Economics Bulletin

Volume 44, Issue 2

A convex mapping for the first order approach: A note

Corrado Benassi
Dipartimento di Scienze Economiche, Alma Mater Studiorum Università di Bologna

Alessandra Chirco
Dipartimento di Scienze dell'economia, Università del Salento

Abstract

The first order approach to solving the standard one-dimensional principal-agent model is conditional upon the relevant stochastic production function obeying two noteworthy restrictions: that the Likelihood Ratio be monotonically increasing in output, and that the distribution function be convex in effort. It is usually claimed that such conditions are very restrictive, as very few of the standard probability distributions satisfy both properties. The purpose of this note is to show that some simple transformations or parametrizations are available, that enable one to work out convenient distributions with the required properties

Citation: Corrado Benassi and Alessandra Chirco, (2024) "A convex mapping for the first order approach: A note", *Economics Bulletin*, Volume 44, Issue 2, pages 698-706
Contact: Corrado Benassi - corrado.benassi@unibo.it, Alessandra Chirco - alessandra.chirco@unisalento.it.
Submitted: December 10, 2023. **Published:** June 30, 2024.

1 Introduction

The aim of this note is to provide a user-friendly methodology to apply the so-called First Order Approach (FOA) for the solution of the one-dimensional principal-agent model with unobservable effort.

The key characteristics of this model is that the outcome (output) of the bilateral contract between the principal and the agent is a stochastic variable, the distribution of which depends on the unobservable effort provided by the agent. In this setup, the optimal contract, which specifies an output-contingent wage schedule, implements the effort desired by the principal only if an incentive compatibility constraint is satisfied, ensuring that such an effort is actually provided by the agent.

The FOA tackles this problem by embodying the first order conditions of the agent's maximization problem with respect to effort as the incentive compatibility constraint. It is well known that a sufficient condition for this strategy to properly identify the optimal contract is that the stochastic production function obeys two noteworthy restrictions: the Monotone Likelihood Ratio Property, and that the distribution function be convex in effort.

While some probability distributions are available which satisfy both conditions, it is usually claimed that the latter are indeed very restrictive, as standard probability distributions, like the Normal, Beta, Chi square, F, Weibull, etc., do not in fact satisfy the above properties (e.g., Jewitt, 1988; LiCalzi and Spaeter, 2003).

We show that some simple transformations are available, which allow to work with suitably parametrized standard distributions, in such a way that some key and desirable properties of the latter (e.g., unimodality) are preserved, and made consistent with those required by the application of the FOA methodology.

The note is organized as follows. In Section 2 we describe the standard setup of a principal-agent model with unobservable effort, and we outline the FOA approach. In Section 3 we suggest a user-friendly procedure to generate probability distributions satisfying the conditions for its application. Section 4 concludes.

2 The basic model

The standard formulation of the continuous principal-agent model with separable utility[1] takes the general form

$$\max_{w(x),a^*} \int_{\mathcal{X}} v\left[x - w(x)\right] f(x, a^*)dx \tag{1.a}$$

subject to

$$\int_{\mathcal{X}} u\left(w(x)\right) f(x, a^*)dx - c(a^*) \geq \overline{u} \tag{1.b}$$

$$a^* \in \arg\max_a \int_{\mathcal{X}} u\left(w(x)\right) f(x, a)dx - c(a) \tag{1.c}$$

where x denotes output, $w(x)$ the agent's compensation and a effort; $v(\cdot)$ and $u(\cdot)$ are respectively the principal's and the agent's Bernoulli utility functions, while $c(a)$ denotes the agent's disutility of effort, which is assumed to be increasing and convex. It is usually assumed that v and u are both increasing concave functions, with u strictly concave. In equation (1.b) $\overline{u}$ denotes the agent's reservation utility. Output is a stochastic variable, distributed according to the density function $f(x, a)$.

Letting subscripts denote derivatives, $f(x, a) = F_x(x, a)$ is the strictly positive density of

$$F : \mathcal{X} \times \mathcal{A} \rightarrow [0, 1]$$

which gives the parametrized distribution $F(x, a)$ of observable (and verifiable) output $x \in \mathcal{X}$, given the agent's hidden action $a \in \mathcal{A} \subset \mathbb{R}_+$. It is standardly assumed that the support $\mathcal{X}$ is a compact interval independent of a; that F is continuously differentiable at least twice; and that the agent's effort exerts a positive effect on output in the sense of first order stochastic dominance, that is

$$F_a(x, a) \leq 0 \tag{2}$$

for all $a \in \mathcal{A}$.[2]

Problem (1) embodies the idea of the principal maximizing her utility subject to the agent's participation (1.b) and incentive compatibility (1.c) constraints. It is well known that in the above formulation the problem turns

---

[1] Non separable utility is much harder to work with. An extension of the FOA to the nonseparable utility case is provided by Alvi (1997).

