Economic Theory **21**, 167–173 (2003)

Economic Theory

Exposita Notes

Distributions for the first-order approach to principal-agent problems⋆

Marco LiCalzi[1] **and Sandrine Spaeter**[2]

[1] Department of Applied Mathematics, University of Venice, 30135 Venice, ITALY
(e-mail: licalzi@unive.it)
[2] BETA-Theme, Louis Pasteur University, 67085 Strasbourg, FRANCE
(e-mail: spaeter@cournot.u-strasbg.fr)

Received: December 4, 2000; revised version: December 10, 2001

Summary. The first-order approach is a technical shortcut widely used in agency problems. The best known set of sufficient conditions for its validity are due to Mirrlees and Rogerson and require that the distribution function is convex in effort and has a likelihood ratio increasing in output. Only one nontrivial example was so far known to satisfy both properties. This note provides two rich families of examples displaying both properties.

Keywords and Phrases: Principal-agent problem, First-order approach, Monotone likelihood ratio, Convexity in effort.

JEL Classification Numbers: D82.

1 Introduction

In the standard formulation of an agency problem, a (male) principal solves a maximization problem where one of the constraints is that the (female) agent privately chooses a level of effort that maximizes her own expected utility. For tractability, this complex constraint is often replaced with the agent's first-order condition. As shown in Mirrlees (1975), this widely used first-order approach is not always valid because the agent's first-order condition may refer to stationary

⋆ The authors thank Kaïs Dachraoui, Georges Dionne and Rodolphe Dos Santos Ferreira for their useful comments. A first version of this note was written during the post-doctoral training of the second author at the Risk Management Chair, HEC Montréal. MURST and RMC_2 are acknowledged for financial support.
Correspondence to: M. LiCalzi

points that are not global maxima. This has motivated the search for classes of problems where the first-order approach is valid.

The best known of such classes is due to Mirrlees (1975) and Rogerson (1985). Its main restriction is that the probability distributions linking output to effort satisfy two properties. The first one is that the density of output with respect to effort has a monotone likelihood ratio (MLR). The second one is that the distribution function of output be convex in effort at each level of output (CDF).

The MLR property implies that more effort tend to yield higher output and is crucial in making the agent's reward increasing in output. Its application in problems of information economics, initiated in Milgrom (1981), usually goes unquestioned. To the contrary, the CDF property, which leads to the agent's reward being concave in effort, is subject to two criticisms raised in Jewitt (1988).

First, very few of the standard statistical distributions have the CDF property: indeed, only one non-trivial positive example is known in the literature. Second, in a simple situation where the output x is the sum of the effort a and a noisy disturbance ε, the CDF property requires the unnatural assumption that the density of the observable output be increasing for any level of effort. This motivated Jewitt (1988) in formulating an alternative set of restrictions that avoids the CDF condition.

Jewitt (1988) itself, however, had to reintroduce the CDF property to deal with the case of two signals. And, more importantly, it features prominently in the elegant extension of the Mirrlees-Rogerson conditions to multi-signal problems that was carried out in Sinclair-Desgagné (1994). Therefore, it seems that the criticisms to the CDF property deserve a closer scrutiny.

This note addresses the criticisms and tries to yield some support to the widespread use of the first-order approach. More precisely, we provide two classes of distributions that display both the MLR and the CDF properties. Contrary to the examples so far known, these classes are generic and encompass a large number of specific functional forms. Due to their genericity, they allows for a large array of possibilities. We exhibit examples such that the density of the observable output is decreasing, unimodal or bimodal for any level of effort.

Section 2 introduces notation and briefly recalls the first-order approach. Section 3 presents the first example and Section 4 deals with the second example.

2 The first-order approach

We follow Jewitt (1988) and consider a continuous formulation of the standard principal-agent problem. The (male) principal agrees to pay a (female) agent a reward $w(x)$ based on the realized (monetary) output of her efforts. The agent chooses an action a unobservable to her principal in the set A. Her choice of a affects the (observable) output X, which is randomly distributed according to a cumulative distribution function $F(x;a)$ with (strictly) positive density $f(x;a)$.

