Reducing incentive constraints in bidimensional screening

Calagua, Braulio

Instituto de Matemática Pura e Aplicada

2018

Reducing incentive constraints in bidimensional screening *

Braulio Calagua[†]

PUCP, Av. Universitaria 1801, Lima, Perú

UDEP, Calle Mártir Olaya 162, Miraflores, Lima, Perú

July 23, 2020

Abstract

This paper studies screening problems with quasilinear preferences, where agents' private information is two-dimensional and the allocation instrument is one-dimensional. A pre-order in the set of types is defined comparing types by their marginal valuation for the instrument, which allows reducing the incentive compatibility constraints that must be checked. With this approach, the discretized problem becomes computationally tractable. As an application, it is numerically solved an example from Lewis and Sappington [Lewis, T. and Sappington, D. E., 1988. Regulating a monopolist with unknown demand and cost functions. The RAND Journal of Economics, 438-457].

Keywords: two-dimensional screening, Spence-Mirrlees condition, incentive compatibility, regulation of a monopoly.

JEL Classification: D82, L51, C69.

---

*I am grateful to Aloisio Araujo, Alexandre Belloni, Carolina Parra, and Mauricio Villalba for their helpful comments. This research was developed during my PhD stage at IMPA, and supported by Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq).

[†] *E-mail addresses:* jcalagua@pucp.edu.pe

1 Introduction

Although the importance of multi-dimensions for modeling agents' characteristics in adverse selection problems is recognized, there are few models analyzed in the literature even when the agents' characteristics are two-dimensional continuously distributed and the principal's instrument is one-dimensional. This is because of the inherent difficulty of the problem for finding explicit solutions and to the challenging task of obtaining numerical approximations.

The purpose of this article is to provide a methodology that can be used for numerically solving a wide variety of such adverse selection problems. The main assumption is that the agent's marginal valuation can be ranked respect to each one of the private information parameters.

In contrast to most of the numerical approaches to approximate the solution, this study does not consider the incomplete or relax problem in which the constraints are obtained from the necessary conditions of the original problem (Wilson (1996), Tarkiainen and Tuomala (1999)), neither focus on the Lagrange multipliers associated with the incentive compatibility constraints (Berg and Ehtamo (2009)), nor it appeals to optimization methods for the complete problem (Judd et al. (2018)). It explores the idea −from the unidimensional case− that a priori eliminates some constraints when the set of types is finite and the Spence-Mirrlees condition holds.

In the multidimensional case, it is well known that the lack of an exogenous order among types makes the problem difficult to solve. That is, there is not a satisfactory generalization of the Spence-Mirrlees condition[1]. One of the novelties of this article is the introduction of a pre-order in the set of types by comparing the marginal valuation for the instrument. With this definition, it is proved that it is sufficient to consider that each type does not want to imitate any type on a subset of the boundary of the type space. As a consequence, a significant number of incentive constraints are ruled out in the discretized problem, thus making it computationally tractable for a relatively fine discretization.

With this methodology, it is numerically solved the regulation model introduced by Lewis and Sappington (1988b) and then reviewed by Armstrong (1999), who showed that Lewis and Sappington's solution was incor-

[1] McAfee and McMillan (1988) proposed a generalization of the single-crossing condition according to which bunching is linear.

rect. Since this is a model with an unknown analytical solution, it might be meaningful to know the numerical solution. Besides, Armstrong (1999) has conjectured that it is optimal to exclude a positive mass of agents, as in the non-linear pricing setting. However, the numerical solution suggests that the exclusion should not be optimal in this case.

The plan of the paper is as follows. The model is described (in the style of Mussa and Rosen (1978)) in Section 2. In Section 3 the endogenously determined isoquants are analyzed through a partial differential equation derived from the incentive compatibility constraints[2]. Section 4 is dedicated to explaining the reduction of incentive constraints, establishes the discretized problem, and study asymptotic properties when discretization becomes finer. The regulation model of Lewis and Sappington (1988b) is numerically solved −for particular parameters− in Section 5, and I provide some considerations about the optimality of exclusion. In Appendix A the method is tested by comparing the numerical solutions with the analytic solutions for some models from the literature. All proofs are relegated to Appendix B.

