On the strategic use of risk and undesirable goods in multidimensional screening

A. Lachapelle[*], F. Santambrogio[†]

August 21, 2011

Abstract

A monopolist sells goods possibly with a characteristic consumers dislike (for instance, he sells random goods to risk averse agents), which does not affect the production costs. We investigate the question whether using undesirable goods is profitable to the seller. We prove that in general this may be the case, depending somehow on the correlation between agents types and aversion. This is due to screening effects that outperform this aversion. We analyze, in a continuous framework, several multidimensional cases.

Keywords: principal-agent problem, adverse selection, calculus of variations, nonlinear pricing.

1 Introduction

Principal-Agent problems and the theory of multidimensional nonlinear pricing (see for instance [3] for an overview of the topic) have known rapid advances at the end of the XXth century with the works of Armstrong [1] and Rochet & Choné [8]. This framework considers a partially informed monopolist (or principal), selling multiproducts to a continuum of agents having multidimensional types. This is the classical adverse selection case of informational asymmetries, involving multidimensional screening. The starting

[*]Operations Research and Financial Engineering, Princeton University, Princeton, NJ, USA, alachape@princeton.edu

[†]Laboratoire de Mathématiques d'Orsay, Université Paris-Sud 11, 91405 Orsay Cedex, FRANCE filippo.santambrogio@math.u-psud.fr

point of the paper, and its motivation, is the following very natural and fundamental question. Suppose that a risk neutral monopolist may "randomize" the goods, facing risk averse agents (so that random contracts are undesirable from their viewpoint), then what will happen in the end? Should the seller actually offer randomized goods (i.e. lotteries) or not? Will adverse selection force him to produce those goods preferred by the consumers, or will it be convenient to use this extra possibility in order to make profit?

Although these questions are not novel, only few answers can be found in the literature. A simple case where risk aversion is introduced in a scalar problem involving two agents can be found in [12] or [3]. In such cases, it is shown that the monopolist benefits by randomizing contracts as soon as the difference between both agent's aversion is large enough. Another work on n risk averse agents facing a not informed monopolist has been developed by Maskin & Riley in [13], where they focus on auctions on a single and indivisible good. We could also mention the work of Arnott & Stiglitz (see [2]) for a comparison between random and deterministic insurance contracts that a monopolist offers - again - to a finite number of buyers with unknown preferences. More recently, and in a more general framework, Carlier proved in [5] an existence result of optimal stochastic contracts in a relaxed version of what we call the Rochet-Choné problem (i.e. the multidimensional screening problem). However, the author of such a paper does not discuss the profitability of introducing lotteries. Let us finally mention a paper by Carlier, Ekeland & Touzi ([6]), where the principal offers random contracts, but an important difference is that he cannot control the level of the risk.

It seems, according to this literature, that this question has been analyzed so far only in discrete models (a finite number of different agents), and often with zero production costs. The novelty of our paper is that we find the same kind of answers in a multidimensional and continuous framework *à la* Rochet & Choné, thanks to a calculus of variations formulation.

Before entering the mathematical details of the principal's optimization problem, we discuss some assumptions in a rather qualitative way. The agents will be supposed to be described by a type θ and the goods the monopolist may produce by a quality x. Both θ and x belong to a same Euclidean space $\mathbb{R}^d$ (or $\mathbb{R}^d_+$) and the utility that an agent of type θ derives from consuming a good x is supposed to be the scalar product $\theta \cdot x$. Each coordinate θ_i can be viewed as a marginal utility to the ith quality of the good. Besides the utility function, the other element of the model is the production cost $C(x)$. This cost is supposed to be convex in x. We consider, as in [8], the purely quadratic production cost, i.e. $C(x) = \frac{1}{2}|x|^2$. Both these choices (linear utility and quadratic production cost) are the most classical and the most significant ones, since they already capture all the qualitative features of the

solution of more general cases.

We then consider the fact that good qualities may include a random part, and we will describe it by introducing a new variable $y \in \mathbb{R}_+$ standing for instance for their variance (or for the standard deviation). More precisely, we think of a good as a random variable X where $x = E[X]$ stands for the mean and $y = E[|X - x|^2] = E[X^2] - |E[X]|^2$ is a scalar variance. We suppose that consumer agents have *mean-variance* preferences; their utility function is of the form $\theta \cdot x + \alpha y$, where $\alpha < 0$ is a new parameter in the description of agents' type, standing for their risk aversion. Notice that we are assuming that the utility is linearly decreasing in y, which means that y must be chosen as a parametrization of the randomness of the good such that agents are linearly averse with respect to it. If by chance in modeling a situation we find that agents are linear with respect to the standard deviation instead of the variance, we should call y the standard deviation. But this is, up to now, just a matter of naming variables.

The role played by y would be more important should one look at the production cost. Due to the random components in this extended model, we will call it $C_R(x, y)$. In this random interpretation we have in mind at least two toy models for the dependence of C_R on y. In some situations C_R depends increasingly on y: it is the case when the unit production cost is quadratic, which means that producing a random variable X costs proportionally to $|E[X]|^2 = |E[X]|^2 + (E[X^2] - |E[X]|^2) = |x|^2 + y$. But there are also cases where C_R depends decreasingly on y: this happens when the producer needs special machines for guaranteeing more precise good qualities, and he does not really choose to introduce variances, even if he can pay for reducing it.

In this paper we will focus on the intermediate case, where $C_R(x, y)$ is constant in y. This case can be obtained as a combination of the two previous effects, but more important reasons suggest to consider it as key case. First of all, we already pointed out that we use a parametrization of the variable y so as to guarantee that the utility functions of agents are linear. If C_R really depends on y, when specifying a dependence (for instance $C_R(x, y) = |x|^2 + const \cdot y$) we would be fixing a very specific case where both production costs and utilities are linear in the same parametrization, while there is no modeling reason to suppose that.
Second, the case where C_R depends in a strictly convex way on both x and y is not a new problem: it is just the Rochet-Choné Problem (see [8]) with an extra dimension. The only point which was not investigated in [8] is the fact that we look at consumers in $\mathbb{R}_+^d \times \mathbb{R}_-$ and goods in $\mathbb{R}_+^d \times \mathbb{R}_+$, i.e. we explicitly account for a characteristic that is negatively perceived by agents,

3

but the formalism would be the same.

Third, we will see that the main results of the paper consist in examples and conditions where the optimal strategy for the principal is exactly to offer goods with nonzero values of y. This makes the case where the cost decreases in y less interesting. Indeed, it is clear that whenever there exist examples where the seller offers costless lotteries, this would be *a fortiori* the case if the cost decreases w.r.t. y.

Moreover, it is clear from the choice of a cost C_R independent of y that the label "variance" does not play any special role any more, in the sense that this last characteristic may be thought of as any extra feature of the good that consumers dislike. It is easy to re-think at the whole problem in these terms : if the monopolist may produce goods with qualities (x, y), and the value of y does not affect the production cost, but negatively affects the utilities of the consumers, will he use this "undesirable" quality y?

It is widely accepted that examples of this kind have been used in real life by several companies. It seems a common strategy of the firms to create low-quality goods just to attract consumers that would otherwise not enter the market. It was for instance the case of third-class seats in some trains, which were just more unconfortable than second-class ones without saving space for the railway company. In the transport domain, the saturday-night stay-over constraint of some tickets fits into this framework as well. From the airline viewpoint, indeed, this ticket has no significant differences w.r.t. to a normal one, but the constraint may be undesirable for the customer, and each customer may valuate this constraint in a different way. We will see that the reason for this kind of contracts to be profitable for the company lies in the fact that the customers that valuate the most this constraint are usually the highest types. For these reasons, we will from now on refer to the quality y as the *undesirable quality* and to α as the *aversion parameter* (and not specifically risk aversion). And, for the sake of generality, we will drop the random terminology.