The agent has a reservation utility u_0 and maximizes the expected value of $u[w(x)] - a$, where u is increasing and (strictly) concave. Taking the agent's behavior into account, the principal maximizes the expected value of $v[x - w(x)]$, where v is increasing and (weakly) concave. Both expected values are computed with respect to the distribution function $F(x;a)$. The principal's maximization problem is thus

$$\max_{w(\cdot),a^*} \quad Ev[X - w(X)] \tag{1}$$

$$\text{s. t.} \quad Eu[w(X)] - a^* \geq u_0 \tag{2}$$

$$a^* \in \arg\max_a Eu[w(X)] - a \tag{3}$$

The first-order approach replaces (3) with the weaker constraint that a^* is a stationary point of $Eu[w(X)] - a$. Mirrlees (1975) and Rogerson (1985) proved that the ensuing simpler but relaxed program is equivalent to the original one if the following two properties hold:

MLR: the likelihood ratio $f(x_1;a)/f(x_2;a)$ is increasing in a for all $x_1 > x_2$;
CDF: the distribution $F(x;a)$ is convex in a for all x.

The MLR property states that larger outputs are more likely for higher levels of effort. An alternate and more expedient characterization of the MLR property requires x and a to be affiliated; that is, $\log f$ is supermodular in (x,a). See Milgrom (1981) and Milgrom and Weber (1982).

The CDF property states that increases in effort improves the distribution of output but at a decreasing rate. To our knowledge, the only nontrivial example of a distribution which satisfies both properties is $F(x;a) = x^a$; see Rogerson (1985), which attributes the example to S. Matthews.

3 A first class of examples

We assume here and in the following that the set of available actions is a (possibly unbounded) interval A and that the support of $F(x;a)$ is a compact interval independent of the choice of a. We define and study $F(x;a)$ only on the open interval $S = (0,1)$, leaving it understood that $F(x;a) = 0$ for $x \leq 0$ and $F(1;a) = 1$ for $x \geq 1$. This is without loss of generality, because the statements in this paper are valid for arbitrary increasing affine transformations of a and x. To keep the proofs as simple as possible, we assume twice differentiability whenever needed; however, the results do not depend on this assumption.

Proposition 1 *Let $A \subseteq \mathbb{R}$ and $S = (0,1)$. The distribution function*

$$F(x,a) = x + \beta(x)\gamma(a)$$

satisfies both the MLR and the CDF properties if the two following conditions hold:

i) *$\beta(x)$ is a positive and concave function on $S = (0,1)$ such that $\lim_{x\downarrow 0}\beta(x) = \lim_{x\uparrow 1}\beta(x) = 0$ and $|\beta'(x)| \leq 1$ for all x in S;*

ii) $\gamma(a)$ *is a decreasing and convex function on A such that* $|\gamma(a)| < 1$.

Proof. First, since $F(x;a)$ must be a cumulative distribution function for all a, we need to check that $\lim_{x\downarrow 0} F(x;a) = \lim_{x\downarrow 0} \beta(x)\gamma(a) = 0$, $\lim_{x\uparrow 1} F(x;a) = 1 + \beta(x)\gamma(a) = 1$, and $F(x;a)$ is increasing in x. This follows from substitution and from $f(x;a) = 1 + \beta'(x)\gamma(a) > 0$.

Second, for $F(x;a)$ to satisfy the MLR property, we need $\log f$ to be supermodular in (x,a). This follows from the necessary and sufficient condition

$$\frac{\partial^2 \log f(x;a)}{\partial a \partial x} = \frac{\beta''(x)\gamma'(a)}{[1+\beta'(x)\gamma(a)]^2} \geq 0.$$

Third, $F(x,a)$ satisfies the CDF property because it is an increasing affine transformation of the convex function $\gamma(a)$ for all x. □

For a specific example, let A be an interval in $(0,+\infty)$. Choosing $\beta(x) = x - x^2$ and $\gamma(a) = 1/(a+1)$, we get

$$F(x,a) = x + \frac{x - x^2}{a+1}$$

for x in $(0,1)$ and a in A. If the output is to take values in a generic (nondegenerate) compact interval $[c,d]$, replace this distribution with $F[(x-c)/(d-c);a]$. If A is the interval $[0,+\infty)$, choose $\gamma(a) = 1/(a+k)$, with $k > 1$.