[2] The set $\mathcal{A}$ is sometimes assumed to be an open interval, so as to characterize optima by interior maxima (e.g., Jewitt, 1988, p.1179)

out to be analytically unmanageable (e.g. Laffont and Martimort, 2002, p.197), and that a possible way out is to invoke the so-called First Order Approach (Mirrlees, 1975; Holmström,1979).[3] This amounts to substituting constraint (1.c) by:

$$\int_{\mathcal{X}} u(w(x)) f_a(x, a^*) dx - c_a(a^*) = 0 \tag{3}$$

i.e., by the requirement that a^* satisfies the first order conditions of the agent's expected utility maximization.

The related literature (e.g. Holmström,1979; Rogerson, 1985) has established that the FOA, by relying upon necessary and not sufficient conditions, cannot clearly be valid in general; but it is so if the following twin conditions are satisfied

$$F_{aa}(x, a) \geq 0 \text{ for all } x \in \mathcal{X} \tag{4.a}$$

$$\alpha_x(x, a) \geq 0 \text{ for all } x \in \mathcal{X} \tag{4.b}$$

Condition (4.a) is usually known as the CDF property (effort-Convexity of the Distribution Function), while (4.b) is known as the MRLP (Monotone Likelihood Ratio Property), where

$$\alpha(x, a) = \frac{f_a(x, a)}{f(x, a)}$$

is the likelihood ratio, assumed to be monotonically increasing in x.

While the MRLP is generally looked at as a non-controversial assumption, which also ensures that in the optimal contract the agent's compensation is increasing in output, the CDF is usually considered very much restrictive. Indeed, though one reasonable implication of the CDF is that of decreasing marginal (expected) productivity of effort,[4] very few distributions seem to share this property.[5]

Examples of distributions satisfying MLRP and CDF are $F(x, a) = x^a$ (Rogerson, 1985), or $F(x, a) = 1 - e^{\frac{x}{a}}$ (Laffont and Martimort, 2002) – which

[3] Approaches not relying on FOA are provided by Grossman and Hart (1983) and Araujo and Moreira (2001); see also Chaigneau *et al.* (2019). This note is concerned with the one-dimensional problem; sufficient conditions for the FOA to be valid in a multidimensional setting are discussed among others by Kirkegaard (2017).

[4] Chaigneau *et al.* (2019) argue that *increasing* marginal productivity might be a more realistic alternative.

[5] A property which moreover does have one disturbing feature highlighted by Jewitt (1988, p.1177): in the simple linear case where realized output is $x = a + \varepsilon$, effort convexity implies a monotonic density.

however imply monotone densities. Other cases are discussed by LiCalzi and Spaeter (2003), who identify two classes of densities obeying MLRP and CDF, one of which allows for non monotone densities. Among the distributions yielding unimodal densities, one could also enlist the Burr type-X distribution (e.g., Johnson *et al.*,1994, p.54)

$$F(x,a) = \left(1 - e^{-x^2}\right)^a$$

defined for $x > 0$ and unimodal for all $a > 1/2$, which obeys the required properties.

The above examples show that the search for suitable distributions can be fruitful and enlarge the scope of application of the FOA methodology. However, in this note we suggest a different approach. We start from commonly used distributions, e.g. the Normal, Beta, Gamma, etc, the parameters of which cannot be related to effort in a such a way as to satisfy both MLRP and CDF. We then apply to these distributions simple transformations involving an effort variable, through which the FOA conditions are met while preserving some of their desirable features. In other words we show that there exist mappings which, starting from some given distribution, may help building one satisfying both the MRLP and the CDF, while retaining some properties of the original one (like unimodality) which may be relevant in a variety of environments.