Actually, thanks to these last considerations, the problem we study and the (partial) answers we give may switch very easily from the random interpretation to the more abstract and general one of an undefined "undesirable" quality. We can now link our work with another main class of studies on undesirable goods, called *damaged goods*. This terminology has been introduced by Deneckere and McAfee in [9] and describes situations in which a portion of a good is intentionally damaged or decreased (i.e. when the monopolist produces a lower quality good from a higher quality good). The authors motivate their work with several examples of what is known as "crimping the product"

in the field of marketing (eg. softwares, microprocessors, etc.). Actually, the saturday night example of the previous paragraph is taken from their work. They adopt a very simplified setting where the seller produces two goods and sell them to two agents or to a continuum (with one-dimensional valuation type). The authors investigate the profitability and the Pareto improvement of introducing a damaged good. In a second paper [11], McAfee gives sufficient and necessary conditions for when it is profitable for the monopolist to offer damaged goods. Nevertheless, the author focuses again on a very simple case with two goods and where the (one-dimensional) valuation of the damaged product is a monotonic function of the high quality good value. In this work the only fundamental difference with Mussa & Rosen's problem is that the production cost is constant for each level of the quality.

We emphasize that in the present paper, we will find again many features of McAfee's conditions but in much more general frameworks, mainly because we consider multidimensional problems and also by considering both endogenized quality reductions and more general correlations between the valuations (traditional type and aversion).

In section 2 we will first discuss more precisely the *maximization problem* for the principal, which we have not yet introduced. We will then jump to the extended problem (i.e. adding the components y and α) and present some considerations on it. Section 3 and 4 will be devoted to the description of some classes of examples where the solution of the extended problem really uses y (hence it does not coincide with the solution of the classical problem), both in $2D$ and in higher dimensions cases. The paper will be closed by a conclusive section with comments on the result we got.

2   Classical and extended problems

We start by recalling the classical problem studied by Rochet and Choné, including its powerful calculus of variations formulation.

General setting Let us see more in details the very famous problem studied by Mussa & Rosen (see [14]) in the one dimensional case and by Armstrong and Rochet & Choné for the case of multidimensional types (see respectively [1, 8]). It is a nonlinear pricing model in a monopolistic situation with adverse selection. The principal (also the monopolist, the firm, the seller...) knows the global distribution f of agents types on a convex open bounded domain $\Omega \subset \mathbb{R}^d_+$. This domain will be supposed for simplicity to be a straight rectangle (i.e. with edges parallel to the axes) and the distribution of agents over it will be described through its density, i.e. a positive function on Ω with integral

equal to 1, that we still denote by f. In this way we consider a continuum of buyers. Any agent with type $\theta \in \Omega$ derives the utility $u_\theta(x, t) := \theta \cdot x - t$ from consuming a good of quality $x \in Q \subset \mathbb{R}^d_+$ at price t. The monopolist has a strictly convex cost $C : x \in Q \to C(x)$ for producing a good of quality x.

The optimization of the monopolist's strategy is performed on possible *pricing* of the goods he sells : the monopolist may choose tariffs $t : Q \to \mathbb{R}_+$, taking into account that any agent θ buys a good $x(\theta)$ selected so as to maximize $\theta \cdot x - t(x)$; she also has the possibility of not buying anything if for every good x she gets $\theta \cdot x - t(x) < 0$; globally each agent buys either zero or one good (or, equivalently, she buys exactly one good but there also exists a *zero good* $0 \in Q$ whose production cost is 0, which gives utility 0 to everybody, and such that the principal may not charge for it, i.e. $t(0) = 0$ is prescribed). In this way the profit for the firm is

$$\int_\Omega \Big(t(x(\theta)) - C(x(\theta)) \Big) f(\theta) d\theta.$$

Thus, the principal's optimization problem has another optimization problem embedded: the agent's optimization problem.

A more standard viewpoint in *incentive theory*, which by the way avoids the problem of possible multiple goods x maximizing the utility (net of price) of the agent, is the following : agents buy either one or zero good and, on the other hand, the firm directly chooses a family of contracts $(x(\theta), t(\theta))_{\theta \in \Omega}$ (i.e. produces $x(\theta)$ and sells it at price $t(\theta)$ to the agent of type θ), and has to take into account some incentive and participative constraints.

The first one is known as the Incentive Compatibility (IC) constraint and basically ensures that agent of type θ will really choose $x(\theta)$ at price $t(\theta)$ when she looks at the product line $(x(\theta'), t(\theta'))_{\theta' \in \Omega}$. The second one is the Individual Rationality constraint (IR), modeling the fact that agents do not buy if they get a utility lower than what they would get if they bought nothing, i.e. they have a reservation utility equal to zero. We define the two constraints as:

(IC) $\theta \in \operatorname{argmax}\{\theta \cdot x(\theta') - t(\theta'), \ \theta' \in \Omega\}$
(IR) $\theta \cdot x(\theta) - t(\theta) \geq 0$.

The classical problem then reads as

$$\sup \left\{ \int_\Omega \Big(t(\theta) - C(x(\theta)) \Big) f(\theta) d\theta, \ (x, t) \text{ satisfies (IC) and (IR)} \right\}.$$

Notice that it is well-known that this problem admits a unique solution.

Calculus of variations A very powerful tool to study the above problem is the following idea : introduce (as in [1, 8]) the potential (or *surplus function*)

$$v(\theta) := \max\{\theta \cdot x(\theta') - t(\theta'), \ \theta' \in \Omega\}$$

and define

$$\mathcal{K}_{\mathcal{D}} = \{v : \Omega \to \mathbb{R}_+, \ v \text{ is convex, and } \nabla v \in \mathbb{R}^d_+\}.$$

Then the problem may be re-stated in terms of v, in the following way :

$$(\mathcal{P}_{\mathcal{D}}) \quad \sup_{v \in \mathcal{K}_D} J_D(v) = \int_\Omega \Big(\theta . \nabla v(\theta) - v(\theta) - C(\nabla v(\theta))\Big) f(\theta) d\theta.$$

It is a calculus of variations formulation subject to convexity constraints. This problem, after being introduced in [8], has been studied by Carlier and Lachand-Robert in [7] as far as regularity results are concerned. Surprisingly, we will really need their C^1 regularity theorem for the purposes of our multi-dimensional results, which is unexpected for a purely analytic result. Notice that this formulation is very useful to compute the solution, either explictely (see section 3) or numerically ([8, 10]). From now on we refer to the solution of $(\mathcal{P}_{\mathcal{D}})$ as the *classical solution*, in opposition to the solution of the extended problem that we introduce in details in the next paragraph.

The extended problem We now add a component to good qualities and to agent types. They are denoted by y and α, respectively, and we impose $y \geq 0$ and $\alpha \leq 0$. In particular, α is non-positive (we will impose that it belongs to a bounded interval of $\mathbb{R}_-$, $\alpha \in A := [-\kappa, 0]$) and it models agent's aversion to the new (one-dimensional) characteristic of the good quality.
In the sequel, we refer to the component α as the *aversion parameter* and to the corresponding one-dimensional characteristic of the good as the *undesirable quality* or, more simply, we call it y. In the present work we look at agents having linear utilities and the case where the y component of the good quality does not lead to any production cost.
In other words, we consider utilities of the type $u_{\theta,\alpha}(x, y, t) = \theta \cdot x + \alpha y - t$, i.e. the utility that any agent of type $(\theta, \alpha) \in \Omega \times A$ derives from buying $(x, y) \in \mathbb{R}^{d+1}_+$ at price t.
As before, the distribution h of agent types on $\Omega \times A$ is given (mathematically, it is an absolutely continuous measure on $\Omega \times A$, and we will denote by h its density as well). When integrating according to this density we will equivalently write $h(\theta, \alpha) d\theta d\alpha$ or $dh(\theta, \alpha)$. We will always tacitly assume that the $\theta-$marginal of h will be the same density f that we consider in the classical problem (i.e. $\int_A h(\theta, \alpha) d\alpha = f(\theta)$ for every $\theta \in \Omega$).

