International Journal of Business and Economics, 2003, Vol. 2, No. 3, 245-261

A Didactic Example of Linear (Multidimensional) Screening Contracts

Ludovic Renou*
Department of Economics, *European University Institute*, *Italy*
and
Economics Division, *University of Southampton*, *U.K.*

Abstract

This short paper proposes a didactic example on how to solve a multidimensional screening problem in the linear case. In the proposed example, shareholders of a cash-constrained firm propose to the firm management a recapitalization in counterpart of the distribution of future dividends. The capacity of the firm to distribute future dividends depends on its production costs and its technology, which are private information of the management. Thus shareholders face a (multidimensional) screening problem. We completely characterize the optimal menus of contracts that shareholders offer. Notably, we show that there always exist optimal menus of contracts with at most two contracts offered: a low dividend, low recapitalization contract and a high dividend, high recapitalization contract. This is an extreme case of bunching.

Key words: multidimensional screening; bunching; adverse selection; shareholders; dividends; recapitalization

JEL classification: C6; D8

1. Introduction

Since the seminal work of Mirrlees (1971), the theory of optimal screening contracts has received considerable attention. The theory has notably been applied to issues such as optimal taxation, public good provision, imperfect competition, and auctions to name a few examples. The vast majority of these applications have made the simplifying assumption that preferences can be ordered by a single dimension of private information. For instance, in the canonical model of Mussa and Rosen (1978), consumers' preferences are ordered by their willingness to pay for additional units of quality. However, in most economic situations, multiple dimensions of pri-

Received August 26, 2003, revised February 11, 2004, accepted February 27, 2004.
*Correspondence to: Economics Division, Office 3007, School of Social Sciences, University of Southampton, SO17 1BJ Southampton, U.K. E-mail: lpr203@socsci.soton.ac.uk. I thank Guillaume Carlier, Vincenzo Denicolo, and three anonymous referees for useful comments.

vate information seems to be a more appropriate assumption. In the model of Mussa and Rosen, we might think of consumers' preferences not only being ordered by their willingness to pay for additional units of quality but also by their opportunity costs. Similarly, in an employer-employee relationship, it is reasonable to assume that employees have private information not only about their productivity but also about their preference for leisure and time.

However, solving a multidimensional screening problem is far less straightforward than solving its one-dimensional counterpart. The essential difficulty lies in the *lack of complete ordering* of preferences. To take an analogy, a subset of the real line is completely ordered, a point a is bigger than, smaller than, or equal to another point b. But a subset of a two-dimensional Euclidean space is not completely ordered (say, with respect to the component-wise order). As an example, the points (1, 2) and (2, 1) are not comparable. To be more precise, this is the lack of complete ordering in multiple-dimensional environments that is at the source of most difficulties, not the multiple dimensionalities by itself. To see this, consider the multidimensional screening model of Courty and Li (2000). They study a problem of refund pricing by airline companies in which information evolves over time. At the time of buying his ticket, a consumer does not know his valuation of the trip but knows the distribution of the possible valuations of the trip. The consumer learns his valuation after having bought the ticket. In their model, the consumer type is clearly multidimensional (in fact, infinite dimensional since a type is a distribution function). However, Courty and Li impose sufficient conditions to guarantee a complete ordering of the consumer preferences (i.e., a sufficient single-crossing condition), which renders the analysis of their model similar to a one-dimensional one. Unfortunately, such sufficient conditions do not translate easily to other multidimensional screening models.

In turn, this lack of complete ordering implies that we are generally uncertain as to which incentive compatibility constraints bind. Hence, we are forced to maximize profits subject to a far larger set of *global* constraints. This problem clearly does not apply in a one-dimensional screening model as the single crossing condition implies a complete ordering of the preferences, and typically the incentive compatibility constraints are determined by *local* conditions: a simple differential equation (a first-order condition) and a monotonicity condition (a second-order condition).

The present paper is part of the burgeoning literature on multidimensional screening contracts; see for instance Armstrong (1996), Basov (2001), Carlier (2001), McAfee and McMillan (1988), Rochet and Chone (1998), and Rochet and Stole (2001) for an excellent survey. An important result of this literature is that bunching (i.e., several types of an agent will be offered the same contract) is a robust phenomenon. This result sharply contrasts with one-dimensional screening problems where perfect discrimination of types is the rule rather than the exception; however, see Julien (2000) for the possibility of bunching in one-dimensional problems. First, perfect screening (i.e., each type of an agent is offered a different contract) might be ruled out by dimensionality considerations, this is bunching of the first type in the terminology of Rochet and Chone (1998). In other words, if the principal has fewer instruments than the dimension of the type space, he is not able to perfectly screen since he does not have enough degrees of freedom. Second, even though the principal has enough instruments, there typically exists a conflict between the desire to extract as much rent as possible from the agent and the need to satisfy the second-order incentive compatibility constraint. As a result of this conflict, several types will be offered the same contract. This is an example of bunching of the second type in the terminology of Rochet and Chone (1998).

Unfortunately, most of the problems studied in the literature do no admit closed-form solutions and thus limit possible applications to interesting economic problems. Notable exceptions are Laffont et al. (1987), Lewis and Sappington (1988), and Carlier and Gaumont (2002). These authors study a problem where the principal has a unique instrument to screen several characteristics of an agent. For instance, in Carlier and Gaumont (2002), the private characteristics are the productivity and the discount rate of a worker and, in Lewis and Sappington (1988), private characteristics are cost and demand functions of a regulated firm. The common approach followed in these papers is first to transform the problem into a one-dimensional problem (the aggregation step) and second to solve the one-dimensional problem (the maximization step). In this paper, we follow the same approach.

