theta))\Pi_\theta d\theta - a. \tag{21}$$

The derivative of the expected payoff with respect to w_Π is $(1 - F(\theta))\Pi_\theta$. The higher the term $(1 - F(\theta))\Pi_\theta$ is, the less we need to increase the slope of the payoff w_Π to provide a given compensation level to the agent. The term $(1-F(\theta))\Pi_\theta$ represents how efficient a state is in providing compensation. Intuitively the lower the revenue and the faster the revenue increases in the state of nature θ, the more efficient the state is in providing a compensation to the agent.

The optimal contract depends on the critical ratio of incentives per unit of compensation. The critical ratio is the product of the hazard rate and the marginal product of the agent's effort over the marginal product of the random shock in producing outcomes,

$$\rho(\theta, a) = \frac{f(\theta)}{1 - F(\theta)} \frac{\Pi_a}{\Pi_\theta}, \tag{22}$$

This ratio measures the relation between incentives and expected payoff when the slope of the contract w_Π changes. The ratio can be interpreted as the rate of return (from the principal's perspective) of increasing the slope of the compensation to the agent in a state θ. The optimal contract is the one that implements the highest level of effort at the minimum cost, since that contract extracts the most surplus from the task. If the ratio is increasing then high states provide more incentives per unit of compensation and the optimal contract is debt. If the ratio is constant then all states are equally efficient and linear contracts are optimal.

According to this intuition, we can generalize our results to cases when the critical ratio is not monotonic. For that we need to make the following definition.

Definition *A contract is of the class L if there exists $\lambda > 0$ so that $w'(\Pi) = 1$ if $\rho > \lambda$ and $w'(\Pi) = 0$ if $\rho < \lambda$.*

The next proposition shows that the L class of contracts is in general better from the principal's perspective.

Proposition 4 *Let l belong to the L class of contracts. Let a be an effort level that satisfies the IC constraint under the contract l and some arbitrary contract $w()$, then $V(w, a) \leq V(l, a)$.*

Proof See the appendix.

This proposition extends the idea that the optimal shape of the contract depends on the behavior of the critical ratio ρ.

4.2 Comparison with the Standard Model.

Of course the model can be expressed in terms of the reduced form distribution as is now standard in the moral hazard literature.(To see a more detailed explanation of the relation between the two representations, see Conlon (2009)). To do this observe first that given an effort level a for the output Π to be less than or equal to $\overline{\Pi}$ it must be the case that the state θ is less than or equal to $\widehat{\theta}(\overline{\Pi}, a)$. Let $G(\Pi, a)$ be the joint distribution, it must be the case that

$$G(\Pi, a) = F(\widehat{\theta}(\Pi, a)) \tag{23}$$

Differentiating

$$g(\Pi, a) = f(\widehat{\theta}(\Pi, a))\frac{\partial\widehat{\theta}(\Pi, a)}{\partial\Pi}, \tag{24}$$

also from the identity

$$\Pi(\widehat{\theta}(\Pi, a), a) = \Pi \tag{25}$$

it is straightforward to check that

$$\frac{\partial\widehat{\theta}(\Pi, a)}{\partial\Pi} = 1/\Pi_\theta \tag{26}$$

Replacing this we can express the critical ratio as

$$\rho(\theta, a) = \frac{f(\theta)}{1 - F(\theta)}\frac{\Pi_a}{\Pi_\theta} = \frac{g(\Pi, a)}{1 - G(\Pi, a)}\Pi_a \tag{27}$$

Unfortunately we can't express the monotonicity of the critical ratio in terms only of the induced distribution and therefore the more structural state space representation is preferred.

In our previous example when $\Pi(\theta, a) = \theta Q(a)$ for some concave function Q and $\theta \sim \exp(\lambda)$, the joint distribution becomes $g(\Pi, a) = \lambda/q(a)e^{-\lambda\Pi/q(a)}$. In this case the optimal contract is debt.

The standard assumptions in the literature using the reduced form just introduced are the

Monotone Likelihood ratio Property (MLRP) and the Convex Distribution Function condition or CDFC. The last condition can sometimes been replaced by implementability assumptions weaker than CDFC. The objective of the next section is to show that our conditions (1-3) are different from standard conditions.

4.2.1 The MLRP Condition

We show by means of an example that our assumptions do not imply the standard MLRP condition on the induced distribution.

Multiplicative examples are particularly relevant since many shocks in economics such as prices and productivity shocks affect revenues multiplicatively. For example, in Stiglitz' (1974) classic model of sharecropping, the production function has the multiplicative form

$$\Pi = \theta Q(a) \tag{28}$$

Let the probability density be given by $f(\theta) = \gamma + \alpha\theta$ with $\alpha < 1$. If $Q(a)$ is concave enough (see the appendix) Π satisfies the assumptions of Proposition 1, but it does not satisfy $MLRP$. A proof of this is presented in the appendix.

4.2.2 The CDFC Condition.

The second standard assumption in the literature is the Convex Distribution Function Condition or CDFC. The condition can be stated as follows

CDFC Condition For every contract $w(\Pi)$, the problem $\max_a \int_0^{\overline{\theta}} w(\Pi) f(\theta) d\theta - a$ is concave.

Innes (1990) imposes an implementability assumption that is weaker than CDFC but stronger than our assumption. Innes (1990) requires the following.