7

As we mentioned before, we consider the case where the production cost (denoted in this case by C_R) does not depend on y, and we take $C_R(x, y) = C(x) = \frac{|x|^2}{2}$. Introducing $w(\theta, \alpha) := \max\{\theta \cdot x(\theta') + \alpha y(\alpha') - t(\theta'), \ (\theta, \alpha') \in \Omega \times A$ and the set

$$\mathcal{K}_\mathcal{R} = \{w : \Omega \times A \to \mathbb{R}_+, \ w \text{ is convex, and } \nabla w \in \mathbb{R}_+^{d+1}\},$$

we can directly jump to the variational formulation of the problem, which is

$$(\mathcal{P}_\mathcal{R}) \sup_{w \in \mathcal{K}_R} J_R(w) = \int_{\Omega \times A} \Big((\theta, \alpha) . \nabla w(\theta, \alpha) - w(\theta, \alpha) - C(\partial_\theta w(\theta, \alpha))\Big) dh(\theta, \alpha).$$

Notice that, due to the fact that the $C-$part of the cost (the strictly convex one, which is given by the production cost) does not depend on $\partial_\alpha w$, this problem does not coincide with a multidimensional version of Mussa-Rosen's problem (i.e. with a Rochet-Choné problem). We will not enter in this paper the problem of existence of a maximizer for $(\mathcal{P}_\mathcal{R})$, which is a little bit trickier than that for $(\mathcal{P}_\mathcal{D})$ or for Rochet-Choné, due to the lack of bounds on $\partial_\alpha w$. Anyway, a maximizer does exist, see for instance [4].

The following - straightforward - proposition states that $(\mathcal{P}_\mathcal{R})$ is really an extension of $(\mathcal{P}_\mathcal{D})$.

Proposition 2.1. *If the solution $\overline{w}$ of $(\mathcal{P}_\mathcal{R})$ is such that $\partial_\alpha \overline{w} = 0$, then it coincides with the solution of $(\mathcal{P}_\mathcal{D})$ and $\sup (\mathcal{P}_\mathcal{R}) = \sup (\mathcal{P}_\mathcal{D})$. On the other hand, if h is a measure concentrated on $\{\alpha = 0\}$, then $(\mathcal{P}_\mathcal{R}) = (\mathcal{P}_\mathcal{D})$.*

In the sequel we analyse some easy features of the maximization problem $(\mathcal{P}_\mathcal{R})$, focusing on the question whether the principal really uses positive values of the new quality y. The two following results show that adverse selection will "control" the use of y, even if not preventing it completely. We will interprete later these results in terms of *discrimination* (or *screening*) and *correlation*.

Proposition 2.2. *If $\overline{w}$ is a solution of $(\mathcal{P}_\mathcal{R})$ such that $\partial_\alpha \overline{w} \geq c$, for some nonnegative constant c, then $c = 0$.*

Proof. Let $\overline{w} \in \mathcal{K}_R$ be a solution of $(\mathcal{P}_\mathcal{R})$ such that $\partial_\alpha \overline{w} \geq c$. We are only concerned with perturbations of the type

$$w_\varepsilon(\theta, \alpha) = \overline{w}(\theta, \alpha) - \varepsilon(\alpha + \kappa),$$

for $0 \leq \varepsilon \leq c$. Since $\partial_\theta w_\varepsilon = \partial_\theta \overline{w}$ and $\partial_\alpha w_\varepsilon = \partial_\alpha \overline{w} - \varepsilon$, it is clear that w_ε is convex and increasing. The constant κ has been inserted in order to

guarantee the other admissibility constraint $w_\varepsilon \geq 0$. To see that it is satisfied it is sufficient to check it in the corner of $\Omega \times A$ where all the coordinates are minimal. In this point we have both $\overline{w} = 0$ and $c(\alpha + \kappa) = 0$, so that $w_\varepsilon \in \mathcal{K}_R$. Then, we easily get

$$J_R(w_\varepsilon) - J_R(\overline{w}) = \varepsilon\kappa \geq 0, \ \forall \varepsilon \in [0, c].$$

This contradicts the optimality of $\overline{w}$ except when $c = 0$ (so that $\varepsilon = 0$). $\quad\square$

An interesting consequence of proposition 2.2 is that the monopolist does not offer products involving nonzero y when he cannot use them to discriminate (which is the case if the value of y is the same for every good he sells), i.e. to screen the agents. If one restricted the offer of the monopolist adding the constraint $y = $ constant, then the solution would be the classical one.

Moreover, the same proposition proves that the minimal value of y among the goods that are really present in the market must be 0. The principal will not increase the value of y for every good just for the sake of consumers' unhappiness, it will not be profitable to him!

There is another general case where we can prove the principal does not offer goods with undesirable qualities. Indeed, the following proposition says that he does not use the undesirable quality as soon as the distribution of types θ and the one of aversions α are independent.

Proposition 2.3. *If $h(\theta, \alpha) = f(\theta)g(\alpha)$, then the classical solution is the unique minimizer of $(\mathcal{P}_\mathcal{R})$.*

Proof. Assume that $\overline{w}$ is a solution of $(\mathcal{P}_\mathcal{R})$. Define $v(\theta) = \int_{\Omega \times A} \overline{w}(\theta, \alpha)g(\alpha)d\alpha$. Since v does not depend on α, we get $J_R(v) = J_D(v)$. On the other hand, using

$$(\nabla v(\theta))^2 \leq \int \left(\partial_\theta \overline{w}(\theta, \alpha)\right)^2 g(\alpha)d\alpha \quad \text{(Jensen's inequality)},$$

$$-\alpha\partial_\alpha\overline{w} \geq 0 \quad \text{(since } \alpha \leq 0 \text{ and } \partial_\alpha\overline{w} \geq 0\text{)},$$

$$\int \theta\nabla v(\theta)f(\theta)g(\alpha)d\theta d\alpha = \int \theta\,\partial_\theta\overline{w}(\theta, \alpha)f(\theta)g(\alpha)d\theta d\alpha \quad \text{(by linearity)},$$

one has:

$$J_R(v) = \int_{\Omega \times A} \left(-\frac{1}{2}(\nabla v(\theta))^2 + \theta\nabla v(\theta) - v(\theta)\right) f(\theta)g(\alpha)d\theta d\alpha$$

$$\geq \int_{\Omega \times A} \left(-\frac{1}{2}(\partial_\theta\overline{w}(\theta, \alpha))^2 + (\theta, \alpha)\cdot\nabla\overline{w}(\theta, \alpha) - \overline{w}(\theta, \alpha)\right) dh(\theta, \alpha) = J_R(\overline{w}).$$

Hence, the maximization may be restricted to those convex functions that only depend on θ, and, among them, the maximizer is $\overline{v}$. Also remark that the inequality $-\alpha\partial_\alpha\overline{w} \geq 0$ is strict for functions truly depending on α which proves that $\overline{v}$ is the unique solution of $(\mathcal{P}_\mathcal{R})$. $\qquad\square$

In the next sections, our study consists in describing situations where the solution of the extended problem is not the classical one. We will also provide a complete example showing the impact of the correlation (and generalizing the previous proposition).

3 From $1D$ to $2D$ screening problems

In this part we provide first an example where the seller offers goods with various undesirable qualities y (hence showing that it is sometimes profitable to use it for screening purposes) and then analyze two classes of examples to enlighten the role of correlation. To be able to deal with these examples, we need before to introduce some computations on the classical solution.

Since Problem $(\mathcal{P}_\mathcal{R})$ is pretty much involved in the general setting, the examples of this section will be presented in the simplest case, i.e. $d = 1$. Notice that this means that the original problem (i.e. without y and α) is one-dimensional, and it actually becomes two-dimensional after adding these new variables.

