

On the Consistency of the First-Order-Approach to Principal-Agent Problems

Óscar Gutiérrez

Universidad Autónoma de Barcelona, Barcelona, Spain
Email: oscar.gutierrez@uab.es

Received December 20, 2011; revised January 13, 2012; accepted January 20, 2012

ABSTRACT

This paper revisits the principal-agent model with moral hazard when its solution is obtained invoking the first-order-approach. We show that the solution can be economically inconsistent even when “sufficient conditions” ensuring its validity ([1,2]) hold. To be more precise, we provide examples where is impossible to find Lagrange multipliers validating the approach. The correct solution to the problem provides a rationale for option-like contracts and minimum payments.

Keywords: Moral Hazard; Principal-Agent Model; First-Order Approach; Likelihood Ratio; Option-Like Incentives

1. Introduction

The principal-agent model under moral hazard represents a cornerstone in the analysis of contractual arrangements intended to achieve mutual gains between agents with asymmetric knowledge; see [3-8], and others. This model pursues the characterization of the second-best contract, *i.e.*, a rule which establishes how to share the output between the employer (principal) and the employee (agent) in order to extract from the last one the optimal amount of an unobservable input. The importance of the agency relationship in economic activities makes the model relevant from a theoretical point of view because it provides a conceptual framework to analyze actual incentives in terms of efficiency. Furthermore, it makes it possible to perform exercises of comparative statics on the model parameters (in order to determine, for example, the effect of uncertainty on performance and welfare).

However, a satisfactory mathematical characterization of the solution to the principal-agent model involves serious technical difficulties, which has prevented the development of simple solutions to the problem. In Haubrich's words: “Unfortunately, most versions [of the principal-agent model] have intractable solutions. Quantitative solutions do not readily emerge from the implicit equations that define the sharing rules, especially in models with a continuum of states” (see [9], p. 259).

An important solution to the problem of determining optimal contracts in realistic agency setups is based on the so-called *first-order approach* (FOA henceforth). The simplification introduced by the FOA is important because it makes it feasible to characterize optimal

incentive contracts ([8,10-15], *et al.*). We center our attention on analyzing the consistency of the solution provided by the FOA.

Technically, the FOA consists in replacing the *incentive compatibility constraint* of the optimization problem by its first-order condition. This procedure has a long history in the incentives literature, but its validity must not be taken for granted. James Mirrlees was the first who observed that substituting the *correct* constraint by its first-order condition may lead to wrong solutions (see [3]). In particular he showed that, in some circumstances, the informational unconstrained solution can be arbitrarily approximated, so local maxima will not coincide with global ones. That is, when the unconstrained solution to the problem can be arbitrarily approximated by means of a *Mirrlees forcing contract*, the FOA fails because global maxima cannot be found within the set of stationary points. The counterexample provided by Mirrlees against the validity of the FOA, rather than a solution to the problem, has been considered a pathological situation to be avoided in more elaborated and realistic models.

In order to avoid this problem, [1,2,5] among others attempted to provide sufficient conditions validating the FOA. Many papers on managerial retribution and agency relationships appeal to the FOA in their analysis of incentive contracts (see references above). Many others appeal to the FOA in contexts where it is well known that it cannot be invoked (e.g., [16,17]).

Although the FOA is a successful shortcut often used in theoretical work, we show that the proposed solutions can lead to inconsistencies that make them invalid from a

theoretical point of view. In particular, we show that some solutions contradict basic economic assumptions (e.g., that marginal utility must be positive). Once the inconsistencies are detected, it is possible to interpret the correct solutions to the problem as actual solutions to moral hazard problems implemented in practice, in particular, non-linear incentives and minimum payments.

The use of non-linear contracts is a common practice in executive compensation. Stock options, bonus contracts or contracts based on performance-standards are paramount examples of non-linear contracts used in firms. However, in spite of the undeniable success that such contracts have in managerial compensation, a fully satisfactory explanation for their use in terms of efficiency in effort-extraction and risk-sharing still remains as a theoretical challenge. In Prendergast words, “the theoretical literature has made little progress in understanding the *observed* (nonlinear) shape of compensation contracts, despite costs associated with nonlinearities” ([18], p. 15). This paper provides a rationale for the use of non-linear incentives. For example, those including a lower bound in compensation. In [8] it is suggested that some bounds may be necessary in incentives design in order to account for the agent’s and principal’s wealth constraints. In [19] the question is analyzed in detail. We instead justify the appearance of lower bounds endogenously, and piecewise-linear contracts implemented in practice can be seen as approximations to second-best contracts with bounds.

