Implementability and Single Crossing in Bidimensional Screening

Perez Vilcarromero, Victor Miguel

Rio de Janeiro

July 2022

Advisor: Aloisio Araujo   Coadvisor: Carolina Parra

Thesis presented as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Mathematics.

Instituto de Matematica Pura e Aplicada

Resumo

Esta tesis estuda problemas de ”screening”otimo onde o instrumento do principal e unidimensional e a informação privada do agente e bidimensional. Em particular, a atenção e focada no problema do monopolista de gerar uma estrutura de preços não linear otima. Nossa contribuição principal e prover condições na utilidade dos agentes que simplificam o problema da mesma forma que a hipotese de single crossing nos modelos unidimensionais. Baixo estas condições, as restrições de compatibilidade de incentivos locais implican a restrições de compatibilidade de incentivos globais e, mais ainda, explotando uma apropiada sustitubilidade das diferentes dimensões de informação asimetrica na utilidade marginal, mostramos que o espaço dos tipos pode ser ordenado endogenamente de acordo com a utilidade marginal.

Damos uma nova caracterização de contratos implementaveis en termos da solução de um problema de Cauchy particular com dado inicial não decrescente. Esta caracterização e mais geral que aquelas que são conhecidas na literatura e não requer que a utilidade do agente seja linear nos tipos ou satisfaça McAfee McMillan (1988)’s ”single crossing generalizado”. Seguindo o trabalho de Araujo et al. (2022) usamos nossa carateterização para reformular o problema de screening bidimensional em um problema variacional simples para o qual proveemos condições de otimalidade. Tambem utilizamos estas condições para resolver varios exemplos.

Key words: Single crossing, Spence-Mirrless, Implementabilidade, compatibilidade de incentivos, Screening, Selecao adversa.

Abstract

This thesis studies optimal screening problems where the principal's instrument is one dimensional and the agent's private information is bidimensional. We focus attention on the monopolist nonlinear pricing setting. Our main contribution is to provide conditions on the utility of the agents that simplify the problem in the same way the single crossing assumption does on unidimensional models. Under these conditions, local incentive compatibility implies global incentive compatibility and, moreover, exploiting an appropriate substitubility of the different dimensions of private information on the marginal utility, we show that the type space can be endogenously ordered according to marginal valuation.

We give a new characterization of implementable contracts in terms of solving a particular Cauchy problem with nondecreasing initial data. This is a more general characterization of what is available on the literature and doesnt require agent's utility function to be linear on types or to satisfy McAfee McMillan (1988)'s Generalized single crossing. We also follow Araujo et al. (2022) and use our characterization to reformulate the bidimensional screening problem into a simpler variational problem for which we provide general optimality conditions. We use this conditions to solve several examples.

Key words: Single crossing, Spence-Mirrless, Implementability, Incentive compatibility, Screening, Adverse selection.

”As no better man advances to take this matter in hand, I hereupon offer my own poor endeavors. I promise nothing complete; because any human thing supposed to be complete, must for that very reason infallibly be faulty.[...]But I now leave my cetological system standing thus unfinished, even as the great cathedral of Cologne was left, with the crane still standing upon the top of the uncompleted tower. For small erections may be finished by their first achitects; grand ones, true ones, ever leave the copestone to posterity. God keep me from ever completing anything. This whole book is but a draught -nay, but the draught of a draught”

Herman Melville

Acknowledgments

The completion of this work would not have been possible without the participation of many people to which I'm thankful. First, I want to thank my advisor Aloisio Araujo for guiding me through this process. He conveyed to me his intuition and faith in the research agenda of screening without single crossing and in particular in the bidimensional methodology that he started developing with Sergei Vieira. Although I must confess that I was initially quite suspicious that such methodology, as he said, took more advantage of the economics of the problem instead of relying in big mathmatical weapons to kill the problem and was therefore a fruitfull terrain for further research; in the end, he was right and it was only thanks to his insistence and encouragement that I managed to get the results I now present in this thesis.

I'm also indebted to Sergei Vieira, Carolina Parra and Braulio Calagua. All of them worked with Aloisio in furthering the same research agenda to which I now make my small contribution. Their contributions served as a basis for a better understanding of the problem and the conversations I had with them were indispensable to further that understanding. Special thanks to Carolina who as my coadvisor made herself available for many conversations and followed more closely my work since the start of my PhD to the very end of the thesis stage.

Beyond the work on the thesis, I also want to thank the teachers who helped me develop into a researcher in mathematical economics. During my undergraduate years, I started to appreciate mathematics through the courses of Abelardo Jordan. Max Perez's workshop on linear algebra for economic students opened a new world to me by showing me the beauty of pure mathmatics and Alejandro Lugon's course microeconomics II made me appreciate for the first time the conjunction of mathmatics with economic theory. My debt to Alejandro goes beyond that. If it wasnt because I was lucky enough to work as his teaching assistant and he took the initiative to inquire into my plans for the future and suggest and encourage me to pursue an academic path I may not have embarked on this journey. Had I not done this, I would probably be working as an economic analyst for some gobernmetal institution and I would be considerably unhappier (and probably considerably richer). For that I'm very thankful. I'm also grateful for the opportunity to learn more about mathmatical economics from Mauricio Villalba and Juan Pablo Gamma at IMPA. Mauricio's kind reply to my questions about the academic path in a time of selfdoubt are also kindly remembered and appreciated.

I also want to thank IMPA for providing an excellent environment for academic growth. The work of all the professors whose classes I was privileged to attend as well as the work of the staff members are kindly appreciated. I'm indebted to CNPq for the PhD scholarship. Finally, I want to express my utmost gratitute to all the special people in my life, friends and family whose support and influence in my life goes beyond this thesis and my academic endeavors.

Contents

Chapter 1

Introduction

Since their origins in the 1970's, the models of adverse selection or optimal screening have become a well established part of microeconomic theory. This is in part due to the wide range of economic phenomenon that can be studied with these models. Indeed, they have been used to study monopolist nonlinear pricing, optimal taxation, auctions, labor contracts, regulation and procurement among many other applications. In all this different economic scenarios there is a common underlying structure that screening models help unveil. In each of these cases we are concerned with studying what kind of contracts may emerge from the interaction of an uninformed party that has all bargaining power (Principal) with an informed party (Agent) which posses private information about one or many of his traits that are relevant to determine the gains of contracting.

Consider for instance the framework of the monopolist nonlinear pricing which is basically the problem of exchange under asymetric information where one of the goods is money. Under perfect information, that is, if the monopolist were able to observe the valuation of its costumers in addition to having monopolist power, its rationality would imply that the monopolist will sell an efficient quantity to the costumer and will charge all the consumer surplus. Having perfect information in addition of monopolist power means that the monopolist decides how to divide the proverbial cake (i.e., the gains of trade) and hence the monopolist achieves an efficient outcome because maximizing his share of the cake is the same as maximizing the size of the cake. However, this behaviour known as first degree price discrimination is not only ilegal in some parts of the world but also not feasible in general since the monopolist doesnt know the exact valuation of the consumer for his good. There are many ”types” of consumers depending on how they value the good and the monopolist doesnt know the type of their costumers.

Screening models allows us to study what happens in the presence of these assimetries of information where the monopolist is trying to take the biggest share of the cake without knowing exactly how big the cake will be. If the monopolist demands a portion bigger than the cake itself, the cake will not be baked (trade will not happen) but it isnt coherent with its rationality and his full bargaining power either to ask for a small piece and leave the consumer with the bigger part of the cake. The main assumption that helps us ground this problem is that the consumer, as a consequence of his own rationality, will act strategically using his informational advantage to deceive the principal and maximize his own share of the cake. A rational monopolist foresees this and therefore has to solve the problem of maximizing his share of the cake but under the constraint that, whatever he does, the consumer too will try to maximize his own share of the cake and therefore minimize the share of the monopolist. This is how, mathematically, screening models reduce to optimization problems constrained by other optimization problems.

Note how our cake analogy is different from a minimax, zero-sum game since two things are going to be determined simultaneously: the size of the cake as well as the distribution of this cake. All parties can potentially benefit from having a bigger cake. However, as is usually the case in these models it may happen that to generate the biggest possible cake the monopolist would have to surrender too big a piece of the cake and therefore, unlike the perfect information scenario, maximizing the size of the cake is no longer the prefered strategy of the monopoly. This is actually one of the conclusions of the standard unidimensional model which tells us that there is indeed a conlict between baking the biggest posible cake and getting away with most of it since the participation and private knowledge of the consumer is required to produce the cake.

More concretely, the solution of the standard unidimensional model tells us that the monopolist will only offer an efficient quantity to the agent that has the highest marginal valuation for the good (i.e, only for this type the size of the cake will be maximized) and unlike the perfect information scenario, to ensure the participation of the agent the monopolist has to surrender a positive rent to the consumer (i.e, the consumer gets a non-zero fraction of the cake). To agents who dont have the highest marginal valuation for the good, the monopolist will offer suboptimal quantities of the good and also give up some positive (but smaller) rent. Agents on the lower extreme of the spectrum of marginal valuation may be even excluded from contracting: no contract will be offered to them even when there may be potential gains from trading or, in our terminology, no cake will be baked even when both parties would be better off baking and sharing a small cake.

We can either think of the monopolist offering contracts for each possible type of Agent and the agents revealing themselves at the solution by choosing the contracts designed for them (Revelation principle) or we can think that the monopolist is actually choosing a nonlinear pricing scheme, that is, he is offering different amounts of the goods at different prices and the consumers by choosing the amount they purchase are also revealing their marginal valuation for the good (Taxation principle). This later interpretation provides an explanation to the widely observed phenomena that for some goods the cost of purchasing them doesnt grow linearly with the quantity as is the case in general equilibrium models where prices are fixed and independent of the amount of units purchased. A discount for purchasing higher amounts is indeed consistent with the conclusion that only those with the highest marginal valuation will get an efficient amount while agents with lower marginal valuation, are only offered subefficient amounts.

To arrive at these conclusions, however, some assumptions are needed. We had to assume that the monopolist's good and the consumer's private information can be modeled as unidimenisonal quantities and we also had to assume that the consumer's preferences satisfies the crucial single-crossing or Spence-Mirlless assumption. However, in many, if not most, of the economic applications we can think of, the nature of the agent's private information is multidimensional rather than unidimensional. It is only due to the technical difficulties that arise from a multidimensional treatment of the type space, that most of the literature treats the informational asymetries as unidimensional. This has been recognized very eloquently by Rochet and Stole (2003) in their excellent survey of the development of the multidimensional theory:

"Unfortunately, the techniques for confronting multidimensional settings are far less straightforward as in the one dimensional paradigm[...]. As a consequence, the results [...] remain uncomfortably restrictive and possibly inaccurate (or at least non-robust) in their conclusions. In this sense, we have been searching under the proverbial street lamp, looking for our lost keys, not because that is where we believe them to lie but because it is apparently the only place where we can see." (2003, pag 151.)

This work is a contribution to the existing literature in multidimensional screening and as such it is part of the common effort to light the terrain beyond the "one-dimensional street lamp". We are interested in extending the conclusions as well as the intuition behind the unidimensional model described above, to multidimensional settings. Notice how there isnt anything inherently unidimensional in our description of the type space as being divided in regions of lower or higher marginal valuation for the good. As long as the good is unidimensional, the marginal valuation of a consumer is a well defined scalar quantity independently of the dimension of the private information. This is the key to the approach we have adopted. At the same time, we are also interested in seeing what is exclusively a consequence of the unidimensional modelling and doesnt hold when going to greater dimensions.

Most of the multidimensional literature, has allowed an arbitrary multidimensional type space only if the good itself is multidimensional and the dimension of the good and the private information are equal. Building upon the groundbreaking work of Rochet and Chone (1998), there has been great progress on this type of multidimensional screening problems in the last years (Manelli and Vincent, 2006, 2007; Figalli, et. al, 2011; Daskalakis et al, 2017; Kleiner and Manelli, 2019). In this work we focus on the less explored case where the dimension of the agent's private information is strictly bigger than the dimension of the monopolist's good.

Especifically, we will consider the case of a one dimensional good and a bidimensional type space. Although this is a rather specific case, there are many important examples that fit within this framework. For instance, this is the case of a regulator who has to determine the price that a firm should charge when the firm has private information about its costs and the demand for its product. It is also the case of a monopolist determining the quantity sold to a customer who has private information about the parameters of his own linear demand curve: the slope (sensitivity to price) and the intercept (intensity of demand). Furthermore, we believe that the results we obtain from this simpler case gives us some insights on what can be expected from other multidimensional adverse selection problems with unequal dimensions.

But even though, this work can be best understood as part of the effort to lift the assumption of unidimensionality, it is historically part of a different research agenda that tried to lift the Single crossing assumption while remaining on the unidimensional framework. It turns out that both assumptions are intimately linked. Single crossing is the reason unidimensional models are straightforward. Unidimensional models without single crossing are much more difficult to solve (Araujo and Moreira 2010; Araujo, Moreira and Vieira 2015; Schotmuller 2015). It would be too difficult to try to solve multidimensional models without an appropriate generalization of Single crossing. Consequently, we first offer a set of conditions that we consider to be the analogous to single crossing for bidimensional models and we then solve the bidimensional screening problem only for the family of cases satisfying these conditions.

This work is organized as follows. In the second chapter we present the formal model of optimal screening or adverse selection and we quickly review the received wisdom from the standard unidimensional model as well as the relevant multidimensional literature. The third chapter contains the bulk of our contribution and this is where we present our methodology for solving bidimensional screening problems. We present a set of assumptions that allows us to give a complete characterization of the implementability of contracts in terms of a cauchy problem and we also give some important interpretations on the economic meaning of our assumptions and how and why our methodology mimicks the way the standard unidimensional model works. Then, building upon the work of Araujo, Calagua and Vieira (2022), we proceed to use the cauchy problem to generate optimality conditions that helps us solve the problem. In the process, we generalize some of the optimality results of the aforementioned authors. Finally, the fourth chapter presents and discusses many examples that show how our methodology is applied in concrete cases.

Chapter 2

Preliminaries

In this chapter we present the basic general framework of our screening problem putting special emphasis on the issue of implementability. First, we will quickly review the standard one dimensional model and see how the single-crossing assumption give us a complete characterization of implementability which results in a straightforward solution method to deal with such problems. Then we will proceed to review the relevant multidimensional literature. We are interested in two different but interconnected kind of results. On one hand, we will review the attempts to characterize implementability in multidimensional settings in connection with the different proposed generalizations of the single crossing assumption. On the other hand, we will also focus on proposed algorithms and methodologies that allows us to arrive at concrete solutions in multidimensional settings where the dimension of the Agent's private information is strictly bigger than the dimension of the principal's instrument. As mentioned in the introduction, this case has been less explored than the case of equality between dimension. However, as follows from our discussion in this chapter we believe that it is the increase of the dimensionality of the type space rather than the good what brings forward greater challenges to the characterization of implementability and may be of greater economic interest.

2.1 The model

Screening problems may be formulated in a variety of different settings. For the sake of concreteness, in this work we will focus on the monopolist nonlinear pricing version of this model. Hence, we will speak of the monopolist and the consumer when referring to the Principal and the Agent. We want to study how the interaction between this two parties results in the formation of a contract on two variables: a physical good or instrument $q \in \mathbb{R}_+^n$ provided by the monopolist and a monetary transfer $t \in \mathbb{R}$ paid by the consumer. The principal's utility is $t - c(q)$ where $c \in C^2$ gives the cost associated with the provision of good q. The agent's utility depends on his type $\theta \in \Theta \subset \mathbb{R}_+^m$ which is privately known. We assume that the type space Θ is a convex subset of $\mathbb{R}_+^m$ and has positive lebesque measure. The agent's utility is quasilinear and given by $u(q, \theta) - t$ where $u \in C^3$

is sometimes refered to as the valuation function. Types are distributed on Θ according to a positive and differentiable density $\rho(\theta)$ which is common knowledge.

A direct contract or simply a contract is a pair of functions (q, t) from Θ to $\mathbb{R}^n_+ \times \mathbb{R}$. A contract is called incentive compatible if and only if

$$u(q(\theta), \theta) - t(\theta) \geq u(q(\hat{\theta}), \theta) - t(\hat{\theta}) \quad \forall \theta, \hat{\theta} \in \Theta$$

or equivalently

$$\theta \in \underset{\hat{\theta} \in \Theta}{\operatorname{argmax}} \{u(q(\hat{\theta}), \theta) - t(\hat{\theta})\} \quad \forall \theta \in \Theta$$

Hence, a contract is incentive compatible when it doesnt give the agents any incentive to misrepresent their true type. An instrument $q : \Theta \to \mathbb{R}^n_+$ is called inplementable if there exists a transfer $t : \Theta \to \mathbb{R}$ such that (q, t) is incentive compatible. We also say that a contract (q, t) is individually rational if it satisfies

$$u(q(\theta), \theta) - t(\theta) \geq 0 \quad \forall \theta \in \Theta$$

In other words, an individually rational contract cannot force types to be worst than their resevation utility which we will consider to be normalized and equal to zero for all types.

By the revelation principle, when looking to maximize his expected profit, the monopolist can restrict attention to direct and incentive compatible contracts that are individually rational. Hence, the monopolist tries to design a contract that solves the following maximization problem:

$$\max_{(q(\theta), t(\theta))} \int_{\Theta} [t(\theta) - c(q(\theta))] \rho(\theta) d\theta$$

$$s.t. \quad (IR) \quad u(q(\theta), \theta) - t(\theta) \geq 0 \quad \forall \theta \in \Theta$$

$$(IC) \quad \theta \in \underset{\hat{\theta} \in \Theta}{\operatorname{argmax}} \{u(q(\hat{\theta}), \theta) - t(\hat{\theta})\} \quad \forall \theta \in \Theta$$

Notice that what makes the screening problem a hard problem is the structure of the feasible set which reflects the informational constraints faced by the uninformed principal. Indeed, there is nothing particularly challenging in the expression of the objective functional while the feasible set cannot be treated by any variational technique without first transforming and reexpressing the conditions defining it.

2.2 The one dimensional street lamp

For this section assume $n = m = 1$ and also $\Theta = [0, 1]$ for concreteness. We say that the agent's utility function satisfies the single crossing or the Spence Mirrless assumption whenever $u_{q\theta}$ keeps a constant sign, lets say $u_{q\theta} > 0$. Hence, the content of this assumption is to order the type space according to marginal valuation: the higher the type the greater their marginal valuation for the good. However, the importance of this assumption comes from the following result:

Proposition 2.1. *Assume u satisfies single crossing:*

(a) q is implementable if and only if q is nondecreasing

(b) A contract (q, t) is incentive compatible if and only if $u_q(q(\theta), \theta)q_\theta(\theta) = t_\theta(\theta)$ and $q_\theta(\theta) \geq 0$ for all $\theta \in [0, 1]$.

Item (b) simply says that to solve the maximization subproblem of the agents the local first and second order conditions are not only necessary but also sufficient. We refer to this phenomenon as local incentive compatibility implying global incentive compatibility. The proof of (a) can be consulted on Basov (2005) on chapter 6, theorem 172 while a proof of (b) can be found on section 2.3.1 in Salanie (1997)[1]. From a computational point of view, single crossing allows us to exchange the complex condition (IC) for the simple condition $q_\theta \geq 0$ which can be easily incorporated on variational methods.

There are many different ways to take advantage of proposition 2.1 and solve the screening problem. One such way that has been proven to be very convenient is the so called "parametric utility approach" (Rochet and Stole 2003) which consists in defining the informational rent $V(\theta) = u(q(\theta), \theta) - t(\theta)$ and change variables from (q, t) to (q, V). If we also assume $u_\theta > 0$ its easy to see using item (b) that $V_\theta(\theta) = u_\theta(q(\theta), \theta)$ and hence the left side of (IR) is increasing in θ. Therefore, we can reduce (IR) to $V(0) \geq 0$ and our problem reduces to

$$\max_{(q(\theta),V(\theta))} \int_{[0,1]} [u(q(\theta), \theta) - c(q(\theta)) - V(\theta)]\rho(\theta)d\theta$$

$$q_\theta(\theta) \geq 0, \quad V(0) \geq 0$$

The simplest and oldest approach to proceed is what Basov (2005) calls the "direct approach" method. This relies in the use of (b) to integrate by parts

$$\int_{[0,1]} V(\theta)\rho(\theta)d\theta = \int_{[0,1]} u_\theta(q(\theta), \theta)[1 - G(\theta)]d\theta + V(0)$$

where $G(.)$ is the cummulative distribution function associated with ρ. Substituting on the objective function its evident that we must have $V(0) = 0$ and the problem can be reformulated solely in terms of the instrument q

$$\max_{q(\theta) s.t. q_\theta \geq 0} \int_{[0,1]} [u(q(\theta), \theta) - c(q(\theta)) - \frac{1 - G(\theta)}{\rho(\theta)} u_\theta(q(\theta), \theta)]\rho(\theta)d\theta$$

[1]This proof relies on assuming differentiability of t and q. Since q is nondecreasing, however, we alredy know that it is at least almost everywhere differentiable.

If we ignore the restriction $q_\theta \geq 0$[2] this reduces to a pointwise maximization of the integrand that give us an optimality condition:

$$u_q(q(\theta),\theta) - c_q(q(\theta)) - \frac{1-G(\theta)}{\rho(\theta)} u_{q\theta}(q(\theta),\theta) = 0$$

We only need to solve this non linear equation on q to find the optimal screening solution q_{SB}. It is inmediate that only the highest type $\theta = 1$ (i.e., the one with highest marginal valuation) will have an efficient contract (where $u_q(q_{SB}(\theta),\theta) = c_q(q_{SB}(\theta))$) and under the usual assumption $u_{qq} < 0$ types $\theta \neq 1$ will have subefficient contracts.

The explanation of this canonical result is quite straightforward. Under single crossing the type space is ordered according to marginal valuation and this results in local incentive compatibility implying global incentive compatibility. Hence, the principal only has to take into account the possbility of two kinds of strategic behaviour: agents may marginally overstate or understate their true type. If they have no incentive to lie locally, they have no incentive to lie globally. But no type actually wants to pretend to have a greater marginal valuation than their real one. They want to pretend to have lower marginal valuation in order to counteract the monopolist power of the principal: they want to get away with a good they value highly at a low price. It is precisely because of this that optimally the monopolist has to worsen the contracts of types with low marginal valuation below the efficient level to ensure that types with high marginal valuation do not have incentive to pretend to be types with low marginal valuation.