2 Model

Consider a monopolistic firm producing a single product $q \in \mathbb{R}_+$ at cost $C(q)$. Customers' characteristics, reflecting their preferences over the products, are captured by a bidimensional vector $(a, b) \in [0, 1] \times [0, 1]$ which is labeled as its *type*. This type is private information for each customer, but the firm knows the probability distribution over $[0, 1]^2$ according to a differentiable density function $\rho(a, b) > 0$. The utility of customer's type (a, b) is quasi-linear $v(q, a, b) - t$ where $v(q, a, b)$ is the value for consumption $q \in \mathbb{R}_+$, and $t \in \mathbb{R}_+$ is the monetary transfer.

The firm designs a menu of options to offer to the agent specifying the quantity and the corresponding payment according with customers' type revealed. By the revelation principle (Myerson (1979)), it is sufficient to restrict attention to contracts where true-telling is the best response for customers. Thus, in order to maximize expected revenue, the monopolist's problem is

$$\max_{q(\cdot),t(\cdot)} \int_0^1 \int_0^1 \big(t(a, b) - C(q(a, b))\big)\rho(a, b)da\, db \tag{MP}$$

[2]This analysis is presented with more detail in Araujo et al. (2019).

subject to

$$v(q(a,b),a,b) - t(a,b) \geq v(q(a',b'),a,b) - t(a',b') \tag{IC}$$

and

$$v(q(a,b),a,b) - t(a,b) \geq 0 \tag{IR}$$

Labels (IC) and (IR) refers to *incentive compatibility* and *individual rationality* constraints. It is assumed that reservation utility is type independent and normalized at zero. For an incentive compatible contract $(q(\cdot), t(\cdot))$ the informational rent is defined as

$$V(a,b) = v(q(a,b),a,b) - t(a,b)$$

This variable V is used to eliminate monetary transfer. The monopolist' problem can now be set as

$$\max_{q(\cdot),V(\cdot)} \int_0^1 \int_0^1 \big(v(q(a,b),a,b) - C(q(a,b)) - V(a,b)\big)\rho(a,b)da\, db \tag{1}$$

subject to

$$\begin{array}{lll} \text{(IC)} & V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) & \forall\ (a,b),(\widehat{a},\widehat{b}) \in [0,1]^2 \\ \text{(IR)} & V(a,b) \geq 0 & \forall\ (a,b) \in [0,1]^2 \end{array}$$

Assumptions. *Agent's valuation function v is three times differentiable, and cost function C is differentiable. Also, it is verified:*

1. *$v(q^{\text{out}},a,b)$ is constant.*
2. *$v_{qa} > 0$ and $v_{qb} < 0$.*

Assumption 1 is usually presented as $v(0,a,b) = 0$ because in nonlinear pricing the exit option is $q^{\text{out}} = 0$, and any agent assigns it zero value. However, in other adverse selection problems q^{out} could be endogenously determined. Assumption 2 is the single-crossing condition in each direction, i.e., agent's marginal valuation can be ranked respect to each one of the private information parameters. It has been assumed those particular signs, but what really matters is the uniform signs of v_{qa} and v_{qb}. As a consequence, it requires that an implementable $q(a,b)$ be non-decreasing with respect to a and non-increasing with respect to b.

If, additionally, the following condition is satisfied:

$$v_a \geq 0 \quad \text{and} \quad v_b \leq 0 \tag{*}$$

then the informational rent V is non-decreasing with respect to a and non-increasing with respect to b (since $V(a,b)$ is the optimal value of agent's maximization problem, by the Envelope Theorem $V_a(a,b) = v_a(q(a,b),a,b)$ and $V_b(a,b) = v_b(q(a,b),a,b)$, hence $V_a \geq 0$ and $V_b \leq 0$). Thus, it will be sufficient to impose $V(0,1) = 0$ and all the IR constraints will be fulfilled.

Unless otherwise stated, condition (*) will not be treated as an assumption for our model because the main issue is about IC constraints.

From now on, let us restrict attention on piecewise twice-differentiable and continuous contracts (q,t).

3 Characteristic Curves

In this section, it is summarized the methodology used by Araujo et al. (2019) to derive a partial differential equation which determines iso-quantity curves.