The purpose of this paper is to offer another example for which we can *completely* characterize the optimal menu of contracts that a principal offers to an agent. The novelty is that the principal problem turns out to be of the *linear programming* type in our example, while, in previous examples, it is strictly convex. To the best of my knowledge, no such an example exists in the literature. In this respect, emphases are put on simplicity, clarity, and rigor, and it is hoped that this didactic example would help applied economists in solving similar problems. In the proposed example, shareholders of a cash-constrained firm propose to the firm management a recapitalization in counterpart of the distribution of future dividends. We assume that the dividends that the firm can distribute are proportional to the revenue the firm generates in its activity, say the production of cars. Moreover, the firm has private information over its production costs, i.e., the prices at which the firm buys its inputs *and* its technology. As both production costs and technology affect the revenue that the firm can generate, shareholders face a (multidimensional) screening problem. Indeed, the firm management can *overstate* its production costs and/or *understate* its productivity in order to distribute fewer dividends to shareholders. Without incentives to tell the truth, the firm management will clearly lie as it increases its payoff.

Two key ingredients are responsible for the linearity: risk-neutrality of shareholders and the firm and a restriction to linear contracts. Risk neutrality is a reasonable assumption to make if we have in mind an economic situation in which shareholders and the firm are sufficiently diversified. For instance, if shareholders are banks and the firm is Virgin Corporation, it is certainly reasonable to assume risk neutrality as banks have diversified portfolios and Virgin is sufficiently diversified in producing many goods and services (cars, books, leisure, flights, etc.). As for the restriction of the contracts space to linear contracts, it is a simplifying assumption that enables us to obtain closed-form solutions. At this point, it is worth stressing that both restrictions are certainly too restrictive in some economic problems, but since the purpose of this paper is to offer a didactic example on how to solve a linear multidimensional screening problem, I see these two assumptions as simply forcing the structure of the model to be linear. The main result is that there always exist optimal menus of contracts with at most two contracts offered: a low dividend, low recapitalization contract and a high dividend, high recapitalization contract. This extreme case of bunching sharply contrasts with previous contributions in which, although several types will be bunched on the same contract, the principal nonetheless offers a continuum of contracts.

In Section 2, the model is presented. Section 3 is devoted to the characterization of admissible contracts, notably the aggregation step. Section 4 addresses the existence of admissible contracts and solves the shareholders' problem. Finally, Section 5 extensively discusses the results.

2. A Simple Model

We present a simple model of conflict between shareholders and the management of a firm when the management can manipulate or divert the revenue the firm generates because of private information. The recent collapse of Parmalat in Italy is a good illustration of the diversion of funds by the management. More than €8 billions have been diverted!

A price-taking firm produces a good, say cars, from N different inputs $(z_j)_{j\in\{1,\ldots,\mathrm{N}\}}$ according to the production function f with:

$$f(z_1,\ldots,z_\mathrm{N})=\sum_{\iota=1}^{\mathrm{N}} b_j \ln(a_j z_j), \tag{1}$$

where $\sum_{i=1}^{N} b_j = 1$ and all b_j's are strictly positive. We denote by $\ln(\cdot)$ the logarithmic neperien function. Notice that the production function f is a monotone increasing transformation of a constant return to scale Cobb-Douglas production function and has constant elasticity of substitution. Here a_j is a productivity factor, which is specific to input z_j; all a_j's are strictly positive. We normalize the price of the good to one and denote the (strictly positive) prices of inputs $w_1,\ldots,w_\mathrm{N}$. However, there are two unusual things about this firm. First, instead of maximizing its profit, we assume that the management maximizes its revenue R. Fluck (1999) has shown that by maximizing its revenue, management minimizes the probability of being dismissed. Or the management might simply want to have bigger sales than any other firm. Second, the firm is cash-constrained. In particular, it has only M dollars on hand before production. For simplicity, we set $M = 0$.

Shareholders propose to the firm management an increase in capital (i.e., a transfer of cash) in counterpart of a future distribution of dividends. For simplicity, we assume that dividends are proportional to the revenue realized; that is, if the firm realizes a revenue of R, $(1-\alpha)R$ of the revenue is distributed to the shareholders and

αR is kept by the management. We also assume that the proposed recapitalization depends on the proportion of revenue the firm is willing to distribute. The higher the proportion of revenue the firm is willing to distribute, the higher the recapitalization. After all, shareholders do not inject money in firms without substantial financial counterparts. Thus, the recapitalization is a function $T(\cdot)$ of the proportion α of revenue distributed as dividends, and we let $T(\alpha) = T^{1/\alpha}$. We also assume that shareholders value the dividends distributed $v[(1-\alpha)R]$. We might think of the difference between the dividends distributed and the valuation of shareholders as a loss due to taxation or gains from the resale of stocks. For instance, in a Lucas asset-pricing model, stock prices are positively correlated with dividends distributed; hence more dividends distributed also implies gains from the resale of stocks. We suppose that v is increasing in dividends distributed and homogenous of degree one (to avoid scale effects).

Finally, we assume that the technology *and* the inputs' costs are private information of the firm. In other words, *two* essential dimensions of the production process are unknown to shareholders. Clearly, because of its private information, the management can manipulate or divert the revenue, and hence the dividends distributed, as in the Parmalat case. For instance, suppose that the agreed-upon proportion α is 10%, and the firm realizes revenue of €10 billions. The revenue realized obviously depends on the production costs and the technology of the firm; hence it is not observable by shareholders. It follows that the firm management has a clear incentive to *understate* its realized revenue, say by announcing €5 billions, and thus to distribute only €0.5 billions as dividends, instead of €1 billions if it tells the truth. Part of the problem of the shareholders is therefore to design a mechanism that forces the firm management to tell the truth.