The main purpose of this work is to generalize everything that was said on this section for multidimensional models where the instrument q is still unidimensional. We want to give Single-crossing-like conditions that allow an ordering of the type space according to marginal valuation and ensure that local IC imply global IC. Furthermore, we want to achieve a characterization of implementability that can practically be used to give simple optimality conditions that allow us to compute concrete examples. Since we will follow the idea of ordering the type space by marginal valuation we do expect to extend the logic of the unidimensional model to more general multidimensional environments.

Unlike other multidimensional works, we do not attempt to raise the unidimensionality of the good $n = 1$ together with the dimension of the private information m. This is because it is the increase on the dimensionality of the type space m rather than the instrument n what presents greater challenges to the characterization of implementability. Take for instance $\Theta = [0,1]^m$ and look at the feasible set in the previous section. When we fix m and increase n the "amount" of constraints remains fixed and the space in which each agent's optimization subproblem is performed is the same. The opposite happens when we fix n and increase m. In this case the principal must now take into account the optimizing behaviour of a "bigger" set of types and each of these agents have greater strategic posibilities that the principal have to consider when designing an incentive compatible contract.

Indeed, even if the analogous of single crossing would hold on a bidimensional type space so that local IC imply global IC, to guarantee implementability we would need to check

[2] Of course, the restriction may bind in which case the optimal solution will be composed by flat pieces and pieces of the unrestricted solution. For a formal treatment see Guesnerie and Laffont (1984)

that it is not advantageous for the agents to deviate along any of the infinite possible directions. In other words, even disregarding global constraints, the posibility to combine the overstatement or understatement of two different sources of asymetric information gives rise to infinitely many strategic posibilities for the agents that the principal must forsee when desinging an implementable contract. This added layer of complexity will not prevent us from ordering the type space according to marginal valuation. It will, however, result in the key difference that, unlike the unidimensional case, this order will be endogenous. But first we will look at the relevant literature for our problem.

2.3 Multidimensional literature

As illustrated by the unidimensional case, the key step to find a succesful solution method for the screening model is to impose some single-crossing-like condition which give us a characterization of implementability that can later be employed to simplify the problem. Attempts to characterize implementability and attempts to solve screening problems, however, not always go together and hence we discuss those issues separately on the next two subsections.

2.3.1 Characterizations of implementability

(a) Rochet's Cyclical monotonicity:

Rochet (1987) has provided the more general characterization of implementability which doesnt rely on any structural assumption such as Single crossing or any corresponding generalization.

Proposition 2.2. *A necessary and sufficient condition for q to be implementable is that for all finite cycles $\theta_0, \theta_1, ..., \theta_{N+1} = \theta_0 \in \Theta$,*

$$\sum_{i=0}^{N} u(q(\theta_i), \theta_i) - u(q(\theta_{i+1}), \theta_i) \geq 0$$

The idea behind the "Cyclical monotonocity" condition is that in an implementable contract any group of agents can not improve by swaping contracts in a cyclical manner (θ_0 doe not win from pretending to be θ_1, θ_1 from θ_2,..., and θ_n doesnt win from pretending to be θ_0). Rochet proved that this condition is indeed sufficient. Despite its generality, however, this characterization of implementability has not been proven to be useful for solving concrete screening problems with the exception of two concrete cases where implementability can be characterized more precisely.

The first case is the usual unidimensional model with single crossing. Rochet (1987) showed that under Single crossing, monotonicity of q is equivalent to the above "cyclical condition" inequality for cycles of order 2 and that if the inequality holds for cycles of order n it holds for cycles of order $n+1$. The second case is a more general case that can be carried over to multidimensional environments but relies on a very specific form for the utility of the agents: they must be linear on types.

(b) Rochet's linear case:

By far the most used assumption in multidimensional screening is that u is linear on types. This is due to the pioneering work of Rochet (1985, 1987) who showed that implementability can be characterized in terms of the convexity and a condition on the subgradient of the informational rent V. More concretely, Rochet (1987) proved:

Proposition 2.3. *Let $u(q,\theta)$ be linear with respect to θ and C^1 with respect to q. Then q is implementable if and only if there exists a convex function $V:\Theta\to\mathbb{R}$ such that*

$$\frac{\partial u}{\partial\theta}(q(\theta),\theta)\in\partial V(\theta)\quad\forall\theta\in\Theta$$

where $\partial V(\theta)$ is the subdifferential of V at θ.

Proposition 2.4. *Let $u(q,\theta)$ be linear with respect to θ and C^2 with respect to q. Then a C^1 instrument q is implementable if and only if*[3]

$$rot(\frac{\partial u}{\partial\theta}(q(\theta),\theta))=0,\quad\forall\theta\in\Theta$$

$$u(q(\theta_0),\theta_0)+u(q(\theta_1),\theta_1)\geq u(q(\theta_1),\theta_0)+u(q(\theta_0),\theta_1)$$

Observe that linearity on types is a very concrete functional form and as such it is a very different assumption than Single crossing. There are utility functions that satisfy single crossing and are not linear on types and utilities that are linear on types but do not satisfy single crossing. There is however a generalization of Rochet's first result that doesnt rely on any assumption on u.

(c) Carlier's Generalized convexity:

First, we give the following definitions for any given u.

(i) A function $V:\Theta\to\mathbb{R}$ is u-convex if and only if there exists a nonempty subset $A\subset\mathbb{R}^n_+\times\mathbb{R}$ such that
$$V(\theta)=sup_{(q,t)\in A}\{u(q,\theta)-t\}\quad\forall\theta\in\Theta$$

(ii) Given $V:\Theta\to\mathbb{R}$, $q\in\mathbb{R}^n_+$ is a u-subgradient of V at θ if and only if

$$V(\hat\theta)\geq V(\theta)+u(q,\hat\theta)-u(q,\theta),\quad\forall\hat\theta\in\Theta$$

(iii) The u-subdifferential of V at θ, denoted $\partial^u V(\theta)$, is the set of all subgradients of V at θ and V is said u-subdifferentiable at θ if $\partial^u V(\theta)\neq\emptyset$

[3] The partial differential equation defining the first condition, first derived by Rochet in the context of utilities that are linear on types, will be central in our approach to implementability

When we assume that u is linear on types the concepts of u-convex functions, u-subgradients and u-subdifferentials turn into the familiar concepts of convex functions, subgradients and subdifferentials. Hence, Carlier (2001) generalized Rochet's result with the following proposition

Proposition 2.5. *An instrument q is implementable if and only if there exists a u-convex, u-subdifferentiable function $V : \Theta \to \mathbb{R}$ such that*

$$\frac{\partial u}{\partial \theta}(q(\theta), \theta) \in \partial^u V(\theta) \quad \forall \theta \in \Theta$$

where $\partial^u V(\theta)$ is the u-subdifferential of V at θ.

Despite its generality, however, this condition makes it much harder to check implementability than the usual single crossing which only required checking the sign of the derivatives. Moreover, at the time of writing this work I am not aware of any application of this characterization to solve screening problems. Carlier only used this characterization to prove the existence of a solution. There is however a "general philosophy" behind this approach and Rochet's linear approach: we have to reexpress the problem completely in terms of V instead of (q, t).

One of the advantages of reexpressing everything in terms of V is that under mild conditions (See Carlier 2001 and Ekeland 2010), locally bounded, u-convex functions are almost everywhere differentiable and u-subdifferentiable everywhere so that

$$\nabla V(\theta) = \frac{\partial u}{\partial \theta}(q, \theta), \quad a.e. \quad \theta \in \Theta \quad \forall q \in \partial^u V(\theta)$$

Hence, the informational rent has a priori better differentiability properties than either q or t and if we manage to find the optimal V we can then use the previous condition to recover q provided our model satisfies a technical condition that generalizes single crossing.

Definition: Generalized Spence Mirrless (GSM)[4] The utility function u satisfies the generalized Spence Mirrless property if $\frac{\partial u}{\partial \theta}(., \theta)$ is injective for all θ , i.e.,

$$\frac{\partial u}{\partial \theta}(q_1, \theta) = \frac{\partial u}{\partial \theta}(q_2, \theta) \Rightarrow q_1 = q_2$$

Hence, under GSM we can reduce the search for (q, t) to the search for V since once we get V we recover q from $\nabla V(\theta) = \frac{\partial u}{\partial \theta}(q(\theta), \theta)$ and also t from $V(\theta) = u(q(\theta), \theta) - t(\theta)$. Observe that on the unidimensional model GSM is only a slightly more general condition than Spence Mirrless since assuming that u_θ is continuous, its injectivity is equivalent to being strictly monotone which implies $u_{q\theta} \geq 0$ or $u_{q\theta} \leq 0$[5]. Hence, marginal utility is increasing (not necessarily strictly) on θ and the type space is still ordered according to marginal utility. However, when we are dealing with a onedimensional instrument and multiple dimensions of private information, $m > 1$, the injectivity of u_θ no longer implies that the type sapce can be ordered according to

[4]This condition is also known as the twist condition and it is also used in multidimensional matching. See Chiappori et al (2016)

[5]There are functions that are continuous and strictly monotone but its derivative is zero a.e.

marginal utility. Indeed, one way to ensure that $u_\theta = (u_{\theta_1}, ..., u_{\theta_m})$ is injective on q is to require $u_{q\theta_i} > 0$ for some $i \in \{1, ..., m\}$ without imposing any restriction on $u_{q\theta_j}$ for $j \neq i$. This means that we can only compare apriori the marginal utility of types with the same value of $\theta_{-i} = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_m)$[6].

(d) McAfee McMillan's Generalized single crossing:

McAfee and Mcmillan (1988) discovered a class of functions u more general than linearity on types for which a general version of Single crossing holds.

Definition: Generalized single crossing (GSC)[7] The utility function u satisfies the generalized single crossing property if for all $\theta, \hat{\theta}$ and q there exists a $\lambda > 0$ such that

$$u_q(q, \theta) - u_q(q, \hat{\theta}) = \lambda u_{\theta q}(q, \hat{\theta})(\theta - \hat{\theta})$$

For the case $n = 1$ we can interpret this condition as saying that whenever marginal utility increases/decreases marginally in a certain direction, it always increases/decreases in that direction. There are two particular cases of great interest. First, if $n = m = 1$ and $u_{q\theta}$ does not change sign, then u satisfies (GSC) with $\lambda = \frac{u_{\theta q}(q, \hat{\hat{\theta}})}{u_{\theta q}(q, \hat{\theta})} > 0$ for some $\hat{\hat{\theta}} \in [\theta, \hat{\theta}]$. Second if $u(q, \theta) = v_0(q) + \sum_{i=1}^m \theta_i v_i(q)$ (i.e. u is linear on types) then u satisfies (GSC) with $\lambda = 1$. This condition, however, is not much more general than linearity on types since it forces the set of types having the same marginal utility to form straight hyperplanes (straight lines for $n = 2$).

For this class of functions however, the authors showed that local IC imply global IC just as in the standard unidimensional case. The local IC can be summarized by the conditions in the following proposition.

Proposition 2.6. *Assume that $n \leq m$, $u_{\theta q}$ has rank n and $u(q, \theta)$ satisfies GSC. A differentiable contract (q, t) is incentive compatible if and only if*

$$t'(\theta) = -u_q(q(\theta), \theta) q'(\theta)$$
$$q'(\theta) = C u_{\theta q}(q(\theta), \theta)$$

for some positive semidefinite $n \times n$ matrix C (which generally depends on θ).

From all the characterizations of implementability given, only (b) and (d) have been used with some success to solve the screening problem. The class of functions satisfying GSC is strictly bigger than the class of functions that are linear on types. However, McAfee McMillan only proposed a general solution method when $n = 1$ while Rochet and Chone (1998) used characterization (b) to propose a solution method for a general multidimensional model where $n = m$. This work was later extended for the cases $n < m$ and $m > n$ by Basov (2001).

[6] Compare this situation with the case in which we require the stronger condition that $u_{q\theta_i} > 0$ for all $i = 1, ..., m$ (a condition we will later use under the name of A2). In this case we can say that $\theta \leq \hat{\theta} \Rightarrow u_q(q, \theta) \leq u_q(q, \hat{\theta})$ so now we are able to compare the marginal utility of many more types based on the order $\theta \leq \hat{\theta} \Leftrightarrow \theta_i \leq \hat{\theta}_i \forall i = 1, ..., m$. Ofcourse this order is not complete and hence we will need to introduced a further assumption A3 to really order the type space in terms of marginal utility.

[7] We state the GSC as applied to a quasilinear utility function. McAfee Mcmillan, however, formulate their condition for a general utility u, not necessarily quasilinear

2.3.2 Solution methods for $n < m$

We now quickly review the proposed solution methods for dealing with multidimensional screening problems where $n < m$. To the best of my knowledge, there are five such methodologies including the one we follow in this work. However, most of them are restrictive in several ways. This is evidenced by the relative few examples that have been solved in the literature. In particular, for our case of interest $1 = n < m$ there has been few concrete examples solved beyond the pioneering example provided by Laffont, Maskin and Rochet (1987). These authors relied on specific arguments to solve a concrete example of nonlinear monopoly pricing where the good was one dimensional but preferences depended on a bidimensional parameter of asymetric information. In practice this example constitutes some kind of benchmark to then propose more general methodologies[8].

The first methodology was proposed by McAfee and Mcmillan (1988) and relies on characterization (d) of implementability. Although this characterization was provided for general utility functions (not necessarily quasilinear) and for $n \leq m$, they provide a solution method only for quasilinear utilities when $n = 1$, $\Theta = [0,1]^m$ and with the aditional assumptions $u_{q\theta_1} > 0$ and $u(q, \theta_1, \theta_{-1}) \leq 0$ for all $q \geq 0$ and $\theta_{-1} \in [0,1]^{m-1}$ where $\theta = (\theta_1, \theta_{-1})$. Then they proof that via a change of variables inspired on Laffont, Maskin and Rochet (1987) the screening problem reduces to a one dimensional calculus of variation problem.

However, they dont proof that the change of variables will always be well defined for the class of functions under consideration. It is our believe that they need more than $u_{q\theta_1} > 0$ to achieve this[9]. Furthermore, the other special hypothesis implies types $(0, \theta_{-1})$ must be excluded from contracting a priori. This limitations add up to the already discussed limitation of their characterization of implementability that relies on the GSC condition. This condition is not generic and it can be hard to check wether a given utility function satisfies it. More importantly, this condition implies that the level curves of the optimal allocation should be straight lines. This means that GSC is a very restrictive assumption since it is already ruling out the possibility that the optimal contract exhibits curvature on its level sets.

A second approach is due to Armstrong (1996). This author showed that applying an "integration by rays" we are able to pass to a simple relaxed problem. If the only binding incentive compatibility constraints (IC) are radial, then the solution of the relaxed problem is indeed the optimal contract. Unlike, other approaches this one doesnt rely on a characterization of implementability. However, when the type space is multidimensional there are infinitely many different paths connecting two agents with different types and there is typically no way to know apriori along which of this paths the IC will be binding. Hence, it is extremely unlikely that the only binding IC will be radial and this methodology is only appropriate for solving very specific examples.

A third approach was presented in Basov (2001). This approach relies on characterization (b) and the fundamental trick is to artificially increase n to make it equal to m and then apply

[8] There is another early multidimensional example with $1 = n < m = 2$ solved in the literature. This time it was a model of regulation proposed by Lewis and Sappington (1988). However, as Armstrong (1999) showed their solution was flawed. This example escapes from the scope of this work because part of the complications comes from the IR constraint rather than the IC constraints.

[9] In particular, see our conjunction of our hypothesis A2 and A3 in the next chapter.

Rochet and Chone (1998)'s methodology and fall into an optimal control problem. This approach relies hence on the hypothesis of linearity on types. In contrast to this three methodologies, the methodology we present allows optimal contracts that exhibit curvature, does not require agent's utility function to be linear and is not restricted to a case in which the IC have to bind in a specific manner.

The approach we follow was pioneered by Araujo, Calagua and Vieira (2022) and studies the case $1 = n < m = 2$. These authors noted that an implementable contract must solve a partial differential equation with nondecreasing initial data and we can use the method of characteristics to reparametrize the problem and generate optimality conditions that often lead us to a solution. The authors, however, dont provide a characterization of implementability which makes it unclear wether their methodology lead us to solution candidates that are implementable. Since their focus is in providing necessary conditions they also dont provide a formal justification of when their change of variables can be expected to work and this leaves open the question of what is the economic reason of the apparent success of their methodology in the examples they present. Finally, they also develop their optimality conditions only for models where the optimal contract exhibits "sufficient exclusion" (i.e., the exclusion set is assumed a priori to be sufficiently big just as was the case with McAfee and McMillan) and the distribution of the different dimensions of asymetric information are independent. We will extend their work in all of this directions[10].

Finally, Deneckere and Severinov (2017) developed a different methodology for the case the case $1 = n < m = 2$ that is based on the same assumptions that we employ together with some additional assumptions that we do not employ such as[11]: $u(q, \theta_1, 0) \leq 0$ for all $q \geq 0$, $\theta_1 \in [0, 1]$ (sufficient exclusion), $lim_{q \to +\infty} u_q(q, \theta_1, \theta_2) < 0$ for all $(q, \theta_, \theta_2)$, $lim_{q \to 0} - \frac{u_{q\theta_2}}{u_{q\theta_1}} = +\infty$ for all $(\theta_1, \theta_2) \in (0, 1]^2$ and $sup\{-\frac{u_{\theta_1}}{u_{\theta_2}}\} < \infty$. Indeed, they were the first to propose the set of conditions that we consider to be the appropriate analog of Single crossing for bidimensional screening models[12]. They however use these conditions very differently both formally and in terms of the economic reasoning behind. Formally, they proceed generalizing the "demand profile approach" while we follow the paradigm of the "parametric utility approach"[13]. In terms of the interpretation behind, they read their assumptions as implying "single crossing of demand" while we read it as providing us with the ability to substitute different dimensions of private information on the marginal utility in such a way that ultimately allows us to order the type space in terms of marginal utility.

[10] There is one aspect in which Araujo et al. (2022)'s treatment is more general than ours. They explore the possibility of a discontinuous jump in the optimal contract at the frontier between the participation and the exclusion region and provide corresponding optimality conditions. We restrict ourselves to continuous contracts.

[11] We rewrite their assumptions in our terminology for the sake of comparison. The common introduction of a negative sign comes from the fact that we will assume $u_{\theta_2}, u_{\theta_2} < 0$ while they assume that these signs are positive which is completely analogous.

[12] The key assumption labeled by Deneckere and Severinov as "single crossing of demand" was discovered independently while proving a simpler version of our characterization of implementability

[13] Rochet and Stole (2003) made this distinction between methodologies used on unidimensional models which have later been tried for multidimensional models. Wilson (1993) is the main reference for applying the demand profile approach to multidimensional problems

Chapter 3

Bidimensional screening

In this chapter we give our main contributions by extending the necessary conditions presented by Araujo, Calagua and Vieira (2022) into a complete solution method to deal with screening problems where $1 = n < m = 2$. We first introduce and comment a set of assumptions that allow us to solve the screening problem in a way that mimicks the standard one dimensional model with single crossing. Then we show that under our assumptions, local IC imply global IC and the type space is endogenously ordered by marginal utility. Moreover, we provide a new characterization of implementability in terms of solving a particular Cauchy problem with nondecreasing initial data. Finally, we use this characterization to reformulate our problem in simpler terms and generate simple optimality conditions that give us a simple and straightforward algorithm to solve bidimensional screening problems. All proofs are relegated to the appendix.

Our starting point is the optimization problem stated on section 2.1. We will follow the ”parametric utility approach” and change variables from (q, t) to (q, V). Given an instrument $q : \Theta \to \mathbb{R}_+$ we will use the identity $V(\theta) = u(q(\theta), \theta) - t(\theta)$ to define $V : \Theta \to \mathbb{R}$ in terms of $t : \Theta \to \mathbb{R}$ or viceversa. Its easy to see that a contract (q, t) satisfies (IR) and (IC) if and only if the pair (q, V), which will also be called a contract, satisfies the corresponding condition

$$(IR') \quad V(\theta) \geq 0 \quad \forall \theta \in \Theta$$

$$(IC') \quad V(\theta) - V(\theta') \geq u(q(\theta^{'}), \theta) - u(q(\theta^{'}), \theta') \quad \forall \theta, \theta^{'} \in \Theta$$

We say that the contract (q, V) is incentive compatible if it satisfies (IC') (i.e., if the associated (q, t) is incentive compatible). In this line we also say that q is implementable if there is a V such that (q, V) is incentive compatible.