Let (q,t) be an incentive compatible contract twice-differentiable at (a,b). Then, (a,b) must solve the problem

$$\max_{(a',b')\in[0,1]^2} \{v(q(a',b'),a,b) - t(a',b')\} \tag{2}$$

The first-order necessary optimality conditions for problem (2) are

$$\begin{aligned} v_q(q(a,b),a,b)q_a(a,b) &= t_a(a,b) \\ v_q(q(a,b),a,b)q_b(a,b) &= t_b(a,b) \end{aligned} \tag{3}$$

From the equations in (3), the cross derivatives t_{ab} and t_{ba} can be calculated. Finally, by using Schwarz's integrability condition $t_{ab}(a,b) = t_{ba}(a,b)$, the following quasi-linear PDE is derived

$$-\frac{v_{qb}}{v_{qa}} q_a + q_b = 0 \qquad \text{a.e. in } [0,1]^2 \tag{4}$$

Define $\Gamma_0 = \{(r,0) : r \in [0,1]\}$, let ϕ be a function defined over $[0,1]$, and consider the initial value problem

$$\begin{aligned} &-\frac{v_{qb}}{v_{qa}} q_a + q_b = 0 \\ &q|_{\Gamma_0} = \phi(r) \end{aligned} \tag{5}$$

Following the method of characteristic curves[3]there is a family of plane characteristics curves $(a(r,s), b(r,s))$ defined by:

$$\begin{aligned} a_s(r,s) &= -\frac{v_{qb}}{v_{qa}}(\phi(r), a(r,s), b(r,s)) \quad , \quad && a(r,0) = r \\ b_s(r,s) &= 1 \quad , \quad && b(r,0) = 0 \end{aligned} \tag{6}$$

That is, for a fix $r \in [0,1]$ the characteristic curve $(a(r,s), b(r,s))$ parametrized by $s \in [0, \overline{s}(r)]$ is an isoquant of $q(\cdot)$ at level $\phi(r)$. Such a function ϕ must be optimally determined, which endogenously defines isoquants of $q(\cdot)$. This idea was developed in Araujo et al. (2019).

Note that characteristic curves are strictly increasing because, by Assumption 2, both entries of the vector tangent (a_s, b_s) are positive.

4 Reduction of Incentive Constraints

When numerically solving the problem, the main difficulty is related to the number of constraints. This is because after discretizing the type set $[0,1]^2$ into a grid of n points over each axis, there are $n^4 - n^2$ IC constraints. Therefore, fine discretizations result in memory storage problems. Next, it is presented a methodology that allows us to reduce the number of IC constraints. It is inspired by the ideas to address IC constraints in the unidimensional case with a finite type set when single-crossing holds (see Laffont and Martimort (2001)).

In bidimensional models, there is not a condition similar to the single-crossing in the unidimensional case where all types can be exogenously ordered by their marginal valuation for consumption. This is because in one-dimension $v_{q\theta} > 0$ is equivalent to $\theta_1 < \theta_2 \implies v_q(q,\theta_1) < v_q(q,\theta_2) \ , \forall q \in Q$. Then, to be able to compare a priori two different types at least partially, it is introduced the following binary relation:

Definition 1. *Given $(a,b), (\widehat{a}, \widehat{b}) \in [0,1]^2$, we say that (a,b) is* worse than *$(\widehat{a}, \widehat{b})$, denoted by $(a,b) \preceq (\widehat{a}, \widehat{b})$, if and only if*

$$v_q(q,a,b) \leq v_q(q,\widehat{a},\widehat{b}) \quad \forall\, q \in Q$$

[3]See Evans (1998) for a description of the method.

Note that $\preceq$ is a pre-order (reflexive and transitive) on $[0,1]^2$. This definition try to capture the idea that when $(a,b) \preceq (\widehat{a},\widehat{b})$, the (a,b)-agent has no incentive to announce the type $(\widehat{a},\widehat{b})$ because for any $q \in Q$, the $(\widehat{a},\widehat{b})-$agent has greater marginal valuation for consumption and is willing to pay more for each additional unit of the product.

As a direct consequence of Assumption 2, type (a,b) is worse than any type in the southeast.

Proposition 4.1. *Given (a,b), if $\widehat{a} > a$ and $\widehat{b} < b$, then $(a,b) \preceq (\widehat{a},\widehat{b})$*

By fixing type (a,b), the IC constraints with any type in the southeast are a priori excluded, as the difficulty comes from better agents willing to claim that they are worse agents rather than the reverse. Specifically, given $(a,b) \in [0,1]^2$, the following IC constraints are omitted:

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) \quad \forall\, \widehat{a} > a\ ,\ \widehat{b} < b \tag{7}$$

Next proposition shows that these constraints are indeed fulfilled when (q,V) satisfies the necessary conditions related to the envelope theorem and the monotonicity of $q(\cdot,\cdot)$ over each axis.