Section 2 introduces the mathematical model and its solution when the FOA is applied. Section 3 presents a number of inconsistencies the approach can yield when conditions “validating” the FOA are met. In Section 4, we justify the use of non-linear contracts and minimum payments in principal-agent relationships. Section 5 concludes.

2. Mathematical Specification

The setup is the principal-agent model in the formulation due to [8] in an extension of [4]: a risk-neutral principal (she) employs a risk-averse and effort-averse agent (he) to carry out an unobservable task that influences an *ex post* signal x (which we identify with output); the agent receives a payment $s(x)$ for his job. The optimization program is stated as follows:

$$\begin{aligned} & \max_{s(x), a} \int [x - s(x)] f(x, a) dx \\ & \text{s.t. } \int U(s(x)) f(x, a) dx - C(a) \geq R \quad (P1) \\ & \text{s.t. } a \in \arg \max_{a'} \int U(s(x)) f(x, a') dx - C(a'), \end{aligned}$$

where $U(\cdot)$ represents the (standard) utility function of the agent; R represents the agent’s reservation utility; $C(a)$ represents the agent’s cost of providing a level of effort a , with $C'(a) > 0$ and $C''(a) > 0$; and $f(x, a)$ represents

the output density function for a given level of effort. The two constraints in (P1) are respectively known as the *participation constraint* and the *incentive compatibility constraint*. In [20, pp. 51-55] the reader can find a pedagogical presentation of the principal-agent model.

The optimization program (P1) is difficult to solve, and few general results can be obtained if additional specific assumptions are not made. A shortcut often used is to replace the incentive compatibility constraint by its first-order condition, *i.e.*, the original constraint

$$a \in \arg \max_{a'} \int U(s(x)) f(x, a') dx - C(a')$$

is replaced by the relaxed constraint

$$\int U(s(x)) f_a(x, a) dx - C'(a) = 0,$$

where $f_a \equiv \partial f / \partial a$. This procedure is known as the first-order approach (FOA). The economic literature provides sets of conditions under which the FOA holds. The importance of this simplification is that makes it feasible the characterization of optimal contracts. In [21, p. 253], an intuitive argument justifying the FOA is presented.

Let us review the sets conditions intended to validate the FOA. The Mirrlees-Rogerson conditions establishes that if the density f satisfies the Monotone Likelihood Ratio Property (*i.e.*, f_a/f is non-decreasing), and the distribution function $F(x, a) \equiv \int^x f(y, a) dy$ satisfies the Convex Distribution Function condition (CDFC, which mathematically means that $\partial^2 F(x, a) / \partial a^2 \geq 0$), then the first-order approach is valid.¹ It has been argued that this set of conditions is not very useful from a practical point of view, because it is not easy to find distributions satisfying the CDFC. In order to avoid this drawback, [22] provides two rich families of examples satisfying the Mirrlees-Rogerson conditions.

[2] provides a second set of conditions, less stringent than the Mirrlees-Rogerson one. In particular, [2] shows that if the density function can be expressed as $f(x, a) = \theta(x) \varphi(a) e^{\alpha(a)\beta(x)}$ with $\beta(x)$ concave, f_a/f is concave, expected output is concave in effort, and $U(U'^{-1}(1/z))$ is concave for $z > 0$, then the FOA holds (see [2], Cor. 1). He also provides a more relaxed set of conditions (see [2], Th. 1).

[21] discusses the results in [1,2], generalizing and comparing them in the multisignal case.