More precisely, we suppose that the vector w of inputs' prices and the vector a of productivity factors is privately known to the firm management. Shareholders have some beliefs about the productivity factors and inputs' prices. They believe that (a, w) is distributed according to a non-degenerate probability measure with (continuous) density $\tilde{\rho}$ with respect to Lebesgue on the rectangle

$$[e, e^{\bar{a}}] \times \ldots \times [e, e^{\bar{a}}] \times [e^{-1}, e^{-\bar{\omega}}] \times \ldots \times [e^{-1}, e^{-\bar{\omega}}] \ .$$

We also assume that shareholders have no budget constraints, and we normalize the opportunity cost of the firm management to zero.

2.1 The Firm Problem

Given a transfer T and an agreed-upon proportion α of revenue to be distributed, the firm problem consists in maximizing its revenue subject to the cash constraint as follows:

$$\max_{\{z_j\}_{j\in\{1,\ldots,N\}}} \quad R = \sum_{j=1}^{\mathrm{N}} b_j \ln(a_j z_j)$$

s.t. $\sum_{j=1}^{N} w_j z_j \leq \mathrm{T}^{\frac{1}{\alpha}}$.

The solution to the firm problem is

$$z_j = \frac{1}{w_j} \frac{b_j}{\sum_{j=1}^{N} b_j} T^{\frac{1}{\alpha}},$$

for all $j \in \{1, \ldots, N\}$, and the optimal revenue R^* is given by

$$R^* = \sum_{j=1}^{N} b_j \left[\ln(a_j) + \ln\left(\frac{1}{w_j}\right) + \ln(b_j) \right] + \frac{1}{\alpha} \ln(T).$$

From the above equation, an increase in the productivity of one input has a positive effect on the revenue, while an increase in the price of this input has a negative effect. Similarly, by relaxing the cash constraint, the recapitalization of the firm has a positive effect on its revenue. It is also worth noting that this log-linear specification is a particularly nice feature of the model as it can easily be estimated by an econometrician. For instance, b_j is the elasticity of the firm revenue with respect to the price w_j of z_j.

Let us denote by θ the following row vector of dimension 2*N*:

$$\theta = \left(\ln(a_1), \ldots, \ln(a_N), \ln(\frac{1}{w_1}), \ldots, \ln(\frac{1}{w_N})\right),$$

with θ_n the *n*th component. θ represents all the private information of the firm, and we call θ the type of the firm. Similarly, we define the row vector ω of dimension 2*N* as follows:

$$\omega = (b_1, \ldots, b_N, b_1, \ldots, b_N),$$

with ω_n the *n*th component. Let $C = \ln(T)$ and $k = \sum_{j=1}^{N} b_j \ln(b_j)$. Notice that $k < 0$.

With this new notation, we can rewrite the optimal revenue of the firm in a more compact way, namely,

$$R^* = \sum_{n=1}^{2N} \theta_j \omega_j + k + \frac{1}{\alpha} C. \tag{2}$$

We also denote by ρ the density function of θ on the rectangle

$$[0, \bar{a}] \times \ldots \times [0, \bar{a}] \times [0, \bar{\omega}] \times \ldots \times [0, \bar{\omega}].$$

The probability density ρ is obviously obtained from the probability density $\tilde{\rho}$

by changes of variables and is easily shown to be continuous on Θ. Its support is also obtained from the support of $\tilde{\rho}$.

From equation (2), it is clear that the firm management can lie in several dimensions. In particular, it can overstate its inputs' prices and understate its productivity to claim lower revenue. But it can also moderately understate its inputs' prices and sufficiently understate its productivity to claim the same lower revenue. Thus, the problem features a lack of complete ordering of the type space.

2.2 Linear Contracts

Without loss of generality, we restrict ourselves to the class of contracts satisfying the direct revelation mechanism (see Myerson, 1979). From the previous section, it is clear that offering a dividend policy α and a recapitalization $T^{1/\alpha}$ is equivalent to offering a dividend policy α and a transfer of cash C since, for a given α and C, there exists a unique T that satisfies $T = e^{\alpha C}$. Thus, without loss of generality, we concentrate on contracts that only specify a dividend policy α and a cash transfer C. As already explained, shareholders face an adverse selection problem (remember that the firm management has private information over its productivity and inputs' prices), hence they cannot offer a first-best contract. Shareholders can nonetheless propose a menu of individually rational and incentive compatible contracts in such the way that the firm management will self-select the contract that corresponds to its type θ.

Definition 1: A *menu of linear contracts* is a pair of mappings $(\alpha(.), C(.)) : \Theta \to [\underline{\alpha}, \overline{\alpha}] \times \mathbb{R}$.

Notice that we impose the bound conditions $\alpha(\theta) \in [\underline{\alpha}, \overline{\alpha}]$ for all $\theta \in \Theta$, with $\underline{\alpha} > 0$ and $\overline{\alpha} < 1$. The lower bound captures the idea that the firm should be left with some money, i.e., it cannot distribute all its revenue, while the upper bound captures the idea that some dividends have always to be distributed. We might think of these constraints as imposed by the legislator.

In the two following sections, we completely characterize the set of incentive compatible and individually rational contracts and solve the shareholders' problem. A reader less interested in these technicalities might skip Sections 3 and 4 at a first reading.

3. Admissible Contracts

We define the set of *admissible* contracts as the set of contracts that are incentive compatible, individually rational, and that satisfy the bound conditions $\alpha(\theta) \in [\underline{\alpha}, \overline{\alpha}]$ for all $\theta \in \Theta$.