In some agency problems such as those typical in insurance, the output is usually interpreted as the amount of damage associated with the effort exerted by the agent. A natural parameterization has that more effort should make higher damage less likely and that this should occur at a decreasing rate. This corresponds respectively to requiring that $F(x;-a)$ satisfies the MLR property and that $-F(x;a)$ displays the CDF property. A trivial modification of Proposition 1 can accomodate this case by imposing that $\gamma(a)$ is increasing and concave.

Note that dropping the assumption that $\gamma(a)$ is convex from Proposition 1 leads to a large class of distributions satisfying only the MLR property. This may be of interest for use in any of the various contexts where only the MLR property is required, such as portfolio theory and auction theory; see for instance Landsberger and Meilijson (1990), Ormiston and Schlee (1993), or Milgrom and Weber (1982).

The genericity of the restrictions imposed on β and γ implies that the class of examples in Proposition 1 is relatively rich. However, all the examples in this class exhibit a density which is monotone in x; or, more precisely, increasing for $\gamma(a) \leq 0$ and decreasing for $\gamma(a) \geq 0$. The monotonicity of the density is also in the Matthews's example and runs counter the plausible requirement that the density of output should be unimodal. The class proposed in the next example allows for a more diverse set of possibilities and shows that unimodality is not ruled out by the MLR and CDF properties.

4 A second class of examples

We maintain the assumptions of Section 3, namely that the set of available actions is a (possibly unbounded) interval A and that the support of $F(x;a)$ is the closure of an interval $S = (0,1)$ independent of the choice of a.

Proposition 2 *Let $A \subseteq \mathbb{R}$ and $S = (0,1)$. The distribution function*

$$F(x,a) = \delta(x)e^{\beta(x)\gamma(a)}$$

satisfies both the MLR and the CDF properties if the three following conditions hold:

i) *$\beta(x)$ is a nonconstant, negative, increasing, and convex function on $S = (0,1)$ such that $\lim_{x\uparrow 1}\beta(x) = 0$;*
ii) *$\gamma(a)$ is a strictly positive, increasing, and concave function on A;*
iii) *$\delta(x)$ is a positive, strictly increasing, and concave function on $S = (0,1)$ such that $\lim_{x\downarrow 0}\delta(x) = 0$ and $\lim_{x\uparrow 1}\delta(x) = 1$.*

Proof. For clarity, we omit arguments when there is no ambiguity. First, we need to check that $\lim_{x\downarrow 0} F(x;a) = 0$, $\lim_{x\uparrow 1} F(x;a) = 1$, and $F(x;a)$ is increasing in x. This follows from substitution and $f(x;a) = \left[\delta' + \beta'\gamma\delta\right] \cdot \exp[\beta\gamma] > 0$. Note that the assumptions imply $\delta' > 0$ and therefore $f(x;a)$ is indeed strictly positive.

Second, $F(x;a)$ satisfies the MLR property because the necessary and sufficient condition

$$\frac{\partial^2 \log f(x;a)}{\partial a \partial x} = \beta'\gamma' + \frac{\beta''\gamma'\delta\delta' + \beta'\gamma'\delta'^2 - \beta'\gamma'\delta\delta''}{[\delta' + \beta'\gamma\delta]^2} \geq 0$$

is satisfied.

Third, $F(x,a)$ satisfies the CDF property because the necessary and sufficient condition

$$\frac{\partial^2 F(x;a)}{\partial a^2} = \delta\beta\left(\gamma'' + \beta\gamma'^2\right)e^{\beta\gamma} \geq 0$$

is satisfied. □

For a specific example, let A be an interval in $(0,+\infty)$. Choosing $\beta(x) = x-1$, $\gamma(a) = a$ and $\delta(x) = x$, we get

$$F(x;a) = xe^{a(x-1)} \tag{4}$$

for x in $(0,1)$ and a in A. If x takes values in a generic (nondegenerate) compact interval $[c,d]$, use $F[(x-c)/(d-c);a]$. If $A = [0,+\infty)$, let $\gamma(a) = a + k$ with $k > 0$.

Note that dropping the assumption that $\gamma(a)$ is concave leads to another large class of distributions satisfying only the MLR property.

The examples in the class of Proposition 1 all had monotone densities over output. The class in Proposition 2 allows a richer set of shapes for its densities.

We present two examples, whose purpose is to show that the MLR and CDF properties do not rule out other and perhaps more intuitive assumptions about the shape of the densities. That most of the distributions commonly assumed in statistics fail to display both properties shows only that they are hard to satisfy simultaneously, and no more than that.