Proposition 4.2. *Assume (q,V) is such that*

$$V_a(a,b) = v_a(q(a,b),a,b)\ ,\ V_b(a,b) = v_b(q(a,b),a,b)\ ,\ q_a \geq 0\ ,\ q_b \leq 0$$

By fixing (a,b), the constraints given in (7) are satisfied.

Denote by $CC(\widehat{a},\widehat{b})$ the planar characteristic curve that contains $(\widehat{a},\widehat{b})$. Additionally, the expression "(a,b) is IC with $(\widehat{a},\widehat{b})$" means that

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q(\widehat{a},\widehat{b}),a,b) - v(q(\widehat{a},\widehat{b}),\widehat{a},\widehat{b})$$

The following proposition shows that it is sufficient to verify the IC constraint with a representative type over each characteristic curve.

Proposition 4.3. *Let $(a,b),(\widehat{a},\widehat{b})$ be such that (a,b) is IC with $(\widehat{a},\widehat{b})$. Then (a,b) is IC with (x,y) , $\forall\ (x,y) \in CC(\widehat{a},\widehat{b})$*

As a consequence, we can focus on the border of $[0,1]^2$. The following proposition is key to reduce the restrictions.

Proposition 4.4. *Let $(x, y), (\widehat{a}, \widehat{b})$, and (a, b) be such that (a, b) is IC with $(\widehat{a}, \widehat{b})$ and $(\widehat{a}, \widehat{b})$ is IC with (x, y). If $(\widehat{a}, \widehat{b}) \preceq (a, b)$ and $q(x, y) \leq q(\widehat{a}, \widehat{b})$, then (a, b) is IC with (x, y).*

Due to the kind of transitivity shown in Proposition 4.4, it is not necessary that type (a, b) verifies the IC constraints with all the types (x, y) on the left of a certain characteristic curve. Indeed, it is sufficient to verify the IC constraint with any type worse than (a, b) over such a curve, but making sure that this type verifies the IC constraints with all of those (x, y).

By taking the characteristic curve as close as possible to type (a, b), most restrictions can be eliminated. Since the characteristic curves are endogenously determined but any of them passing through (a, b) intersects the border of the square $[0, 1]^2$ on the northeast of that point, previous propositions suggest that it would be sufficient to verify that (a, b) is IC with all the points over the set

$$F^{(a,b)} := \{(s, 1) \mid a \leq s \leq 1\} \cup \{(1, s) \mid b < s \leq 1\} \tag{8}$$

which is formalized it in the following theorem.

Theorem 4.1. *Assume $\frac{d}{dq}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ and $\frac{d}{da}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$. Let (q, V) be such that*

$$\forall\ (a, b) \in [0, 1]^2\ ,\ (a, b) \text{ is IC with } (x, y)\ \ \forall\ (x, y) \in F^{(a,b)}$$

Then, (q, V) satisfies all the incentive compatibility constraints.

Technical assumptions $\frac{d}{dq}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ and $\frac{d}{da}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ are given to avoid pathological cases. This result could be understood as analogous to the claim *local IC constraints implies global IC constraints*, which is right in the unidimensional case when single-crossing holds.

4.1 Particular valuation function

The following proposition could allow us to reduce even more the IC constraints when the valuation function v has a special structure.

Proposition 4.5. *Assume that v_q is concave in a and convex in b. Let (a, b) and $(\widehat{a}, \widehat{b})$ be in $[0, 1]^2$ with $a < \widehat{a}$, $b < \widehat{b}$. We have*

1. *if* $(a,b) \preceq (\widehat{a},\widehat{b})$ *then* $\dfrac{\widehat{b}-b}{\widehat{a}-a} \leq \dfrac{-v_{qa}(q,a,\widehat{b})}{v_{qb}(q,a,\widehat{b})}$

2. *if* $\dfrac{\widehat{b}-b}{\widehat{a}-a} \leq \dfrac{-v_{qa}(q,\widehat{a},b)}{v_{qb}(q,\widehat{a},b)}$ *then* $(a,b) \preceq (\widehat{a},\widehat{b})$

This proposition says that, in order to $(a,b) \preceq (\widehat{a},\widehat{b})$, it is necessary that $CC(a,b)$ be on the left of $CC(\widehat{a},\widehat{b})$, because at the point $(a,\widehat{b})$ the slope of $CC(a,\widehat{b})$ is greater than the slope between (a,b) and $(\widehat{a},\widehat{b})$.