3.1 Incentive Compatible Contracts and Aggregation

We define the payoff $U(\theta, \alpha, C)$ of the firm management of type θ associated with the contract (α, C) as

$$U(\theta,\alpha,C)=\alpha\sum_{n=1}^{2N}\theta_n\omega_n+k\alpha+C=\alpha R^*. \tag{3}$$

Moreover, we denote by $Z:[\underline{\alpha},\bar{\alpha}]\to\mathbb{R}^{2N}$ the map

$$\alpha\mapsto Z(\alpha)=(\alpha\omega_1,...,\alpha\omega_{2N}),$$

and by $V:\Theta\to\mathbb{R}$ the potential associated with the menu (α,C) defined by

$$V(\theta)=U(\theta,\alpha(\theta),C(\theta)),\ for\ all\ \theta\in\Theta. \tag{4}$$

Definition 2: A menu (α,C) is *incentive compatible* if for all θ, $\theta^{'}\in\Theta$,

$$V(\theta)\geq U(\theta,\alpha(\theta^{'}),C(\theta^{'})). \tag{5}$$

In the sequel, we denote by $\partial_n V=\partial V/\partial\theta_n$ the partial derivative of V with respect to θ_n.

Proposition 1: A menu (α,C) is incentive compatible if and only if the potential V defined by (4) satisfies:

1. V is convex,
2. $\partial_n V(\theta)=\omega_n\alpha(\theta)$ *a.e.* for n = 1, ..., $2N$.

Proof: *Necessity*. Assume (α,C) is incentive compatible. Then for all $\theta\in\Theta$,

$$V(\theta)=\sup_{\theta^{'}\in\Theta}U(\theta,\alpha(\theta^{'}),C(\theta^{'})), \tag{6}$$

and U is linear in θ, therefore V is convex as the supremum of convex functions. Hence V is differentiable a.e. For a.e. $\theta\in\Theta$, the Envelope Theorem then yields

$$\nabla V(\theta)=\frac{\partial}{\partial\theta}U(\theta,\alpha(\theta),C(\theta)), \tag{7}$$

which means $\partial_n V(\theta)=\omega_n\alpha(\theta)$ a.e. for n = 1, ..., $2N$.
Sufficiency. Let V be a potential satisfying 1 and 2 of Proposition 1 and define

$$C(\theta)+k\alpha(\theta)=V(\theta)-\alpha(\theta)\sum_{n=1}^{2N}\theta_n\omega_n .$$

Then the convexity of V implies that for all θ, $\theta'\in\Theta$,

$$\begin{aligned}V(\theta)&\geq V(\theta')+(\theta-\theta')\nabla V(\theta')\\&\geq V(\theta')+\alpha(\theta')\sum_{n=1}^{2N}(\theta_n-\theta'_n)\omega_n\\&\geq\alpha(\theta')\sum_{n=1}^{2N}\theta_n\omega_n+k\alpha(\theta')+C(\theta'),\end{aligned}$$

so that (α, C) is an incentive compatible contract.

Proposition 1 gives a standard characterization of incentive compatible contracts (for similar results, see Carlier, 2001, Rochet, 1985, and Rochet and Chone, 1998). A contract (α, C) is incentive compatible if the potential associated with it is convex and its gradient belongs to the image of *Z*, i.e., $\nabla V \in \{Z(\alpha), \alpha \in [\underline{\alpha}, \overline{\alpha}]\}$. This set is a manifold of dimension one, and thus we might expect that shareholders discriminate different types of the firm in only one dimension. Two additional remarks are worth making. First, observe that the mapping *Z* is linear and thus injective. Therefore it satisfies the generalized Spence-Mirrlees condition of Carlier (2002). As in the one-dimensional case, an explicit characterization of incentive compatible contracts crucially rests upon the injectivity of *Z*, and thus our linear specification plays the role of a single crossing type condition. Second, the linearity of *U* in θ is crucial, as otherwise, the convexity of the potential is a necessary but not sufficient condition for incentive compatibility (see Rochet, 1987).

Before turning to the aggregation step, we introduce new notation. We denote by *x.y* the usual inner product of *x* and *y* in $\mathbb{R}^{2N}$.

A simple inspection of (3) should convince the reader that $\omega.\theta = \sum_{n=1}^{2N} \theta_n \omega_n$ is the relevant aggregate type. In other words, the revenue net of cash transfer $\omega.\theta$ is a sufficient statistic that summarizes all relevant information on the unobservable heterogeneity of types. It is indeed fairly intuitive since shareholders' payoff depends only on the fraction of revenue distributed, i.e., the dividends, and not on a particular profile of inputs' prices and productivity factors. While this intuition is correct, I nonetheless give a proof that illustrates the general strategy to find an aggregate type (equivalently, a sufficient statistic) in more complex problems. Observe that equation (7) forms a system of partial differential equations (PDEs) and finding an aggregate type crucially rests upon the solution of this system of PDEs. Due to the linearity of the system (7), characteristics are simply hyperplanes and Proposition 1 can be easily simplified. Although it is a simple application of the method of characteristics, we state and prove the following:

Proposition 2: *V* satisfies the requirements of Proposition 1 if and only if there exists a function $f : [0, \overline{a} + \overline{\omega}] \to \mathbb{R}$ such that:

$$V(\theta) = f(\omega.\theta) \text{ for all } \theta \in \Theta \tag{8}$$

and *f* is convex.

Proof: Assume first that *V* is Lipschitz and satisfies (7). Let θ and θ' be two points of Θ such that $\omega.\theta = \omega.\theta'$. We have:

$$V(\theta) - V(\theta') = (\theta - \theta').\int_0^1 \nabla V(\theta' + s(\theta - \theta'))ds.$$

Since (7) implies that $\nabla V(\theta' + s(\theta - \theta'))$ is collinear to ω, we obtain $V(\theta) = V(\theta')$, and thus $V(\theta)$ only depends on $\omega.\theta$. Hence, if *V* satisfies the requirements of Proposition 1, there exists a function f such that (8) is satisfied, and since the convexity of V is equivalent to that of f, we get the desired characterization.