When the first-order approach can be invoked, the optimal sharing rule $s^*(x)$ satisfies the two following conditions (see [8] or [2]):

$$\frac{1}{U'(s^*(x))} = \mu_1 + \mu_2 \frac{f_a(x, a^*)}{f(x, a^*)} \quad (1a)$$

¹The validity of the FOA under such conditions has not been shown in a model where output is continuous, although it is usually taken as valid. [1], with some more requirements, shows it in a model with discrete output.

$$\int (x - s^*(x)) f_a(x, a^*) dx + \mu_2 [U(s^*(x)) f_{aa}(x, a^*) dx - C''(a)] = 0, \tag{1b}$$

where μ_1 and μ_2 represent the Lagrange multipliers associated to the participation constraint and the incentive compatibility constraint respectively. The optimal contract can thus be derived from (1a) as a function of μ_1 , μ_2 and a^* , which are obtained from the two constraints in (P1) together with (1b). Thus, the principal can design a contract that extracts the (second-best) level of effort from the agent.

However, the papers devoted to justify the FOA are centered on showing that the agent's utility is concave in effort, ignoring that in some cases the existence of the two (positive) multipliers may not be guaranteed. Let us analyze this question in detail.

3. Consistency of the Solution

This section shows that the contract defined by Equation (1a) can lead to situations where the FOA does not hold even when the Mirrlees-Rogerson or Jewitt's conditions hold. In particular, situations where the existence of the Lagrange multipliers implies that the agent's marginal utility is negative for low outcomes.

First, let us see what happens when the likelihood ratio f_a/f is a non-decreasing and unbounded-from-below function of output. Unbounded-from-below likelihood ratios have been considered in [4, p.125], and ruled out in [19, ft.2]. However, none of them mention that when f_a/f is not bounded-from-below, the solution provided by the FOA is not admissible. The reason is that multipliers μ_1 and μ_2 have constant values, so they must be independent of realized output. Given that μ_2 is positive when the FOA holds (see [2, Lemma 1]), if the likelihood ratio is unbounded from below, for any pair of multipliers there exist low values of x such that $\mu_1 + \mu_2 f_a(x, a^*)/f(x, a^*)$ becomes negative. Then Equation (1a) can not be satisfied because marginal utility must be positive all along the output range. Consequently, the likelihood ratio must be bounded from below.

However, neither the Mirrlees-Rogerson nor the Jewitt sets of sufficient conditions guarantee such a requirement. A counterexample to Jewitt's "sufficiency" can be found in his own paper (see [2], example c, p. 1183). The example provided by [1] (footnote 4) also has an unbounded-from-below likelihood ratio. As Rogerson restricts output to take values in a discrete and bounded set, the likelihood ratio is not unbounded in his example. However, the continuous version of Rogerson's conditions can lead to unbounded-from-below likelihood ratios.

[22] provides two families of examples (both continu-

ous) satisfying the Mirrlees-Rogerson conditions, but as has been shown, such requirements do not guarantee that the likelihood ratio is well-behaved. In fact, in both families of examples it is easy to impose certain requirements under which likelihood ratios are unbounded.

An example that satisfies both sets of conditions with an unbounded likelihood ratio corresponds to the case where output has a density function equal to

$$f(x, p) = p(p+1)x^{p-1}(1-x), x \in [0,1]$$

and where the agent influences the parameter p with his actions in such a way that $p = a + 2$.

Furthermore, if the agent's utility function is unbounded-from-below, a Mirrlees forcing contract can be constructed; see [3]. However, it is worth observing that this kind of forcing contracts, which produces an arbitrary approximation to the first-best solution, can only arise when the FOA does not work (because the FOA does not provide well-defined contracts when f_a/f is unbounded from below). Then, there is a more fundamental argument invalidating the FOA than the appearance of Mirrlees forcing contract because, contrary to [3], an unbounded utility function is not required.