Strong Implementability Condition "*When $w(\Pi)$ takes the standard debt functional form $w(\Pi) = \min\{\Pi, r\}$ for some $r > 0$, there is a unique solution to the agent's effort problem a.*"

We now show that the CDFC condition or the strong implementability condition cannot hold in the present model if the effort of the agent is needed to produce output and the shock has finite support. The assumption that effort is needed for output is used by Innes (1990) and others and is consistent with many applications used in previous literature like the multiplicative production function used by Stiglitz (1974).

Proposition 5 *The CDFC condition or the strong implementability condition does not hold if* $\Pi(\theta, 0) = 0$ *and the distribution of* θ *has finite support.*

Proof. See the appendix.

Jewitt (1988) observes that few distributions satisfy both the MLRP and CDFC conditions. One distribution was provided by Rogerson (1985) (attributed to Steve Matthews) and later two classes of differentiable examples were provided by Licalzi and Spaeter (2003). None of these examples satisfy all the conditions assumed in Innes(1990). [5]

4.3 Comparison with Adverse Selection

In agency models with moral hazard, the principal observes the outcome that results from the agent's effort and random shocks. The standard approach is to examine the probability distribution over outcomes induced by the agent's effort. In contrast, the critical feature of our modeling approach is that we explicitly consider random shocks rather than working with an induced distribution over outcomes. By examining the interaction between random shocks and the agent's effort, we can exploit a similarity between moral hazard models and adverse selection models. The agent has private information in both types of models. The agent has private information about his action in a moral hazard model and the agent in an adverse selection model has private information about his type but not his action. The key similarity between the two models is the correspondence between the random shock in

[5] Innes assumes the density is $g(\Pi|z)$ where Π satisfy all the following conditions: 1) $\frac{\partial}{\partial \Pi}\left(\frac{g_z(\Pi|z)}{g(\Pi|z)}\right) > 0$;2) Unique implementation with debt (Which is implied by CDFC); 3) $g(\Pi|z) > 0 \forall \Pi, z \geq 0$. 4) $\int \Pi g(\Pi|0) d\Pi = 0$.

a moral hazard model and the agent's private information in an adverse selection model. The main difference between the two types of models is that the agent in a moral hazard model moves before observing the outcome of the random shock whereas the agent in an adverse selection model moves after observing his type. The hazard rate and complementarity conditions play similar roles in both models.

Formally in the standard adverse selection, or "hidden information" model the agent has a type θ unknown by the principal. The agent can choose an output Π at a cost $c(\Pi|\theta)$. To induce the agent to pick the best possible output, the principal designs a contract $w(\Pi)$ that determines the payment to the agent contingent on the output Π.The agent's problem is given by

$$\max_{\Pi} w(\Pi) - c(\Pi|\theta) \tag{29}$$

A different interpretation is that the agent chooses the cost $c(\Pi|\theta)$ depending on his type θ and the contract w. Let $a = c(\Pi|\theta)$. Then the output that agent θ obtains if the cost of effort chosen is a is given by $\Pi(\theta, a) = c^{-1}(a|\theta)$.

The problem for the agent can be re-written as

$$\max_{a} w(\Pi(a, \theta)) - a. \tag{30}$$

If the agent were to choose an action before he becomes aware of his own type then the problem would be

$$\max_{a} \int_0^{\overline{\theta}} w(\Pi(a, \theta)) f(\theta) d\theta - a,$$

This is precisely the agent's problem in the moral hazard problem. In this context the only difference between the moral hazard and the adverse selection model is that in the adverse selection version we assume the type θ is known to the agent but the principal only observes the distribution $f(\theta)$, while in the moral hazard version the type of the agent θ is unknown by both the principal and the agent. The fact that the type is unknown transforms the hidden information (the type) into a hidden action (the effort). We could also interpret the difference between the models as a problem of timing. In the moral hazard model, the agent chooses an action before observing his type, while in the adverse selection model, the agent chooses an after observing his type.

The hazard rate plays a similar role in the moral hazard and the adverse selection problem. In the adverse selection problem this rate represents the trade off between providing incentives to the agent of type θ and the cost of providing payments to agents with higher types. In the moral hazard version the hazard rate represents the trade off between providing the agent with incentives for the case when the state of nature turns out to be θ and the fact that the payment in higher states will increase.

The complementarity condition also plays a similar role in the two models. Complementarity in adverse selection implies that agents with higher types are more suitable to exert effort because they have a higher return to effort. Complementarity in Moral Hazard implies that between equally likely states higher states are more suitable to induce effort because the expected return on effort is greater in higher states.

5 Conclusion

The paper develops a basic model of agency with risk neutrality and unobservable effort. Moral hazard inefficiencies arise as a consequence of the agent's limited liability. We characterize a critical ratio that determines the form of the optimal contract. This ratio is equal to the hazard rate of the shock times the marginal return on effort over the marginal return of the shock. When the critical ratio is increasing in the random shock, higher states are better for providing incentives. Therefore, when the critical ratio is increasing in the random shock, it follows that the optimal contract is debt. When the critical ratio is strictly decreasing in the random shock, lower states are more efficient at providing incentives and the optimal contract is a call option. When the critical ratio is linear in the random shock, all states are equally efficient and only in that case are linear contracts efficient. Applying these types of conditions in a moral hazard setting yields a consistent characterization of the optimal contract based on assumptions that can be easily verified and occur in most applications in economics and finance.