Proposition 2 states that incentive compatible contracts are *one-dimensional* in the sense that they only depend on $\omega.\theta$, the revenue realized net of the cash transfer. Heuristically, this is not surprising since the dimension of the firm's private information is $2N$ whereas shareholders have only two instruments to screen the type of the firm.

Finally, since $(1/\omega_1)\partial_1 V(\theta) = f'(\omega.\theta) = \alpha(\theta)$, the bound conditions impose that $f'(t) \in [\underline{\alpha}, \overline{\alpha}]$ for a.e. $t \in [0, \overline{a} + \overline{\omega}]$, with t standing for the sufficient statistic $\omega.\theta$.

3.2 Individual Rationality

Definition 3: A menu (α, C) is *individually rational* if for all $\theta \in \Theta$,

$$V(\theta) \geq 0. \tag{9}$$

The following Proposition expresses the participation constraint in terms of the function f satisfying $V(\theta) = f(\omega.\theta)$ for all $\theta \in \Theta$.

Proposition 3: Let $f : [0, \overline{a} + \overline{\omega}] \to \mathbb{R}$ be a Lipschitz function such that $f' \geq \underline{\alpha}$. Then

$$\min_{\theta \in \Theta} f(\omega.\theta) \geq 0 \textit{ if and only if } f(0) \geq 0 .$$

The proof is trivial since f is increasing.

4. Optimal Contracts

Without loss of generality, we suppose that shareholders offer a contract that is accepted. The shareholders' profit consists of the dividends they received plus gains or losses from the sales of assets (including taxation) minus the initial transfer of cash. The shareholders' program consists of maximizing its total profit over the set of *admissible* contracts as follows:

$$\sup\left\{\Pi(\alpha, C),\ (\alpha, C) \textit{ is admissible}\right\}, \tag{10}$$

where

$$\Pi(\alpha, C) = \int_\Theta \left[(1-\alpha) v\left(\sum_{n=1}^{2N} \theta_n \omega_n \right) - C \right] \rho(\theta) d\theta. \tag{11}$$

4.1 Rewriting the Problem

Our first step in solving the principal's program consists of rewriting the profit (11) in function of *f*. On the one hand, we have for all $\theta \in \Theta$,

$$\alpha(\theta) = f'(\omega.\theta) \tag{12}$$
$$V(\theta) = f(\omega.\theta), \tag{13}$$

and, on the other,

$$V\left(\theta\right) = \alpha(\theta)\sum_{n=1}^{2N} \theta_n \omega_n + k\alpha(\theta) + C(\theta). \tag{14}$$

It follows that the cash transfer *C* could be as a function of *f*, namely:

$$C(\theta) = f(\omega.\theta) - (\omega.\theta) f'(\omega.\theta) - kf'(\omega.\theta)\,. \tag{15}$$

Note that *C* only depends on $\omega.\theta$, and slightly abusing notation we write $C(\omega.\theta)$ instead of $C(\theta)$ in the sequel. Let us define the probability measure μ on $[0, \overline{a} + \overline{\omega}]$ as the image of $\rho(\theta)\, d\theta$ by the linear form $\theta \mapsto \omega.\theta$, that is, for every continuous function φ on $\left[0, \overline{a} + \overline{\omega}\right]$, we have

$$\int_0^{\overline{a}+\overline{\omega}} \varphi(t) d\mu(t) = \int_\Theta \varphi(\omega.\theta)\rho(\theta) d\theta.$$

Substitution of (12) and (15) in (11) then enables us to write the shareholders' profit as a function of *f*, namely,

$$\int_0^{\overline{a}+\overline{\omega}} \left[(t + k - v(t)) f'(t) - f(t) + v(t)\right]\, d\mu(t)\,. \tag{16}$$

It is worth pointing out that the shareholders' profit is linear with respect to *f*. This linearity is precisely the main difference with previous models, which have explicitly solved optimal (multidimensional) screening contracts.

We are now ready to solve the shareholders' problem. First, observe that μ is absolutely continuous with respect to the Lebesgue measure, that μ has a has a density g_μ, and that g_μ is continuous. We denote by G_μ the cumulative function of μ. Moreover, it can be checked that this density vanishes at endpoints $g_\mu(0) = g_\mu(\overline{a} + \overline{\omega}) = 0$. (A proof is available upon request.) Second, an integration by parts in (16) yields

$$\Pi = -f(0) + \int_0^{\overline{a}+\overline{\omega}} \left[G_\mu(t) - 1 + \left(t + k - v(t)\right) g_\mu(t)\right] f'(t) dt$$
$$+ \int_0^{\overline{a}+\overline{\omega}} v(t) g_\mu(t) dt.$$

Note that the last term is a constant; it is the expected valuation of shareholders. Using Propositions 2 and 3, the shareholders' problem is then equivalent to maximizing the previous quantity (linear in *f*) in the set of functions

$$\left\{f:[0,\overline{a}+\varpi]\to\mathbb{R}, f \text{ is convex}, f'\in\left[\underline{\alpha},\overline{\alpha}\right] a.e., and\ f(0)\geq 0\right\}.$$

Any solution *f* obviously is such that the *participation constraint is binding at the bottom*: *f*(0)=0, and taking $u=f'$ as a new unknown (which is natural since $u(\omega.\theta)=f'(\omega.\theta)=\alpha(\theta)$), we have the following:

Proposition 4: The principal program is equivalent to

$$max_{u\in K} L(u), \tag{17}$$

where L is the linear form:

$$L(u)=\int_0^{\overline{a}+\varpi} h(t)u(t)dt,$$

with

$$h(t)=G_\mu(t)-1+(t+k-v(t))g_\mu(t) \tag{18}$$

and

$$K=\left\{u:[0,\overline{a}+\varpi]\to\left[\underline{\alpha},\overline{\alpha}\right], u \text{ is non-decreasing}\right\}.$$

The proof immediately follows from previous computations and Propositions 2 and 3. Note that *K* is convex and compact, say for instance in the weak (or even strong) L^P topology ($p\in(1,\infty)$). And since *h* is continuous and bounded, the maximum of (17) is achieved. This proves the existence of at least one solution.