The simplest example we can provide is the case of a quadratic production cost and a uniform distribution of type θ on segment $\Omega = [\theta_{min}, \theta_{max}]$. The next proposition is well-known and describes the explicit classical solution in a class of cases, which includes this one. We denote by F the primitive of f, i.e. $F(\theta) = \int_{\theta_{min}}^{\theta} f(s)ds$.

Proposition 3.1. *Suppose that f is a positive regular density on $[\theta_{min}, \theta_{max}]$, with $\int_{\theta_{min}}^{\theta_{max}} f(\theta)d\theta = 1$, that the production cost is quadratic, i.e. $C(x) = \frac{1}{2}x^2$. Then, the solution $\overline{v}$ of problem $(\mathcal{P}_\mathcal{D})$ is characterized by*

$$\overline{v}(\theta_{min}) = 0, \quad \overline{v}'(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)},$$

provided the inequalities $\theta_{min}f(\theta_{min}) \geq 1$ and $f'(\theta)(1 - F(\theta)) \geq -2f(\theta)^2$ hold on $[\theta_{min}, \theta_{max}]$. In particular, for the uniform density on $[1, 2]$, these assumptions are satisfied and $\overline{v}(\theta) = (\theta - 1)^2$.

Although the computation of $\overline{v}$ is standard, we provide it for the sake of completeness.

Proof. The key point for solving $(\mathcal{P}_{\mathcal{D}})$ is the integration by parts

$$\int_{\theta_{min}}^{\theta_{max}} vf = \int_{\theta_{min}}^{\theta_{max}} v'(1 - F),$$

where we used $F(\theta_{max}) = 1$ and $v(\theta_{min}) = 0$ (we may assume $\inf v = 0$ and, v being increasing, the minimum si realized at θ_{min}). We set $u = v'$ and we have:

$$J_D(v) = \tilde{J}(u) = \int_{\theta_{min}}^{\theta_{max}} \left(\frac{1}{2} u(\theta)^2 - \theta u(\theta) \right) f(\theta) + u(\theta) \Big(1 - F(\theta) \Big) d\theta.$$

Then we easily get that the minimal value would be realized by choosing:

$$u(\theta) = \theta - \frac{1 - F(\theta)}{f(\theta)}, \ \theta \ a.e.,$$

provided this choice is possible. This means that $u(\theta)$ must be the derivative of a function v satisfying the constraints. Since v must be increasing and convex, then u must be positive and increasing. It is sufficient to compute its derivative, to require that it is positive on $[\theta_{min}, \theta_{max}]$, and to impose $u(\theta_{min}) \geq 0$. This last condition gives $\theta_{min} - \frac{1}{f(\theta_{min})} \geq 0$, and the condition on u' gives

$$0 \leq u'(\theta) = 1 - \frac{-f(\theta)^2 + (1 - F(\theta))f'(\theta)}{f(\theta)^2} = 2 + \frac{(1 - F(\theta))f'(\theta)}{f(\theta)^2},$$

which is known as the *hazard condition*. The assumptions on F exactly ensure that these conditions are satisfied. For the particular case of the uniform density on $[1, 2]$, just replace $F(\theta) = \theta - 1$ and $f(\theta) = 1$ and deduce from this equation that $\overline{v}'(\theta) = 2(\theta - 1)$. We then obtain the expression of $\overline{v}$ above and finally check that $\overline{v} \in \mathcal{K}_D$. $\qquad \square$

Note that it is a simple case without bunch (or equivalently perfect screening i.e. the monopolist offers different contracts to different types).

Let us now look at the extended problem $(\mathcal{P}_{\mathcal{R}})$ for the quadratic production cost. Our goal is to find some cases where the solution of $(\mathcal{P}_{\mathcal{R}})$ is not the classical solution $\overline{v}$ of $(\mathcal{P}_{\mathcal{D}})$.

The strategy starts by introducing, for all $\varepsilon > 0$, the perturbations

$$w_\varepsilon(\theta, \alpha) = \overline{v}(\theta + \varepsilon l(\alpha)), \ \forall (\theta, \alpha) \in \Omega \times A,$$

where $l(\alpha) = (\alpha + a)_+ = \max(\alpha + a, 0)$ and $a \in (-\kappa, 0)$ is given. If necessary, one can extend $\overline{v}$ outside Ω by affine extensions before defining w_ε. Notice that $w_0 = \overline{v}$ and that, due to the convexity and monotonicity properties of l, w_ε is convex and increasing as well.

Proposition 3.2. *Let $\overline{v}$ be the classical solution of $(\mathcal{P}_\mathcal{D})$ for a certain density of types f. Assume that $\overline{v}$ is C^2 and that there exists $B \subset \Omega$ such that $\int_B f(\theta)d\theta > 0$ and $\forall \theta \in B$, $\overline{v}''(\theta)(\theta - \overline{v}'(\theta)) - \overline{v}'(\theta) > 0$. Then there exists a joint distribution h on $\Omega \times A$, with $\theta-$marginal equal to f, such that $J_R(w_\varepsilon) > J_R(\overline{v})$ for small $\varepsilon > 0$. In particular $\sup(\mathcal{P}_\mathcal{D}) < \sup(\mathcal{P}_\mathcal{R})$ and the solution of $(\mathcal{P}_\mathcal{R})$ is not $w(\theta, \alpha) = \overline{v}(\theta)$.*

Since $J_R(\overline{v}) = J_D(\overline{v}) = \sup(\mathcal{P}_\mathcal{D})$ and using proposition 2.1 we immediately deduce the next corollary.

Corollary 3.3. *For some joint distributions of types and aversions, the principal offers goods with nonzero undesirable quality y.*
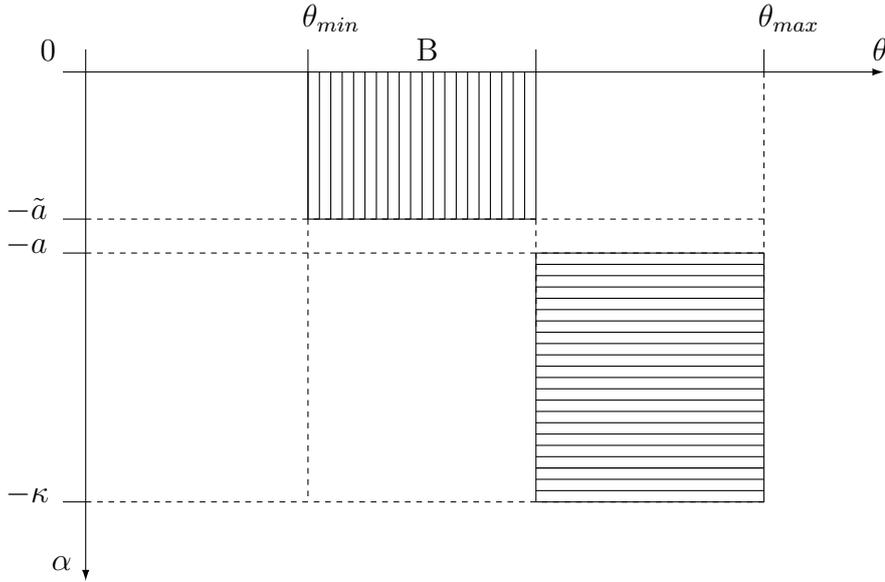

Figure 1: Structure of the example

We now prove Proposition 3.2.