The result that the optimal contract takes the form of debt when monotonicity holds, has far reaching implications. Debt-style contracts help to explain the use of performance targets and rewards in a wide variety of economic situations. Therefore, debt-style contracts are optimal for sharecropping contracts, employee performance contracts, procurement contracts, and regulatory incentives. Moreover, debt-style contracts are at the heart of financial contracts and performance rewards for entrepreneurs and managers. Debt contracts are extremely simple to design and apply, and have important properties that allow for market pricing and trading.

Our analysis emphasizes explicit uncertainty by considering shocks to the output function. By directly examining the random variable that affects the outcome, we can apply two conditions that are familiar from the adverse selection model of agency. The two conditions are that the random variable has a monotonic hazard rate, and the agent's effort and the random variable are complements. They are easily verified for most economics and finance models simply by checking the distribution of the random variable and the form of the production technology. What is most important is that these conditions hold for a very wide range of applications in economics and finance. Our approach should help researchers to derive optimal contracts within many economic and financial models. The assumptions allow for tractable economic analysis that should yield comparative statics results pertaining to uncertainty, technology, and wealth effects.

The present analysis suggests that incentive contracts can have critical performance levels rather than more complex performance schedules. It may be useful to reevaluate many standard analyses of performance incentives by explicitly modeling uncertainty. Our analysis contrasts with the standard results based on an induced probability distribution. The standard results generally apply the MLRP and CDFC conditions which tend not to be satisfied by most models in economics and finance. Such models with implicit uncertainty tend to suggest that performance rewards take the form of piece-rate compensation and linear sharing rules. The present analysis suggests instead that agency contracts can feature critical performance targets, bonus schemes, and performance guarantees.

An important aspect of our analysis is that it identifies a close connection between moral hazard and adverse selection. The uncertainty in the moral hazard setting corresponds to the unobservable type in the adverse selection setting. What is critical is the timing of the agent's decision. Before choosing his effort level, the agent does not observe the outcome of uncertainty in the moral hazard model while the agent observes his type in the adverse selection model. The agent's effort in the moral hazard setting depends on the probability distribution on states of the world, which can be interpreted as the agent's future type. In the moral hazard setting, the agent's action thus depends on anticipating what will be his future type.

References

Conlon, J, 2009 "Two New Conditions Supporting the First-Order Approach to Multisignal Principal-Agent Problems" Econometrica, 77, pp 249-278.

Cox, J. C. and M. Rubinstein, 1985, "Options Markets", Englewood Cliffs, N.J.: Prentice Hall.

Dixit, A. K. and R. S. Pindyck, 1994,"Investment under Uncertainty", Princeton: Princeton University Press.

Diamond, D. W., 1994, "Financial Intermediation and Delegated Monitoring," Review of Economic Studies, 51, pp.393-414.

Dunford, N. and J.T. Schwartz, 1988, Linear Operators Part I: General Theory, Wiley Interscience, New York.

Hall, P. and I. Van Keilegom, 2005, Testing for Monotone Increasing Hazard Rate, The Annals of Statistics, 33 (3), pp. 1109-1137.

Harris, M. and A. Raviv, Optimal Incentive Contracts with Imperfect Information, Journal of Economic Theory Volume 20, Issue 2, April 1979, Pages 231-259

Holmstrom, B. "Moral Hazard and Observability." Bell Journal of Economics, Vol. 19 (1979), pp. 74-91.

Holmstrom, B. and L. Weiss, 1985, "Managerial Incentives, Investment and Aggregate Implications: Scale Effects", Review of Economic Studies, 52, July, pp.403-425.

Holmstrom, Bengt and Paul Milgrom, 1987, "Aggregation and linearity in the provision of intertemporal incentives" Econometrica 55, March, pp 303-328

Innes, R. D., 1990, "Limited Liability and Incentives Contracting with Ex-ante Action Choices," Journal of Economic Theory, 52, pp.45-67.

Jensen, M., and W. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure," Journal of Financial Economics, 3, pp. 305-360.

Jewitt I., 1988, Justifying the First-Order Approach to Principal-Agent Problems",Econometrica 56, 1177-90.

Licalzi, M. and S. Spaeter ,2003, "Distributions for the first-order approach to principal-agent problems" Economic Theory, 21, 167-73

Milgrom, P. and C. Shannon (1994), "Monotone Comparative Statics," Econometrica, Vol. 62, No. 1., January, pp. 157-180

Mirrlees, J. "The Optimal Structure of Incentives and Authority within an Organization." The Bell Journal of Economics, Vol. 7, No. 1 (Spring 1976), pp. 105-131.

Mirrlees, J., "Notes on Welfare Economics, Information, and Uncertainty," in Balch, McFadden, and Wu, eds., Essays on Economic Behavior under Uncertainty, Amsterdam: North Holland Publishing Co., 1974.

Otsuka, K. , H. Chuma, and Y. Hayami, 1992, "Land and Labor contracts in Agrarian Economies. Theories and Facts" Journal of Economic Literature, 30

Rogerson W. P., (1985), The First-Order Approach to Principal-Agent Problems", Econometrica 53, 1357-67.

Ross, S. "The Economic Theory of Agency: The Principal's Problem." The American Economic Review, Vol. 63 (1973), pp. 134-139.

Spence, M. and Zeckhauser, R. "Insurance, Information, and Individual Action." The American Economic Review, Vol. 61 (1971), pp. 380-387.

Stiglitz, J,E and and A. Weiss, 1981 "Credit Rationing in Markets with Incomplete Information" The American Economic Review, Vol 71, N3 pp 393-410.