4.2 The Geometry of Optimal Solutions

Since the shareholders' program is a simple linear program, the maximum is achieved in at least one extreme point of *K*. Moreover, Krein-Millman's Theorem (see Royden, 1988) and compactness of *K* in L^P imply that the set of solutions of (17) (which is a *face* of *K*) is the closed convex hull of the set of extreme points which also solve (17). We shall therefore focus on solutions in the set of extreme points of *K*.

The next result characterizes extreme points of *K*—these are the non-decreasing functions which take values only in $\{\underline{\alpha},\overline{\alpha}\}$. Without loss of generality, we normalize non-decreasing functions so as to be right-continuous.

Lemma 1: The set of extreme points of *K*, ext(*K*), is given by:

$$\text{ext}(K)=\left\{t\in[0,\overline{a}+\varpi],\ \underline{\alpha}\mathbf{1}_{[0,t)}+\overline{\alpha}\mathbf{1}_{[t,\overline{a}+\varpi]}\right\}.$$

Proof: First it is obvious that if *u* is of the form $\underline{\alpha}\mathbf{1}_{[0,t)}+\overline{\alpha}\mathbf{1}_{[t,\overline{a}+\varpi]}$, *u* is an extreme point of *K*. To prove the converse inclusion let us proceed as follows. Let $u\in$ ext(*K*). Define *c* and *d* by $c=u(0)$, $c+d=u(\overline{a}+\varpi)$ ($\underline{\alpha}\leq c\leq c+d\leq\overline{\alpha}$), and define

$$K_{c,d} = \left\{x \in K,\ x(0) = c,\ x(\overline{a} + \overline{\omega}) = c + d\right\}.$$

It is trivial to see that $K_{c,d}$ can also be parameterized with probability measures:

$$K_{c,d} = \left\{x, x(t) = c + d\int_0^t d\upsilon, \textit{for some probability measure}\ \upsilon\ \textit{on}\,[0, \overline{a} + \overline{\omega}]\right\}.$$

Hence, we write *u* in the form

$$u(t) = c + d\int_0^t d\upsilon. \tag{19}$$

Obviously *u* is an extreme point of $K_{c,d}$. We claim that this implies that υ in (19) is a Dirac mass δ_t for some $t \in [0, \overline{a} + \overline{\omega}]$. If not, υ would not be an extreme point of the set of probability measures, hence there would exist probabilities υ_1 and υ_2 with $\upsilon_1 \neq \upsilon_2$ and $\upsilon = 1/2(\upsilon_1 + \upsilon_2)$. Defining for $i \in \{1, 2\}$:

$$u_i(t) = c + d\int_0^t d\upsilon_i,$$

we would have $u = 1/2(u_1 + u_2)$ with $u_1 \neq u_2$ and $(u_1, u_2) \in K \times K$, a contradiction with the extremality of *u*. We have therefore proved that *u* is of the form $u = c\mathbf{1}_{[0,t)} + (c+d)\mathbf{1}_{[t,\,\overline{a}+\overline{\omega}]}$ for some $t \in [0, \overline{a} + \overline{\omega}]$. Finally, it is easy to prove that extremality of *u* implies that either $t \in (0, \overline{a} + \overline{\omega})$ and $(c, c+d) = (\underline{\alpha}, \overline{\alpha})$, or *u* is constant with value $\underline{\alpha}$ or $\overline{\alpha}$. This ends the proof.

As an immediate consequence, we have the following.

Corollary 1: Program (17) admits at least one solution which only takes values $\underline{\alpha}$ and $\overline{\alpha}$.

Thus there always exists an optimal menu of contracts with at most two contracts offered. Let us define

$$F(t) = \underline{\alpha}\int_0^t h(t)dt + \overline{\alpha}\int_t^{\overline{a}+\overline{\omega}} h(t)dt,$$

so that the extreme function $u = \underline{\alpha}\mathbf{1}_{[0,t)} + \overline{\alpha}\mathbf{1}_{[t,\overline{a}+\overline{\omega}]}$ is a solution of (17) if and only if *t* maximizes *F*. Finding the solutions of (17) that belong to ext(*K*) reduces then to solving the one-dimensional problem

$$\max\{F(t),\ t \in [0, \overline{a} + \overline{\omega}]\}. \tag{20}$$

Let us denote by *A* the set of solutions of (20). Since *h* is continuous, *A* is a nonempty compact subset of $[0, \overline{a} + \overline{\omega}]$. The set of solutions of (17), hence of optimal contracts is fully determined by *A* as expressed by the following statement:

Proposition 5: The set of solutions of (17) is the closed convex hull (say in the L^p topology, $p \in (1, \infty)$) of $\{\underline{\alpha}\mathbf{1}_{[0,t)} + \overline{\alpha}\mathbf{1}_{[t,\overline{a}+\overline{\omega}]}, t \in A\}$.

If the set A where F achieves its maximum is not reduced to a singleton, say $(t,t') \in A^2$ with $t < t'$ then both step functions $\underline{\alpha}\mathbf{1}_{[0,t)} + \overline{\alpha}\mathbf{1}_{[t,\overline{a}+\overline{\omega}]}$ and $\underline{\alpha}\mathbf{1}_{[0,,t')} + \overline{\alpha}\mathbf{1}_{[t',\overline{a}+\overline{\omega}]}$ are solutions of (17), and any convex combination of those step functions is also optimal. Taking convex combinations amounts to adding an intermediate third value to the function. In the case where a menu with three or more contracts yields the same profit to shareholders as a simpler menu, shareholders are more likely to offer the simplest one. Indeed, remember that we abstract from lawyers' costs, the cost of drafting contracts, etc. Hereafter, we only consider those simplest menus.