Proof. First we compute the derivatives of w_ε : $\partial_\theta w_\varepsilon = \overline{v}'$ and $\partial_\alpha w_\varepsilon = \varepsilon l' \overline{v}'$, where l is differentiable (i.e. almost everywhere). We already noticed the admissibility of w_ε, which follows from the convexity and monotonicity of l and $\overline{v}$. Now, define $j(\varepsilon) := J_R(w_\varepsilon)$ and compute the first variation

$$j'(0) = \int_{\Omega \times A} \left[\left(\overline{v}''(\theta)(\theta - \overline{v}'(\theta)) - \overline{v}'(\theta) \right) l(\alpha) + \alpha \overline{v}'(\theta) l'(\alpha) \right] dh(\theta, \alpha). \quad (3.1)$$

Assume that there exists a set B satisfying the same assumptions as in the proposition. Then we can choose a density h of the following form : $h = \frac{1}{\tilde{a}}f(\theta)1_B(\theta)1_{[-\tilde{a},0]}(\alpha) + \frac{1}{\kappa-a}f(\theta)1_{B^c}(\theta)1_{[-\kappa,-a]}(\alpha)$, for some small value $\tilde{a} \in (0,a]$. In this case it is clear that the integral in (3.1) will be immediately restricted to the set $B \times (-\tilde{a},0]$, where the first term is positive. Since the goal is to get $j'(0) > 0$, we only need to choose $\tilde{a}$ small enough. In this case, the term involving $\alpha \overline{v}'(\theta)l'(\alpha)$ will be negligible (since the density h tends to concentrate around the set $\{\alpha = 0\}$), and

$$\lim_{\tilde{a} \to 0} \int \left(\overline{v}''(\theta)(\theta - \overline{v}'(\theta)) - \overline{v}'(\theta) \right)l(\alpha)dh = l(0)\int \left(\overline{v}''(\theta)(\theta - \overline{v}'(\theta)) \right)f(\theta)d\theta > 0.$$

Thus, we finally get $j'(0) > 0$.

$\square$

Remark 3.4. Note that the assumption that $\overline{v}$ should be C^2 is not really needed to get such a result. Indeed we could suppose that $\overline{v}$ is only C^1 and apply a similar technique as the one we use for proving proposition 4.1 in the next section.

As we announced before, after the very first example of Proposition 3.2, we would like to analyze some families of examples showing the importance of correlation.

We consider in this framework a density f on a segment $\Omega = [\theta_{min}, \theta_{max}]$. We also consider a joint distribution h of the following form : $h(\theta, \alpha) = \frac{1}{L}f(\theta)$ if $\theta \in \Omega$ and $\alpha \in [r(\theta) - L, r(\theta)]$, $h(\theta, \alpha) = 0$ otherwise, where r is a given affine function $r(\theta) = C + \lambda\theta$. This means that h is concentrated on a strip and that the value of α is uniformly distributed, conditionally to θ, on a segment depending on θ.

Proposition 3.5. *Given Ω and h as before, suppose $\lambda \geq 0$ (i.e. the strip is directed upwards or horizontally). Suppose that the classical solution $\overline{v}$ of $(\mathcal{P}_\mathcal{D})$ satisfies $\overline{v}'(\theta) \leq \theta$ for all $\theta \in \Omega$. Then $\overline{v}$ also minimizes $(\mathcal{P}_\mathcal{R})$.*

Proof. Consider any admissible function $w(\theta, \alpha)$, and define a function $g = g(\theta)$ in the following way: $g(\theta) = \int w(\theta, \alpha)h(\theta, \alpha)d\alpha$. In this way g is the mean value of w on the segment (that we will denote by S_θ) corresponding to θ. It is easy to check that, thanks to the structure of the density h, g is convex and increasing. Hence, it is an admissible competitor for $(\mathcal{P}_\mathcal{D})$ and the optimality of $\overline{v}$ gives (from a first-variation formula)

$$\int_\Omega \left((\theta - \overline{v}'(\theta))(g'(\theta) - \overline{v}'(\theta)) - (g(\theta) - \overline{v}(\theta)) \right)f(\theta)d\theta \leq 0. \qquad (3.2)$$

13

On the other hand, we can compute the value of $g'(\theta)$, by differentiating $g(\theta) = \frac{1}{L} \int_{r(\theta)-L}^{r(\theta)} w(\theta, \alpha) d\alpha$:

$$g'(\theta) = \frac{1}{L} \int_{S_\theta} w_\theta(\theta, \alpha) d\alpha + r'(\theta)\big(w(\theta, r(\theta)) - w(\theta, r(\theta) - L)\big) = \int_{S_\theta} \frac{1}{L} w_\theta + \lambda w_\alpha.$$

To prove the optimality of $\overline{v}$, the goal would be to prove that the first order optimality condition for $(\mathcal{P}_\mathcal{R})$ is satisfied, which implies optimality because the criterion is concave. We would like to get

$$\int_\Omega \big((\theta - \overline{v}')(w_\theta - \overline{v}') + \alpha w_\alpha - (w - \overline{v})\big) dh(\theta, \alpha) \leq 0.$$

Yet we can re-write the left hand side in terms of g, by using its definition and the computation of g', and get

$$\int_\Omega \big((\theta - \overline{v}')(w_\theta - \overline{v}') + \alpha w_\alpha - (w - \overline{v})\big) dh(\theta, \alpha)$$
$$= \int_\Omega \left((\theta - \overline{v}')(\frac{1}{L} \int w_\theta - \overline{v}') + \alpha \frac{1}{L} \int w_\alpha - (g - \overline{v})\right) f(\theta) d\theta$$
$$= \int_\Omega \big((\theta - \overline{v}')(g' - \overline{v}') - (g - \overline{v})\big) f(\theta) d\theta + \int \left(-\lambda(\theta - \overline{v}') + \frac{\alpha}{L})w_\alpha\right) dh.$$

In the last line, the first term is negative from (3.2) and the second is negative because $\overline{v}'(\theta) \leq \theta$ and $\lambda, w_\alpha \geq 0$.

As a consequence, we have checked the optimality of $\overline{v}$ in the direction of $w - \overline{v}$ for an arbitrary admissible w and this implies that $\overline{v}$ is a solution of $(\mathcal{P}_\mathcal{R})$. $\qquad\square$

Notice that the assumption $\overline{v}'(\theta) \leq \theta$ is satisfied whenever $\overline{v}$ is obtained through the formula given by Proposition 3.1. In particular, it is the case for the uniform density on $[1, 2]$.

The next class of examples will be particular to this last case, i.e. $f = 1$ on $[1, 2]$.

Proposition 3.6. *Take once more h as before, for $\Omega = [1, 2]$ and $f = 1$. Suppose $\lambda < 0$ and $C > 0$ (notice that $C = r(0) = r(1) + |\lambda|$; the strip is directed downwards but the intersection of the straight line giving its upper border with the vertical line $\theta = 0$ is above the level 0). In this case the classical solution $\overline{v}$ of $(\mathcal{P}_\mathcal{D})$ does not minimize $(\mathcal{P}_\mathcal{R})$.*

14

Proof. We will perform the same kind of variation as in Proposition 3.2. The same computations give

$$j'(0)=\int_{\Omega\times A}\left[\left(\overline{v}''(\theta)\left(\theta-\overline{v}'(\theta)\right)-\overline{v}'(\theta)\right)l(\alpha)+\alpha\overline{v}'(\theta)l'(\alpha)\right]dh(\theta,\alpha). \quad (3.3)$$

The computation becomes easier if one replaces $\overline{v}$ with its expession: $\overline{v}(\theta)=(\theta-1)^2$, $\overline{v}'(\theta)=2(\theta-1)$ and $\overline{v}''(\theta)=2$. In this case the integral function becomes $(6-4\theta)l(\alpha)+2(\theta-1)\alpha l'(\alpha)$. Moreover, if one notices $\alpha l'(\alpha)=(\alpha l(\alpha))'-l(\alpha)$, the same expression becomes $(8-6\theta)l(\alpha)+2(\theta-1)(\alpha l(\alpha))'$.