Stiglitz, J,E. 1974. “Incentives and Risk Sharing in Sharecropping.”Review of Economic Studies, 41 (2), 219-255

Topkis, D, 1998. "Supermodularity and Complementarity", Princeton University Press, Frontiers of Economic Research Series.

Appendix

Sufficient conditions for the multiplicative separable example to satisfy the weak implementability assumption.

Consider a multiplicatively separable model,

$$\Pi(\theta, a) = \theta Q(a).$$

Under these circumstances the agent gets paid as long as $\theta > \frac{r}{Q(a)}$ and therefore the agent's problem can be expressed as:

$$\max_{a} \int_{r/Q(a)}^{\overline{\theta}} [\theta Q(a) - r] f(\theta) d\theta - a. \tag{31}$$

The first order condition of the problem is:

$$Q_a \int_{r/Q(a)}^{\overline{\theta}} \theta f(\theta) d\theta = 1. \tag{32}$$

The second order condition is

$$Q_{aa} \int_{r/Q(a)}^{\overline{\theta}} \theta f(\theta) d\theta + r \frac{Q_a^2}{Q^3} f\left(\frac{r}{Q}\right) \leq 0. \tag{33}$$

The second order condition can be re-written as:

$$\frac{Q_{aa} Q^3}{r Q_a^2} + \frac{f\left(\frac{r}{Q}\right)}{1 - F\left(\frac{r}{Q}\right)} \frac{1}{E\left[\theta | \theta \geq \frac{r}{Q}\right]} \leq 0. \tag{34}$$

The first term $\frac{Q_{aa} Q^3}{Q_a^2}$ is decreasing for a large family of positive concave functions such as for example, the function a^{γ} with any $\gamma < 1$, and logarithmic functions. The second term is the multiplication of the hazard rate times one over a conditional expectation. The hazard rate decreases in a by assumption, however the conditional expectation also decreases. The complete term is decreasing for common distributions such as the uniform.

PROOF of Proposition 2

Consider a call option contract with strike price r, the return of the agent is given by:

$$U = \int_0^{\widehat{\theta}(r,a)} \Pi(\theta,a)f(\theta)d\theta + \int_{\widehat{\theta}(r,a)}^{\overline{\theta}} rf(\theta)d\theta - a \tag{35}$$

Equation (35) is strictly concave in a since it is the composition of two concave functions. Therefore the first order condition is necessary and sufficient for optimality. Moreover since for each r there is a unique optimum effort level $a(r)$, by the theorem of maximum $a(r)$ is continuous in r. Finally since $\partial^2 U/\partial r \partial a > 0$ it is also the case that $a(r)$ is increasing in r by standard monotone comparative statics arguments.

Now to prove the proposition suppose to the contrary that the optimal contract is $w()$ different from an option contract and implements an effort level $\widehat{a}$. Consider the strike price $\widehat{r}$ that at a_0 satisfies

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)d\theta = \int_0^{\widehat{\theta}(\widehat{r},a)} \Pi_a(\theta,a)f(\theta)d\theta$$

It is clear that $\widehat{a}$ is incentive compatible with the option contract $\widehat{r}$. We claim that this contract gives the principal a higher expected utility to the principal. To prove this consider the problem:

$$\min_w \int_0^{\overline{\theta}} w(\Pi(\theta,a))f(\theta)d\theta \tag{36}$$

$$s/t \int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)d\theta = 1 \tag{37}$$

$$w' \leq 1 \tag{38}$$

$$w \geq 0 \tag{39}$$

Integrating by parts $\int_0^{\overline{\theta}} w(\Pi)f(\theta)d\theta = w(0,a) + \int_0^{\overline{\theta}} w'(\Pi)(1-F(\theta))\Pi_\theta d\theta$;

Therefore we can express the problem with the Lagrangian:

$$\min_{w_0,w'} \pounds = w(0,a) + \int_0^{\overline{\theta}} w'(\Pi)(1-F(\theta))\Pi_\theta - \lambda f(\theta)\Pi_a)d\theta \tag{40}$$

subject to $w' \leq 1$ and $w' \geq 0$. First it is clear that it is optimal to set $w(0,a) = 0$.

The problem is linear in w' and the variation with respect to w' is the term $(1-F(\theta))\Pi_\theta - \lambda f(\theta)\Pi_a)$ proportional to $(1/\lambda - \rho)$. The solution will have the smaller possible w' for higher levels of θ and the lowest possible for low levels of θ. This is the call option contract with strike price $\widehat{r}$ that implements $\widehat{a}$. We therefore know that among all contracts that could potentially implement $\widehat{a}$ (all the contracts for which $\widehat{a}$ is a critical point) the option contract $\widehat{r}$ minimizes the utility of the agent, and therefore maximizes the expected utility for the principal.

From the previous result it follows that $V(\widehat{a},\widehat{r}) > V(\widehat{a},w)$. By continuity we can increase r untill $V(a(r^*),r^*)$=$V(\widehat{a},w)$. Because effort is increasing in the strike price and since $a(\widehat{r}) = \widehat{a}$ it follows that $a(r^*) > \widehat{a}$. Finally since expected surplus $E(\Pi(\theta,a) - a)$ is quasiconcave in a and $a^{FB} > a(r^*) > \widehat{a}$ we know that more surplus is generated with the contract r^* than with w . Finally since more surplus is created, both agents are risk neutral and the principal is not better off, it must be the case that the agent is better off under the option contract r^*. So we have proven that if there exists an optimal contract it must be an option contract.