More important and of greater concern is the case where likelihood ratios are *bounded*-from-below. Next, we show that, even when f_a/f is bounded from below, expression $\mu_1 + \mu_2 f_a/f$ can be negative if output is low enough, which implies that FOA does not produce a satisfactory candidate for a second-best contract. Consider a principal-agent setting where output has a gamma distribution with parameters p and a , where p is an integer and a is a positive number. The density function is:

$$f(x) = \frac{a^{-p}}{(p-1)!} x^{p-1} e^{-x/a}. \tag{3}$$

The expected output is pa ; the output variance is pa^2 . The likelihood ratio is linear in output:

$$f_a/f = (x - pa)/a^2.$$

This model has been often used in the incentives literature: in the seminal article [8] it is used with illustrative purposes. [11] discusses the advantages of replacing curvilinear second-best bonus plans by their piecewise linear approximations (quota-based sales compensation schemes). [12] analyzes the advantages of introducing *stock options* in managerial incentives.

Assuming that the agent has a cost of effort equal to $C(a) = ka^2/2$, the multipliers are:

$$\mu_1 = [R + C(a)]/2 = \frac{R + ka^2/2}{2},$$

$$\mu_2 = \frac{C'(a)}{2 \int (f_a/f)^2 f dx} = \frac{ka}{2p/a^2}.$$

If the agent's reservation utility R is set equal to zero (for example because the market for executives' services is competitive), simple calculations yield:

$$\mu_1 + \mu_2 f_a/f = kax/(2p) - ka^2/4,$$

which is negative if output x is sufficiently low. It must be mentioned that $\mu_1 + \mu_2 f_a/f$ can also be negative when the reservation utility R is strictly positive.² Then, none of the two examples provided by Jewitt with continuous output is *in general* valid (see [2], p. 1183, examples (a) and (c)).

4. Minimum Payments and Kinked Incentives

In this section we link the inconsistencies that can appear when the FOA is invoked to actual incentives used in practice. In particular, we suggest that solving correctly the optimization program P1 provides a rationale for the use of non-linear and option-like incentives, which represents a popular practice today (stock-options, bonus based on performance standards, and others), and a theoretical foundation for the existence of minimum payments in executive compensation, as [13,14] suggest. The following result formalizes the argument:

Proposition: Assume that the agent's utility function is concave in effort. If no positive multipliers can be found ensuring that Equation (1a) holds for the entire output range, the solution to P1 will include a minimum wage.

Proof: Consider first that the likelihood ratio is unbounded from below.

As $\lambda + \mu f_a/f$ must be non-negative, the agent's fee can not be below $(U')^{-1}(\infty)$, which becomes the minimum wage (whenever finite).³

(An alternative reasoning can be taken from [4, p. 125, Equations (69)-(71)]: if $s(x)$ admits a minimum wage, the maximization of the Lagrangean (L) with respect to $s(x)$ in order to obtain Equation (1a) leads to:

$$\partial L/\partial s = -1 + U'(s(x))(\lambda + \mu f_a/f) \leq 0.$$

²To see this, take into account that: 1) $\mu_2 > 0$ (see [2]); 2) f_a/f is negative for low outputs (and positive for large ones) because

$$\int (f_a/f) f dx = \frac{d}{da} \int f dx = \frac{d}{da}(1) = 0 \text{ and } f_a/f \text{ increases with output.}$$

Then, $\mu_1 + \mu_2 f_a/f$ will be negative if output is low (unless multiplier μ_1 is sufficiently large). If μ_1 is small (for example because the reservation utility R is small), then $\mu_1 + \mu_2 f_a/f$ is negative for small outputs.

³Holmstrom restricts $s(x)$ to lie in $[c, d + x]$, which can be justified because agent's and principal's wealth put bounds on $s(x)$. In our context, c is naturally given by $(U')^{-1}(\infty)$ as long as it remains finite. For example, if $U(x) = \log x$, $(U')^{-1}(\infty) = 0$ and $U(U')^{-1}(\infty) = -\infty$. Observe that in this case both the agent's utility and the likelihood ratio are both unbounded from below functions, so a Mirrlees forcing contract can be constructed.

If $\lambda + \mu f_a/f \leq 0$, then $\partial L/\partial s < 0$; for those output values that correspond to $\lambda + \mu f_a/f \leq 0$, $s(x)$ must be equal to the minimum wage.)