Let us now choose the function l of the same form as in Proposition 3.2, i.e. $l(\alpha)=(\alpha+a)_+$, for a value of a sufficiently close to $-r(1)$, so that the line $\alpha=-a$ intersects, in the region $1\leq\theta\leq 2$, the line $\alpha=r(\theta)$ but not $\alpha=r(\theta)-L$. In such a case the integral is performed over a triangle (the region R_1 in Figure 3 below), and l is a linear function on such a triangle, vanishing on its lower boundary $\{\alpha+a=0\}$.

Computing first the integral w.r.t. α on every vertical line of this triangle, one gets (for δ such that $r(1+\delta)+a=0$)

$$\int_1^{1+\delta}\left[(8-6\theta)\frac{(r(\theta)+a)^2}{2}+2(\theta-1)r(\theta)(r(\theta)+a)\right]d\theta.$$

It is not difficult to compute that, if $-a\to r(1)$, i.e. $\delta\to 0$, one has

$$\int_1^{1+\delta}(8-6\theta)\frac{(r(\theta)+a)^2}{2}d\theta=\frac{\lambda^2\delta^3}{3}+o(\delta^3),$$

$$\int_1^{1+\delta}2(\theta-1)r(\theta)(r(\theta)+a)d\theta=-\frac{\lambda r(1)\delta^3}{3}+o(\delta^3).$$

Hence, the condition $\lambda^2-\lambda r(1)>0$, i.e. $\lambda-r(1)=-C<0$, is sufficient to guarantee that $\overline{v}$ cannot be optimal for $(\mathcal{P}_{\mathcal{R}})$, since for a close enough to $-r(1)$, the derivative $j'(0)$, corresponding to the variation w_ε, is positive. $\square$

The two previous classes of examples give rise to the following:

Corollary 3.7. *Take a density h of the form presented before, for $\Omega=[1,2]$ and $f=1$. Under the additional assumption that $C>0$ if $\lambda<0$, the fact that the classical solution $\overline{v}$ also solves $(\mathcal{P}_{\mathcal{R}})$ is equivalent to $\lambda\geq 0$. In particular, in this situation, using the undesirable quality y is profitable for the seller if and only if the correlation of the type θ and the aversion $|\alpha|$ is nonnegative.*
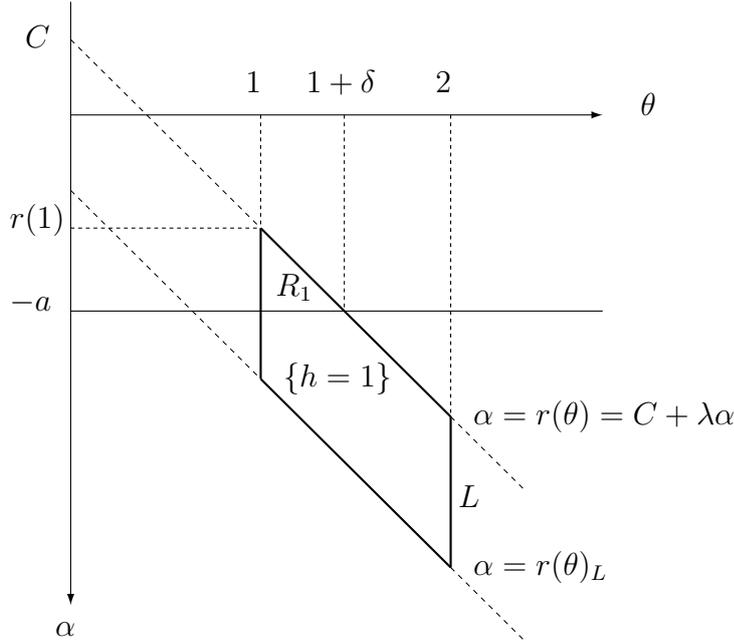
15

Figure 2: Structure of the downwards strip example

Notice also that, in all the situations above, when the optimal solution for the monopolist is not the classical one, it is profitable to offer goods with an undesirable quality y to the lowest type agents (those in region R_1, who are also, under this correlation assumption, the least averse). The higher aversion of the high-type agents prevents them from mimicking the behavior of these low-type agents.

4   Higher dimensional cases

In this section we start from the multidimensional case for the classical problem, which has been mainly studied in [1, 8]. Since less information on the classical solution in this case is known than in the 1D one, we will try to use the minimal regularity and the minimal structure results on the optimal $\overline{v}$ that are known so far.

Proposition 4.1. *Suppose that f is a density on a convex domain $\Omega \subset \mathbb{R}^d$ such that the optimal function $\overline{v}$ in the Rochet-Choné problem $(\mathcal{P}_{\mathcal{D}})$ is a C^1 convex function, with $\Omega_0 := \{\overline{v} = 0\}$ having non-empty interior. Then there exists a density h on $\Omega \times A$ such that the solution of $(\mathcal{P}_{\mathcal{R}})$ is not $w(\theta, \alpha) = \overline{v}(\theta)$.*

For the sake of this proposition, we need two things: C^1 regularity of $\overline{v}$ and the fact that there exists a flat zone $\{\overline{v} = 0\}$.
General C^1 regularity results are proven in [7] for the solutions of a wide

class of variational problems under convexity constraints, including the one of Rochet & Choné. As far as the structure of $\overline{v}$ is concerned, [8] proves that if $d = 2$, $\overline{v}$ has the shape described in Fig.3 when f is an exponential density, and checks numerically the same behavior for the uniform density. For numerical computations showing this kind of behavior, we also refer to [10]. Anyway the crucial feature we need is that the region $\Omega_0 := \{\overline{v} = 0\}$ has non-empty interior. This statement means that the seller excludes a set of consumers. Actually, it has been proven by Armstrong in [1] that, for any dimension $d \geq 2$, the set Ω_0 has positive measure as soon as Ω is strictly convex; yet, the same proof easily generalizes to other classes of domains, for instance to any straight rectangle (i.e. a rectangle whose edges are parallel to the coordinate axes). We thus easily derive the following corollary.

Corollary 4.2. *For every regular density f of types on a domain Ω which is either strictly convex or a straight rectangle in $\mathbb{R}^d$, for $d \geq 2$, there exists a density h on $\Omega \times A$ such that the optimal contract contains undesirable goods.*

We will concentrate the proof of proposition 4.1 in the case of a straight rectangle in 2D, which is the one that has been most investigated so far, but the reader may easily notice that the procedure may be extended to higher dimensions and to different shapes of Ω, since we only perform local variations.

Proof. The strategy for looking at the higher dimensions case is similar to the previous one (i.e. that of Proposition 3.2). We suppose that the solution is a function $\overline{v}$ depending on θ only, and we produce a function w_ε through $w_\varepsilon(\theta, \alpha) = \overline{v}(\theta + \varepsilon \mathbf{e} l(\alpha))$, where l is the usual function $l(\alpha) = (\alpha + a)_+$ and $\mathbf{e}$ is a given unit vector with positive components. In the following we choose $\mathbf{e} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. The function w_ε is convex and increasing with respect to each variable, provided that $\overline{v}$ is convex and increasing. We notice that

$$\nabla_\theta w_\varepsilon(\theta, \alpha) = \nabla \overline{v}(\theta + \varepsilon \mathbf{e} l(\alpha)), \ \partial_\alpha w_\varepsilon = \varepsilon \nabla \overline{v}(\theta + \varepsilon \mathbf{e} l(\alpha)) \cdot \mathbf{e} l'(\alpha), \ \text{ for a.e. } (\theta, \alpha).$$