To show existence of an optimal contract note that without loss of optimality we can restrict attention to option contracts, and also that we can restrict attention to a compact subset of those contracts. To show this define $\overline{r}$ as the unique value that satisfies $V(a^*,\overline{r}) = 0$. Since $a(r) < a^{FB}$ for every r , a contract with $r > \overline{r}$ will not satisfy the principal IR constraint and thus we can restrict attention to contracts where $r \in [0,\overline{r}]$ □

PROOF of proposition 3. (The proof does not assume that contracts need to be non-decreasing for the agent)

The proof of Proposition 3 follows from 4 properties of contracts. To prove these properties we first need to show that given any feasible contract $w(\Pi)$, the agent's utility $U(a,w)$ is differentiable with respect to a.

Notice that for a contract to be feasible, the return to the investor $v(\Pi) = \Pi - w(\Pi) - K$ must be increasing in Π ; and therefore $v(\Pi)$ has at most countably many jumps and it is differentiable $a.e.$ Let k_i be each point of discontinuity and Δ_i be the size of every discontinuity; each Δ_i is strictly positive to keep $v(\Pi)$ increasing in Π. There exists an increasing continuous function $\widetilde{v}(\Pi)$ such that the following equality holds $a.e.$

$$v(\Pi)+K=\widetilde{v}(\Pi)+\sum_{i:k_i\leq\Pi}\Delta_i. \tag{41}$$

Because the agent gets $w(\Pi)=\Pi-v(\Pi)+K$ then it is also the case that $a.e.$

$$w(\Pi)=\Pi-\widetilde{v}(\Pi)-\sum_{i:k_i\leq\Pi}\Delta_i. \tag{42}$$

Because the equality holds $a.e$ and the distribution does not have any mass points; the utility of agents is the same if two contracts are equivalent $a.e$. Therefore without loss of generality we can restrict attention to the set of functions that can be expressed as the sum of a continuous function, and a step function $w(\Pi)=\widetilde{w}(\Pi)-\sum_{i:k_i\leq\Pi}\Delta_i$ where $\widetilde{w}$ is differentiable $a.e$ and $\widetilde{w}'\leq 1$.

Lemma 1 *The utility of the agent $U(a,w)$ is differentiable with respect to a* for any feasible w.

PROOF By the definition of an integral $\int\sum_{i:k_i\leq\Pi}\Delta_i dF=\sum_i\Delta_i\left(1-F(\widehat{\theta}(k_i,a)\right)$. The utility of the agent can therefore be written as:

$$U(w,a)=\int_\theta(\widetilde{w}(\Pi(\theta,a)))f(\theta)d\theta-\sum_i\Delta_i\left(1-F(\widehat{\theta}(k_i,a)\right)-a. \tag{43}$$

And therefore $U_a=\int(\widetilde{w}_\Pi)\Pi_a f(\theta)d\theta+\sum_i\Delta_i f(\widehat{\theta}(k_i,a))\widehat{\theta}_a-1$.Where we have taken the derivative inside the integral by the Lebesgue dominated convergence theorem.□

Observe that by the implicit function theorem, $\widehat{\theta}_a=-\frac{\Pi_a}{\Pi_\theta}$, and thus any contract that implements $a>0$ must satisfy the condition:

$$U_a=\int(\widetilde{w}_\Pi)\Pi_a f(\theta)d\theta-\sum_i\Delta_i f(\widehat{\theta}(k_i,a))\frac{\Pi_a}{\Pi_\theta}-1=0. \tag{44}$$

Definition 1 *Let $Z(r)$ be a set such that a belongs to $Z(r)$ if and only if a satisfies the agent IC constraint in a debt contract with face value r.*Let $a(r)$ be the largest element in $Z(r)$.

Property 1 *$a(r)$ is decreasing in the face value of the debt r.*

PROOF In a debt contract the agent solves:

$$\max_a U(a,R) = \int_0^{\overline{\theta}} \max\{\Pi(\theta,a) - r, 0\} f(\theta)d\theta - a, \tag{45}$$

U has increasing differences in $\{a, -r\}$ since $\frac{\partial^2 U}{\partial a \partial r} = -\frac{1}{\Pi_{\widehat{\theta}}(r,a)} \leq 0$ and therefore the result follows from Theorem 2.8.5 in Topkis (1998).□

Property 1 does not imply uniqueness or continuity of the effort level a.

Property 2 *If* $V(a, r_0) = \overline{V}$ *and* $a \in Z(r_0)$ *then given any* $0 \leq V' \leq \overline{V}$ *there exists* $r' \leq r_0$ *such that* $V(a', r') = V'$ *for some* $a' = a(r')$.

PROOF Define $\widetilde{V}(r) = \min_{\widetilde{r}:\widetilde{r}\in[r,r_0)} V(a(\widetilde{r}), \widetilde{r})$. $\widetilde{V}(r)$ is increasing and continuous in r in all the interval $[0, r_0)$. Increasing follows by definition and continuity follows because either $V(a(r), r)$ is continuous in r or if discontinuous it is decreasing in r. Moreover $\widetilde{V}(0) = 0$ and $\widetilde{V}(r_0) > V'$. By continuity there exists $r' \in [0, r_0)$ such that $\widetilde{V}(r') = V'$ which by definition implies that there exists $r'' \in [r', r_0]$ such that $V(a(r''), r'') = V'$ □

The next Property is fundamental in proving the main result of the paper.