When the likelihood ratio is *bounded* from below, we have shown that $\mu_1 + \mu_2 f_a/f$ can also be negative for low enough outputs. The argument given for the existence of a minimum wage in the unbounded case keeps working in the bounded case, and the correct solution to program P1 will contain a corner.□

The unbounded case has been considered in [4,19]. [19] rules it out in order to avoid Mirrlees forcing contracts, while [4] uses it to justify the existence of minimum payments. Our interpretation is that when effort is highly productive for very low outcomes (*i.e.*, the likelihood ratio is unbounded from below), the optimal fee has a minimum value, which induces a kinked point in the incentive system if salaries are positive and the minimum wage is above zero.

The bounded case is more worrying, and the conclusions on it are novel, underlining that the underlying problem in the unbounded case is deeper than it seems at first sight. From the discussion in Section 3, we conjecture that non-linear salaries are more likely to be observed in moral hazard environments where agents have low outside opportunities. For example, if executives offer their services in a competitive market.

Finally, it is worth mentioning that [19] also justifies the appearance of option-like incentives when minimum wages are imposed. However, [19] imposes minimum payments exogenously, while we have endogenously justified the existence of such lower bounds on fees. [13] also mentions that kinked non-linearities may appear in incentives when a minimum wage is binding, something implicit in Holmstrom's seminal work (the contract solution is restricted to lie in the closed domain $[c, d + x]$). [13] appeals to the manager's *limited liability* in order to justify such a minimum fee, which is unnecessary in our context. They argue that stock options can be a better alternative than shares in creating incentives for CEOs, but their argument is not centered on the appearance of corners in incentives. On the other hand, and in answer to their claim that "the power class utility functions provides additional motivation for the requirement of limited liability in the contract", we point out that $\mu_1 + \mu_2 f_a/f$ must be positive, so no problem involving power utilities can arise.

5. Conclusions

The widespread use of incentives under moral hazard makes it important the characterization of second-best contracts. This paper analyzes the consistency of the solutions obtained appealing to the first-order approach.

We have shown that, in some contexts, the solutions

are devoid of economic sense, because they lead to situations compatible with negative marginal utility. This can happen both with bounded and unbounded likelihood ratios. Then, although the first-order approach is a powerful and elegant tool, it is crucial to check whether the contracts obtained are well-defined indeed.

Once analyzed the consistency of the approach, we interpret the correct solutions to the agency problem as salaries containing non-linearities, which represent actual arrangements implemented in practice.

We obtain that if the likelihood ratio is unbounded from below (*i.e.*, agent's effort is extremely productive for low outputs), a minimum payment must be imposed. If the likelihood ratio is bounded-from-below, corners can also arise in optimal contracts; in this case, we conjecture that kinked incentives are more likely to be used when agents have low outside opportunities.

We thus offer a rationale for the existence of minimum wages and kinked contracts without departing from a canonical principal-agent framework, nor imposing additional constraints on agents' wealth.