3 Useful transformations for the First Order Approach

The basic idea relies on modelling the relationship between output and effort by appropriately parametrizing in terms of effort some "core" output distribution $G(x) : \mathcal{X} \to [0,1]$. If one lets $\varphi : [0,1] \times \mathcal{A} \to [0,1]$, any given a induces a distribution F such that $F : \mathcal{X} \times \mathcal{A} \to [0,1]$:

$$F(x,a) = \varphi(G(x), a)$$

The following can then be established:

Proposition *Let φ have the following properties for all $x \in X$ and $a \in A$: (i) $\varphi(0,a) = 0$, $\varphi(1,a) = 1$; (ii) $\varphi_1(\cdot,a) \geq 0$; (iii) $\frac{\partial}{\partial x}\frac{\varphi_{12}}{\varphi_1} \geq 0$; (iv) $\varphi_{22} \geq 0$. Then F is a distribution obeying both the MRLP and the CDF property.*

Proof Conditions *(i)* and *(ii)* ensure that F is a distribution, while conditions *(iii)* and *(iv)* ensure that F satisfies the MRLP and the CDF property.■

Through this Proposition, we suggest that it is possible to associate the effort variable a to any "core" distribution G, in such a way that F is endowed with the required features, while some relevant properties of G carry over to F. In this sense, our suggestion is to move from the search of distributions satisfying the MRLP and CDF property, to the search for suitable convexifying transformations of any distribution G.

In the sequel we show how the separation between the distribution G and the effort variable a allows a straightforward representation of the way in which effort affects the output distribution. To this end, we present three examples:

Example 1 One natural example is the power function. Take *any* distribution $G : \mathcal{X} \to [0,1]$: then if one defines $F(x,a) = G(x)^{\lambda(a)}$, F is a proper distribution which satisfies both the CDF and the MRLP for any (weakly) concave positive λ function increasing in a.[6] An instance where the "core" distribution G is the quadratic Beta distribution and $\lambda(a) = a$ is given in Fig. 1; Fig.2 presents the same core density with $\lambda(a) = \sqrt{a}$, in both cases with $a = 1$ (thick), $a = 2$ (thin) and $a = 3$ (dashed).

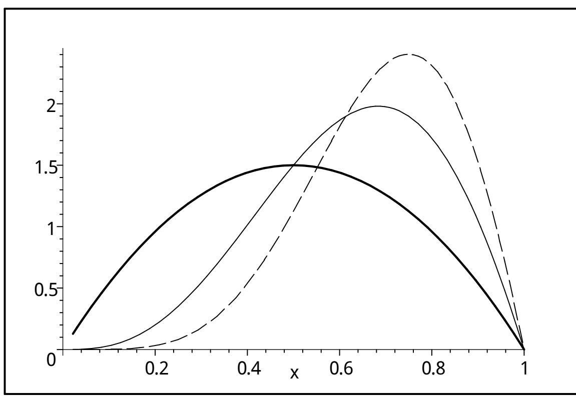


Fig. 1: Beta density with $\lambda(a) = a$

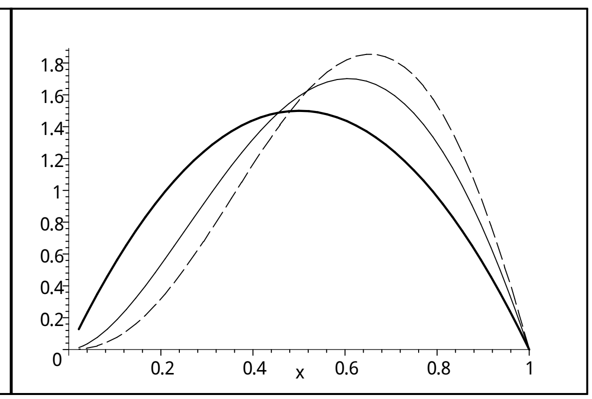


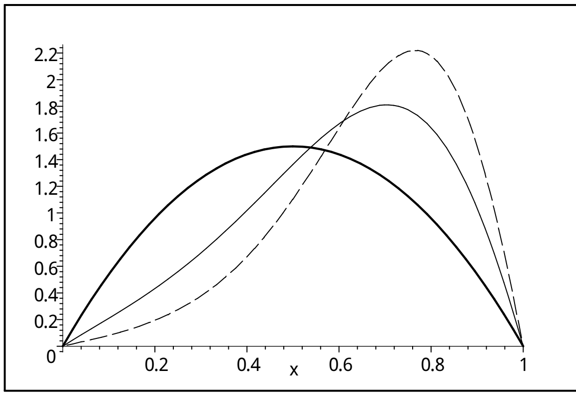
Fig.2: Beta density with $\lambda(a) = \sqrt{a}$