Property 3 *If* $a_0 \in Z(r_0)$, $a_0 > a(\widetilde{r})$ *and* w *is a contract different from debt that implements* a_0; *then* $V(a_0, w) < V(a_0, r_0)$.

The proof shows that among all the contracts that satisfy the first order condition at a_0, $V(a_0, w) < V(a_0, r_0)$.

Consider the following problem

$$\max_w V(a_0, w) \text{ s/t } U_a(a_0, w) = 1; \; v(\Pi) \text{ is non decreasing} \tag{46}$$

A solution to this problem must exist since the problem can be stated as finding the optimal payment for the principal, $\Pi - w(\Pi)$, which is weakly increasing and bounded in a bounded interval. Since increasing bounded functions in a bounded interval are compact in the relevant metric ($\mathcal{L}_1$) the problem must have a solution. [6]

[6] For a proof of this see Dunford and Schwartz Corollary 11 page 294.

Since maximizing the expected return to the principal is equivalent to minimizing it for the agent the problem can be rewritten as:

$$\min_{w,k_i,\Delta_i} \int_0^{\theta} \left(w\left(\Pi(\theta,a)\right) + \sum_{i:k_i \leq \Pi} \Delta_i \right) f(\theta) d\theta, \tag{47}$$

subject to:

$$\int_0^{\overline{\theta}} w'(\Pi(\theta,a))\Pi_a(\theta,a)f(\theta)d\theta - \sum_i \frac{\Pi_a}{\Pi_\theta} f(\widehat{\theta}(k_i,a))\Delta_i = 1, \tag{48}$$

$$w(\cdot) \geq 0, \tag{49}$$

$$w'(\cdot) \leq 1. \tag{50}$$

We first claim that the optimal contract has no discontinuities. To the contrary suppose that the optimal w has a $\Delta_i > 0$ at some $\Pi = k_i$

Consider the alternative contract $\widehat{w}$ that only differs from w in two ways,

1) The discontinuity at k_i decreases in δ

2) A new discontinuity of $\delta \frac{f(\widehat{\theta}(k_i,a)\Pi_a/\Pi_\theta(k_i.a)}{f(\widehat{\theta}(k_i-\varepsilon,a)\Pi_a/\Pi_\theta(k_i-\varepsilon.a)}$ is created at $\Pi(\widehat{\theta}-\varepsilon,a)$.

First observe that by construction the constraint 48 is still satisfied.

Second observe that the objective function changes in:

$$\delta \left[\frac{1-F(\widehat{\theta})}{f(\widehat{\theta})} \frac{\Pi_\theta(\widehat{\theta})}{\Pi_a(\widehat{\theta})} - \frac{1-F(\widehat{\theta}-\varepsilon)}{f(\widehat{\theta}-\varepsilon)} \frac{\Pi_\theta(\widehat{\theta}-\varepsilon)}{\Pi_a(\widehat{\theta}-\varepsilon)} \right] \tag{51}$$

This is strictly negative, since we have assumed that $\frac{f(\theta)}{1-F(\theta)}$ and Π_a/Π_θ are increasing in θ . This contradicts the assumption that the optimal contract was w; and therefore the optimal contract cannot have any discontinuities.

Therefore without loss of optimality we can restrict attention to continuous $a.e$ differentiable functions w. The problem then becomes:

$$\min_w \int_0^{\overline{\theta}} w(\Pi(\theta,a))f(\theta)d\theta \tag{52}$$

subject to

$$\int_0^{\overline{\theta}} w'(\Pi)\Pi_a f(\theta)d\theta = 1, \tag{53}$$

$$w' \leq 1, \tag{54}$$

$$w \geq 0. \tag{55}$$

Integrating by parts $\int_0^{\overline{\theta}} w(\Pi)f(\theta)d\theta = w(0) + \int_0^{\overline{\theta}} w'(\Pi)(1 - F(\theta))\Pi_\theta d\theta$;

Therefore we can express the problem with the Lagrangian

$$\min_{w_0, w'} \pounds = w(0) + \int_0^{\overline{\theta}} w'(\Pi)(1 - F(\theta))\Pi_\theta - \lambda f(\theta)\Pi_a)d\theta \tag{56}$$

subject to $w' \leq 1$ and $w' \geq 0$ if $w = 0$. First it is clear that it is optimal to set $w(0) = 0$.

The problem is linear in w' and the variation with respect to w' is the term $(1 - F(\theta))\Pi_\theta - \lambda f(\theta)\Pi_a)$ proportional to $1/\lambda - \frac{\Pi_a}{\Pi_\theta}\frac{f(\theta)}{1-F(\theta)}$. The second term is the hazard rate times $\frac{\Pi_a}{\Pi_\theta}$; both increasing in θ by complementarity and the hazard rate condition. The optimal contract will have the smaller possible w' for small levels of θ and the highest possible for high levels of θ. Because for low levels of θ, $w = 0$ the optimal contract is $w' = 0$ if $\widehat{\theta}(\Pi, a) < \theta^*$ and $w' = 1$ otherwise. Because Π is increasing in θ, this is precisely the form of a debt contract.