When we are dealing with an incentive compatible contract we then have that V is the value function of the subproblem of the agents: $V(\theta) = \max_{\hat{\theta} \in \Theta} \{u(q(\hat{\theta}), \theta) - t(\hat{\theta})\}$. Thus, by the envelope theorem[1] we have $\nabla_\theta V(\theta) = \nabla_\theta u(q(\theta), \theta)$. This condition is implied by (IC) but it will be important to single out for the purpose of relaxation. Hence, our screening problem can be

[1]See Milgrom and Segal (2002)

equivalently reformulated as:

$$\max_{(q(\theta),V(\theta))} \int\limits_{\Theta} [u(q(\theta),\theta) - c(q(\theta)) - V(\theta))]\rho(\theta)d\theta$$

$$s.t. \quad (IR) \quad V(\theta) \geq 0 \quad \forall\theta \in \Theta$$

$$(IC) \quad V(\theta) - V(\theta') \geq u(q(\theta^{'}),\theta) - u(q(\theta^{'}),\theta') \quad \forall\theta,\theta^{'} \in \Theta$$

$$(E) \quad \nabla_{\theta}V(\theta) = \nabla_{\theta}u(q(\theta),\theta) \quad \forall\theta \in \Theta$$

3.1 Assumptions:

We will assume now that $n = 1$ and for simplicity we will focus on the case $m = 2$. We also fix $\Theta = [0,1]^2$ and restrict attention to contracts (q,V) that are continuous and where q is continuously differentiable a.e. on Θ and V is twice continuously differentiable a.e. on Θ^2. As usual, assume $u_q > 0$, $u_{qq} < 0$ and $u(0,\theta_1,\theta_2) = 0, \forall\theta \in [0,1]^2$. In line with the discussion on the previous chapter we make the following fundamental assumptions that help us simplify the structure of the feasible set:

A1. $u_{\theta_1} > 0$ and $u_{\theta_2} < 0$

A2. $u_{q\theta_1} > 0$ and $u_{q\theta_2} < 0$

A3. $\frac{d}{dq}(-\frac{u_{q\theta_2}}{u_{q\theta_1}}) \geq 0$

The first two assumptions are familiar. The first assumption is often used to simplify the rol of participations constraints in the structure of the feasible set. In our case, when $\Theta = [0,1]^2$, A1 together with the envelope condition (E) implies that for all θ_1, $\theta_2 \in [0,1]$

$$V(\theta_1,\theta_2) = V(0,1) - \int_{\theta_2}^{1} u_{\theta_2}(q(0,s),0,s)ds + \int_{0}^{\theta_1} u_{\theta_1}(q(t,\theta_2),t,\theta_2)dt \geq V(0,1)$$

So the informational rent assumes a minimum at the vertix $(0,1)$ and therefore (IR) reduces to $V(0,1) \geq 0$. The second assumption requires that Spence Mirrless holds for each dimension of private information. When $m = 1$, A2 is enough to guarantee that the only incentive constraint that matters are local ones and it also reduces (IC) to a monotonicty constraint on the instrument[3].

We will see that when $m > 1$ this is no longer the case and an additional assumption relating the different dimensions of private information is needed in order to restrict the feasible set on the same way Spence-Mirrless does on unidimensional models. This assumption is A3. In

[2]It is known that the informational rent of an incentive compatible contract is a generalized convex function (See Carlier 2001 or Ekeland 2010) and as such under mild conditions it possesses good regularity properties.

[3]What matters in A1 and A2 is not whether a sign is positive or negative but rather that it always keeps the same sign.

words, A3 says that the marginal rate of substitution of θ_2 by θ_1 on the marginal utility is increasing on q or, to put it more simply, that whenever we are confronted with the purchase of a higher level of the instrument q a small increase in θ_2 needs to be compensated for a bigger increase in θ_1 to keep the marginal utility constant.

We will see that A3 works in a similar way to the single crossing assumption for bidimensional models: it provides an appropriate way of substituting the different dimensions of asymetric information on the marginal utility in such a way that it will allows us to naturally order the type space according to the levels of the marginal valuation just as Spence Mirrless does on unidimensional models. We should point out that unlike the unidimensional case this order will be endogenous. Therefore, A3 together with A2 will constitute a bidimensional vesion of the Spence Mirrless assumption.

We should point out however that unlike the conditions GSC or GSM, A3 is not a generalization of the Spence-Mirrless assumption since this condition is meaningless when $m = 1$. However, and more importantly, A3 works in the same way as the Spence Mirrless condition by guaranteeing that local IC imply global IC and by giving a natural ordering of the type space in terms of marginal valuation. It is in this sense that we consider A2 together with A3 a bidimensional version of the Spence Mirrless assumption. A simple way to see the unidimensional case as a particular case of our bidimensional model satisfying A2 and A3 is suggested on the first example solved on the next chapter.

3.2 Necessary conditions for implementability:

We start by presenting our necessary conditions for implementability and discussing their economic meaning. These conditions are not new and proofs are only presented for the sake of completeness. In particular, the derivation of the PDE satisfied by an implementable contract q presented on proposition 3.1 comes from proposition 2.4 due to Rochet (1987) while the alternative derivation in proposition 3.3 comes from Araujo et al (2022). The nondecreasingness of the initial data was first noted by these last authors although their parametrization of the initial condition differs from ours since they assume "sufficient exclusion"[4].

Proposition 3.1 (Necessary Conditions A). *Under A2 an implementable contract q satisfies the following:*

(i) q solves the partial differential equation

$$-u_{q\theta_2}q_{\theta_1} + u_{q\theta_1}q_{\theta_2} = 0, \quad a.e. \quad on \quad [0,1]^2$$

(ii) $q(\theta_1, .)$ is monotone decreasing for each $\theta_1 \in [0,1]$

(iii) $q(., \theta_2)$ is monotone increasing for each $\theta_2 \in [0,1]$

[4] See section 2.3.2

Define

$$\Gamma(r) = \begin{cases} (0, 1-2r) : 0 \le r \le \frac{1}{2} \\ (2r-1, 0) : \frac{1}{2} \le r \le 1 \end{cases}$$

Then, it follows inmediately from proposition 3.1 that:

Corollary 3.2. *Under A2 an implementable contract q satisfies the Cauchy problem:*

$$-u_{q\theta_2} q_{\theta_1} + u_{q\theta_1} q_{\theta_2} = 0, \quad a.e. \quad on \quad [0,1]^2$$

$$q(\Gamma(r)) = \phi(r)$$

for some $\phi(r)$ nonnegative and nondecreasing.

Although this derivation is quite straightforward, it somewhat obscures what is the economic meaning behind the fact that an implementable q must solve a particular Cauchy problem with nondecreasing initial data and why out of all the information contained on the incentive compatibility constraints we can expect this Cauchy problem to summarize all relevant information to solve the monopolist problem. To this end consider an alternative derivation where we see explicitly that the information contained on our Cauchy problem is nothing else than the first and second order local necessary conditions of the maximization subproblems of the agents (the local IC).

Proposition 3.3 (Necessary Conditions B). *Consider a contract $(q(\theta), t(\theta))$ that is incentive compatible and twice continuously differentiable a.e. on Θ[5]. Then, this contract satisfies a.e. on Θ the following:*

$$(1) \quad u_{q\theta_1}(q(\theta), \theta) q_{\theta_2}(\theta) = u_{q\theta_2}(q(\theta), \theta) q_{\theta_1}(\theta)$$

$$(2) \quad u_{q\theta_i}(q(\theta), \theta) q_{\theta_i}(\theta) \ge 0 \quad \forall i = 1, 2.$$

We did not use any structural assumption to derive (1) and (2). However, A2 gives economic meaning to these conditions. From (1) we see that for a.e. $\theta \in \Theta$ there exists a $\lambda \ge 0$[6] such that

$$\nabla_\theta q(\theta) = \lambda \nabla_\theta u_q(q(\theta), \theta)$$

In words, the normal planes to the allocation and the marginal utility hypersurfaces are parallel. This means that whenever $\lambda > 0$ along a parametrized curve $\theta(s)_{s\in[0,1]}$ that satisfies $q(\theta(s)) = k$ then $u_q(q(\theta(s)), \theta(s))$ is also constant:

$$\frac{d}{ds} u_q(q(\theta(s)), \theta(s)) = \nabla_\theta u_q(k, \theta(s)) \cdot \theta'(s) = \frac{1}{\lambda}(\nabla_\theta q(\theta) \cdot \theta'(s)) = 0$$

Thus, the first order condition tells us that in incentive compatible contracts, types choosing the same amount of the instrument must have the same marginal utility. This is best understood as a

[5] Observe how on the previous proposition we only required q to be continuously differentiable a.e. and here we require it to be twice continuously differentiable a.e. Moreover, before we required V to be twice continuously differentiable and now we ask the same for t. Since an implementable V is a generalized convex function, it usually has greater regularity properties than t (See Carlier (2001) and Ekeland (2010)). Hence, overall, the hypothesis of proposition 3.3 and the ensuing discussion are stronger than the ones of proposition 3.1 although the content is the same.

[6] Of course, $\lambda = \frac{q_{\theta_i}}{u_{q\theta_i}}$ for all i and we know it is non negative by condition (2)

consequence of the "taxation principle": our problem is equivalent to that of choosing an optimal nonlinear pricing scheme for the good and letting the consumers optimize in q. Therefore, since each consumer will face the same nonlinear tariff and equate the marginal tariff of their purchase to their marginal valuation, types choosing the same amount of the good must also exhibit the same level of marginal valuation for it.

On the other hand, using A2 on the second order condition (2) we see that:

$$u_{q\theta_1} > 0 \Rightarrow q_{\theta_1} \geq 0$$

$$u_{q\theta_2} < 0 \Rightarrow q_{\theta_2} \leq 0$$

This is a phenomenon well known from the unidimensional case: Whenever a dimension of private information orders the types so that types with higher/lower level of that dimension exhibit higher marginal valuation for the good, any implementable allocation q must be increasing/decreasing in that dimension. Moreover, on unidimensional models ($m = 1$) that satisfy Spence Mirrless, this monotonicty of contracts is enough to characterize implementability: we can expect to force the types to completely reveal themselves by offering larger amounts of the good to the types with higher/lower levels of the parameter of asymetric information (since those are the ones with higher marginal utility).

In our case ($m > 1$), of course, this expectation fails due to the existence of multiple dimensions of private information. An agent which chooses a high level of the instrument reveals a high level of marginal utility but we do not know wether that high level of marginal utility is due to a high level of the parameter θ_1 o a low level of the parameter θ_2 (or some mixture). Therefore, we do not know its position on the type space and we do not know which are the types that should be offered higher amounts to ensure truth-telling. Even when we are assuming A2, the lack of an exogenous order on the type space when $m > 1$ makes it hard to identify in general which types should be offered greater amounts.

This is why we need an additional assumption. A3 imposes some order on the substitubility of different dimensions of private information on the marginal utility[7]. We will next see that this substitubility allows us to endogenously order the type space according to marginal valuation. Moreover, we will also see that under A3 satisfaction of the Cauchy problem stated in corollary 1.2 is both necessary and sufficient to guarantee implementability. As noted above, this Cauchy problem is nothing more than the necessary first and second order conditions of the agents subproblem, also known as the local IC.

[7]Observe that this sustitutibility of different dimensions of private information is common knowledge (part of the structure of u) and not part of the asymetric information (not included on θ). Otherwise the principal would not be able to exploit this sustitubility.

3.3 A closer look to the Cauchy problem:

First, lets take a closer look at the Cauchy problem that an implementable q must solve. Observe that the PDE can be written equivalently as:

$$(-u_{q\theta_2}, u_{q\theta_1}) \cdot \nabla_\theta q = 0$$
$$(1, -\frac{u_{q\theta_1}}{u_{q\theta_2}}) \cdot \nabla_\theta q = 0$$
$$(-\frac{u_{q\theta_2}}{u_{q\theta_1}}, 1) \cdot \nabla_\theta q = 0$$

Any of the vectors appearing on the left gives a direction along which q must remain constant. Based on this information we can recover curves corresponding to the level sets of q. For a fixed nonnegative and nondecreasing function ϕ, we can try to reconstruct the solution to the PDE q that assumes the values ϕ at the southwestern boundary by following the level sets of q. This is precisely what the method of characteristics does[8].

By following the method of characteristics, a solution to the Cauchy problem for this ϕ is characterized by solving the following system of ODE's for all $r \in [0, 1]$:

If $0 \le r \le 1/2$, solve:

- $a_s(r, s) = 1$, $a(r, 0) = 0$
- $b_s(r, s) = -\frac{u_{q\theta_1}}{u_{q\theta_2}}(c(r, s), a(r, s), b(r, s))$, $b(r, 0) = 1 - 2r$
- $c_s(r, s) = 0$, $c(r, 0) = \phi(r)$

If $1/2 \le r \le 1$, solve:

- $a_s(r, s) = -\frac{u_{q\theta_2}}{u_{q\theta_1}}(c(r, s), a(r, s), b(r, s))$, $a(r, 0) = 2r - 1$
- $b_s(r, s) = 1$, $b(r, 0) = 0$
- $c_s(r, s) = 0$, $c(r, 0) = \phi(r)$

Therefore, a solution to the Cauchy problem is then characterized by solving for each $r \in [0, 1]$ a single ODE:

[8]For an exposition of the method of characteristtics to solve first order Cauchy pde problems see either John (1981) or Evans (1998)

If $0 \leq r \leq 1/2$, solve:

$$B_s(\phi, r, s) = -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), s, B(\phi, r, s)), \quad B(\phi, r, 0) = 1 - 2r$$

If $1/2 \leq r \leq 1$, solve:

$$A_s(\phi, r, s) = -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), A(\phi, r, s), s), \quad A(\phi, r, 0) = 2r - 1$$

A local solution to the Cauchy problem always exists since the vector $(a_s, b_s) \parallel (-u_{q\theta_2}, u_{q\theta_1})$ is nonparallel to the boundary: $(a_r, b_r) = (0, -2)$ at $(r, 0)$ for $0 \leq r \leq 1/2$ or $(a_r, b_r) = (2, 0)$ at $(r, 0)$ for $1/2 \leq r \leq 1$[9].

For a nonnegative s, lets define $\gamma_r(s)$, by:

$$\gamma_r(s) = \begin{cases} (s, B(\phi, r, s)) : 0 \leq r \leq \frac{1}{2} \\ (A(\phi, r, s), s) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

Lets now define[10]:

$$U(\phi, r) = sup\{s \in [0,1] : \gamma_r(s) \in [0,1]^2\}$$

$$\gamma_r = \{\gamma_r(s) : 0 \leq s \leq U(\phi, r)]\}$$

A global solution will exist as long as $\{\gamma_r\}_{r\in[0,1]}$ does not intersect before reaching the edge of the box $[0,1]^2$. Otherwise, we can only define a local solution up to the region where the characteristics $\{\gamma_r\}_{r\in[0,1]}$ intersect[11]. When a global solution exists it is given implicitly by:

- The level curves of q: $\{\gamma_r\}_{r\in[0,1]}$
- The value of q at each level curve: $\phi(r)$

$$q(\gamma_r(s)) = \phi(r) \quad \forall s \in [0, U(\phi, r)], \forall r \in [0,1]$$

For any implementable contract q lets define $r_0 = sup\{r \in [0,1] : \phi(r) = 0\}$ where $\phi(r) = q(\Gamma(r))$ as usual. Then, since ϕ is nonnegative and nondecreasing, $\phi(r) = 0$ for all $0 \leq r \leq r_0$, and γ_{r_0} determines the frontier between the participation and the exclusion region (where q, t, V and the objective functional vanish). Hence, when looking at the behaviour of an

[9]More formally, what guarantees that the method of characteristics works at least locally to recover the solution of the PDE is the implicit function theorem and the nonparallel condition just checked verifies the hypothesis needed to use the theorem. See John (1981).

[10]Observe that U is well defined since the curve $\gamma_r(s)$ starts at the southwestern edge ($\gamma_r(0) = \Gamma(r)$) and always travels northeast thanks to A2 ($\dot{\gamma_r}(s) \parallel (-u_{q\theta_2}, u_{q\theta_1}) >> 0$)

[11]Sometimes a global solution can be defined going beyond what the characteristics method gives but this solution will typically be discontinuous.

implementable contract at the boundary $\Gamma([0,1])$ we only need to consider $\phi(r)$ for $r_0 \leq r \leq 1$ and we will only solve the ODE's for that range of r.

Definition: We will say that our model is nested[12] if for every nonnegative, nondecreasing function ϕ the characteristics do not cross on $[0,1]^2$.

Hence, nestedness implies that the partial differential equation has a global solution for any possible nonnegative, nondecreasing initial data and hence there is a one to one relation between nonnegative, nondecreasing functions ϕ and solutions q of the Cauchy problem. Since we will later identify implementable contracts with solutions to the Cauchy problem, nestedness will be fundamental in effectively reducing the dimension of our problem from searching among implementable contracts q (a surface satisfying complex restrictions) to searching among nonnegative nondecreasing functions ϕ (a curve with simple properties)[13]. Observe that as nestedness is a property of our quasilinear partial differential equation it ultimately depends on its coefficients, that is, it is ultimately a property of the valuation function u.

Proposition 3.4 (**Non-crossing of characteristics**). *When u satisfies A2 and A3 the model is nested. Moreover, we have*

$$\frac{d}{dr}A(\phi, r, s) = A_\phi \phi' + A_r > 0, \quad \forall 1/2 \leq r \leq 1$$
$$\frac{d}{dr}B(\phi, r, s) = B_\phi \phi' + B_r < 0, \quad \forall 0 \leq r \leq 1/2$$

3.4 Sufficient conditions for implementability:

Now we give our main result which together with corollary 2 fully characterize implementable contracts as solutions to a Cauchy problem with nondecreasing initial data.

Proposition 3.5 (**Sufficent conditions**). *Under A1-A3, any q which solves the Cauchy problem*

$$-u_{q\theta_2} q_{\theta_1} + u_{q\theta_1} q_{\theta_2} = 0 \quad a.e. \quad on \quad [0,1]^2$$
$$q(\Gamma(r)) = \phi(r) \quad \forall r \in [0,1]$$

for some ϕ nonnegative and nondecreasing, is implementable.

This result has a natural interpretation. Our assumptions A2 and A3 are revealed to work as Spence-Mirrless does on unidimensional models because they guarantee that the local incentive compatibility contained in the cauchy problem implies global incentive compatibility. Moreover, it allows us to impose an order on the type space $[0,1]^2$ according to the marginal valuation of the

[12]We borrow this terminology from Chiappori et al (2016) who employ a somewhat similar methodology for dealing with multidimensional matching problems when one of the sides is unidimensional.

[13]Nestedness is also convenient because it implies that as long as our initial data ϕ is continuous then any solution to the Cauchy problem with initial data ϕ must also be continuous. That is, it partially justifies our restriction of attention to continuous contracts although discontinuous implementable (and even optimal) contracts may still exist if the contract is discontinuous at the edge $\Gamma([0,1])$

types. More specifically, since implementable contracts are the same as solutions to the Cauchy problem and any such solution is described parametrically by $\{(\gamma_r, \phi(r))\}_{r\in[r_0,1]}$ each of these elements can be interpreted as playing a distinctive rol: The contour lines $\{\gamma_r\}_{r\in[r_0,1]}$ order the types on $[0,1]^2$ according to their marginal valuation while the values $\phi(r)$ along these contour lines set the proper incentives accross this order so that types do not pretend to have a lower marginal valuation.

To see that the level sets of any solution of the Cauchy problem orders the type space according to marginal utility its enough to notice that any two types on the same level set $\theta = \gamma_r(s_2)$ and $\theta' = \gamma_r(s_1)$ have the same marginal utility

$$u_q(\phi, \gamma_r(s_2)) - u_q(\phi, \gamma_r(s_1)) = \int_{s_1}^{s_2} \nabla_\theta u_q(\phi, \gamma_r(t)).\dot{\gamma_r}(t)dt = 0$$

We could therefore set an equivalence relation on the type space whose equivalence classes are the level sets γ_r and taking as representatives the types on the southwestern edge of the type space we can see that the quotient space is effectively unidimensional and ordered by marginal utility since $u_{q\theta_1} > 0$ and $u_{q\theta_2} < 0$. The requirement for ϕ to be nondecreasing has therefore the usual interpretation of setting the proper incentives across this order.

Observe how A3 plays a key rol in the proof. The PDE captures the pattern of binding incentive compatibility constraints (claim 1) which follows from first order necessary conditions in the subproblem of the agent. On the other hand, the nondecreasingness of the initial data which follows from the second order necessary conditions guarantees that incentive compatibility holds on the edge $\Gamma([0,1])$ (claim 2). But what allows us to disregard nonlocal incentive compatibility constraints is that A3 implies a transitivity-like property. For any two types $\theta \in \gamma_r$ and $\theta' \in \gamma_{r'}$ we know that θ wins nothing from pretending to be $\alpha \in \Gamma([0,1]) \cap \gamma_r$, α wins nothing from pretending to be $\alpha' \in \Gamma([0,1]) \cap \gamma_{r'}$ and α' wins nothing from pretending to be θ'. It is A3 what allows us to conclude as if we were using transitivity that θ also doesnt have any incentive to pretend to be θ' (claim 3).

3.5 Optimality

The characterization of implementability done in the previous sections allows us to rewrite our monopolist's problem as:

$$\max_{(q(\theta),V(\theta))} \int_{[0,1]^2} [u(q(\theta),\theta) - c(q(\theta)) - V(\theta))]\rho(\theta)d\theta$$

$$s.t. \quad (IR) \quad V(0,1) \geq 0$$

$$(E) \quad \nabla_\theta V(\theta) = \nabla_\theta u(q(\theta),\theta)$$

$$(IC) - u_{q\theta_2} q_{\theta_1} + u_{q\theta_1} q_{\theta_2} = 0 \quad a.e. \quad on \quad [0,1]^2$$

$$\forall r \in [0,1]: \quad q(\Gamma(r)) = \phi(r) \quad nondecreasing \quad and \quad nonnegative$$

Condition (IC) is just expressing the fact that q solves the Cauchy problem for some nonnegative and nondecreasing initial data. In order to be able to compute a solution our next step will be to reexpress the problem only in terms of q. It should be noted how at this point our approach diverges from other authors who also follow the parametric utility approach, such as Rochet and Chone (1998) or Carlier (2001), but proceed by reexpressing the problem only in terms of the informational rent V. This is because they rely on very different characterizations of implementability such as characterizations (b) and (c) reviewed on section 2.3.1. This approach has an interesting connection with the optimal transport literature (See Ekeland 2010) where it is part of the common wisdom that it may be more simple to look for an object $V : \Theta \to \mathbb{R}$ rather than an object $q : \theta \to \mathbb{R}^n$. However, in our case $n = 1$ and V isnt necessarily a simpler object than q. Here we follow the older "direct" approach of integrating away the informational rent since we have characterized implementability completely in terms of q.