Example 2 Consider the formulation:

$$F(x,a) = G(x)e^{[G(x)-1]\lambda(a)} \qquad (5)$$

[6] Indeed, it is readily seen that $F_{aa}(x,a) = G(x)^{\lambda}\left[(\lambda_a \ln G(x))^2 + \lambda_{aa} \ln G(x)\right] > 0$ for all positive x, while the likelihood ratio is $\alpha(x,a) = \lambda_a/\lambda + \lambda_a \ln G(x)$, obviously increasing in x. As is well known, Rogerson (1985, p.1362) presents as an example the distribution $F(x,a) = x^a$ (in our notation), which in fact can be looked as the specific case where $\lambda(a) = a$, and the distribution being exponentially effort-parametrized is uniform (see also LiCalzi and Spaeter, 2003, p.169). One should notice that some lower bound to $\alpha(x,a)$ should be imposed, which is not considered here as $\lim_{x\to 0} \alpha(x,a) = -\infty$. In Rogerson's original formulation this issue does not arise as he is using a discrete distribution. See Gutierrez (2012) for a discussion of this boundary requirement.

where again λ is an increasing weakly concave function of a, and again $G : \mathcal{X} \to [0,1]$ is *any* distribution. It is easily seen that F satisfies both CDF and MLRP,[7] and we present in Fig.s 3 and 4 the densities obtained when $G(x)$ is again the quadratic Beta distribution, with $\lambda(a) = a$ and $\lambda(a) = \sqrt{a}$; we plot these for $a = 0$ (thick), $a = 1$ (thin) and $a = 2$ (dashed).



Fig,3: Beta density with $\lambda(a) = a$

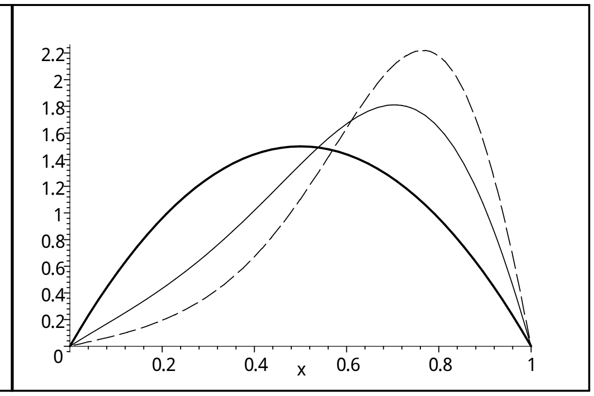


Fig.4: Beta density with $\lambda(a) = \sqrt{a}$

It may be worth noting that LiCalzi and Spaeter (2003, p.171) identify an example such that $F(x,a) = xe^{(x-1)a}$ which, given $\mathcal{X} = [0,1]$, can be seen as a parametrization of $G(x) = x$, the uniform distribution, with $\lambda(a) = a$ and $\mathcal{A} = \mathbb{R}_{++}$. This yields a monotone density,[8] but a different choice of G would deliver unimodality.

Example 3 Another example is given by the following parametrization. Take again *any* $G : \mathcal{X} \to [0,1]$ and define

$$F(x,a) = \frac{a^{G(x)} - 1}{a - 1} \tag{6}$$

for $\mathcal{A} = (1,\infty)$. One can check that $F_{aa}(x,a)$ is positive for all $x \in \mathcal{X}$ and $a > 1$ and the likelihood ratio is increasing in x.[9] Also in this case we present

[7]Indeed, $F_{aa}(x,a) = G(x)\left[G(x)-1\right]e^{[G(x)-1]\lambda(a)}\left[G(x)-1\right]\lambda_a^2 + \lambda_{aa}$ which is positive for any positive x, while $\alpha(x,a) = \{G(x) - 1 + G(x)/\left[1 + G(x)\lambda(a)\right]\}\lambda_a$, increasing in x.

[8]LiCalzi and Spaeter (2003, p.171) discuss this case within the class considered in their Proposition 2: any distribution of the form $F(x,a) = \delta(x)e^{\beta(x)\gamma(a)}$ satisfies both CDF and MLRP, provided the functions β, γ, and δ obey the restrictions these authors impose. Suitable specifications of these function can ensure the existence of a mode.