If there exists an optimal contract, it is a debt contract r, that satisfies the first order condition. By Lemma 3, the contract must be the debt contract that implements a_0.□

Property 4: There exists an effort level $\widehat{a}$ that $i)$ is implementable with a debt contract $\widehat{r}$ and $U(\widehat{r}, \widehat{a}) = 0$ and $ii)$ if a is implementable by some contract w but not with a debt contract then $a < \widehat{a}$.

Proof: This Property follows from two Lemmas

Lemma 2 *Feasible contracts always induce an effort level $a \leq a^{FB}$.*

PROOF Remember that we can restrict attention to contracts of the form $w(\Pi) = \widetilde{w}(\Pi) - \sum_{i:k_i \leq \Pi} \Delta_i$. Where $w' \leq 1$. The level of effort a is either 0 or satisfies the first order condition:

$$\int_0^{\overline{\theta}} w'(\Pi(\theta, a))\Pi_a(\theta, a)f(\theta)d\theta - \sum_i \frac{\Pi_a}{\Pi_\theta} f(\widehat{\theta}(k_i, a))\Delta_i = 1 \tag{57}$$

If the contract implements $a > a^{FB}$, then by the definition of a^{FB} we know that $\int \Pi_a f(\theta)d\theta <$

1. Moreover we know that $w' \leq 1$ and therefore $\int_0^{\overline{\theta}} w'(\Pi(\theta, a))\Pi_a(\theta, a)f(\theta)d\theta < 1$ which contradicts the fact that the first order condition is satisfied.□

The utility of the agent in a debt contract is given by $U(r) = \max_a \int \max\{\Pi(\theta, a) - r, 0\} - a$, decreasing and continuous in r. Let $\widetilde{r}$ be the smallest face value of the debt such that $a = 0$ is optimal. Whenever $r < \widetilde{r}$;then $a(r) > 0$ and therefore must be a local interior maximum. The weak implementability assumption implies that whenever $r < \widetilde{r}$ there exists a unique a optimum in the agent problem, since any optimum is interior, and there exists a unique local interior optimum. Moreover for any $r_1 > \widetilde{r}$, $a = 0$ is optimal. Since by assumption $\lim_{a \to 0} \Pi_a(\theta, a) = \infty$ then $a = 0$ is optimal only if the agent is not the residual claimant at any state and therefore it must be the case that $U(r_1) = 0$, by continuity it must be the case that $U(\widetilde{r}) = 0$.

Lemma 3 *Any effort level* $a \in [a(\widetilde{r}),\ a^{FB}]$ *can be induced with a debt contract.*

PROOF First observe that in $[0, \widetilde{r})$ the effort $a(r)$ is unique and, by the Theorem of the Maximum, it is also continuous. Moreover $\lim_{r \to \widetilde{r}^-} a(r) = a(\widetilde{r})$ since the objective function is continuous in a and r and thus $Z(r)$ is an upper semi-continuous correspondence. Therefore $a(r)$ is continuous and decreasing in $[0, \widetilde{r}]$ and the lemma follows by the intermediate value theorem.□

Property 4 is satisfied by defining $\widehat{a} = a(\widetilde{r}\ )$.

We are now ready to prove the main proposition of the paper.

PROOF of Proposition 3 To the contrary assume the optimal contract w is different from a debt contract and implements an effort level a_0.

i) If $a_0 \geq a(\widetilde{r})$ then by Property 4 there exists a debt contract r_0 that implements a_0. By Property 3, $V(a_0, r_0) > V(a_0, w)$. By Property 2, there exists $r_1 < r_0$ such that $V(a(r_1), r_1) = V(a_0, w)$. Therefore r_1 satisfies the principal's IR constraint. Finally since $a_0 < a_1 \leq a^{FB}$; $E(\Pi(\theta, a_0)) = V(a_0, w) + U(\ a_0, w) < V(a_1, r_1) + U(\ a_1, r_1) = E(\Pi(\theta, a_1))$ and because by definition of r_1, $V(a(r_1), r_1) = V(a_0, w)$ then $U(\ a_0, w) < U(\ a_1, r_1)$ which violates the optimality of w.

ii) If $a_0 < a(\widetilde{r})$; then consider the debt contract with face value $\widetilde{r}$. Since $a_0 < a(\widetilde{r})$, ($V(a_0, w) + U(a_0, w) < V(a(\widetilde{r}), \widetilde{r}) + U(a(\widetilde{r}), \widetilde{r})$ and since $U(a(\widetilde{r}), \widetilde{r}) = 0$, $V(a(\widetilde{r}), \widetilde{r}) > V(a_0, w)$. Moreover, by Property 2, there exists $r_1 < \widetilde{r}$ such that $V(a(r_1), r_1) = V(a(\widetilde{r}), \widetilde{r})$. Therefore r_1 satisfies the principal's IR constraint. Finally since $a_0 < \widetilde{r} \leq a^{FB}$; $V(a_0, w) + U(a_0, w) < V(a_1, r_1) + U(a_1, r_1)$ and because $V(a(r_1), r_1) = V(a_0, w)$ then $U(a_0, w) < U(a_1, r_1)$ which violates the optimality of w.