REFERENCES

- [1] W. Rogerson, "The First-Order Approach to Principal-Agent Problems," *Econometrica*, Vol. 53, No. 6, 1985, pp. 1357-1367. [doi:10.2307/1913212](https://doi.org/10.2307/1913212)
- [2] I. Jewitt, "Justifying the First-Order Approach to Principal-Agent Problems," *Econometrica*, Vol. 56, No. 5, 1988, pp. 1177-1190. [doi:10.2307/1911363](https://doi.org/10.2307/1911363)
- [3] J. Mirrlees, "Notes on Welfare Economics, Information and Uncertainty. In: M. Balch, D. McFadden and S.-Y. Wu, Eds., *Essays on Economic Behavior under Uncertainty*, North-Holland, Amsterdam, 1974.
- [4] J. Mirrlees, "The Optimal Structure of Incentives and Authority within an Organization," *Bell Journal of Economics*, Vol. 7, No. 1, 1976, pp. 105-131. [doi:10.2307/3003192](https://doi.org/10.2307/3003192)
- [5] J. Mirrlees, "The Implications of Moral Hazard for Optimal Insurance," *Seminar Given at Conference Held in Honor of Karl Borch*, Bergen, 1979.
- [6] M. Spence and R. Zeckhauser, "Insurance, Information and Individual Action," *American Economic Review*, Vol. 61, 1971, pp. 380-387.
- [7] S. Ross, "The Economic Theory of Agency: The Principal's Problem," *American Economic Review*, Vol. 63, No. 2, 1973, pp. 134-139.
- [8] B. Holmstrom, "Moral Hazard and Observability," *Bell Journal of Economics*, Vol. 10, No. 1, 1979, pp. 74-91. [doi:10.2307/3003320](https://doi.org/10.2307/3003320)
- [9] J. Haubrich, "Risk Aversion, Performance Pay and the Principal-Agent Problem," *Journal of Political Economy*, Vol. 102, No. 2, 1994, pp. 258-276. [doi:10.1086/261931](https://doi.org/10.1086/261931)
- [10] D. Diamond and R. Verrecchia, "Optimal Managerial Contracts and Equilibrium Security Prices," *Journal of Finance*, Vol. 37, No. 2, 1982, pp. 275-287. [doi:10.2307/2327326](https://doi.org/10.2307/2327326)
- [11] J. Raju and V. Srinivasan, "Quota-Based Compensation Plans for Multiterritory Heterogeneous Salesforces," *Management Science*, Vol. 42, No. 10, 1996, pp. 1454-1462. [doi:10.1287/mnsc.42.10.1454](https://doi.org/10.1287/mnsc.42.10.1454)
- [12] T. Hemmer, O. Kim and R. Verrecchia, "Introducing Convexity into Optimal Compensation Contracts," *Journal of Accounting and Economics*, Vol. 28, No. 3, 2000, pp. 307-327. [doi:10.1016/S0165-4101\(00\)00008-2](https://doi.org/10.1016/S0165-4101(00)00008-2)
- [13] L. Lambert and D. Larcker, "Stock Options, Restricted Stock, and Incentives," *SSRN Working Paper*, 2004.
- [14] O. Kadan and J. Swinkels, "Stock or Options? Moral Hazard, Firm Viability, and the Design of Compensation Contracts," Washington University, St. Louis, 2005. <http://www.olin.wustl.edu/faculty/swinkels/>
- [15] O. Gutiérrez and V. Salas-Fumás, "Performance Standards and Optimal Contracts," *Journal of Accounting and Economics*, Vol. 45, No. 1, 2008, pp. 139-152. [doi:10.1016/j.jacceco.2007.11.001](https://doi.org/10.1016/j.jacceco.2007.11.001)
- [16] G. Feltham and M. Wu, "Incentive Efficiency of Stock versus Options," *Review of Accounting Studies*, Vol. 6, No. 1, 2001, pp. 7-28. [doi:10.1023/A:1011377902967](https://doi.org/10.1023/A:1011377902967)
- [17] X. Zhou and P. Swan, "Performance Thresholds in Managerial Incentive Contracts," *Journal of Business*, Vol. 76, No. 4, 2003, pp. 665-696. [doi:10.1086/377035](https://doi.org/10.1086/377035)
- [18] C. Prendergast, "The Provision of Incentives in Firms," *Journal of Economic Literature*, Vol. 37, No. 1, 1999, pp. 7-63. [doi:10.1257/jel.37.1.7](https://doi.org/10.1257/jel.37.1.7)
- [19] I. Jewitt, O. Kadan and J. Swinkels, "Moral Hazard with Bounded Payments," *Journal of Economic Theory*, Vol. 143, No. 1, 2008, pp. 59-82. [doi:10.1016/j.jet.2007.12.004](https://doi.org/10.1016/j.jet.2007.12.004)
- [20] J. Tirole, "The Theory of Industrial Organization," The MIT Press, Cambridge, 1988.
- [21] J. Conlon, "Two New Conditions Supporting the First-Order Approach to Multi-Signal Principle-Agent Problems," *Econometrica*, Vol. 77, No. 1, 2009, pp. 249-278. [doi:10.3982/ECTA6688](https://doi.org/10.3982/ECTA6688)
- [22] M. LiCalzi and S. Spaeter, "Distributions for the First-Order Approach to Principal-Agent Problems," *Economic Theory*, Vol. 21, 2003, pp. 167-173. [doi:10.1007/s00199-001-0250-y](https://doi.org/10.1007/s00199-001-0250-y)