To this end, define the functions:

$$F_1(\theta_1, \theta_2) = \int_{\theta_1}^{1} \rho(\alpha, \theta_2) d\alpha$$

$$F_2(\theta_1, \theta_2) = \int_{\theta_2}^{1} \rho(\theta_1, \alpha) d\alpha$$

$$G(q, \theta_1, \theta_2) = [u(q, \theta_1.\theta_2) - c(q) - \frac{F_1(\theta_1, \theta_2)}{\rho(\theta_1, \theta_2)} u_{\theta_1}(q, \theta_1, \theta_2)]\rho(\theta_1, \theta_2)$$

Then we have the following result:

Proposition 3.6. *The objective functional can be reexpressed as:*

$$\int_0^1 \int_0^1 G(q(\theta_1, \theta_2), \theta_1, \theta_2) d\theta_1 d\theta_2 - \int_0^1 u_{\theta_2}(q(0, \theta_2), 0, \theta_2)[\int_0^1 F_2(\theta_1, \theta_2) d\theta_1 - 1] d\theta_2$$

This proposition allows us to get rid of restriction (IR) and (E) by incorporating them into the objective function. The next step will be to do the same for the restriction (IC). This can be done by using the change of variables provided by the characteristics method. This is the key idea of Araujo, Calagua and Vieira (2022). Our situation, however, is more complex since the aforementioned authors consider models that force the exclusion of a sizeable subset of the type space at the optimum. More concretely, their proposition 1 is a particular case of our proposition 3.6 when $\rho(\theta_1, \theta_2) = f_1(\theta_1) f_2(\theta_2)$ (i.e., the distribution of the dimensions of asymetric information are independent) and the solution satisfies $q(0, \theta_2) = 0$ for all $\theta_2 \in [0, 1]$ at the optimum (i.e., there is sufficient exclusion). Moreover, the hypothesis that there is sufficient exclusion is also present implicitly in their theorem 1 since they adopt a parametrization of the initial condition in the cauchy problem that reflects the belief that all types $(0, \theta_2)$ will be excluded at the optimum and hence, need not be considered.

It is interesting to note that an exogenous condition guaranteeing $q(0, \theta_2) = 0$ for all $\theta_2 \in [0, 1]$ at the optimum is the condition $v(q, 0, \theta_2) \leq 0$ for all $q \geq 0$ and $\theta_2 \in [0, 1]$ which was used by Mcafee and Mcmillan (1987) in their proposed methodology to solve screening problems.

In any case, either of the conditions guaranteeing enough exclusion not only restricts the family of models that could be solved but also doesnt allow us to study issues such as the genericity of exclusion proposed by Armstrong (1996) which could only be explored by a solution method that doesnt forbid a priori complete participation.

To incorporate restriction (IC), we can use the fact that there is a one to one correspondence between implementable contracts q and the set of nonnegative nondecreasing functions ϕ as long as the model is nested. Corollary 3.2 tells us that every implementable contract has nonnegative, nondecreasing "southwest edge behaviour" ϕ. Reciprocally, proposition 3.4 guarantees that for any nonnegative, nondecreasing ϕ the characteristics method reconstructs the unique solution of the Cauchy problem q with initial value ϕ and then proposition 3.5 implies that this solution is implementable.

Hence, we will reexpress everything in terms of ϕ by changing variables on the integrals that define the objective functional. For the second part of the objective functional we use $\theta_2 = 1 - 2r$:

$$\int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)[\int_0^1 F_2(\theta_1,\theta_2)d\theta_1 - 1]d\theta_2 = 2\int_0^{1/2} u_{\theta_2}(\phi(r),0,1-2r)[\int_0^1 F_2(t,1-2r)dt-1]dr$$

and for the first part of the objective functional we use $(\theta_1, \theta_2) = \gamma_r(s)$:

$$\int_0^1\int_0^1 G(q(\theta_1,\theta_2),\theta_1,\theta_2)d\theta_1 d\theta_2 = \int_{r_0}^1\int_0^{U(\phi(r),r)} G(\phi(r),\gamma_r(s))|\frac{\partial(\theta_1,\theta_2)}{\partial(r,s)}|dsdr$$

Where we have:

$$|\frac{\partial(\theta_1,\theta_2)}{\partial(r,s)}| = \begin{cases} -B_\phi\phi' - B_r : 0 \le r \le \frac{1}{2} \\ A_\phi\phi' + A_r \quad : \frac{1}{2} \le r \le 1 \end{cases}$$

Observe that the last change of variables is well defined as long as the Jacobian keeps a strict positive sign. This is guaranteed by proposition 3.4. Combining both change of variables, the objective functional becomes[14]:

$$\int_{r_0}^1 (\int_0^{U(\phi(r),r)} G(\phi(r),\gamma_r(s))|\frac{\partial(\theta_1,\theta_2)}{\partial(r,s)}|ds) - 2u_{\theta_2}(\phi(r),0,1-2r)[\int_0^1 F_2(t,1-2r)dt-1])1_{r\le 1/2}dr$$

Then if we define:

$$H(r,\phi,\phi') = (\int_0^{U(\phi(r),r)} G(\phi(r),\gamma_r(s))|\frac{\partial(\theta_1,\theta_2)}{\partial(r,s)}|ds) - 2u_{\theta_2}(\phi(r),0,1-2r)[\int_0^1 F_2(t,1-2r)dt-1]1_{r\le 1/2}$$

We end up with a simple maximization problem which is, nonetheless, equivalent to our original problem:

$$\max_{\phi(r), r_0\in[0,1]} \int_{r_0}^1 H(r,\phi(r),\phi'(r))dr$$

[14] Observe that we need to introduce the indicator function because $[\int_0^1 F_2(t,1-2r)dt - 1]$ is not defined for $r > 1/2$. This, however, introduces no discontinuity since this term vanishes at $r = 1/2$.

$$s.t. \quad \phi(r_0) = 0$$
$$\phi \quad is \quad nonnegative \quad and \quad nondecreasing$$

The integration starts in r_0 since for $0 \leq r \leq r_0$ we are in the exclusion region where the profit of the principal is zero. Following our assumptions, we look for ϕ in the class of continuous and piecewise differentiable functions over $[0, 1]$. Hence, we have effectively reduce the dimension of our problem. Instead of looking for the optimal implementable contract q (a surface) we only need to look for its optimal southwest edge behaviour ϕ (a curve).

Now we can apply traditional calculus of variations methods. For example, if we disregard the restriction that ϕ is nonnegative and nondecreasing[15] we can derive from the euler equation the following necessary condition for optimality:

Proposition 3.7. *The optimal southwest edge behaviour of the contract $\phi(r)$ satisfies:*

$$\int_0^{U(\phi(r),r)} \frac{G_q}{u_{q\theta_2}}(\phi(r), s, B(\phi(r), r, s))ds = \int_0^1 F_2(t, 1-2r)dt - 1,$$
$$\int_0^{U(\phi(r),r)} \frac{G_q}{u_{q\theta_1}}(\phi(r), A(\phi(r), r, s), s)ds = 0$$

where the first condition is for all $r_0 \leq r \leq 1/2$ and the second for $\max\{r_0, 1/2\} \leq r \leq 1$.

Hence, we have a very simple algorithm for solving bidimensional screening problems that satisfy A1-A3. First, for each nonnegative, nondecreasing ϕ we find the characteristics curves by solving the ODE's:

$$\forall r_0 \leq r \leq 1/2 : B_s(\phi, r, s) = -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), s, B(\phi, r, s)), B(\phi, r, 0) = 1 - 2r$$

$$\forall \max\{r_0, 1/2\} \leq r \leq 1 : A_s(\phi, r, s) = -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), A(\phi, r, s), s), A(\phi, r, 0) = 2r - 1$$

These characteristics provides us with a change of variables that allows us to pass to a simpler variational problem in ϕ. To find the solution of this variational problem, we only need to solve the optimality conditions on proposition 3.7. Hence we can find the solution of a bidimensional screening model by solving some ODE's, computing certain integrals and then solving an algebraic nonlinear equation for ϕ. Although this is in principle a very simple procedure, we should note that each of these steps may turn out to be nontrivial for a particular example[16].

The algorithm described above gives us the optimal southwest edge behaviour of the optimal contract ϕ along with r_0. Observe however that the behaviour of the optimal contract q on

[15]Observe that, since we are ignoring the nondecreasing restriction, the following result is analogous to solving unidimensional screening problems without ironing. However, unlike the unidimensional case, we dont fall on a pointwise maximization of the integrand and the euler equation depends on both ϕ and ϕ'.

[16]This is only to be expected since we haven't get rid of all the complexities inherent in a bidimensional screening problem, we have only transformed the problem in a manner that allows us to deal with these complexities with familiar tools.

the participation set then follows as the solution to the Cauchy problem with initial condition given by ϕ and it is described parametrically by its level sets $\{\gamma_r\}_{r\in[r_0,1]}$ and the values $\phi(r)$ that q takes on its level sets. The exclusion region is determined by r_0 since γ_{r_0} is the frontier between the participation and the exclusion region. Observe also that $\gamma_r(s)$ depends on $\phi(r)$ and therefore it is only completely determined once we find the optimal ϕ. This is the reason that the optimal ordering of types according to marginal valuation is endogenous unlike what happens in the unidimensional case.

Our algorithm gives us a solution q in parametric form $(r_0, \phi, \{\gamma_r\}_{r\in[r_0,1]})$. If we want to express the solution in terms of the original variables $(\theta_1.\theta_2)$ we only need to use again the change of variables $(\theta_1, \theta_2) = \gamma_r(s)$ in the opposite direction. This last step however is not necessary since the parametric form of the solution is well defined for nested models and a lot of useful information can be extracted from manipulating the solution in parametric form. As usual, once we have $q(\theta)$ we can also obtain $V(\theta)$ from the envelope condition $\nabla_\theta V(\theta) = \nabla_\theta u(q(\theta), \theta)$ together with $V(0,1) = 0$ and we can also obtain $t(\theta)$ from $t(\theta) = u(q(\theta), \theta) - V(\theta)$. In the next chapter we present a variety of examples to illustrate how this algorithm works in practice.

Appendix: Proofs

Proof of proposition 3.1

Proof. From (E) we have $V_{\theta_1}(\theta) = u_{\theta_1}(q(\theta), \theta)$ and $V_{\theta_2}(\theta) = u_{\theta_2}(q(\theta), \theta)$. Differentiating both equations we get:

$$V_{\theta_2\theta_1}(\theta) = u_{q\theta_1}(q(\theta), \theta)q_{\theta_2}(\theta) + u_{\theta_2\theta_1}(q(\theta).\theta)$$
$$V_{\theta_1\theta_2}(\theta) = u_{q\theta_2}(q(\theta), \theta)q_{\theta_1}(\theta) + u_{\theta_1\theta_2}(q(\theta).\theta)$$

Hence, if θ is a point where V is twice continuously differentiable we have that

$$u_{q\theta_1}(q(\theta), \theta)q_{\theta_2}(\theta) = u_{q\theta_2}(q(\theta), \theta)q_{\theta_1}(\theta)$$

Therefore, we have that

$$-u_{q\theta_2}q_{\theta_1} + u_{q\theta_1}q_{\theta_2} = 0, \quad a.e. \quad on \quad [0,1]^2$$

Since q is implementable $\theta \in \operatorname{argmax}_{\hat{\theta}\in[0,1]^2}\{u(q(\hat{\theta}), \theta) - t(\hat{\theta})\} \quad \forall\theta \in [0,1]^2$. In particular, for all $(\theta_1, \theta_2) \in [0,1]^2$

$$\theta_1 \in \operatorname*{argmax}_{\hat{\theta_1}\in[0,1]}\{u(q(\hat{\theta_1}, \theta_2), \theta_1, \theta_2) - t(\hat{\theta_1}, \theta_2)\} \quad \forall\theta_2 \in [0,1]$$

$$\theta_2 \in \operatorname*{argmax}_{\hat{\theta_2}\in[0,1]}\{u(q(\theta_1, \hat{\theta_2}), \theta_1, \theta_2) - t(\theta_1, \hat{\theta_2})\} \quad \forall\theta_1 \in [0,1]$$

Then A2 together with the monotone maximum theorem imply: $u_{q\theta_1} > 0 \Rightarrow q(., \theta_2)$ is monotone increasing for each $\theta_2 \in [0,1]$ and $u_{q\theta_2} < 0 \Rightarrow q(\theta_1, .)$ is monotone decreasing for each $\theta_1 \in [0,1]$. $\square$

Proof of proposition 3.3

Proof. By the first and second order conditions for the maximization problems of the agents, given $V(s,t) = u(q(s),t) - t(s)$, the contract satisfies:

$$(1') \quad V_s(t,t) = u_q(q(t),t)\nabla q(t) - \nabla t(t) = 0$$

$$(2') \quad V_{ss}(t,t) = u_{qq}(q(t),t)\nabla q(t).(\nabla q(t))^T + u_q(q(t),t)D^2q(t) - D^2t(t) \leq 0$$

Differentiate the $1st$ coordinate of (1') with respect to t_2 and the $2nd$ coordiate with respect to t_1, we get:

$$(3') \quad u_{qq}(q(t),t)q_{t_2}(t)q_{t_1}(t) + u_{qt_2}(q(t),t)q_{t_1}(t) + u_q(q(t),t)q_{t_2t_1}(t) - t_{t_2t_1}(t) = 0$$

$$(4') \quad u_{qq}(q(t),t)q_{t_1}(t)q_{t_2}(t) + u_{qt_1}(q(t),t)q_{t_2}(t) + u_q(q(t),t)q_{t_1t_2}(t) - t_{t_1t_2}(t) = 0$$

If t is a point where (q,t) is twice continously differentiable then $t_{t_1t_2}(t) = t_{t_2t_1}(t)$ and $q_{t_1t_2}(t) = q_{t_2t_1}(t)$ so from (3') and (4') we get (1):

$$(1) \quad u_{q\theta_1}(q(\theta),\theta)q_{\theta_2}(\theta) = u_{q\theta_2}(q(\theta),\theta)q_{\theta_1}(\theta)$$

To get (2) differentiate (1') in relation to t to get:

$$0 = V_{ss}(t,t) + V_{st}(t,t)$$

$$= [u_{qq}(q(t),t)\nabla q(t).(\nabla q(t))^T + u_q(q(t),t)D^2q(t) - D^2t(t)] + u_{tq}(q(t),t)\nabla q(t)$$

Then (2') is equivalent to:

$$u_{tq}(q(t),t)\nabla q(t) \geq 0$$

Since this is a product of a 2×1 by a 1×2 this is equivalent to:

$$(2) \quad u_{qt_i}(q(\theta),\theta)q_{\theta_i}(\theta) \geq 0 \quad \forall i = 1,2.$$

□

Proof of proposition 3.4

Proof. By definition of γ_r and of A and B as solutions of their respective ODE's we have

$$\dot{\gamma_r}(t) = \begin{cases} (1, -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r),\gamma_r(t))) : 0 \leq r \leq \frac{1}{2} \\ (-\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r),\gamma_r(t)),1) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

from where the following fundamental equation that we will use follows:

$$u_q(\phi,\gamma_r(s)) - u_q(\phi,\gamma_r(0)) = \int_0^s \nabla_\theta u_q(\phi,\gamma_r(t)).\dot{\gamma_r}(t)dt = 0$$

First lets see that $\{\gamma_r\}_{r\in[1/2,1]}$ never cross each other and are on the right of $\gamma_{1/2}$. Since we have

$$\gamma_r = \{(A(\phi,r,s),s) : 0 \leq s \leq U(\phi,r)\}$$

it is enough to prove that $\frac{d}{dr}A(\phi,r,s) = A_\phi\phi' + A_r > 0$ which implies that whenever $1/2 \leq r_1 < r_2 \leq 1$ we have $A(\phi(r_2),r_2,s) > A(\phi(r_1),r_1,s)$ for all s and hence $\gamma_{r_1} \cap \gamma_{r_2} = \emptyset$.

By partially differentiating our fundamental equation $u_q(\phi,A(\phi,r,s),s) = u_q(\phi,2r-1,0)$ with respect to ϕ and r we get.

$$u_{qq}(\phi,A,s) + u_{q\theta_1}(\phi,A,s)A_\phi + = u_{qq}(\phi,2r-1,0)$$
$$u_{q\theta_1}(\phi,A,s)A_r = 2u_{q\theta_1}(\phi,2r-1,0)$$

Then, it is clear that $A_r = \frac{2u_{q\theta_1}(\phi,2r-1,0)}{u_{q\theta_1}(\phi,A,s)} > 0$ and, since we are considering only nondecreasing ϕ, to conclude its enough to show that

$$A_\phi = \frac{u_{qq}(\phi,2r-1,0) - u_{qq}(\phi,A,s)}{u_{q\theta_1}(\phi,A,s)} \geq 0$$

This last inequality follows from examining the numerator:

$$u_{qq}(\phi,\gamma_r(o)) - u_{qq}(\phi,\gamma_r(s)) = \int_s^0 \nabla_\theta u_{qq}(\phi,\gamma_r(t)).\dot{\gamma_r}(t)dt$$
$$= \int_0^s \frac{u_{qq\theta_1}u_{q\theta_2} - u_{qq\theta_2}u_{q\theta_1}}{(u_{q\theta_1})^2}(\phi,\gamma_r(t))u_{q\theta_1}(\phi,\gamma_r(t))dt \geq 0$$

since by A3 $\frac{u_{qq\theta_1}u_{q\theta_2}-u_{qq\theta_2}u_{q\theta_1}}{(u_{q\theta_1})^2} = \frac{d}{dq}(-\frac{u_{q\theta_2}}{u_{q\theta_1}}) \geq 0$ holds identically.

To finish we show that $\{\gamma_r\}_{r\in[0,1/2]}$ never cross each other and are on the left of $\gamma_{1/2}$. The procedure is the same. Now we have

$$\gamma_r = \{(s,B(\phi,r,s)) : 0 \leq s \leq U(\phi,r)\}$$

so it is enough to prove that $\frac{d}{dr}B(\phi,r,s) = B_\phi\phi' + B_r < 0$. By partially differentiating our fundamental equation $u_q(\phi,s,B(\phi,r,s)) = u_q(\phi,0,1-2r)$ we get

$$B_r = \frac{-2u_{q\theta_2}(\phi,1-2r,0)}{u_{q\theta_2}(\phi,s,B)} < 0$$
$$B_\phi = \frac{u_{qq}(\phi,0,1-2r) - u_{qq}(\phi,s,B)}{u_{q\theta_2}(\phi,s,B)}$$

Since ϕ must be nondecreasing and $u_{q\theta_2} < 0$ we only need to establish $u_{qq}(\phi,0,1-2r) - u_{qq}(\phi,s,B) \geq 0$ which follows from

$$u_{qq}(\phi,\gamma_r(o)) - u_{qq}(\phi,\gamma_r(s)) = \int_s^0 \nabla_\theta u_{qq}(\phi,\gamma_r(t)).\dot{\gamma_r}(t)dt$$
$$= \int_0^s \frac{u_{qq\theta_2}u_{q\theta_1} - u_{qq\theta_1}u_{q\theta_2}}{(u_{q\theta_2})^2}(\phi,\gamma_r(t))u_{q\theta_2}(\phi,\gamma_r(t))dt \geq 0$$

since by A3 $\frac{u_{qq\theta_2}u_{q\theta_1}-u_{qq\theta_1}u_{q\theta_2}}{(u_{q\theta_2})^2} = \frac{d}{dq}(-\frac{u_{q\theta_1}}{u_{q\theta_2}}) \leq 0$ holds identically. $\square$

Proof of proposition 3.5

Proof. Observe that we only need to proof that incentive compatibility holds on the closure of the participation region since every agent in the exclusion region is offered the same contract as every agent on the frontier γ_{r_0}. Agents on the participation region dont find this contract attractive and since agents $\theta \in \gamma_{r_0}$ dont find the contracts of participating agents attractive and $u_{\theta_1} > 0$, $u_{\theta_2} < 0$ then agents on the exclusion region dont find it attractive either.

Given q we choose $V(\theta)$ s.t. $\nabla_\theta V(\theta) = \nabla_\theta u(q(\theta), \theta)$ and $V(0,1) = 0$. We proceed on three steps to prove that (q, V) satisfies

$$(IC) \quad V(\theta) - V(\theta') \geq u(q(\theta^{'}), \theta) - u(q(\theta^{'}), \theta') \quad \forall \theta, \theta^{'} \in [0,1]^2$$

Claim 1: (IC) are binding along characteristics

Let $\theta_1, \theta_2 \in \gamma_r$ for some $r \in [r_0, 1]$, then for some $0 \leq s_1 \leq s_2 \leq U(\phi, r)$ we can write, according to the value of r,

$$\theta_i = \gamma_r(s_i) = \begin{cases} (s_i, B(\phi(r), r, s_i)) : 0 \leq r \leq \frac{1}{2} \\ (A(\phi(r), r, s_i), s_i) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

Using condition (E) $\nabla_\theta V(\theta) = \nabla_\theta u(q(\theta), \theta)$ together with the fundamental theorem of calculus (FTC):

$$V(\theta_2) - V(\theta_1) = \int_{\gamma_r \upharpoonright_{[s_1, s_2]}} \nabla_\theta V(\alpha) d\alpha = \int_{\gamma_r \upharpoonright_{[s_1, s_2]}} \nabla_\theta u(q(\alpha), \alpha) d\alpha$$

$$= \int_{\gamma_r \upharpoonright_{[s_1, s_2]}} \nabla_\theta u(q(\theta_1), \alpha) d\alpha = u(q(\theta_1), \theta_2) - u(q(\theta_1), \theta_1)$$

Where we have use that for all $s \in [s_1, s_2]$, we have $q(\alpha) = q(\gamma_r(s)) = q(\gamma_r(s_1)) = q(\theta_1) = \phi(r)$, a constant along γ_r.