Hence we can write $J_R(w_\varepsilon)$ and we have

$$J_R(w_\varepsilon) = \int \left(\theta \cdot \nabla \overline{v}(z_{\theta,\alpha}) + \varepsilon \alpha l'(\alpha) \nabla \overline{v}(z_{\theta,\alpha}) \cdot \mathbf{e} - \overline{v}(z_{\theta,\alpha}) - \frac{|\nabla \overline{v}|^2}{2}(z_{\theta,\alpha}) \right) dh(\theta, \alpha)$$

$$= \int \left(-k(z_{\theta,\alpha}) - \varepsilon(l(\alpha) - \alpha l'(\alpha)) \nabla \overline{v}(z_{\theta,\alpha}) \cdot \mathbf{e} \right) dh(\theta, \alpha)$$

where $z_{\theta,\alpha}$ denotes $\theta + \varepsilon \mathbf{e} l(\alpha)$ and $k(z) := \frac{1}{2}|\nabla \overline{v}(z)|^2 - z \cdot \nabla \overline{v}(z) + \overline{v}(z)$.
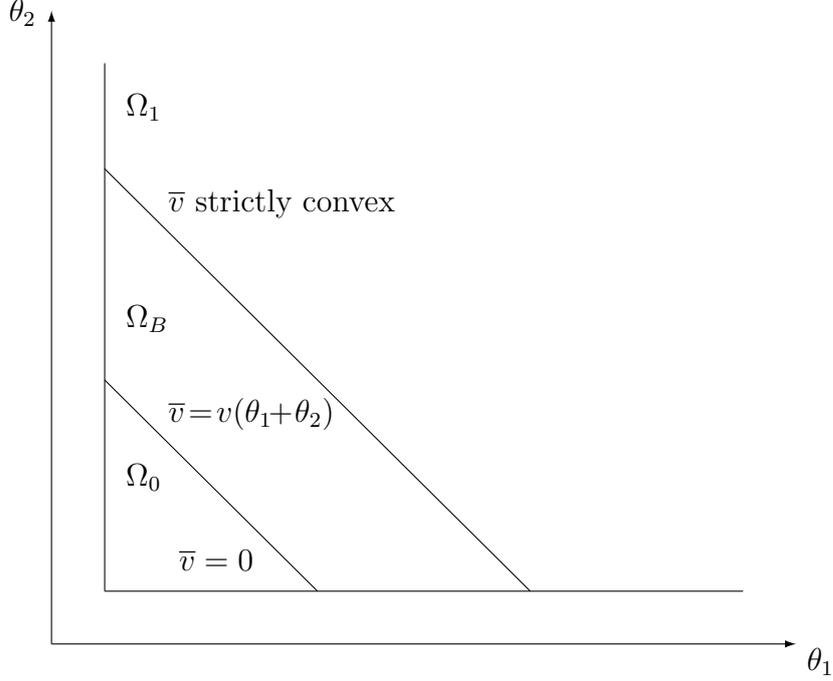
17

Figure 3: The shape of the classical solution for the exponential distribution

Simplifying assumptions on h

Since our goal is just to provide sufficient conditions in order to show an example where $\overline{v}$ is not optimal (and w_ε gives a better result than $\overline{v}$), we can make some simplifying assumptions: we suppose that $h(\theta, \alpha) = 1_K(\theta)1_S(\alpha)$ for those pairs (θ, α) such that $l(\alpha) \neq 0$, where $K \subset \Omega$ is a non-negligible given subset of Ω and $S \subset [-a/2, 0]$ is an interval where l is positive. This is not so restrictive in terms of the density f, provided f is bounded from below by a positive constant on K: it is actually always possible to express f as the $\theta-$marginal of a density h of this kind, if we adjust the rest of the density with the behavior of h on $\{l = 0\}$. This reduces the integral above to

$$-\int_S d\alpha \int_K \left(k(z_{\theta,\alpha}) + \varepsilon(l(\alpha) - \alpha l'(\alpha))\nabla\overline{v}(z_{\theta,\alpha}) \cdot \mathbf{e}\right) d\theta.$$

Simplifying assumptions on K

Now suppose that K is actually a rectangle with two sides which are parallel to $\mathbf{e}$ and two orthogonal sides denoted by ∂K_+ and ∂K_- (see Fig. 4). If we denote by d_1 the one-dimensional integration on ∂K_+ and ∂K_-, it is not

difficult to check that

$$
\begin{aligned}
\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_K k(\theta + \varepsilon \mathbf{e}l(\alpha))d\theta &= \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \int_{K+\varepsilon \mathbf{e}l(\alpha)} k(\theta)d\theta \\
&= l(\alpha)\left(\int_{\partial K_+} k(\theta)\, d_1\theta - \int_{\partial K_-} k(\theta)\, d_1\theta \right)
\end{aligned}
$$

provided that k is continuous (which is the case if $\overline{v} \in C^1$). Moreover, we have

$$
\int_K \nabla \overline{v} \cdot \mathbf{e}\, d\theta = \int_{\partial K_+} \overline{v}\, d_1\theta - \int_{\partial K_-} \overline{v}\, d_1\theta.
$$

Now, suppose that $\overline{v} = \nabla \overline{v} = 0$ on ∂K_-. This simplifies the computation and yields to

$$
\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} J_R(w_\varepsilon) = \int_S l(\alpha)d\alpha \left(\int_{\partial K_+} k\, d_1\theta \right) + \int_S (l(\alpha) - \alpha l'(\alpha))\, d\alpha \left(\int_{\partial K_+} \overline{v}\, d_1\theta \right).
$$

Since $\overline{v}$ is nonnegative and $l(\alpha) - \alpha l'(\alpha) = (\alpha + a)_+ - \alpha = a \leq 2l(\alpha)$ on S, we can estimate the last term through

$$
\int_S (l(\alpha) - \alpha l'(\alpha))\, d\alpha \left(\int_{\partial K_+} \overline{v}\, d_1\theta \right) \leq 2 \int_S l(\alpha)d\alpha \left(\int_{\partial K_+} \overline{v}\, d_1\theta \right),
$$

and get

$$
\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} J_R(w_\varepsilon) \geq - \int_S l(\alpha)d\alpha \left(\int_{\partial K_+} (k + 2\overline{v})\, d_1\theta \right).
$$

Choice of a point $\overline{\theta}$

It is now sufficient to prove that one can choose K so that $k + 2\overline{v} = \frac{1}{2}|\nabla \overline{v}(\theta)|^2 - \theta \cdot \nabla \overline{v}(\theta) + 3\overline{v}(\theta)$ is negative on ∂K_+ in order to have a positive derivative of $J_R(w_\varepsilon)$.

Here the key point is to suppose that $\overline{v}$ has a flat part Ω_0 with non-empty interior. Thanks to the convexity of $\overline{v}$, this set is convex and we suppose that it includes a neighborhood of 0. Let $\overline{\theta}$ be a point on the boundary of Ω_0 which is in the interior of Ω. Then, thanks to the monotonicity properties of $\overline{v}$, the vector $\mathbf{e}$ is an outwards vector for the set Ω_0, in the sense that $\overline{v}(\overline{\theta} + s\mathbf{e}) > 0$ for $s > 0$. The vector $\frac{1}{2}\nabla \overline{v}(\overline{\theta}) - \overline{\theta}$ will have negative coordinates, and the same is true in a neighborhood $V \subset \Omega$ of $\overline{\theta}$. If $-c$ is a constant bounding from above the components of $\frac{1}{2}\nabla \overline{v}(\theta) - \theta$ for $\theta \in V$, we can write