In the first example, all densities are genuinely unimodal or, more precisely, strictly quasiconcave (and nonmonotone). In the second example they are all bimodal or, more precisely, strictly quasiconvex (and nonmonotone).

For the first example, assume $A \subseteq (0,1]$ and $S = (0,1)$. Similar to (4), let $\beta(x) = x - 1$ and $\gamma(a) = a$. Choose $\delta(x) = kx - (k-1)x^3$, where the parameter k belongs to the interval $K = [13/10, 15/10]$. This gives the class of distributions

$$F(x;a) = \left[kx - (k-1)x^3\right] e^{a(x-1)}, \tag{5}$$

for x in $(0,1)$ and a in A. Since (5) satisfies the assumptions in Proposition 2 for all k in K, any of these distributions displays the MLR and the CDF properties. It remains to show that their densities are increasing first and then decreasing. Note that the derivative of the density

$$\frac{\partial f(x;a)}{\partial x} = \left[-a^2(k-1)x^3 - 6a(k-1)x^2 + (a^2k - 6k + 6)x + 2ak\right] e^{a(x-1)}$$

has the same sign as

$$h(x;a,k) = \left[-a^2(k-1)x^3 - 6a(k-1)x^2 + (a^2k - 6k + 6)x + 2ak\right], \tag{6}$$

which is a polynomial in x. The first two coefficients are always negative and the last one is always positive. Whaetever is the sign of the third coefficient, there is exactly one change of sign in the coefficients. Therefore, by Descartes' rule, (6) has exactly one positive real root x^*.

Note that $\lim_{x\uparrow 1} h(x;a,k)$ is strictly decreasing in k and strictly increasing in a for all pairs (a,k) in $A \times K$; therefore, $\lim_{x\uparrow 1} h(x;a,k) < \lim_{x\uparrow 1} h(x;1,13/10) = 0$. Since $\lim_{x\downarrow 0} h(x;a,k) = 2ak > 0$, by continuity x^* belongs to the open interval $(0,1)$. Moreover, (6) is positive for $x < x^*$ and negative for $x > x^*$ for all a and k. Therefore, the density $f(x;a,k)$ is unimodal (but not monotone).

For our second example, assume $A \subseteq (0,+\infty)$ and $S = (0,1)$. As above, let $\beta(x) = x - 1$ and $\gamma(a) = a$. Choose $\delta(x) = x^k$, for k in the interval $K = (0,1)$. This gives the class of distributions

$$F(x;a) = x^k e^{a(x-1)}, \tag{7}$$

for x in $(0,1)$ and a in A. Each of them satisfies the assumptions in Proposition 2, so this class displays the MLR and the CDF properties for all k in K.

The derivative of the density has the same sign as

$$\left[a^2x^2 + 2akx + (k^2 - k)\right], \tag{8}$$

which is a polynomial associated with a convex parabola. Its only positive real root is $x^* = (\sqrt{k} - k)/a$, which belongs to K. Since (8) is negative for $x < x^*$ and positive for $x > x^*$ for all a and k, the density $f(x;a,k)$ is strictly quasiconvex (and nonmonotone).

References

Jewitt, I.: Justifying the first-order approach to principal-agent problems. Econometrica **56**, 1177–1190 (1988)

Landsberger, M., Meilijson. J.: Demand for risky financial assets: a portfolio analysis. Journal of Economic Theory **50**, 204–213 (1990)

Milgrom, P.R.: Good news and bad news: representation theorems and applications. Bell Journal of Economics **12**, 380–391 (1981)

Milgrom, P.R., Weber, R.J.: A theory of auctions and competitive bidding. Econometrica **50**, 1089–1122 (1982)

Mirrlees, J.: The theory of moral hazard and unobservable behavior: Part I. Nuffield College, Oxford, mimeo (1975)

Ormiston, M.B., Schlee, E.: Comparative statics under uncertainty for a class of economic agents. Journal of Economic Theory **61**, 412–422 (1993)

Rogerson, W.P.: The first-order approach to principal-agent problems. Econometrica **53**, 1357–1367 (1985)

Sinclair-Desgagné, B.: The first-order approach to multi-signal principal-agent problems. Econometrica **62**, 459–465 (1994)