Claim 2: (IC) holds along the edge $\Gamma([0,1])$

First lets consider $\Gamma([0, 1/2])$. Take $a, b \in [0,1]$ then using FTC:

$$V(0,b) - V(0,a) = \int_a^b V_{\theta_2}(0, \beta) d\beta = \int_a^b u_{\theta_2}(q(0,\beta), 0, \beta) d\beta$$

$$u(q(0,a), 0, b) - u(q(0,a), 0, a) = \int_a^b u_{\theta_2}(q(0,a), 0, \beta) d\beta$$

By substracting both equations we get on the right:

$$\int_a^b (u_{\theta_2}(q(0,\beta),0,\beta) - u_{\theta_2}(q(0,a),0,\beta)d\beta = \int_a^b \int_{q(0,a)}^{q(0,\beta)} u_{q\theta_2}(q',0,\beta)dq'd\beta \geq 0$$

Since $u_{q\theta_2} < 0$ and $q(0,y) = \phi(\frac{1-y}{2})$ is nonincreasing in y[17]. Therefore:

$$V(0,b) - V(0,a) \geq u(q(0,a),0,b) - u(q(0,a),0,a) \quad \forall a,b \in [0,1]$$

Now consider $\Gamma([1/2,1])$. Take $a,b \in [0,1]$ then using FTC:

$$V(b,0) - V(a,0) = \int_a^b V_{\theta_1}(\alpha,0)d\alpha = \int_a^b u_{\theta_1}(q(\alpha,0),\alpha,0)d\alpha$$

$$u(q(a,0),b,0) - u(q(a,0),a,0) = \int_a^b u_{\theta_1}(q(a,0),\alpha,0)d\alpha$$

By substracting both equations we get on the right:

$$\int_a^b (u_{\theta_1}(q(\alpha,0),\alpha,0) - u_{\theta_1}(q(a,0),\alpha,0))d\alpha = \int_a^b \int_{q(a,0)}^{q(\alpha,0)} u_{q\theta_1}(q',\alpha,0)dq'd\alpha \geq 0$$

Since $u_{q\theta_1} > 0$ and $q(x,0) = \phi(\frac{1+x}{2})$ is nondecreasing in x[18]. Therefore:

$$V(b,0) - V(a,0) \geq u(q(a,0),b,0) - u(q(a,0),a,0) \quad \forall a,b \in [0,1]$$

Finally take $a,b \in [0,1]$ and lets proof that $(0,b)$ has no incentive to pretend to be $(a,0)$. The previous results gives us:

$$V(0,b) - V(0,0) \geq u(q(0,0),0,b) - u(q(0,0),0,0)$$

$$V(0,0) - V(a,0) \geq u(q(a,0),0,0) - u(q(a,0),a,0)$$

Therefore, adding up, we get:

$$V(0,b) - V(a,0) \geq u(q(0,0),0,b) - u(q(0,0),0,0) + u(q(a,0),0,0) - u(q(a,0),a,0)$$

To proof that $(0,b)$ doesnt win from pretending to be $(a,0)$ it is enough to check that the right hand side is greater than $u(q(a,0),0,b) - u(q(a,0),a,0))$ and that is equivalent to:

$$u(q(0,0),0,b) - u(q(0,0),0,0) \geq u(q(a,0),0,b) - u(q(a,0),0,0)$$

$$\Leftrightarrow \int_0^b u_{\theta_2}(q(0,0)0,\beta)d\beta \geq \int_0^b u_{\theta_2}(q(a,0)0,\beta)d\beta$$

[17] We are using our parametrization $(0,y) = (0,1-2r)$ for $0 \leq r \leq 1/2$ and $y \in [0,1]$

[18] We are using our parametrization $(x,0) = (2r-1,0)$ for $1/2 \leq r \leq 1$ and $x \in [0,1]$

$$\Leftrightarrow \int_0^b \int_{q(a,0)}^{q(0,0)} u_{q\theta_2}(q,0,\beta)dqd\beta \geq 0$$

This last inequality holds because $u_{q\theta_2} < 0$ and $q(x,0)$ is nondecreasing in x.

We should also check that $(a,0)$ doesnt win from pretending to be $(0,b)$. In a similar way, use the previous results to get:

$$V(a,0) - V(0,0) \geq u(q(0,0),a,0) - u(q(0,0),0,0)$$

$$V(0,0) - V(0,b) \geq u(q(0,b),0,0) - u(q(0,b),0,b)$$

Therefore, adding up, we get:

$$V(a,0) - V(0,b) \geq u(q(0,0),a,0) - u(q(0,0),0,0) + u(q(0,b),0,0) - u(q(0,b),0,b)$$

To finish the claim is enough to show that the right hand side is greater than $u(q(0,b),a,0) - u(q(0,b),0,b)$ and this is equivalent to:

$$u(q(0,0),a,0) - u(q(0,0),0,0) \geq u(q(0,b),a,0) - u(q(0,b),0,0)$$

$$\Leftrightarrow \int_0^a u_{\theta_1}(q(0,0),\alpha,0)d\alpha \geq \int_0^a u_{\theta_1}(q(0,b),\alpha,0)d\alpha$$

$$\Leftrightarrow \int_0^a \int_{q(0,b)}^{q(0,0)} u_{q\theta_1}(q,\alpha,0)dqd\alpha \geq 0$$

This last inequality holds becauuse $u_{q\theta_1} > 0$ and $q(0,y)$ is nonincreasing.

Thus (IC) holds between any two types on $\Gamma([0,1])$

Claim 3: (IC) holds on the closure of the participation region.

Pick any θ, θ' in such a region and take $r, r' \in [0,1]$ such that $\theta = \gamma_r(s)$ and $\theta' = \gamma_{r'}(s')$ for some $0 \leq s \leq U(\phi, r)$ and $0 \leq s' \leq U(\phi, r')$. Denote also $\alpha = \gamma_r(0)$ and $\alpha' = \gamma_{r'}(0)$ as being the only points on $\gamma_r \cap \Gamma([0,1])$ and on $\gamma_{r'} \cap \Gamma([0,1])$.

By claim 1 we have:

- $V(\theta) - V(\alpha) = u(q(\alpha),\theta) - u(q(\alpha),\alpha)$
- $V(\alpha') - V(\theta') = u(q(\theta'),\alpha') - u(q(\theta'),\theta')$

By claim 2 we have:

- $V(\alpha) - V(\alpha') \geq u(q(\alpha'),\alpha) - u(q(\alpha'),\alpha')$

Adding up and using $q(\theta) = q(\gamma_r(s)) = \phi(r) = q(\gamma_r(0)) = q(\alpha)$ and $q(\theta') = q(\gamma_{r'}(s')) = \phi(r') = q(\gamma_{r'}(0)) = q(\alpha')$ we get:

$$V(\theta) - V(\theta') \geq u(q(\alpha), \theta) - u(q(\alpha), \alpha) + u(q(\alpha'), \alpha) - u(q(\theta'), \theta')$$

To finish claim 3 is enough to show that the right hand side is greater or equal than $u(q(\theta'), \theta) - u(q(\theta'), \theta')$ but this is equivalent to:

$$\begin{aligned} & u(q(\alpha), \theta) - u(q(\alpha), \alpha) \geq u(q(\theta'), \theta) - u(q(\theta'), \alpha) \\ \Leftrightarrow & \int_{\gamma_r\restriction_{[0,s]}} \nabla_\theta u(q(\theta), \lambda) d\lambda \geq \int_{\gamma_r\restriction_{[0,s]}} \nabla_\theta u(q(\theta'), \lambda) d\lambda \\ \Leftrightarrow & \int_0^s \frac{d}{dt} u(x, \gamma_r(t)) dt \geq \int_0^s \frac{d}{dt} u(y, \gamma_r(t)) dt \\ \Leftrightarrow & \int_0^s \frac{d}{dt} \{ \int_y^{\phi(r)} u_q(z, \gamma_r(t)) dz \} dt \geq 0 \end{aligned}$$

Where we have set $x = q(\theta) = \phi(r)$ and $y = q(\theta') = \phi(r')$ which are fixed through the integration. By Leibnitz rule we can differentiate under the integral:

$$\Leftrightarrow \int_0^s \int_y^{\phi(r)} \nabla_\theta u_q(z, \gamma_r(t)) \cdot \dot{\gamma_r}(t) dz dt \geq 0$$

But since we have:

$$\dot{\gamma_r}(t) = \begin{cases} (1, -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), \gamma_r(t))) : 0 \leq r \leq \frac{1}{2} \\ (-\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), \gamma_r(t)), 1) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

Then the integrand becomes:

$$\nabla_\theta u_q(z, \gamma_r(t)) \cdot \dot{\gamma_r}(t) = \begin{cases} u_{q\theta_1}(z, \gamma_r(t)) - \frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), \gamma_r(t))) u_{q\theta_2}(z, \gamma_r(t)) \ : 0 \leq r \leq \frac{1}{2} \\ -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), \gamma_r(t)) u_{q\theta_1}(z, \gamma_r(t)) + u_{q\theta_2}(z, \gamma_r(t)) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

By using A2, $u_{q\theta_1} > 0$ and $u_{q\theta_2} < 0$, we see that the integrand is nonnegative as long as the following holds:

$$\frac{u_{q\theta_2}}{u_{q\theta_1}}(z, \gamma_r(t)) \geq \frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), \gamma_r(t))$$

But by using A3, $\frac{d}{dq}(\frac{u_{q\theta_2}}{u_{q\theta_1}}) \leq 0$, we see that the above inequality holds when $\phi(r) \geq z$ and its reversed when $\phi(r) \leq z$. This implies that the inner integral is always nonnegative and hence the condition that guarantees incentive compatibility between θ and θ' is verified. □

Proof of proposition 3.6

Proof. Integrating by parts and using (E):

$$\int_0^1 V(\theta)\rho(\theta)d\theta_1 = -V(1,\theta_2)F_1(1,\theta_2) + V(0,\theta_2)F_1(0,\theta_2) + \int_0^1 u_{\theta_1}(q(\theta),\theta)F_1(\theta)d\theta_1$$

Since $F_1(1,\theta_2) = 0$ we get

$$\int_0^1\int_0^1 V(\theta)\rho(\theta)d\theta_1 d\theta_2 = \int_0^1\int_0^1 u_{\theta_1}(q(\theta),\theta)F_1(\theta)d\theta_1 d\theta_2 + \int_0^1 V(0,\theta_2)F_1(0,\theta_2)d\theta_2$$

Integrating by parts again and using (E):

$$\int_0^1 V(0,\theta_2)F_1(0,\theta_2)d\theta_2 = \int_0^1\int_0^1 V(0,\theta_2)\rho(\theta_1,\theta_2)d\theta_1 d\theta_2$$

$$= \int_0^1 (-V(0,1)F_2(\theta_1,1) + V(0,0)F_2(\theta_1,0) + \int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)F_2(\theta)d\theta_2)d\theta_1$$

Moreover $F_2(\theta_1,1) = 0$ and by the FTC together with (E) we have:

$$V(0,0) = V(0,1) - \int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)d\theta_2$$

Hence the previous expression reduces to:

$$\int_0^1 V(0,\theta_2)F_1(0,\theta_2)d\theta_2 = V(0,1) + \int_0^1\int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)[F_2(\theta) - F_2(\theta_1,0)]d\theta_2 d\theta_1$$

$$= V(0,1) + \int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)[\int_0^1 F_2(\theta)d\theta_1 - 1]d\theta_2$$

and finally we get:

$$\int_0^1\int_0^1 V(\theta_1,\theta_2)\rho(\theta_1,\theta_2)d\theta_1 d\theta_2 = \int_0^1\int_0^1 u_{\theta_1}(q(\theta),\theta)F_1(\theta)d\theta_1 d\theta_2$$

$$+V(0,1) + \int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)[\int_0^1 F_2(\theta)d\theta_1 - 1]d\theta_2$$

Replacing this on the objective function:

$$\int_{[0,1]^2} [u(q(\theta),\theta) - c(q(\theta)) - V(\theta))]\rho(\theta)d\theta$$

and taking into account (IR) we see that we must have $V(0,1) = 0$ and the objective function becomes:

$$\int_0^1\int_0^1 G(q(\theta_1,\theta_2),\theta_1,\theta_2)d\theta_1 d\theta_2 - \int_0^1 u_{\theta_2}(q(0,\theta_2),0,\theta_2)[\int_0^1 F_2(\theta_1,\theta_2)d\theta_1 - 1]d\theta_2$$

□

Proof of proposition 3.7

Proof. We only give the proof of the first condition since the proof of the second is analogous and simpler. The proof of the second condition can also be found on Araujo, Calagua and Vieira (2022)[19] who first discovered this type of optimality condition. As these authors noted, this optimality condition has a close similarity with the optimality condition found for unidimensional screening problems without Spence-Mirrless (Araujo and Moreira 2010).

First we compute the terms involved in the euler equation when $r_0 \leq r \leq 1/2$. In this case we have that $\gamma_r(s) = (s, B(\phi, r, s))$ and $|\frac{\partial(\theta_1,\theta_2)}{\partial(r,s)}| = -B_\phi\phi' - B_r$ and differentiating we get:

$$H_\phi = \int_0^U -(G_q + G_{\theta_2}B_\phi)(B_\phi\phi' + B_r) - G(B_{\phi\phi}\phi' + B_{\phi r})ds$$

$$-G(\phi, U, B(\phi, r, U))(B_\phi\phi' + B_r)U_\phi - 2u_{q\theta_2}(\phi, 0, 1-2r)[\int_0^1 F_2(t, 1-2r)dt - 1]$$

$$H_{\phi'} = \int_0^U -GB_\phi ds$$

$$\frac{d}{dr}H_{\phi'} = \int_0^U -(G_q\phi' + G_{\theta_2}(B_\phi\phi' + B_r))B_\phi - G(B_{\phi\phi}\phi' + B_{\phi r})ds$$
$$-G(\phi, U, B(\phi, r, U))B_\phi(\phi, r, U)(U_\phi\phi' + U_r)$$

Then, the euler equation $H_\phi - \frac{d}{dr}H_{\phi'} = 0$ reduces to:

$$-G(\phi, U, B)[B_rU_\phi - B_\phi U_r] - 2u_{q\theta_2}(\phi, 0, 1-2r)[\int_0^1 F_2(t, 1-2r)dt - 1] = \int_0^U G_qB_rds$$

By definition $U(\phi, r) = sup\{s \in [0,1] : \gamma_r(s) \in [0,1]^2\}$ and since $r_0 \leq r \leq 1/2$ we know that γ_r is a curve that starts at the western edge of $[0,1]^2$ ($\gamma_r(0) = (0, 1-2r)$) and travels northeast ($\dot{\gamma_r}(s) = (1, -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), \gamma_r(s)))$). Therefore, for all r (except for at most one[20]) there is a neighbourhood of r where either $U(\phi, r) = 1$ or $B(\phi, r, U(\phi, r)) = 1$ holds identically.

In the first case, this implies $U_\phi = 0$ and $U_r = 0$. In the second case, the identity implies $B_\phi + B_sU_\phi = 0$ and $B_r + B_sU_r = 0$ which together imply $B_\phi U_r = B_rU_\phi$. Therefore, in both cases Euler's equation reduces to:

$$\int_0^U G_qB_rds = 2u_{q\theta_2}(\phi, 0, 1-2r)[1 - \int_0^1 F_2(t, 1-2r)dt]$$

[19]The proof is essentially the same with some minor adjustments in notation due to our different parametrization of the initial condition.

[20]This brings no difficulty since we are proving the condition for every $r \neq \hat{r}$ where $\hat{r}$ is the possible exception (i.e, $(U(\phi(\hat{r}), \hat{r}), B(\phi(\hat{r}), \hat{r}, U(\phi(\hat{r}).\hat{r}))) = (1,1)$). Taking a sequence of $r_n \neq \hat{r}$ converging to $\hat{r}$ we see that by continuity the condition also holds for $\hat{r}$.

Now we exploit the fact that types on the same characteristic have the same marginal utility:

$$u_q(\phi,\gamma_r(s)) - u_q(\phi,\gamma_r(0)) = \int_0^s \nabla_\theta u_q(\phi,\gamma_r(t)).\dot{\gamma_r}(t)dt = 0$$

$$\Rightarrow u_q(\phi,0,1-2r) = u_q(\phi,s,B(\phi,r,s))$$

$$\Rightarrow -2u_{q\theta_2}(\phi,0,1-2r) = u_{q\theta_2}(\phi,s,B(\phi,r,s))B_r$$

Dividing the euler equation by $-2u_{q\theta_2}(\phi,0,1-2r)$ and replacing $B_r$we get:

$$\int_0^U \frac{G_q}{u_{q\theta_2}}(\phi,s,B(\phi,r,s))ds = [\int_0^1 F_2(t,1-2r)dt - 1]$$

$\square$

Chapter 4

Examples

In this chapter we show how the simple algorithm developed at the end of the last section can be applied to particular examples to get closed form solutions. One of the main advantages of our methodology is its generality. In particular, we do not require the level sets of the optimal allocation to be straight lines as requiring linearity on types (Basov 2001) or satisfaction of GSC (McAfee, McMillan 1988) would imply. An optimal allocation exhibiting curvature on its level sets can be found on example 5.2 in Araujo et al. (2022). These authors consider a firm with cost function $C(q) = \lambda q$ where $\lambda \in (0, 1)$ and consumers uniformly distributed on the unit square according to a valuation function

$$u(q, \theta_1, \theta_2) = (c - \theta_2) log(\theta_1 q + 1), \quad c > 1$$

Since Araujo et al (2022) focus on deriving optimality conditions they can only verify numerically the implementability of their solution candidate. However, its easy to verify that assumptions A1-A3 are satisfied.

- **A1** $u_{\theta_1} = \frac{(c-\theta_2)q}{\theta_1 q+1} > 0$ and $u_{\theta_2} = -log(\theta_1 q + 1) < 0$
- **A2** $u_{q\theta_1} = \frac{c-\theta_2}{(\theta_1 q+1)^2} > 0$ and $u_{q\theta_2} = -\frac{\theta_1}{\theta_1 q+1} < 0$
- **A3** $\frac{-u_{q\theta_2}}{u_{q\theta_1}} = \frac{\theta_1(\theta_1 q+1)}{c-\theta_2} \Rightarrow \frac{d}{dq}\frac{-u_{q\theta_2}}{u_{q\theta_1}} = \frac{\theta_1^2}{c-\theta_2} \geq 0$

Hence, all of our results apply. Moreover, Araujo et al (2022) optimality condition for this case (i.e., their theorem 1) is a particular case of our optimality condition (proposition 3.7). Indeed, this valuation function satisfies

$$u(q, 0, \theta_2) = 0 \leq 0, \quad \forall \theta_2 \in [0, 1], \forall q \geq 0$$

which implies that at the optimum $q(0, \theta_2) = 0, \forall \theta_2 \in [0, 1]$. Then, according to our parametrization of the initial condition $\phi(r) = 0, \forall 0 \leq r \leq 1/2 \Rightarrow r_0 \geq 1/2$ and we only need to solve for $\forall \max\{r_0, 1/2\} \leq r \leq 1$:

$$A_s(\phi, r, s) = -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), A(\phi, r, s), s), A(\phi, r, 0) = 2r - 1$$

$$\int_0^{U(\phi(r),r)} \frac{G_q}{u_{q\theta_1}}(\phi(r), A(\phi(r), r, s), s)ds = 0$$

which is a reparametrized version of Araujo et al (2022)'s optimality condition and hence their solution candidate is indeed the optimal solution to the screening problem[1].

In the next two sections we focus on examples that do not exhibit curvature in the level sets of their optimal allocation to show how unidimensional models and examples satisfying GSC or linearity on types can be embedded into our framework. In both of these cases the ODE'S can readily be solved and our algorithm reduces to computing the solution by solving the optimality conditions. It is worth noting, again, that all of these examples are the analogous of unidimensional examples without ironing since our optimality condition was derived under the assumption that the nondecreasingness of ϕ does not bind.

As such, all of these examples exhibit a classic "separation" result: consumer with different marginal utilities at equilibrium do not bunch together (unlike what happens in the unidimensional ironing situation). There is however a kind of "bunching" that is purely a consequence of the difference in dimensions of the instrument or allocation on one hand and the private information on the other hand. That is, different consumers $\theta \neq \hat{\theta}$ may bunch together $q(\theta) = q(\hat{\theta})$ only if they have the same marginal utility $u_q(q(\theta), \theta) = u_q(q(\hat{\theta}), \hat{\theta})$. In this case, however, unlike what happens in the unidimensional Ironing situation, from the point of view of the monopolist θ and $\hat{\theta}$ are equivalent: since they both exhibit the same marginal utility at equilibrium the monopolist has no incentive to treat them differently[2].

The existence of this equivalency is a consequence of the difference of dimensions since when $n = 1$ marginal utility is a scalar quantity and can only order a bidimensional type space by bunching together different types. More concretely, taking constants $k, c \in \mathbb{R}$, since $u_{q\theta_1} < 0$ then given $(\hat{\theta_1}, \hat{\theta_2})$ such that $u_q(k, \hat{\theta_1}, \hat{\theta_2}) = c$ by the implicit function theorem we can always find a local solution $\theta_1 = \theta_1(\theta_2)$ s.t. $\hat{\theta_1} = \theta_1(\hat{\theta_2})$ and $u_q(k, \theta_1(\theta_2), \theta_2) = c$ for all θ_2 in a neighbourhood of $\hat{\theta_2}$[3]. Hence, by appropriately substituting θ_2 by θ_1 according to $\theta_1(\theta_2)$ we can keep marginal utility constant[4] and the monopoly has no incentive to treat any type in the image of $\theta_2(\theta_1)$ differently. If $n = 2$ and $q = (q_1, q_2)$ the story would be different since, in principle, marginal utility $u_q = (u_{q_1}, u_{q_2})$ could be employed to perfectly separate types on a bidimensional space such as $[0, 1]^2$[5].

[1] The reparametrization comes from making $\hat{r} = 2r - 1$ for $1/2 \leq r \leq 1$ in the initial condition of the ODE.

[2] In the unidimensional case since $u_{q\theta} > 0$ if two different types $\theta_1 < \theta_2$ bunch together $q(\theta_1) = q(\theta_2)$ then $u_q(q(\theta_1), \theta_1) < u_q(q(\theta_2), \theta_2)$ and the monopolist would like to treat them differently since there is a greater surplus to be extracted from θ_2. The reason to treat θ_1 and θ_2 equally is not profit maximization but rather the need to avoid strategic behaviour by the agents (i.e., the local second order condition of the maximization subproblem of the agents is binding)

[3] Since $u_{q\theta_2} > 0$ we can also write this function as $\theta_2 = \theta_2(\theta_1)$

[4] Moreover, observe the rate at which this substitution occurs is given by $\frac{d\theta_1}{d\theta_2} = -\frac{u_{q\theta_2}}{u_{q\theta_1}}$

[5] Observe that imposing a constant non zero sign for $u_{q_i\theta_j} >$ for each $i, j = 1, 2$ would imply that around each point (θ_1, θ_2) there are two curves along which u_{q_1} and u_{q_2} are constant. However this two curves need not coincide and hence in general different types will exhibit different marginal utilities.