It is worth noting that if t is an interior solution, then the step function $\underline{\alpha}\mathbf{1}_{[0,t)} + \overline{\alpha}\mathbf{1}_{[t,\overline{a}+\overline{\omega}]}$ is a solution of the principal program and thus $\alpha(.)$, i.e., the fraction of revenue distributed as dividends as a function of types is discontinuous in the aggregate type $\omega.\theta$ in the interior of the participation region.

Since h is explicitly given by (18), we can further characterize the solution of (20). Note first that F is differentiable and its derivative can be computed explicitly as follows

$$F'(t) = (\underline{\alpha} - \overline{\alpha})h(t) = (\underline{\alpha} - \overline{\alpha})(G_\mu(t) - 1 + (t + k - v(t))g_\mu(t)) . \tag{21}$$

Hence, since $g_\mu(0) = g_\mu(\overline{a}+\overline{\omega}) = 0$, $G_\mu(0) = 0$, and $G_\mu(\overline{a}+\overline{\omega}) = 1$, we have $F'(0) = (\overline{\alpha} - \underline{\alpha}) > 0$, and $F'(\overline{a}+\overline{\omega}) = 0$. This proves indeed that $0 \notin A$, i.e., the constant function $u \equiv \overline{\alpha}$ is not a solution of (17). Note also that if the condition

$$t + k \le v(t) \ \textit{for all}\ t \in [0, \overline{a}+\overline{\omega}] , \tag{22}$$

is satisfied, then F is increasing; hence $A = \{\overline{a}+\overline{\omega}\}$ and the only optimal menu of contracts is a single contract such that $\alpha \equiv \underline{\alpha}$.

Proposition 6: It is not optimal for the shareholders to offer a unique contract, in which they ask for the minimal proportion $1-\overline{\alpha}$ of revenue to be distributed as dividends. If (22) is satisfied, it is optimal for the shareholders to offer a unique contract, in which they ask for the maximal proportion $1-\underline{\alpha}$ of revenue to be distributed as dividends. If (22) does not hold, then it is optimal for the shareholders to offer two contracts, one in which they ask for the minimal proportion $1-\overline{\alpha}$ of revenue to be distributed as dividends, and one in which they ask for the maximal proportion $1-\underline{\alpha}$.

Proof: Only the last statement has not been established yet. First, we can easily show that g_μ is non-increasing in a neighborhood of $\overline{a}+\overline{\omega}$ since ρ is smooth and strictly positive (proof is available upon request). Second, assume then that for t sufficiently close to $\overline{a}+\overline{\omega}$,

$$k - v(t) > 0.$$

It is enough to prove that F does not achieve its maximum at $\overline{a}+\overline{\omega}$. For $0 < t < \overline{a}+\overline{\omega}$, we have

$$\frac{F'(t)}{(\overline{\alpha}-\underline{\alpha})g_{\mu}(t)}=\left(\frac{1-G_{\mu}(t)}{g_{\mu}(t)}-(t+k-v(t))\right).$$

Since

$$0\leq\frac{1-G_{\mu}(t)}{g_{\mu}(t)}=\frac{\int_{t}^{\overline{a}+\overline{\omega}}g_{\mu}(s)ds}{g_{\mu}(t)},$$

and g_{μ} is non-increasing in a neighborhood of $\overline{a}+\overline{\omega}$, we have that for t close to $\overline{a}+\overline{\omega}$

$$0\leq\frac{1-G_{\mu}(t)}{g_{\mu}(t)}\leq t,\ \textit{hence}\ \lim_{t\to(\overline{a}+\overline{\omega})}\frac{F'(t)}{(\overline{\alpha}-\underline{\alpha})g_{\mu}(t)}<0.$$

This implies that $F'(t)<0$ for t close to $\overline{a}+\overline{\omega}$ so that $\overline{a}+\overline{\omega}\notin A$. This ends the proof.

5. Economic Interpretations

In this section, we summarize and interpret the results we have obtained in sections 3 and 4, paying special attention to bunching.

Our first result (Proposition 2) is that shareholders are able to discriminate different types of the firm in only one dimension. The only dimension in which screening may occur is the *revenue net of the cash transfer*. More precisely, we have shown that any two types θ and θ' such that $\omega.\theta=\omega.\theta'$ are offered the same contract. This result is fairly intuitive as the only dimension that really matters for shareholders is the dividends distributed, and they are proportional to the net revenue. The net revenue is thus a sufficient statistic for shareholders; they do not want to know all the fine details about the firm technology and its production costs. In other words, the net revenue summarizes everything shareholders need to know. Note that this property only follows from the incentive compatibility constraint. One-dimensional discrimination reflects the fact that perfect screening is ruled out by dimensionality considerations. Indeed, in our model, the type space is $2N$ dimensional (i.e., the N inputs' prices and the N productivity factors are private information of the firm) while the instrument space is essentially one-dimensional. Indeed, shareholders have essentially a unique degree of freedom to screen the firm, namely, the fraction α of revenue to be distributed; the cash transfer C being used to satisfy the participation constraints. Hence the dimensionality of the problem implies that perfect discrimination is impossible: *bunching of the first type* occurs, in the terminology of Rochet and Chone (1998). To avoid such bunching of the first type, shareholders would have to offer more sophisticated contracts, including, for example, a requirement to issue new stocks, distribution of stocks options or other derivatives, etc.