$$
\left(\frac{1}{2}\nabla \overline{v}(\theta) - \theta \right) \cdot \nabla \overline{v}(\theta) \leq -c\mathbf{e} \cdot \nabla \overline{v}(\theta),
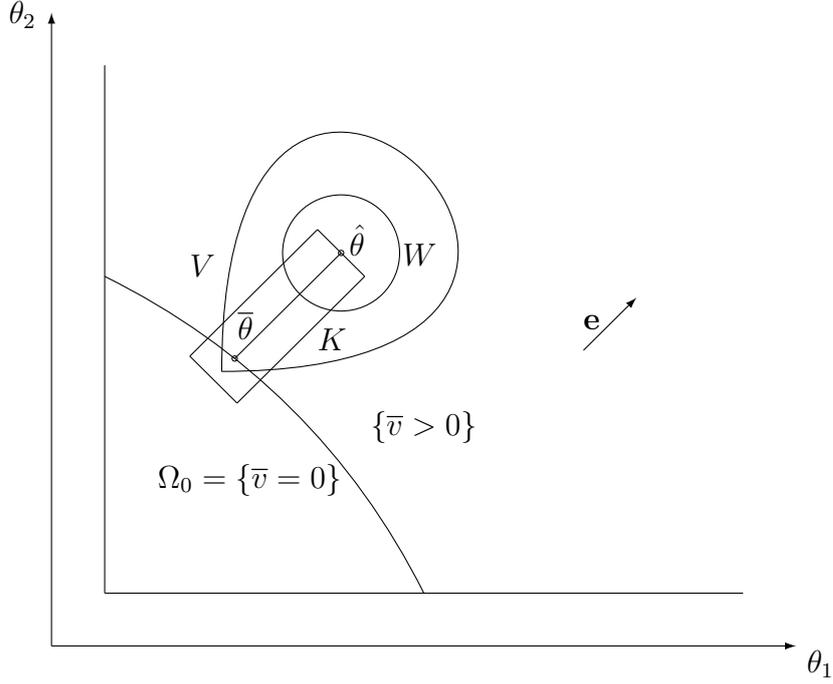$$

Figure 4: Structure of the multidimensional example

and the term $\mathbf{e} \cdot \nabla \overline{v}(\theta)$ is the derivative of $\overline{v}$ in the direction $\mathbf{e}$.

Choice of ∂K_+ so that $k + 2\overline{v} < 0$

Now, let us consider the values of $k + 2\overline{v}$ on a line segment stemming from $\overline{\theta}$ in the direction $\mathbf{e}$, i.e. we consider $(k + 2\overline{v})(\overline{\theta} + s\mathbf{e})$ for $s > 0$ such that $\overline{\theta} + s\mathbf{e} \in V$. Let us suppose that this quantity always stays positive : this means that we have the inequality

$$- c\frac{d}{ds}\overline{v}(\overline{\theta}+s\mathbf{e}) + 2\overline{v}(\overline{\theta}+s\mathbf{e})$$
$$\geq \left(\frac{1}{2}\nabla\overline{v}(\overline{\theta} + s\mathbf{e}) - (\overline{\theta}+s\mathbf{e})\right)\cdot\nabla\overline{v}(\overline{\theta}+s\mathbf{e})+2\overline{v}(\overline{\theta}+s\mathbf{e}) \geq 0 ,$$

which implies

$$\frac{d}{ds}\overline{v}(\overline{\theta} + s\mathbf{e}) \leq \frac{2}{c}\overline{v}(\overline{\theta} + s\mathbf{e}).$$

By Gronwall Lemma, this inequality implies that $\overline{v}(\overline{\theta} + s\mathbf{e}) \leq e^{\frac{2}{c}s}\overline{v}(\overline{\theta}) = 0$, which is in contradiction with the fact that $\overline{\theta}$ is on the boundary of Ω_0, i.e. that $\overline{v}(\overline{\theta} + s\mathbf{e}) > 0$.

This implies that there exists a point $\hat{\theta}$ of the form $\overline{\theta} + s\mathbf{e}$ in the neigborhood V where $k + 2\overline{v}$ is negative. The same will be true for all the points in a neighborhood W of $\hat{\theta}$. In this way, if one chooses a rectangle K such

20

that one of the two sides orthogonal to $\mathbf{e}$ is completely included in Ω_0 and the other is included in W, the derivative of $J_R(w_\varepsilon)$ is positive. $\qquad\square$

Notice that in this proof, we use again profitable (from the seller viewpoint) contracts where goods with positive y are sold to rather low type agents (i.e. the lowest type agents among those who enter the market). Actually, as one can see from Fig. 4, the agents who are concerned are those who are close to the region Ω_0, i.e. to the consumers who stay out of the market.

5   Summary and conclusions

In this paper, we have been concerned with providing partial progress on the possible profitability of undesirable contracts, for a monopolist facing averse buyers with multidimensional types. Our main contribution in comparison with the existing literature is to investigate multidimensional problems with a continuum of buyers.

We first defined an extended version of the multidimensional screening problem mainly studied by Rochet & Choné. We discussed several interpretations of the addition of an undesirable quality and, by means of a fine analysis of this model, we highlighted situations where the seller either can or cannot improve his profit by offering undesirable goods.

The optimal policy does not use undesirable goods in the case where agents type θ and aversion α are independent. The correlation between the distributions of agents type and aversion is at the heart of the answer, as we showed with a family of examples depending on a parameter λ. When $\lambda < 0$ the correlation between θ and $|\alpha|$ is positive, low-type consumers are less averse and it is profitable for the seller to use the undesirable quality y. When $\lambda \geq 0$ the correlation is negative, and the conclusion is the same as for the independent case. A similar behavior is also seen in the higher dimensional example we provided, where we made use of the presence of poorly averse agents close to the "out-of-the-market" zone Ω_0. By the way, this low-type zone of agents is also likely to be the perfect target for offering undesirable goods, as seen at the end of Section 4. What we get is consistent with the results presented in [3, 9, 11, 12]. Yet, the novelty of this paper is to analyze these same questions with a continuum of agents, in a Rochet & Choné framework.

The non profitability conclusion is also true when the monopolist is forced to offer a constant level of the undesirable quality, i.e. removing the possibil-

ity to screen the agents via such a quality (or discriminate). We saw that, if the principal really uses undesirable contracts (and he only does use them for pure screening purposes), then the undesirable characteristic y he proposes will take different values, including 0.

By providing examples in both the scalar and the multidimensional cases, we emphasized the role of the undesirable quality to screen the agents, and we have seen that the screening considerations sometimes outperform aversion.

Acknowledgements

References

[1] M. Armstrong, Multiproduct nonlinear pricing, Econometrica, 64, 51–75, 1996.

[2] R. Arnott, J. E. Stiglitz, Randomization with asymmetric information, The Rand Journal of Economics, vol.9, No.3, pp.344-362, 1988.

[3] P. Bolton, M. Dewatripoint, Contract Theory, MIT Press, Cambridge, 2005.

[4] G. Carlier, Calculus of variations with convexity constraint, J. Nonlinear Convex Anal, 3(2), 125–143, 2002.

[5] G. Carlier, Problèmes de calcul des variations issus de la théorie des contrats, PhD thesis, Université Paris-Dauphine, 2000.

[6] G. Carlier, I. Ekeland, N. Touzi, Optimal derivatives design for mean-variance agents under adverse selection, Math. Finan. Econ. 1, 57–80, 2007.

[7] G. Carlier, T. Lachand-Robert, Régularité des solutions d'un problème variationnel sous contrainte de convexité. (French) C. R. Acad. Sci. Paris Sér. I Math. 332, n.1, p. 79-83, 2001.

[8] P. Choné, J.-C Rochet, Ironing, Sweeping and Multidimensional Screening, Econometrica, vol. 66, issue 4, p. 783-826, 1998.

[9] R. J. Deneckere, R. P. McAfee, Damaged Goods. Journal of Economics & Management Strategy, 5: 149–174, 1996.

[10] I. Ekeland, S. Moreno, An algorithm for computing solutions of variational problems with global convexity constraints, Numerische Mathematik, 115, n.1, pp.45-69, 2010.

[11] R. P. McAfee, Pricing Damaged Goods, Economics, 2007, (1) ISSN 1864-6042.

[12] E. Maskin, J. Riley, Monopoly with incomplete information, The Rand Journal of Economics, 15, pp.171-196, 1984.

[13] E. Maskin, J. Riley, Optimal auctions with risk averse buyers, Econometrica, 52, pp.1473-1518, 1984.

[14] M. Mussa, S. Rosen, Monopoly and product quality, Journal of Economic Theory, vol. 18, issue 2, p. 301-317, 1978.