4.1 Embedding the standard unidimensional model:

In this section, we are first going to test our algorithm for bidimensional screening problems by solving ”fake” bidimensional models. More concretely, we are going to desguise standard unidimensional models as bidimensional models to see how our algorithm generates the same solution that can be obtained by the usual unidimensional tools. We will also see that this ”fake” bidimensional models can be recognized because they exhibit perfect sustitubility of the different dimensions of private information on the marginal utility. This implies that characteristics curves will be straight parallel lines which means that the optimal order of types according to marginal utility is exogenous (i.e, the order is the same for all implementable contracts).

The idea is quite straightforward. Start with a given unidimensional valuation function $\hat{u}(q,\beta)$ that satisfies the usual assumptions $\hat{u}_\beta > 0$ and $\hat{u}_{q\beta} > 0$ (Unidimensional Spence-Mirrless). Now lets assume that β is a linear combination of different dimensions of asymetric information, i.e.,

$$\beta = \lambda_1\theta_1 + \lambda_2(1-\theta_2), \quad \lambda_1, \lambda_2 > 0$$

Then all assumptions are satisfied for the bidimensional model $u(q,\theta_1,\theta_2) = \hat{u}(q,\beta)$

- **A1** $u_{\theta_1} = \lambda_1\hat{u}_\beta > 0$ and $u_{\theta_2} = -\lambda_2\hat{u}_\beta < 0$
- **A2** $u_{q\theta_1} = \lambda_1\hat{u}_{q\beta} > 0$ and $u_{q\theta_2} = -\lambda_2\hat{u}_{q\beta} < 0$
- **A3** $\frac{-u_{q\theta_2}}{u_{q\theta_1}} = \frac{\lambda_2}{\lambda_1} \Rightarrow \frac{d}{dq}\frac{-u_{q\theta_2}}{u_{q\theta_1}} = 0$

Observe how unidimensional Spence-Mirless implies A2 and A3 which together form a bidimensional version of the Spence Mirless assumption. In particular, A3 is satisfied in a very weak sense. The marginal rate of sustitution between dimensions of private information on marginal utility is constant: independent of q. We call this perfect susitubility of different dimensions of private information on the marginal utility or simply perfect sustitubility. Since in this case the different dimensions of private information collapse into a scalar index $\beta(\theta_1,\theta_2)$ it is easy to see that the sustitubility that matters is $-\frac{\beta_{\theta_2}}{\beta_{\theta_1}} = \frac{\lambda_2}{\lambda_1} = \frac{-u_{q\theta_2}}{u_{q\theta_1}}$. On a general bidimensional model there may not be an index that allows us to reduce the model to a unidimensional but we can always study the sustitubility of different dimensions of private information on the marginal utility.

Computationally, perfect sustitubility implies that the ODE's that define the characteristic curves have inmediate solutions:

$$B(\phi,r,s) = 1-2r+\frac{\lambda_1}{\lambda_2}s$$

$$A(\phi,r,s) = 2r-1+\frac{\lambda_2}{\lambda_1}s$$

Therefore, characteristic curves are straight parallel lines with slope $\frac{\lambda_1}{\lambda_2}$:

$$\gamma_r(s) = \begin{cases} (s, 1-2r+\frac{\lambda_1}{\lambda_2}s) : r_0 \le r \le \frac{1}{2} \\ (2r-1+\frac{\lambda_2}{\lambda_1}s, s) : \frac{1}{2} \le r \le 1 \end{cases}$$

As we saw while we were studying implementability, every implementable contract q must come from solving a Cauchy problem with a certain initial data ϕ. In this case, however, the solution that emerges from the ODE's do not depends on ϕ. Hence, every implementable contract has the exact same level sets $\{\gamma_r\}_r$ which order the types according to marginal utility. That is, in this case there is an exogenous order on the type space according to marginal utility.

Since the order of types is exogenous, the only thing left to do in this case is to test different nondecreasing contracts q across this order. That is, to find the optimal ϕ by solving the optimality condition on proposition 3.7. We illustrate the procedure with a concrete example:

Example I

Lets consider consumers uniformly distributed on the unit square with valuation function:

$$u(q,\theta_1,\theta_2) = (\theta_1 + 1 - \theta_2)q - \frac{1}{2}q^2, \quad c > 0$$

Assume for simplicity's sake that the firm has zero cost.

By creating a variable $\beta = \theta_1 + 1 - \theta_2$, we see that this can be reduced to a unidimensional model $u(\beta, q) = \beta q - \frac{q^2}{2}$ where β is distributed on $[0,2]$ according to the derived distribution:

$$F(\beta) = \begin{cases} \frac{\beta^2}{2} & : 0 \leq \beta \leq 1 \\ 1 - \frac{(2-\beta)^2}{2} & : 1 \leq \beta \leq 2 \end{cases}$$

By using standard unidimensional tools we can solve for the optimal contract which is given by:

$$q(\beta) = \begin{cases} 0 & : 0 \leq \beta \leq \sqrt{2/3} \\ \frac{3}{2}\beta - \frac{1}{\beta} & : \sqrt{2/3} \leq \beta \leq 1 \\ \frac{3}{2}\beta - 1 & : 1 \leq \beta \leq 2 \end{cases}$$

Alternatively, we show on the appendix that using our bidimensional tools we get the exact same solution:

$$q(\theta_1.\theta_2) = \begin{cases} 0 & : 0 \leq \theta_1 + 1 - \theta_2 \leq \sqrt{2/3} \\ \frac{3}{2}(\theta_1 + 1 - \theta_2) - \frac{1}{\theta_1 + 1 - \theta_2} & : \sqrt{2/3} \leq \theta_1 + 1 - \theta_2 \leq 1 \\ \frac{3}{2}(\theta_1 + 1 - \theta_2) - 1 & : 1 \leq \theta_1 + 1 - \theta_2 \leq 2 \end{cases}$$

4.2 Examples with linearity on types and GSC

There are two simple classes of utility functions that have been employed in the literature with relative success because of their simplicity. These are utilities that are linear on types and utilities satisfying GSC. They are indeed the simpler cases we can consider once we leave behind perfect substitubility because now the marginal rate of substitution between dimensions of private information on the marginal utility $-\frac{u_{q\theta_2}}{u_{q\theta_1}}(q,\theta_1,\theta_2)$ is no longer constant but unlike the general case it will not depend on s once we set $(q, \theta_1, \theta_1) = (\phi, \gamma_r(s))$. This implies that the characteristics curves are straight lines but not necessarily parallel. This is also the first instance in which assumption A3 is really needed. Since the characteristic lines are not parallel their non-crossing are not guaranteed unless we can ensure that their slopes vary in the right direction.

Definition: We say that the utility function u is linear on types if it can be written as

$$u(q, \theta_1, \theta_2) = \theta_1 v_1(q) - \theta_2 v_2(q) + v_3(q)$$

For a definition of a utility function satisfying GSC see section 2.3.1 where we also show that every utility that is linear on types satisfies GSC.

Proposition 4.1. *For any utility function that satisfies GSC, $-\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi, \gamma_r(s))$ doesnt depend on s. Moreover, the characteristics are given by*

$$\gamma_r(s) = \begin{cases} (s, 1-2r - \frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), 0, 1-2r)s) : 0 \leq r \leq \frac{1}{2} \\ (2r-1 - \frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), 2r-1, 0)s, s) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

If aditionally u is linear on types then the characteristics are

$$\gamma_r(s) = \begin{cases} (s, 1-2r + \frac{v_1'(\phi)}{v_2'(\phi)}s) : 0 \leq r \leq \frac{1}{2} \\ (2r-1 + \frac{v_2'(\phi)}{v_1'(\phi)}s, s) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

Unlike the previous case, now the level sets of an implementable contract depend on the initial data to the Cauchy problem ϕ. We no longer have an exogenous order according to marginal utility on the type space and we say that the optimal order of types according to marginal utility $\{\gamma_r\}_r$ is endogenous (i.e, depends on ϕ and hence not all implementable contracts exhibit the same ordering of types)[6]. In this case the optimal order of types $\{\gamma_r\}_r$ is determined together with the appropriate incentives across this order ϕ via the optimality conditions.

We could further simplify this conditions for the class of valuation functions that are linear on type since $u_{q\theta_1}(\phi, \gamma_r(s)) = v_1'(\phi)$, $u_{q\theta_2}(\phi, \gamma_r(s)) = -v_2'(\phi)$ and

$$G_q(\phi, \gamma_r(s)) = [\theta_1 v_1'(\phi) - \theta_2 v_2'(\phi) + v_3'(\phi) - c_q(\phi) - \frac{F_1(\gamma_r(s))}{\rho(\gamma_r(s))} v_1'(\phi)]\rho(\gamma_r(s))$$

Then the optimality conditions can be rewritten as:

$$\int_0^{U(\phi(r),r)} [-(1-2r)v_2'(\phi) + v_3'(\phi) - c_q(\phi) - \frac{F_1(s,B)}{\rho(s,B)} v_1'(\phi)]\rho(s,B)ds = v_2'(\phi)[1 - \int_0^1 F_2(t, 1-2r)dt],$$

$$\int_0^{U(\phi(r),r)} [(2r-1)v_1'(\phi) + v_3'(\phi) - c_q(\phi) - \frac{F_1(A,s)}{\rho(A,s)} v_1'(\phi)]\rho(A,s)ds = 0$$

[6]Of course, this is what a "truly bidimensional" example looks like. The previous case of perfect sustitubility was really a unidimensional model desguised as bidimensional

Now we present several examples that can be solved with these conditions:

Example II: A simple Loglinear utility function

Consider the case in which connsumers exhibit the following valuation function:

$$u(q, \theta_1, \theta_2) = \theta_1 log(q+1) + (c - \theta_2)q, \quad c \in (-1, 0)$$

For simplicity, lets assume that the firm has Zero cost and consumers are uniformly distributed on the unit square. Observe that the inverse demand function is given by:

$$p = \frac{\theta_1}{q+1} + c - \theta_2$$

So demands are hiperbolas and θ_2 is a dimension of asymetric information that only affects the intercept of the demand function (intensity of the demand) while θ_1 affects both the intercept and the curvature of the demand (the intensity of the demand and the sensitivity to price). The parameter c, which is known to the principal, can be thought as a displacement of the support of θ_2. We could regard $\hat{\theta_2} = c - \theta_2$ as the "real" second dimension of asymetric information and then, since θ_2 is distributed in $[0, 1]$, $\hat{\theta_2}$ is distributed (uniformly) on $[c-1, c]$.

The restriction of c to $(-1, 0)$ has a simple reason. On one hand, if $c \leq -1$ none of the types $(\theta_1, \theta_2) \in [0, 1]^2$ would have a positive demand for $q \geq 0$. Even though the principal doesnt know the exact types of the consumers, he knows that none of the consumers is willing to pay anything for the good and hence the solution is the same as with perfect information: full exclusion due to the low valuation of the types. On the other hand if $c > 0$, the demand of (θ_1, θ_2) has an horizontal asintota at $c - \theta_2$ which is greater than zero for at least one $\theta_2 \in [0, 1]$. For this θ_2, the principal can make unbound profits by selling larger and larger quantities of the good q. With perfect information there is no solution for this types because a solution would need to set $q = \infty$. With imperfect information, the presence of types who offer the posibility of infinitely large profit also distorts the problem in a similar way[7].

We can easily check that all assumptions are satisfied:

A1. $u_{\theta_1} = v_1 = log(q+1) > 0$ and $u_{\theta_2} = -v_2 = -q < 0$

A2. $u_{q\theta_1} = v_1' = \frac{1}{q+1} > 0$ and $u_{q\theta_2} = -v_2' = -1 < 0$

A3. $\frac{d}{dq}(-\frac{u_{q\theta_2}}{u_{q\theta_1}}) = \frac{d}{dq}(\frac{v_2'}{v_1'}) = \frac{d}{dq}(q+1) = 1 \geq 0$

and then we proceed on the appendix to find the solution which is given implicitly by $r_0 = \frac{2-c}{3}$, the level sets

$$\gamma_r(s) = \{(2r - 1 + \frac{-3r+2}{c}s, s) : 0 \leq s \leq \frac{2c(1-r)}{-3r+2}\}$$

[7]Observe that for $c > 0$ the set of types who offer this posibility has nonzero measure and the monopolist is tempted to offer a contract that sellls an infinite amount to this types and exclude the others.

and the values q assumes on its level sets for $r \in [r_0, 1]$

$$\phi(r) = \frac{-3r+2}{c} - 1, \quad \frac{2-c}{3} \leq r \leq 1$$

Quite often this implicit representation is enough to study the qualitative behaviour of the solution. For instance, it is inmediate that the participation region is a triangle with vertices $(\frac{1-2c}{3}, 0)$, $(1,0)$ and $(1, \frac{2}{3}(1+c))$. Therefore as $c \to -1$, i.e. as the support of the parameter that measures the intensity of demand moves to regions of lower marginal valuation, the participation set shrinks until full exclusion is achieved on $c = -1$ as expected.

We can also compute $q(1,0) = \phi(1) = -\frac{1+c}{c}$, which solves $u_q(q(1,0),1,0) = 0$, i.e. the equation marginal utility equals marginal cost which guarantees the efficiency of the contract holds for type $(1,0)$. Therefore, we obtain the usual conclusion that types "at the top of the distribution" get an efficient contract[8]. To see that this type is indeed the only one to get an efficient contract and all the other types get subefficient contracts we can return to the original variables (θ_1, θ_2). From $(\theta_1, \theta_2) = (2r - 1 + (\phi(r)+1)s, s)$ we get $s(\theta_1, \theta_2) = \theta_2$ and $r(\theta_1, \theta_2) = \frac{c\theta_1 - 2\theta_2 + c}{-3\theta_2 + 2c}$ so that finally we get the solution in the participation region $q(\theta_1, \theta_2) = \phi(r(\theta_1, \theta_2))$ which reduces to[9]:

$$q_{SB}(\theta_1, \theta_2) = \frac{-3\theta_1 + 1}{-3\theta_2 + 2c} - 1, \quad \forall (\theta_1, \theta_2) \in [0,1]^2 \quad s.t. \quad \theta_1 + \frac{2c-1}{3} \geq \theta_2$$

On the other hand the solution in the case of perfect information is given by:

$$q_{FB}(\theta_1, \theta_2) = \frac{\theta_1}{\theta_2 - c} - 1, \quad \forall (\theta_1, \theta_2) \in [0,1]^2 \quad s.t. \quad \theta_1 + c \geq \theta_2$$

Where again we specify the solution only in the participation region. Observe how, as expected, the introduction of asymetric information reduces the participation region from the triangle with vertices $\{(-c, 0), (1,0), (1, 1+c)\}$ to the one with vertices $(\frac{1-2c}{3}, 0)$, $(1,0)$ and $(1, \frac{2}{3}(1+c))$

Observe also that $q_{SB}(\theta_1, \theta_2) \leq q_{FB}(\theta_1, \theta_2)$ if and only if:

$$\frac{-3\theta_1 + 1}{-3\theta_2 + 2c} \leq \frac{\theta_1}{\theta_2 - c} \Leftrightarrow \theta_1 c + \theta_2 \geq c$$

Which holds for all $\theta_1, \theta_2 \in [0,1]$ and holds with equality when $(\theta_1, \theta_2) = (1,0)$ and with strict inequality in all the other cases. Hence, $(1,0)$ is the only to get an efficient contract and all the other types get subefficient contract. The usual onedimensional interpretation applies: the principal lowers the contract of types with lower marginal valuation below the efficient level to make it unatractive to types with higher marginal valuation[10]. Hence, only for the type at the top of the distribution the principal doesnt have any incentive to lower his contract below the efficient level.

[8] The type $(1,0)$ is the one "at the top" if we remember our interpretation that $\{\gamma_r\}_{r \in [r_0, 1]}$ orders the types according to their marginal valuation and in this case $\gamma_1 = \{(1,0)\}$

[9] We use the subscript SB to denote that we have found the solution to the adverse selection problem also known as the "second best" in contrast to the solution with complete information or first best solution

[10] The difference, of course, is that now a type has lower or higher marginal valuation according to which level set γ_r contains the type. The higher r, the higher the marginal valuation of $\theta \in \gamma_r$.

Example III: On lack of genericity of exclusion

In the previous example the valuation function was such that the demand of the agents had an horizontal asintota. This implied that we could not raise too much the support of the distribution of the intercepts since we would eventually enter regions where some agents are willing to buy an infinite amount of the good for a positive price and since we assume the monopolist has no cost this implied the posibility of an infinite profit. Now we introduce a linear quadratic model in order to study how much should we raise the support of the distribution of the intercepts in order to attain full participation. That is, we want to examine Armstrong (1996)'s hypothesis of the genericity of exclusion using our tools to compute solutions.

Lets consider again a firm with zero cost and consumers uniformly distributed on the unit square but with valuation function:

$$u(q,\theta_1,\theta_2) = (\theta_1 + c)q - (\theta_2+1)\frac{q^2}{2}, \quad c \geq -1$$

$$\Rightarrow p(q) = (\theta_1 + c) - (\theta_2 + 1)q$$

This is the famous Laffont-Maskin-Rochet (1987)'s example when $c = 0$. Observe that demands are linear, $\theta_1 + c$ is the intercept, $\theta_2 + 1$ the slope and as c increases we are displacing upwards the support of the distribution of the intercepts $[c, c+1]$. We place the restriction $c \geq -1$ since otherwise we would have a solution with full exclusion. We are interested in the behaviour of the participation region as $c \to +\infty$.

As a reference, lets see what happens on a unidimensional model. Fix $\theta_2 = 0$ and let $\theta_1 = \beta$ to consider the resulting unidimensional model. If we use standard unidimensional tools we obtain that the second best unidimensional solution is given by

$$-1 \leq c \leq 1 \Rightarrow q(\beta) = \begin{cases} 0 & : 0 \leq \beta \leq \frac{1-c}{2} \\ 2\beta + (c-1) : & \frac{1-c}{2} \leq \beta \leq 1 \end{cases}$$

$$1 \leq c \Rightarrow q(\beta) = 2\beta + (c-1), \quad 0 \leq \beta \leq 1$$

Hence, if we just let the support of the distribution of the intercepts be sufficiently high ($c > 1$) there is full participation. Armstrong (1996), however, provided and intuition and a result precluding this phenomenon if there is at least two dimensions of private information. However, one of his key hypothesis was that the type space must be strictly convex and in our case we have $\Theta = [0,1]^2$. Hence, we can not rule out the possibility of full participation for c large enough without solving the bidimensional model for each $c \geq 0$. We will do so in order to examine the validity of Armstrong's intuition for this example.

First we check that all assumptions are satisfied:

A1. $u_{\theta_1} = v_1 = q > 0$ and $u_{\theta_2} = -v_2 = -\frac{q^2}{2} < 0$

A2. $u_{q\theta_1} = v_1' = 1 > 0$ and $u_{q\theta_2} = -v_2' = -q < 0$

A3. $\frac{d}{dq}(-\frac{u_{q\theta_2}}{u_{q\theta_1}}) = \frac{d}{dq}(\frac{v_2'}{v_1'}) = \frac{d}{dq}(q) = 1 \geq 0$

Then we proceed to apply our bidimensional tools on the appendix where we show that the solution varies according to c and is given by:

$$-1 \leq c \leq 1 \Rightarrow \phi(r) = \begin{cases} 0 & : 0 \leq r \leq \frac{3-c}{4} \\ 8r - 6 + 2c & : \frac{3-c}{4} \leq r \leq \frac{4-c}{5} \\ 3r - 2 + c & : \frac{4-c}{5} \leq r \leq 1 \end{cases}$$

$$1 < c \leq 3/2 \Rightarrow \phi(r) = \begin{cases} 0 & : 0 \leq r \leq \frac{1}{2c} \\ \frac{1-2rc}{6r^2-4r} & : \frac{1}{2c} \leq r \leq 1/2 \\ 8r - 6 + 2c & : 1/2 \leq r \leq \frac{4-c}{5} \\ 3r - 2 + c & : \frac{4-c}{5} \leq r \leq 1 \end{cases}$$

$$3/2 < c \Rightarrow \phi(r) = \begin{cases} 0 & : 0 \leq r \leq \frac{1}{2c} \\ \frac{1-2rc}{6r^2-4r} & : \frac{1}{2c} \leq r \leq \frac{3}{3+2c} \\ \frac{2c-1}{6-8r} & : \frac{3}{3+2c} \leq r \leq 1/2 \\ 3r - 2 + c & : 1/2 \leq r \leq 1 \end{cases}$$

In particular, when $c = 0$ we recover the solution found by Laffont, Maskin, Rochet. Observe that for all values of c, $q(1,0) = \phi(1) = 1 + c$ which implies efficiency of the contract at the top (i.e., $u_q(q(1,0), 1, 0) = c_q(q(1,0))$). Moreover, as we displace upward the support of the distribution of the intercepts, as c increases, the participation region expands and when $c \to +\infty$ the participation region converges to $[0,1]^2$.

More concretely, lets look at the frontier between the participation and the exclusion region. For $-1 \leq c \leq 1$ we have $r_0 = \frac{3-c}{4}$ and therefore

$$\gamma_{\frac{3-c}{4}} = \{(\frac{1-c}{2}, s) : 0 \leq s \leq 1\}$$

So the frontier is a vertical line that moves from the eastern edge at $c = -1$ (fulll exclusion) to the western edge at $c = 1$ (full participation with the exception of the types at the frontier). Now for $c > 1$ we also have full participation with the exception of a subset of zero lebesque measure. To see this note that we have $r_0 = \frac{1}{2c}$ and now the frontier is given by

$$\gamma_{r_0} = \{(s, 1 - 2r_0 + \frac{1}{\phi(r_0)} s) : 0 \leq s \leq U(\phi(r_0), r_0)\}$$

where $U(\phi, r) = 2r\phi$[11] we have $\gamma_{\frac{1}{2c}}(0) = (0, 1 - \frac{1}{c})$ and $\gamma_{\frac{1}{2c}}(U(\phi(r_0), r_0)) = (0, 1)$ So $\gamma_{\frac{1}{2c}}$ is a vertical segment joining both points[12]. Hence, for $c \geq 1$ the exclusion set has zero lebesque measure contrary to Armstrong's intuition.