[9]Indeed, $F_{aa} = \left(a^{G-2}G^2(x)\left(a-1\right)^2 - \left(3a-1\right)a^{G(x)-2}G(x)\left(a-1\right) + 2\left(a^{G(x)}-1\right)\right)/\left(a-1\right)^3 > 0$ for $a > 1$, while $\alpha_x = \{(\ln a)\,a\,(G(x)-1) - 1 + a - G(x)\ln a\}/\left((a-1)(\ln a)\,a\right)$, obviousy increasing in x.

an example, where the basic Gbx9 distribution is the Cauchy distribution normalized over the unit interval, with a D æ(thick), a D O(thin) and a D) (dashed).

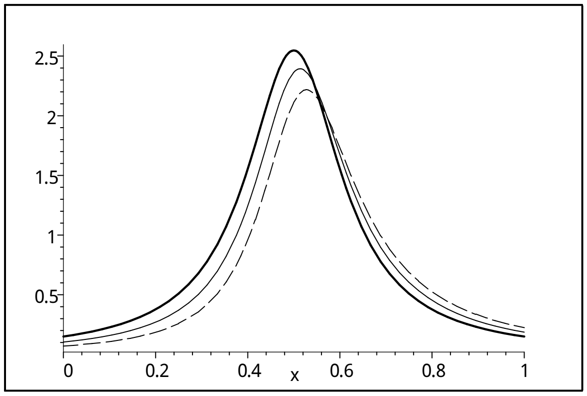


Fig.5: Cauchy density

One can observe that all the above cases a ‘neutral’ effort level $\widehat{a}$ can be identified (given as a thick density in the numerical examples), which gives back the core distribution G (which of course will not in general satisy CDF and MRLP).[10] This may be conveniently seen as the minimum effort level: in our examples, minimum effort in this sense implies an output distribution symmetric around æ⁄O with mean output independent of effort, while higher-than-minimum effort distorts G in the desired way.

4 Concluding remarks

The twin conditions of Monotone Likelihood Ratio Property and Convexity of the Distribution are often claimed to be too stringent for the First Order Approach to the solution of the standard principal-agent problem to be viable in many applications, as most distributions currently in use do not satisfy them. In this short note we have shown that simple mappings are available, which allow one to work with suitably transformed standard distributions.

References

Alvi, E. (1997) “First-Order Approach to Principal-Agent Problems: A Generalization” *The Geneva Papers on Risk and Insurance Theory* **22**, 59-65.

[10] This ‘neutral’ level is $\lambda\left(\widehat{a}\right) = 1$ in example 1, $\lambda(\widehat{a}) = 0$ in example 2, and $\widehat{a} = 1$ in the third case, where $\lim_{a\to 1} F(x,a) = G(x)$.

Araujo, A. and H.Moreira (2001) “A general Lagrangian approach for non-concave moral hazard problems” *Journal of Mathematical Economics* **35**, 17-39.

Chaigneau, P., A.Edmans and D.Gottlieb (2019) “The informativeness principle without the first-order approach” *Games and Economic Behavior* **113**, 743-755.

Grossman, S. and O.Hart (1983) “An Analysis of the Principal-Agent Problem” *Econometrica* **51**, 7-45.

Gutierrez, O. (2012) “On the consistency of the first-order approach to principal-agent problems” *Theoretical Economics Letters* **2**, 157-161.

Holström, B. (1979) “Moral Hazard and Observability” *The Bell Journal of Economics* **10**, 74-91.

Jewitt, I. (1988) “Justifying the first-order approach to principal-agent problems” *Econometrica* **56**, 1117-1190.

Johnson, N.L., S.Kotz and N.Balakrishnan (1994) *Continuous Univariate Distributions* (vol.1), Wiley & Sons: New York and Toronto.

R. Kirkegaard (2017) “A unifying approach to incentive compatiblity in moral hazard problems” *Theoretical Economics* **12**, 25-51.

Laffont, J.-J. and D.Martimort (2002) *The Theory of Incentives: The Principal-Agent Model*, Princeton University Press: Princeton and Oxford.

LiCalzi, M. and S.Spaeter (2003) “Distributions for the first-order approach to principal-agent problems” *Economic Theory* **21**, 167-173.

Mirrlees, J. (1975) “The Theory of Moral Hazard and Unobservable Behaviour: Part 1” *The Review of Economic Studies* (1999) **66**, 3-21.

Rogerson, W.P. (1985) “The First-Order Approach to Principal-Agent Problems” *Econometrica* **53**, 1357-1367.