Second, Proposition 3 implies that if a menu $(\alpha(\cdot),C(\cdot))$ is incentive compatible, then $\alpha(\cdot)$ is non-decreasing in the aggregate type $\omega.\theta$, this feature allows a natural interpretation. The lower the production costs and the more productive the technology, the higher the net revenue is (i.e., the higher $\omega.\theta$). It follows that to secure a certain amount of dividends, the proportion $(1-\alpha)$ of revenue distributed to shareholders needs to be lower, hence α is higher. To take a concrete example, suppose that shareholders want to secure €1 billion. Then if the net revenue is €10 billions, the fraction $(1-\alpha)$ of revenue to be distributed is 10%, but if the firm is much more productive and realizes net revenue of €20 billions, the fraction needs only to be 5%. Hence the fraction of revenue kept by the firm is non-decreasing in the revenue realized. However, the magnitude of the recapitalization is ambiguous, as an increase in α has an ambiguous effect on the capitalization C and hence on the initial cash transfer T. Moreover, Proposition 3 expresses that the participation constraint is binding at the bottom, i.e., the less productive and the less efficient firm receives no informational rent. This is natural since the less productive and the less efficient firm realizes zero revenue, hence it is optimal to not recapitalize it.

Third, since shareholders' profit is linear with respect to the instrument α, the shareholders' program turns out to be of the linear programming type. Solving such a problem amounts to finding extreme points of the admissible set. This argument together with Lemma 1 implies that there always exist very degenerate optimal menus. Our second important result indeed establishes the existence of optimal menus with at most *two* contracts offered by the principal. Therefore, our specific case highlights an extreme case of bunching of the *second type* in the terminology of Rochet and Chone (1998). In fact, this extreme case of bunching is a direct consequence of the linearity of shareholders' profit; as is well-known for such types of problem, we have bang-bang solutions.

Finally, Proposition 6 gives a necessary and sufficient condition under which it is optimal to offer a single contract $\alpha \equiv \underline{\alpha}$ (complete bunching). Condition (22) means that shareholders value sufficiently dividends. However, it does not necessarily mean that they value dividends more than their face value t since k is negative. Actually, depending on the technology parameters b_j , k might be extremely negative, hence the condition can easily accommodate for $v(t) < t$. Intuitively, if more dividends does not have a disproportionate negative effect on shareholders' valuation either through taxation or losses in stocks' resale, then it is clearly optimal to offer a unique contract in which they ask for the maximal amount of dividends to be distributed. If condition (22) is violated, then the optimal menu includes both the *low* $(\alpha \equiv \underline{\alpha})$ and the *high* contract ($\alpha \equiv \overline{\alpha}$). In that case, the type space is split into two regions: productive and efficient firms with high revenue obtaining the high contract, and less productive and less efficient firms with low revenue obtaining the low contract. In fact, this result is fairly intuitive. Shareholders are willing to recapitalize a firm only if the firm is very productive and its production costs are sufficiently low, guaranteeing high revenues and hence a substantial amount of dividends redistributed. Clearly, if the firm is not extremely productive and its production costs high, then it is not really worth recapitalizing. And indeed, the worth type of a firm even gets nothing.

References

Armstrong, M., (1996), "Multiproduct Nonlinear Pricing," *Econometrica*, 64(1), 51-75.

Basov, S., (2001), "Hamiltonian Approach to Multidimensional Screening Problem," *Journal of Mathematical Economics*, 77-94.

Carlier, G., (2001), "A General Existence Result for the Principal-agent Problem with Adverse Selection," *Journal of Mathematical Economics*, 35(1), 129-150.

Carlier, G., (2002), "Duality and Existence for a Class of Mass Transportation Problems and Economic Applications," forthcoming, *Advances in Mathematical Economics*.

Carlier, G. and D. Gaumont, (2002), "A Note on Wage Differentials, Fixed-Wages and Adverse Selection," *Economics Letters*, 77(3), 349-356.

Courty, P. and H. Li, (2000), "Sequential Screening," *Review of Economic Studies*, 67(4), 697-718.

Fluck, Z., (1999), "The Dynamics of the Management-Shareholder Conflict," *Review of Financial Studies*, 12, 347-377.

Jullien, B., (2000), "Participation Constraints in Adverse Selection Models", *Journal of Economic Theory*, 93(1), 1-47.

Laffont, J. J., E. Maskin, and J. C., Rochet, (1987), "Optimal Non Linear Pricing with Two Dimensional Characteristics," in *Information, Incentives and Economic Mechanisms*, T. Groves, R. Radner, and S. Reiter eds., Minneapolis: University of Minnesota Press, 256-266.

Lewis, T. R. and D. M. Sappington, (1988), "Regulating a Monopolist with Unknown Demand and Cost Functions," *Rand Journal of Economics*, 19, 438-457.

McAfee, R. and J. McMillan, (1988), "Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory*, 46(2), 335-354.

Mirrless, J., (1971), "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175-208.

Mussa, M. and S. Rosen, (1978), "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301-317.

Myerson, R., (1979), "Incentive compatibility and the Bargaining Problem," *Econometrica*, 47, 61-73.

Rochet, J. C., (1985), "The Taxation Principle and Multi-Time Hamilton-Jacobi Equations," *Journal of Mathematical Economics*, 14, 113-128.

Rochet, J. C., (1987), "A Necessary and Sufficient Condition for Rationalizability Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191-200.

Rochet, J. C. and P. Chone, (1998), "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66(4), 783-826.

Rochet, J. C. and L. Stole, (2001), "The Economics of Multidimensional Screening," forthcoming, *Conference Volume of the Eighth World Congress of the Econometric Society*.

Royden, H. L., (1988), *Real Analysis*, Third Edition, Prentice Hall.