Example IV: Demand with constant Arrow-Pratt Index

[11] See the appendix for the careful derivation of γ_r as well as $U(\phi, r)$ corresponding to every r for each possible value of c.

[12] Observe that from the parametric form we see that γ_{r_0} should be a straight line with an "infinite" slope $\frac{1}{\phi(r_0)} = +\infty$.

Now consider the case in which consumers exhibit the following valuation function:

$$u(q,\theta_1,\theta_2) = \theta_1 \frac{1}{\alpha}(1-e^{-\alpha q}) + (c-\theta_2)q, \quad \alpha > 0, \quad c \in (-1,0)$$

We assume again that the firm has Zero cost and consumers are uniformly distributed on the square. Observe that the inverse demand function is given by:

$$p = \theta_1 e^{-\alpha q} + (c-\theta_2)$$

Again θ_2 only affects the intercept of the demand while θ_1 also affects the curvature of the demand and c acts as a displacement of the support of θ_2. The reason for restricting the values of c are the same as with example II. If $c \geq 0$ there are types with excesively high valuations that offer the principal the posibility of infinite profit and hence there is no solution. On the contrary, if $c \leq -1$ the valuation of the consumers are too low and full exclusion is the only solution.

The new feature is that the parameter α controls the convexity of the inverse demand function. To see this we can compute the Arrow-Pratt index of $p(q)$ and see that it is constant and equal to α:

$$\begin{aligned} p'(q) &= -\alpha\theta_1 e^{-\alpha q} < 0 \\ p''(q) &= \alpha^2\theta_1 e^{-\alpha q} > 0 \\ \Rightarrow \frac{-p''(q)}{p'(q)} &= \alpha \end{aligned}$$

All assumptions are satisfied:

A1. $u_{\theta_1} = v_1 = \frac{1}{\alpha}(1-e^{-\alpha q}) > 0$ and $u_{\theta_2} = -v_2 = -q < 0$

A2. $u_{q\theta_1} = v_1' = e^{-\alpha q} > 0$ and $u_{q\theta_2} = -v_2' = -1 < 0$

A3. $\frac{d}{dq}(-\frac{u_{q\theta_2}}{u_{q\theta_1}}) = \frac{d}{dq}(\frac{v_2'}{v_1'}) = \frac{d}{dq}(e^{\alpha q}) = \alpha e^{\alpha q} \geq 0$

On the appendix we derive the solution which is given by $r_0 = \frac{2-c}{3}$ and

$$\phi(r) = \frac{1}{\alpha} log(\frac{-3r+2}{c}), \quad \frac{2-c}{3} \leq r \leq 1$$

The participation region is exactly the same triangle as in example II and it also shrinks as $c \to -1$ until it reaches full exclusion. We can also verify that type $(1,0)$ gets an efficient contract since $q(1,0) = \phi(1) = \frac{1}{\alpha} log(-\frac{1}{c})$ implies $u_q(q(1,0),1,0) = 0$.

This example exhibits an additional interesting feature: the degree of convexity of the demand function doesnt affect neither the shape of the participation set[13] nor the optimal order of

[13] The participation set only depends on c. If we had chosen instead $u(q,\theta_1,\theta_2) = \theta_1(1-e^{-\alpha q}) + (c-\theta_2)q$ as our valuation function we would get that the degree of convexity of the demand as measured by α doesnt affect the level sets of q but it does change the participation set which increases as α grows. This happens because now α also affects the intercept of the demand function. As α grows demands are shifted upwards and the principal is willing to attend a bigger set of types.

types in the participation region, that is, the level sets of q. It only affects the values q assumes on each level set but if two types receive the same contract for a given value of α they still receive the same contract under any value of α although the contract itself may be different. To see this denote by $(\{\gamma_r^\alpha\}_{r\in[r_0,1]}, \phi_\alpha)$ the solution to the adverse selection problem for a fixed value of α, then:

$$\gamma_r^\alpha(s) = (A_\alpha(\phi_\alpha, r, s), s) = (2r - 1 + e^{\alpha\phi_\alpha} s, s) = (2r - 1 + \frac{-3r+2}{c} s, s), \quad 0 \leq s \leq \frac{2c(1-r)}{-3r+2}$$

Hence, $\{\gamma_r^\alpha\}_{r\in[r_0,1]}$ is the same for all α and we could say that the optimal order of types is in this case independent of the degree of convexity of the demand as measured by α[14]. Returning to the original variables we may express the solution as:

$$q_{SB}(\theta_1, \theta_2) = \frac{1}{\alpha} log(\frac{-3\theta_1 + 1}{-3\theta_2 + 2c}), \quad \forall(\theta_1, \theta_2) \in [0,1]^2 \quad s.t. \quad \theta_1 + \frac{2c-1}{3} \geq \theta_2$$

In comparison the solution with perfect information is given by:

$$q_{FB}(\theta_1, \theta_2) = \frac{1}{\alpha} log(\frac{\theta_1}{\theta_2 - c}), \quad \forall(\theta_1, \theta_2) \in [0,1]^2 \quad s.t. \quad \theta_1 + c \geq \theta_2$$

As before we verify that $q_{SB}(\theta_1, \theta_2) \leq q_{FB}(\theta_1, \theta_2)$ if and only if $\theta_1 c + \theta_2 \geq c$ Which holds with strict inequality for all $\theta_1, \theta_2 \in [0, 1]$ except for $(1, 0)$ who gets an efficient contract. All other types get subefficient contracts.

Example V: The basic linear case and its extensions

So far we have always assumed that types are uniformly distributed and the firm has zero cost. We now show that our methodology allow us to easily perform comparative statics where we vary preferences, cost functions or distributions. For this purpose we use Araujo et al. (2022)'s generalization of the classic example by Laffont, Maskin and Rochet (1987)[15]. The difference with our generalization in example III is that they consider a general displacement of the dimension of asymetric information that measures sensitivity to price instead of intensity of the demand.

Lets consider consumers exhibiting a valuation function given by:

$$u(q, \theta_1, \theta_2) = \theta_1 q - (\theta_2 + c)\frac{q^2}{2}, \quad c > 1/2$$

The firm has Zero cost and consumers are uniformly distributed on the unit square. It can readily be checked that all assumptions A1-A3 are satisfied. Solving first for $1/2 \leq r \leq 1$ the formula for the characteristics gives us:

$$A(\phi(r), r, s) = 2r - 1 + \frac{v_2'(\phi)}{v_1'(\phi)} s = 2r - 1 + \phi(r)s$$

[14] Moreover, the solution $\{\gamma_r\}_{r\in[r_0,1]}$ is the same as in example II although the optimal ϕ is different. If we had introduced instead the valuation function discussed in the previous footnote we would get a solution with $\{\gamma_r^\alpha\}_{r\in[r_0,1]}$ different than example II.

[15] The same variations in costs and distributions can also be applied for the previous examples. While the variation in preferences is analogous to what was done in example IV

Using $v_1'(\phi) = 1$, $v_3'(\phi) = -c\phi$, $c_q = 0$ $\rho = 1$, $F_1(A,s) = 1 - A = 2 - 2r - \phi(r)s$ in the optimality condition we get:

$$(4r - 3 - c\phi)U + \phi\frac{1}{2}U^2 = 0$$

Solving for $A(\phi, r, U) = 1$ we get $U = \frac{2(1-r)}{\phi}$ and $\phi_1(r) = \frac{3r-2}{c}$ which vanishes at $r = 2/3$. However, now $U(0, 2/3) = +\infty$. Hence, we must stop this solution at $U = 1$ or equivalently at $r_1 = \frac{2c+2}{2c+3}$. Now solving for $U = 1$ we get $\phi_2(r) = \frac{8r-6}{2c-1}$ which vanishes at $r_0 = 3/4$. Hence, the solution to the adverse selection problem is given implicitly by:

$$\phi(r) = \begin{cases} \frac{8r-6}{2c-1} & \frac{3}{4} \leq r \leq \frac{2c+2}{2c+3} \\ \frac{3r-2}{c} & \frac{2c+2}{2c+3} \leq r \leq 1 \end{cases}$$

together with $\gamma_r(s) = \{(2r - 1 + \phi(r)s, s) : 0 \leq s \leq U(\phi, r)\}$.

Observe that the participation set is now a rectangle with vertices $(1/2, 0)$, $(1, 0)$, $(1, 1)$ and $(1/2, 1)$ which doesnt depend on c[16]. However, c does affect the order of types inside the participation region. Observe that γ_{r_1} divides the participation region in two subregions. In the first region $\phi = \phi_2$ while in the second $\phi = \phi_1$ and since $\phi_2' = \frac{8}{2c-1} > \frac{3}{c} = \phi_1'$ we will refer to the first region as the one where the contract q grows faster with marginal utility[17]. Now observe that $\frac{d}{dc}r_1 = \frac{2}{(2c+3)^2} > 0$ and as $c \to 1/2$ we have $r_1 \to 3/4$ while as $c \to +\infty$ we have $r_1 \to 1$. Therefore, as consumers demand become more price insensitive, i.e. $c \to +\infty$ and the support of the demand's slopes is moved toward higher values, the region where the contract grows faster with marginal utility expands. On the contrary, as consumers demand become more price sensitive the region where the contract grows slower with marginal utility expands. At the limit, when $c = 1/2$, the expansion of this region is so big that it is inconsistent with the continuity of the contract: the contract must grow optimally so slowly that efficiency at the top ($r = 1$) and exclusion at the bottom ($r \leq 3/4$) can only be reconciled by a jump discontinuity in the contract[18].

As before we can easily check that $q(1, 0) = \phi(1) = 1/c$ which implies $u_q(q(1, 0), 1, 0) = 0$. Hence, $(1, 0)$ gets an efficient contract and indeed it is the only one to do so since returning to the original variables we have:

$$q_{SB}(\theta_1, \theta_2) = \begin{cases} 0 & \theta_1 \leq 1/2 \\ \frac{4\theta_1 - 2}{4\theta_2 + 2c - 1} & 1/2 \leq \frac{(2c-1)\theta_1 + 2\theta_2}{4\theta_2 + 2c - 1} \leq \frac{2c+1}{2c+3} \\ \frac{3\theta_1 - 1}{3\theta_2 + 2c} & \frac{2c+1}{2c+3} \leq \frac{2c\theta_1 + \theta_2}{2c + 3\theta_2} \leq 1 \end{cases}$$

$$q_{FB}(\theta_1.\theta_2) = \frac{\theta_1}{\theta_2 + c} \quad \forall \theta_1.\theta_2 \in [0, 1]$$

Observe that $q_{SB}(\theta_1, \theta_2) \leq q_{FB}(\theta_1, \theta_2)$ since we have:

$$\frac{3\theta_1 - 1}{3\theta_2 + 2c} \leq \frac{\theta_1}{\theta_2 + c} \Leftrightarrow (\theta_1 - 1)c - \theta_2 \leq 0$$

[16] This is because we are displacing now the sensitivity to price rather than the intercept. The maximum and minimum valuation for $q = 0$ remains the same as we vary c.

[17] Remember that implementability tells us that the contract always grows with marginal utility

[18] An analysis of the example when $0 < c \leq 1/2$ can be found on Araujo, Calagua and Vieira (2022) where general necessary conditions for optimality are derived for cases like this where discontinuities arise at the frontier of the participation region. Here, as noted above, we restrict attention to continuous contracts.

Which holds for all $\theta_1, \theta_2 \in [0, 1]$ and holds with equality only when $(\theta_1, \theta_2) = (1, 0)$.

Furthermore, we also have:

$$\frac{4\theta_1 - 2}{4\theta_2 + 2c - 1} \leq \frac{3\theta_1 - 1}{3\theta_2 + 2c} \Leftrightarrow \frac{(2c-1)\theta_1 + 2\theta_2}{4\theta_2 + 2c - 1} \leq \frac{2c+1}{2c+3}$$

Hence, $q_{SB}(\theta_1, \theta_2) \leq q_{FB}(\theta_1, \theta_2)$ everywhere. Observe that this inequality is an equality only at $(1, 0)$ where the contract is efficient or at $\theta_1 = 0$ where in both solutions we have non-participation.

V.a: Different preferences[19]

Consider now a simple perturbation of the basic linear case. The cost is still zero and consumers are still uniformly distributed but now the valuation function is:

$$u(q, \theta_1, \theta_2) = \theta_1 q - (\theta_2 + c)\frac{q^\alpha}{\alpha}, \quad c > 1/2, \quad \alpha \in (1, +\infty)$$

Observe that the inverse demand function is given by:

$$p = \theta_1 - (\theta_2 + c)q^{\alpha - 1}$$

The parameter α controls the convexity/concavity of the demand. If $\alpha = 1$ we would have perfectly elastic demand, typically graphed as an L. If $1 < \alpha < 2$ we have a convex demand which is increasingly less convex as $\alpha \to 2$[20]. If $\alpha = 2$ we return to the basic case of linear demands. Finally for $\alpha > 2$ we get a concave demand, increasingly more so as α grows. We can easily verify that assumptions A1-A3 are still satisfied.

Solving for $1/2 \leq r \leq 1$, this time we have $A(\phi(r), r, s) = [\phi(r)]^{\alpha-1}s + 2r - 1$ and solving the optimality condition gives us $r_0 = 3/4$ and

$$\phi(r) = \begin{cases} (\frac{8r-6}{2c-1})^{1/(\alpha-1)} & 3/4 \leq r \leq \frac{2c+2}{2c+3} \\ (\frac{3r-2}{c})^{1/(\alpha-1)} & \frac{2c+2}{2c+3} \leq r \leq 1 \end{cases}$$

The rol of c is exactly the same as before and we can also verify that $(1, 0)$ gets an efficient contract since $q(1, 0) = \phi(1) = (1/c)^{1/(\alpha-1)}$ implies $u_q(q(1, 0), 1, 0) = 0$.

Now for any $\alpha \in (1, +\infty)$ consider its solution q_α with level sets $\gamma_r^\alpha = \{(A_\alpha(\phi_\alpha(r), r, s), s) : 0 \leq s \leq U(\phi_\alpha(r), r)\}$ and values $\phi_\alpha(r)$. Observe that

$$\phi_\alpha = [\phi_2]^{1/(\alpha-1)} \Rightarrow A_\alpha = [\phi_\alpha]^{\alpha-1}s + 2r - 1 = \phi_2 s + 2r - 1 = A_2$$

Hence, γ_r^α does not change with α. Therefore, we have the same feature as in example IV: the optimal order of types is invariant by the degree of convexity/concavity of the demands.

[19] This example was first analyzed by Basov (2005). It is interesting to contrast how much more simpler our solution method is.

[20] Observe that the arrow pratt index this time is $-\frac{p''(q)}{p'(q)} = \frac{2-\alpha}{q}$ which decreases to zero as $\alpha \uparrow 2$

On the other hand when demands are concave/convex, ϕ_α is concave/convex and as $\alpha \downarrow 1$, ϕ_α becomes increasingly more convex. Therefore, optimality requires that as the demands become more convex ($\alpha \downarrow 1$) the rate at which the contract grows with marginal utility (i.e., ϕ'_α) must increase: that is the contract grows faster close to the point $(1,0)$ and slower near the frontier of the participation region. In more informal words, for two neighbouring types θ and $\hat{\theta}$, they are treated more similarly if they are "low"[21] than when they are "high". This is to be expected since convexity of the demand means that u_q is more sensitive to changes in q for low values of q rather than for high values q. Hence, the principal only needs to worsen the contract slightly to cause the desired effect[22] when dealing with low types while it needs to worsen the contract substantially when dealing with a high type to cause the same effect.

V.b: Different costs

Consider again types uniformly distributed on the square and keep the same valuation function as in the basic linear case but now consider the cases of:

- Linear cost $C(q) = \lambda q$, for $0 \leq \lambda \leq 1$
- Quadratic cost $C(q) = \lambda \frac{q^2}{2}$, for $\lambda \geq 0$

Since the characteristics are solely determined by the valuation function of the consumers, we still have $A(\phi(r), r, s) = \phi(r)s + 2r - 1$. The difference arises when we consider $c_q = \lambda$ or $c_q = \lambda q$ in the optimality condition.

For the linear cost the solution is given by:

$$\phi_L(r) = \begin{cases} \frac{8r-6-2\lambda}{2c-1} & \frac{3+\lambda}{4} \leq r \leq \frac{\lambda+2c+2}{2c+3} \\ \frac{3r-2-\lambda}{c} & \frac{\lambda+2c+2}{2c+3} \leq r \leq 1 \end{cases}$$

Observe how the constant marginal cost of the monopoly affects only the intercepts but not the slopes of the linear parts of ϕ. Indeed, as λ rises from 0 to 1 the monopoly excludes a bigger and bigger set of consumers until at $\lambda = 1$ full exclusion is achieved. This is because we are uniformly raising the cost of producing one unit independently of how many units are produced. For $\lambda \geq 1$ the marginal valuations of the consumers are too low to cover our constant marginal cost. We can also check that for this case $q(1,0) = \phi(1) = (1-\lambda)/c$, which implies $u_q(q(1,0), 1, 0) = \lambda$. Hence, type $(1,0)$ still gets an efficient contract.

For the quadratic cost the solution is given by:

$$\phi_Q(r) = \begin{cases} \frac{8r-6}{2\lambda+2c-1} & 3/4 \leq r \leq \frac{2\lambda+2c+2}{2\lambda+2c+3} \\ \frac{3r-2}{c+\lambda} & \frac{2\lambda+2c+2}{2\lambda+2c+3} \leq r \leq 1 \end{cases}$$

[21] By this we mean that they belong to some γ_r and $\gamma_{r'}$ where r and r' are low. Of course we assume that r and r' are arbitrarily close.

[22] The desired effect is to worsen the contract of lower types in order to make it unattractive to higher types. Starting at $(1,0)$ who gets an eficiente contract the principal moves progressively to lower types worsening their contracts.

In this case, the effect of the increase in λ is the opposite as in the previous case. As λ grows the participation set remains the same and only the slopes of the linear parts of ϕ are affected but they still pass through $(1/2, 0)$ and $(1/3, 0)$. Now there is no constant marginal cost and no matter how small is the valuation of a consumer, we could profitably serve those consumers with small quantities of q as long as the valuation is nonnegative. On the other hand, as $\lambda \uparrow$ the rate at which marginal cost grows increases and optimality requires that the rate at which the contract grows with marginal utility should decrease so that we sell high values of q only for the very high types. Observe that we still have $q(1,0) = \phi(1) = 1/(c+\lambda)$, which implies $u_q(q(1,0),1,0) = \frac{\lambda}{c+\lambda}$. Hence, we still achieve an efficient contract at the top of the distribution[23].

Hence, the introduction of a linear or quadratic cost function have very different effects on the shape of the optimal contract. A linear cost affects only the participation region and not the rate at which the contract grows with marginal utility and the quadratic cost affects this rate and not the participation region.

V.c: Different distributions

Consider again that the firm has zero cost and keep the same valuation function as in the basic linear case with $c = 1$ but now we want to consider the effects of having a distribution that makes higher values of one of the parameters of asymetric information more likely than lower values. Consider the cases where θ_1 and θ_2 are independently distributed with:

- θ_1 distributed according to the density $f(\theta_1) = 2\theta_1 \Rightarrow F(\theta_1) = \theta_1^2$ and θ_2 is distributed uniformly on $[0,1]$.
- θ_2 distributed according to the density $f(\theta_2) = 2\theta_2 \Rightarrow F(\theta_2) = \theta_2^2$ and θ_1 is distributed uniformly on $[0,1]$.

Again, we still have $A(\phi(r), r, s) = \phi(r)s + 2r - 1$ but now the difference comes from setting either $\rho(A,s) = 2A = 2\phi(r)s + 4r - 2$ and $F_1(A,s) = 1 - A^2 = 1 - (2r-1)^2 - \phi^2(r)s^2 - 2\phi(r)(2r-1)s$ or setting $\rho(A,s) = 2s$ and $F_1(A,s) = (1-A)2s = 4(1-r)s - 2\phi(r)s^2$ in the optimality condition.

For the nonuniform θ_1 the solution is given by:

$$\phi(r) = \begin{cases} \sqrt{\frac{9}{2}(2r-1)^2 - \frac{3}{2}} & \frac{1+\sqrt{3}}{2\sqrt{3}} \leq r \leq \frac{10+\sqrt{156}}{28} \\ \frac{8}{3}r - \frac{4}{3} - \frac{1}{3r} & \frac{10+\sqrt{156}}{28} \leq r \leq 1 \end{cases}$$

As usual, we can show that $(1,0)$ gets an efficient contract since $q(1,0) = \phi(1) = 1$ implies $u_q(q(1,0),1,0) = 0$. More importantly, the effect of introducing a distribution that makes types with higher θ_1 more likely is that the participation set shrinks excluding more types with comparatively lower values of θ_1. Aditionally, the contract ϕ is no longer linear but strictly concave. Concavity implies that if both θ and $\hat{\theta}$ are "high" they are treated more similarly than when they

[23] Observe that this is compatible with the slower growth of the contract since $\frac{8}{2\lambda+2c-1} > \frac{3}{\lambda+c}$ and even though both slopes decrease when λ increases, the region where the contract has a bigger slope also increases $\frac{d}{d\lambda}\left(\frac{2\lambda+2c+2}{2\lambda+2c+3}\right) > 0$

are "low". In particular, since the highest type $(1,0)$ gets an efficient contract and we know that the contract grows with marginal utility, types with high θ are given contracts closer to their efficient levels and types with low θ more subefficent contracts.

This is to be expected. The principal's problem arises from its inability to simultaneously maximize the social surplus and minimize the share of this surplus given to the consumers. Since consumers with a high valuation may always pretend to have a lower valuation to extract a bigger share of the social surplus, the principal must restrict himself to incentive compatible contracts. For the highest type, however, there is no need to introduce distortions in the contract, i.e. the maximization of social surplus becomes more important. The lower types, on the contrary, are frequently excluded from the market even when they offer the possibility of generating a positive social surplus[24]. Hence making types with higher θ_1 more likely gives incentive to the principal to offer those types contracts that are closer to their efficient level[25].