We have shown that if an optimal contract exists, it must be a debt contract. To show existence of an optimal debt contract, note that the face value of debt contracts can be bounded above by $\overline{r}$ where $\overline{r} = \min_r$ such that $a(\overline{r}) \leq \underline{a}$ and $V(\underline{a}\ , \overline{\theta}) = 0$.Higher values of r will not satisfy the investor's IP constraint. Finally, the set of all debt contracts is $[0, \overline{r}]$, compact, and since i) $U(r, a(r))$ is continous in r, and ii) the set of contracts that satisfy $V(r, a(r)) \geq 0$ is also compact, an optimal debt contract exist by the Weiestrass theorem. □

Proof of proposition 4

Consider the problem

$$\min_w \int_0^{\overline{\theta}} w(\Pi(\theta, a)) f(\theta) d\theta \tag{58}$$

Subject to the constraints

$$\int_0^{\overline{\theta}} w'(\Pi) \Pi_a f(\theta) d\theta = 1 \tag{59}$$

$$w' \leq 1$$

$$w \geq 0 \tag{60}$$

Integrating by parts $\int_0^{\overline{\theta}} w(\Pi) f(\theta) d\theta = w(0, a) + \int_0^{\overline{\theta}} w'(\Pi)(1 - F(\theta)) \Pi_\theta d\theta$;

Therefore we can express the problem with the Lagrangian:

$$\min_{w_0, w'} \pounds = w(0, a) + \int_0^{\overline{\theta}} w'(\Pi)(1 - F(\theta)) \Pi_\theta - \lambda f(\theta) \Pi_a) d\theta \tag{61}$$

subject to $w' \leq 1$ and $w' \geq 0$. The optimal contract in this problem is of the class L.

Finally observe that if any contract implements a then it satisfies the first order condition , and therefore it must give the agent a smaller utility under the effort level a than a contract of the class L.□

PROOF that MLRP is not implied by an increasing critical ratio

To see that the multiplicative form does not satisfy the standard affiliation property or MLRP observe that the induced probability distribution over outcomes is as follows,

$$g(\Pi|a) = f(\widehat{\theta}(\Pi, a)) = f(\Pi/Q(a)) \tag{62}$$

Given the induced probability distribution, we can state the standard MLRP condition.

MLRP condition

$$\frac{\partial}{\partial \Pi}\left(\frac{g_a(\Pi|a)}{g(\Pi|a)}\right) > 0 \tag{63}$$

Applying the induced probability distribution on outcomes, observe that

$$\frac{g_a(\Pi|a)}{g(\Pi|a)} = -\frac{f'(\widehat{\theta}(\Pi, a))}{f(\widehat{\theta}(\Pi, a))}\widehat{\theta}(\Pi, a)\frac{Q'(a)}{Q(a)} = -\frac{f'((\Pi/Q(a)))\Pi Q'(a)}{f(\Pi/Q(a))Q^2(a)} \tag{64}$$

Then, the following inequality must hold,

$$\frac{\partial}{\partial \Pi}\left(\frac{g_a(\Pi|a)}{g(\Pi|a)}\right) = -\left(\frac{f'(\theta)}{f(\theta)} + \left[\theta\frac{f''(\theta)f(\theta) - (f'(\theta))^2}{f^2(\theta)}\right]\right)\frac{Q_a}{Q(a)^2} > 0 \tag{65}$$

This inequality implies that

$$f'(\theta) + \theta f''(\theta) - \theta(f'(\theta))^2/f(\theta) < 0 \tag{66}$$

From the form of the probability density, this inequality can be expressed as

$$\alpha - \theta\left(\alpha\right)^2/\left(\gamma + \alpha\theta\right) < 0 \tag{67}$$

If θ is close to zero, this expression converges to $\alpha < 0$. So the MLRP condition fails to hold.[7]

PROOF of proposition 5

Proof. Because the CDFC condition is stronger than the strong implementability condition , it is sufficient to show that the strong implementability condition does not hold. Suppose to the contrary that implementability holds. Then, by the theorem of the maximum there exists a continuous function $a(r)$ that maps debt contracts with face value r to effort levels induce by the contracts $a(r)$ This implies that any effort level in $[0, a(0)]$ is implementable with a debt contract.(Clearly $a = 0$ can be implemented with a debt with face value $r > \Pi(\overline{\theta}, a^{FB})$).

Note that in a debt contract, the agent's expected utility has increasing differences in $\{r, a\}$. Therefore, by standard monotone comparative statics analysis, the function $a(r)$ must be non-increasing in r. Let $a(\widehat{r})$, be the effort level implemented by an arbitrary debt contract $\widehat{r}$ with $\widehat{r} > \Pi(\overline{\theta}, 0)$. Let $a_1 < a(\widehat{r})$ be an effort level such that $\Pi(\overline{\theta}, a_1) < \widehat{r}$. The level a_1 exists since $\widehat{r} > \Pi(\overline{\theta}, 0)$. Because $a_1 < a(\widehat{r})$ it must be implementable by some debt contract r_1 and because $a(r)$ is non-increasing we know that $r_1 \leq \widehat{r}$. Finally, by definition, it must be the case that

$$a_1 = \arg \max_{a \in [0,\infty)} \int_0^{\overline{\theta}} \max \{\Pi(a, \theta) - r_1, 0\} f(\theta) d\theta - a \tag{68}$$

But since $\Pi(\overline{\theta}, a_1) < \widehat{r}$ for every θ it is the case that $\max \{\Pi(a, \theta) - r_1, 0\} = 0$ and therefore it means that

$$a_1 = \arg \max_{a \in [0,\infty)} -a \tag{69}$$

This is a clear contradiction since $a = 0$ is optimal in (24).□

[7] As an example that satisfies MLRP but fails to satisfy the hazard rate condition consider the case where $\gamma/\overline{\theta} < a < 0$.