For the nonuniform θ_2 the solution is given by:

$$\phi(r) = \begin{cases} 12r - 9 & 3/4 \leq r \leq 11/14 \\ \frac{8}{3}r - \frac{5}{3} & 11/14 \leq r \leq 1 \end{cases}$$

As usual, we can show that since $q(1,0) = \phi(1) = 1$ implies $u_q(q(1,0),1,0) = 0$, the type $(1,0)$ gets an efficient contract. Unlike the previous case, the participation set here remains the same. This is because the shape of the participation region is not determined by the distribution of types (i.e., the shape of the characteristics an in particular γ_{r_0} is solely determined by the shaoe of the valuation function). Indeed, in this case the frontier of the participation region is given by $\gamma_{r_0} = \{(2r_0 - 1, s) : 0 \leq s \leq 1\}$ where $\phi(r_0) = 0$. The distribution affects the value of r_0 via the optimality condition but given the shape of the participation region we cannot force the participation region to focus on types with higher (or lower) values of θ_2, only of θ_1 (whose distribution hasnt change). What does change, however, is the rate at which the contract grows with marginal utility. It initially grows faster ($12 > 8$) for a shorter time ($11/14 < 4/5$) but for higher types grow slower ($8/3 < 3$). This means that now "high" types (i.e., those θ such that $\theta \in \gamma_r$ for some high r) have less subefficient contracts.

Appendix

Derivation of example I

In this case we have $u_{q\theta_1} = 1$, $u_{q\theta_2} = -1$, $\rho = 1$, $F_1(\theta_1,\theta_2) = 1-\theta_1$, $F_2(\theta_1,\theta_2) = 1-\theta_2$ and $G = (\theta_1 + 1 - \theta_2)q - \frac{q^2}{2} - (1-\theta_1)q \Rightarrow G_q = 2\theta_1 - \theta_2 - q$. From the optimality conditions we get:

[24] Of course, the reason they are excluded is because any contrat given to them would tempt the higher types out of their more efficient contracts.

[25] Observe that the complete information case may be seen as a limiting case where we have progressively move all the probability mass to the point $(1,0)$ and all the other types have probability zero. In this case, of course, the principal offers the efficient contract inside the participation region.

For the case $1/2 \leq r \leq 1$:

$$(4r-2-\phi)U + \frac{U^2}{2} = \int_0^{U(\phi(r),r)} 2(2r-1+s) - s - \phi ds = 0$$

and because the characteristics are $\gamma_r(s) = (2r-1+s, s)$, its clear by definition of U that $A(\phi, r, U) = 1$ and then $U = 2-2r$. Substituting on the optimality condition:

$$\phi(r) = 3r - 1, \quad 1/2 \leq r \leq 1$$

Returning to the original variables $(\theta_1, \theta_2) = (2r-1+s, s)$ we get $s = \theta_2$ and $r = \frac{\theta_1+1-\theta_2}{2}$ so the solution can be reexpressed as:

$$q(\theta_1, \theta_2) = \frac{3}{2}(\theta_1 + 1 - \theta_2) - 1, \quad 1 \leq \theta_1 + 1 - \theta_2 \leq 2$$

For the case $0 \leq r \leq 1/2$:

$$\frac{U^2}{2} - (1-2r+\phi)U = \int_0^{U(\phi(r),r)} 2s - (1-2r+s) - \phi ds = 1-2r,$$

and because the characteristics are $\gamma_r(s) = (s, 1-2r+s)$, its clear by definition of U that $B(\phi, r, U) = 1$ and then $U = 2r$. Substituting on the optimality condition and solving for ϕ:

$$\phi(r) = 3r - \frac{1}{2r}, \quad r_0 = \sqrt{1/6} \leq r \leq 1/2$$

Returning to the original variables we have $(\theta_1, \theta_2) = (s, 1-2r+s)$ and hence $r = \frac{\theta_1+1-\theta_2}{2}$ again. Putting it all together, we obtain the desired solution.

Proof of proposition 4.1

Proof. Consider a utility function satisfying GSC and take $q = \phi(r)$, $\theta = \gamma_r(t)$ and $\theta' = \gamma_r(0)$ then by GSC there is a $\lambda_{rt} > 0$ such that:

$$u_q(\phi(r), \gamma_r(t)) - u_q(\phi(r), \gamma_r(0)) = \lambda_{rt} \nabla_\theta u_q(\phi(r), \gamma_r(0)) \cdot (\gamma_r(t) - \gamma_r(0))$$

$$\int_0^t \nabla_\theta u_q(\phi(r), \gamma_r(s)) \cdot \dot{\gamma_r}(s) ds = \lambda_{rt} \nabla_\theta u_q(\phi(r), \gamma_r(0)) \cdot (\gamma_r(t) - \gamma_r(0))$$

But since we have:

$$\dot{\gamma_r}(s) = \begin{cases} (1, -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), \gamma_r(s))) : 0 \leq r \leq \frac{1}{2} \\ (-\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), \gamma_r(s)), 1) : \frac{1}{2} \leq r \leq 1 \end{cases}$$

Then, using $\lambda_{rs} > 0$, we get that $\nabla_\theta u_q(\phi(r), \gamma_r(s)) \cdot \dot{\gamma_r}(s) = 0$ implies

$$\nabla_\theta u_q(\phi(r), \gamma_r(0)) \cdot (\gamma_r(t) - \gamma_r(0)) = 0$$

Then making use of:

$$\gamma_r(t) - \gamma_r(0) = \begin{cases} (t, B(\phi, r, t) - (1-2r)) : 0 \le r \le \frac{1}{2} \\ (A(\phi, r, t) - (2r-1), t) : \frac{1}{2} \le r \le 1 \end{cases}$$

we get that for $0 \le r \le 1/2$:

$$u_{q\theta_1}(\phi(r), \gamma_r(0))t + u_{q\theta_2}(\phi(r), \gamma_r(0))[B(\phi, r, t) - (1-2r)] = 0$$

$$\Rightarrow B(\phi, r, t) = (1-2r) - \frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), 0, 1-2r)t$$

$$\Rightarrow -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), \gamma_r(t)) = \frac{d}{dt}B(\phi, r, t) = -\frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), 0, 1-2r)$$

and for $1/2 \le r \le 1$:

$$u_{q\theta_1}(\phi(r), \gamma_r(0))[A(\phi, r, t) - (2r-1)] + u_{q\theta_2}(\phi(r), \gamma_r(0))t = 0$$

$$\Rightarrow A(\phi, r, t) = (2r-1) - \frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), 2r-1, 0)t$$

$$\Rightarrow -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), \gamma_r(t)) = \frac{d}{dt}A(\phi, r, t) = -\frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), 2r-1, 0)$$

So the marginal rate of substitution does not depend on t and we have that the curves $\gamma_r(s)$ are straight lines:

$$\gamma_r(s) = \begin{cases} (s, 1-2r - \frac{u_{q\theta_1}}{u_{q\theta_2}}(\phi(r), 0, 1-2r)s) : 0 \le r \le \frac{1}{2} \\ (2r-1 - \frac{u_{q\theta_2}}{u_{q\theta_1}}(\phi(r), 2r-1, 0)s, s) : \frac{1}{2} \le r \le 1 \end{cases}$$

For the particular case when $u(q, \theta_1, \theta_2) = \theta_1 v_1(q) - \theta_2 v_2(q) + v_3(q)$ its enough to notice that $u_{q\theta_1}(\phi, \gamma_r(0)) = v_1'(\phi)$ and $u_{q\theta_2}(\phi, \gamma_r(0)) = -v_2'(\phi)$ □

Derivation of example II

We start solving for $1/2 \le r \le 1$. In this case we already found that the characteristics are given by

$$\Rightarrow \gamma_r(s) = (A(\phi(r), r, s), s) = (2r - 1 + \frac{v_2'(\phi)}{v_1'(\phi)}s, s) = (2r - 1 + (\phi(r)+1)s, s)$$

Next, we use the simplified optimality condition taking into account that for this example $v_1'(\phi) = \frac{1}{\phi+1}$, $v_3'(\phi) = c$, $c_q = 0$ $\rho = 1$, $F_1(A, s) = 1 - A = 2 - 2r - (\phi(r)+1)s$. After a straightforward calculation we get:

$$(\frac{4r-3}{\phi+1} + c)U + \frac{U^2}{2} = 0$$

Since $U(\phi, r) = sup\{s \in [0,1] : (A(\phi, r, s), s) \in [0,1]^2\}$, we have that either $U = 1$ or $A(\phi, r, U) = 1$[26]. Starting with this last case we see that $U = \frac{2-2r}{\phi(r)+1}$. Plugging this in the previous equation and solving for ϕ we get:

[26]Lets remember that the curve $\gamma_r(s) = (A(\phi, r, s), s)$ starts at the lower edge of the square $[0,1]^2$ and always travels northeast in the direction of the vector $(-u_{q\theta_2}, u_{q\theta_1})$.

$$\phi(r) = \frac{-3r+2}{c} - 1, \quad \frac{2-c}{3} \le r \le 1$$

Where the value $r_0 = \frac{2-c}{3}$ was chosen since $\phi(\frac{2-c}{3}) = 0$ and $U(0, \frac{2-c}{3}) = \frac{2}{3}(1+c) < 1$ which indicates that γ_{r_0} determines the frontier of the participation region and we do not need to analyze the case $U = 1$ since the exclusion set includes the upper edge of $[0,1]^2$. We also dont need to analyze the optimality conditions for $0 \le r \le 1/2$ since the exclusion area contains the western edge of $[0,1]^2$.

Derivation of example III

We start solving for $1/2 \le r \le 1$. In this case we already found that the characteristics are given by

$$\Rightarrow \gamma_r(s) = (A(\phi(r), r, s), s) = (2r - 1 + \phi(r)s, s)$$

Next, we use the simplified optimality condition using $\rho = 1$, $F_1(A,s) = 1 - A$, $v_1'(\phi) = 1$, $v_3'(\phi) = c - \phi$ and $c_q = 0$ to get:

$$\int_0^{U(\phi(r),r)} [(2r-1) + c - \phi - (2 - 2r - \phi s)]ds = 0$$

$$(4r - 3 + c - \phi)U + \phi \frac{U^2}{2} = 0$$

From here either $U = 1$ or $A(\phi, r, U) = 1$. Starting with the last case we get $U = \frac{2-2r}{\phi}$ which after substitution in the optimality condition give us:

$$\phi(r) = 3r - 2 + c$$

This solution holds for $1/2 \le r \le 1$ such that $\phi \ge 0 \Leftrightarrow r \ge \frac{2-c}{3}$ and $U \le 1 \Leftrightarrow \frac{4-c}{5} \le r$. Since $-1 \le c \Rightarrow \frac{2-c}{3} \le \frac{4-c}{5}$ the solution holds on $[\max\{1/2, \frac{4-c}{5}\}, 1]$

On the other hand $U = 1$ give us from the optimality condition:

$$\phi(r) = 8r - 6 + 2c$$

which holds on $1/2 \le r \le 1$ such that $\phi \ge 0 \Leftrightarrow r \ge \frac{3-c}{4}$ and $A(\phi, r, 1) \le 1 \Leftrightarrow r \le \frac{4-c}{5}$. Hence the domain of this solution is $[\max\{1/2, \frac{3-c}{4}\}, \frac{4-c}{5}]$.

Now lets look at $0 \le r \le 1/2$. In this case we already found that the characteristics are given by

$$\Rightarrow \gamma_r(s) = (s, B(\phi(r), r, s)) = (s, 1 - 2r + \frac{1}{\phi(r)}s)$$

Using $\rho = 1$, $F_1(s,B) = 1 - s$, $F_2(s,B) = 1 - B$, $v_1'(\phi) = 1$, $v_2'(\phi) = \phi$, $v_3'(\phi) = c - \phi$ and $c_q = 0$ on the optimality condition give us:

$$\int_0^{U(\phi(r),r)} [-(1-2r)\phi + c - \phi - (1-s)]ds = \phi[1-2r]$$

$$[-(2-2r)\phi + c - 1]U + \frac{U^2}{2} = \phi(1-2r)$$

There is again, two possibilities $U = 1$ or $B = 1$. In the first case, we get

$$\phi(r) = \frac{2c-1}{6-8r}$$

which holds on $0 \leq r \leq 1/2$ such that $\phi \geq 0 \Leftrightarrow c \geq 1/2$ and $B(\phi, r, 1) \leq 1 \Leftrightarrow \frac{3}{3+2c} \leq r$. Hence the domain is $[\frac{3}{3+2c}, 1/2]$ as long as $c \geq 3/2$ (otherwise is empty).

On the other hand when $B(\phi, r, U) = 1$ we have $U = 2r\phi$ which on the optimality condition give us:

$$\phi(r) = \frac{1-2rc}{6r^2-4r}$$

This holds on $0 \leq r \leq 1/2$ such that $\phi \geq 0 \Leftrightarrow r \geq \frac{1}{2c}$ (provided $c > 0$) and $U \leq 1 \Leftrightarrow r \leq \frac{3}{3+2c}$. Hence the domain is $[\frac{1}{2c}, 1/2]$ on $1 < c \leq 3/2$, $[\frac{1}{2c}, \frac{3}{3+2c}]$ on $c > 3/2$ and empty otherwise. Unlike previous results, it is less evident that ϕ is nondecreasing however this holds for $c > 3/4$ because we have

$$\phi'(r) = \frac{3r^2c - 3r + 1}{(3r^2-2r)^2} > 0, \quad r \geq 1/2c$$

Since $g(r) = 3r^2c - 3r + 1$ is a strictly convex function with $g'(1/2c) = 0$ and $g(1/2c) = \frac{4c-3}{4c} \geq 0 \Leftrightarrow c \geq 3/4$. Putting all together we get that the optimal ϕ varies with c and is given by the formulas shown on the main text.

Derivation of example IV

We start solving for $1/2 \leq r \leq 1$. We already found the characteristics which in this case are:

$$\Rightarrow A(\phi(r), r, s) = 2r - 1 + \frac{v_2'(\phi)}{v_1'(\phi)}s = 2r - 1 + e^{\alpha\phi(r)}s$$

Using $v_1'(\phi) = e^{-\alpha\phi}$, $v_3'(\phi) = c$, $c_q = 0$ $\rho = 1$, $F_1(A, s) = 1 - A = 2 - 2r - e^{\alpha\phi(r)}s$ in the optimality condition we get:

$$[(4r-3)e^{-\alpha\phi} + c]U + \frac{1}{2}U^2 = 0$$

Assuming first $A(\phi, r, U) = 1$ we get $U = \frac{2(1-r)}{e^{\alpha\phi}}$ and end up with the solution:

$$\phi(r) = \frac{1}{\alpha}log(\frac{-3r+2}{c}), \quad \frac{2-c}{3} \leq r \leq 1$$

Where $r_0 = \frac{2-c}{3}$ was chosen so that $\phi(r_0) = 0$ and since $U(0, r_0) = \frac{2}{3}(1+c) < 1$, γ_{r_0} determines the participation region and we dont need to check the case $U = 1$ or $0 \leq r \leq 1/2$.

Chapter 5

Conclusions

In this work we have showed that conditions A2 (having each dimension of private information be ordered in terms of marginal utility) and A3 (allowing for an appropriate sustitubility of different dimensions on the marginal utility) work together as a Bidimensional Spence-Mirrless. That is, local incentive compatibility implies global incentive compatibility and the type space is endogenously oredered according to marginal valuation. Hence, bidimensional screening problems satisfying this conditions are qualitatively very similar to unidimensional models. The main difference being that now, in general, the order is not exogenous. We have given for this class of models a general characterization of implementability and a simple algorithm to compute solutions.

There is of course still much to be studied in regards to multidimensional screening problem where the instrument's dimension is smaller than the dimension of the private information. Basov (2001) and Araujo et al. (2022) have provided examples where the optimal mechanisms are discontinuous. This seems to be a peculiarity of this kind of screening problems that we have bypassed by restricting ourselves to continuous contracts. However, the wide variety of examples considered here and in Araujo et al (2022) seems to suggest that this discontinuous solution are not generic as Basov (2001) hypothesised but rather special phenomenon. For example, in the basic linear case the discontinuities seem to arise as a consequence of pushing the support of the distribution of the slopes of the demands too far into regions of greater elasticity.

We have also showed in example III that the issue of genericity of exclusion can be studied by means of particular examples and contrary to established wisdom optimal multidimensional contracts need not exclude a set of consumers with positive measure. Example IV and example V.a seem to imply that the optimal ordering of types, although endogenous, is invariant with respect to the degree of convexity of the demand. Another possible generalization of our work would be to allow for more general participation constraints by relaxing A1 while keeping A2 and A3. This would allow us to give a solution to the regulation model posed by Lewis and Sappington (1988). Finally, we believe that it could be possible to generalize our characterization of implementability for an arbitrary dimension of the type space while the instrument remains unidimensional. For example, we hypothesized that for $\Theta = [0,1]^n$ its enough to consider the assumptions[1]

A1. $u_{\theta_i} > 0$ for all $i = 1, ..., n$.

A2. $u_{q\theta_i} > 0$ for all $i = 1, ..., n$

A3. $\frac{d}{dq}(\frac{u_{q\theta_{i+1}}}{u_{q\theta_i}}) \geq 0$ for all $1 = 1, ..., n-1$.

where A2 guarantees that each dimension is positively ordered according to marginal utility and A3 establishes a hierarchy of dimensions that allows an appropriate substitubility betwen them: $j > i \Rightarrow \frac{d}{dq}(\frac{u_{q\theta_j}}{u_{q\theta_i}}) \geq 0$.

[1] Proposition 3.1, 3.3 as well as the analog of corollary 2 are inmediately generalized for an arbitrary n. It is proposition 3.4 and 3.5 that offer some obstacles for their generalization.

References

[1] ARAUJO A, MOREIRA H (2010) Adverse selection problems without the Spence-Mirrless condition. Journal of Economic Theory 145(3):1113-1141

[2] ARAUJO A, MOREIRA H, VIEIRA S (2015) The marginal tariff approach without single-crossing. Journal of Mathematical Economics 61:166-184

[3] ARAUJO A, CALAGUA B, VIEIRA S (2022) A necessary optimality condition in two-dimensional screening. Economic Theory 73:781-806

[4] ARMSTRONG M (1996) Multiproduct nonlinear pricing. Econometrica 64(1):51-75

[5] ARMSTRONG M (1999) Optimal regulation with Unknown Demand and Cost Functions. Journal of Economic Theory 84(2):196-215

[6] BASOV S (2001) Hamiltonian approach to multi-dimensional screening. Journal of Mathematical Economics 36(1): 77-94

[7] BASOV S (2005) Multidimensional screening. Springer-Verlag

[8] BARON D, MYERSON R (1982) Regulating a Monopolist with Unknown Costs. Econometrica 50(4):911-930

[9] CARLIER G (2001) A general existence result for the principal-agent problem with adverse selection. Journal of Mathematical Economics 35(1): 129-150

[10] CHIAPPORI PA, MCCANN R, PASS B (2016) Multidimensional Matching. https://doi.org/10.48550/arXiv.1604.

[11] DASKALAKIS C, DECKELBAUM A, TZAMOS C (2017) Strong duality for a multiple good monopolist. Econometrica 85(3):735-767.

[12] DENECKERE R, SEVERINOV S (2017) A solution to a class of multi-dimensional screening problems: isoquants and clustering. Mimeo.

[13] EKELAND I (2010) Notes on Optimal Transportation. Economic Theory 42(2):437-459

[14] EVANS LC (1998) Partial Differential Equations, Graduate studies in mathematics, vol 19. American Mathematical Society.

[15] FIGALLI A, Y.-H KIM, R.J. McCann (2011) When is Multidimensional Screening a Convex Program? Journal of Economic Theory 146(2):454-478

[16] GUESNERIE R. LAFFONT JJ (1984) A Complete Solution to a Class of PrincipalAgent Problems with an Applications to the Control of a Self-Managed Firm. Journal of Public Economics 25(3):329-369

[17] JOHN F (1981) Partial differential equations. Springer Verlag.

[18] KLEINER A , MANELLI A (2019) Strong duality in monopoly pricing. Econometrica 87(4):1391-1396

[19] LAFFONT JJ, MASKIN E, ROCHET JC (1987) Optimal nonlinear pricing with two dimensional characteristics. In: Groves T, Radner R, Reiter S (eds) Information, Incentives and Economic mechanisms: Essays in honor of Leonid Hurwicz, University of Minessota Press, pp 256-266

[20] LEWIS T, SAPPINGTON D (1988) Regulating a Monopolist with Unknown Demand and Cost Functions. Rand Journal of Economics 19(3):438-457

[21] MANELLI A, VINCENT DR (2006) Bundling as an Optimal Selling Mechanism for a Multiple-Good Monopolist. Journal of Economic Theory 127(1):1-35

[22] MANELLI A, VINCENT DR (2007) Multidimensional Mechanism Design: Revenue Maximixation and the Multiple-Good Monopoly. Journal of Economic Theory 137(1):153-185

[23] MCAFEE RP, MCMILLAN J (1988) Multidimensional incentive compatibility and mechanism design. Journal of Economic Theory 46(2):335-354

[24] MILGROM P, SEGAL I (2002) Envelope theorems for arbitrary choice sets. Econometrica 70(2):583-601

[25] MIRRLEES J (1971) An Exploration in the Theory of Optimum Income Taxation. Review of Economic Studies 38(114):175-208

[26] MUSSA M, ROSEN S (1978) Monopoly and product quality. Journal of Economic Theory 18(2):301-317

[27] ROCHET JC (1985) The taxation principle and multi-time Hamilton-Jacobi equations. Journal of Mathematical Economics 14(2):113-128

[28] ROCHET JC (1987) A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context. Journal of Mathematical Economics 16(2):191-200

[29] ROCHET JC, CHONE P (1998) Ironing, sweeping and multidimensional screening. Econometrica 66(4):783-826

[30] ROCHET JC, STOLE LA (2003) The economics of multidimensional screening. In: Dewatripont M, Hansen LP, Turnovsky SJ (eds) Advances in economics and econometrics. Theory and applications, Eight world congress, Volume I, Cambridge University Press, pp 150-197

[31] SALANIE B (1997) The economics of contracts. MIT press

[32] SCHOTTMULLER C (2015) Adverse selection without single crossing: Monotone solutions. Journal of Economic Theory 158(A):127-164,

[33] WILSON RB (1993) Nonlinear pricing. Oxford University Press