Besides, a sufficient condition for $(a,b) \preceq (\widehat{a},\widehat{b})$ is that the slope of $CC(\widehat{a},b)$ at the point $(\widehat{a},b)$ be greater than the slope between (a,b) and $(\widehat{a},\widehat{b})$. Thus, $CC(\widehat{a},\widehat{b})$ will be at the right of $CC(a,b)$.

4.2 Discretized problem

By Theorem 4.1, it is sufficient that each point satisfies the IC constraints with all points over a unidimensional set instead of the whole square. Now, the solution of the continuous problem can be approximated by discretizing the set of types. This section is devoted to establishing this discrete problem and discussing its limitations.

Let $X_n = \{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, 1\} \times \{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, 1\}$ be the grid of n^2 points on $[0,1]^2$. For a fixed (a,b) with $a<1$ and $b<1$, let $\widetilde{F}^{(a,b)} := F^{(a,b)} \cap X_n$, where $F^{(a,b)}$ is defined in (8). Because for points over the line $x=1$ or $y=1$ the constraints with the points on the northeast cannot be written, it is equivalently considered

$$\widetilde{F}^{(a,1)} = \big(\{(0,s) : 0 \leq s \leq 1\} \cup \{(s,0) : 0 \leq s < a\}\big) \cap X_n$$

$$\widetilde{F}^{(1,b)} = \big(\{(0,s) : 0 \leq s \leq b\} \cup \{(s,0) : 0 \leq s < 1\}\big) \cap X_n$$

Thus, $\widetilde{F}^{(a,b)}$ is the set of types with which (a,b) must satisfy an IC constraint.

The integral in the monopolist's objective it is approximated by the trapezoidal rule. Thus, consider the associated weights $w(i,j)$ for each point $(a_i,b_j) \in X_n$. Also, denote $q_{i,j} = q(a_i,b_j)$ and $V_{i,j} = V(a_i,b_j)$. The discretized problem we are interested in solving is the following:

$$
\begin{aligned}
&\max_{\{q_{i,j},V_{i,j}\}} \sum_{i=1}^{n}\sum_{j=1}^{n} w(i,j)(v(q_{i,j},a_i,b_j) - V_{i,j} - C(q_{i,j}))\rho(a_i,b_j) \\
&\text{subject to} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (9)\\
&\qquad \text{(IC)} \quad (a_i,b_j) \text{ is IC with } (\widehat{a}_i,\widehat{b}_j)\ ,\ \forall\ (\widehat{a}_i,\widehat{b}_j) \in \widetilde{F}^{(a_i,b_j)} \\
&\qquad \text{(IR)} \quad V_{i,j} \geq 0
\end{aligned}
$$

Remarks:

1. In the original discretized problem, there are $n^4 - n^2$ (maybe nonlinear) IC constraints. After our methodology, the number of IC constraints is of order n^3.

2. In case condition (*) is verified, all IR constraints can be replaced by $V_{1,n} = 0$.

3. In order to obtain better accuracy of the solution, it can be considered the monotonicity constraints $q_{i,j} \leq q_{i+1,j}$ and $q_{i,j} \leq q_{i,j-1}$. These $2n^2$ linear restrictions do not represent large numerical costs.

4. When the valuation function has the special *multiplicative separable* form $v(q,a,b) = \psi(q) + \alpha(a,b)q + \beta(a,b)$, the IC constraints become linear in $q_{i,j}$. Therefore, since the IC constraints are linear in $V_{i,j}$ (regardless of v), if the objective function is strictly concave[4], there exists a unique solution and we can rely on numerical approximations.

Due to discretization, it is impossible to ensure that for each type $(a,b) \in X_n$ all IC constraints are fulfilled. This is because there could be some points between $CC(a,b)$ and the closer to the left characteristic curve intersecting $\widetilde{F}^{(a,b)}$. Figure 1 illustrates this issue.

Note that the premises of Proposition 4.4 may not be true for types at the right of $CC(a,b)$ and violations may propagate. Nevertheless, following Belloni et al. (2010) it is proved that the violations of the IC constraints (i.e., the terms for which these constraints are not satisfied) uniformly converge

[4]This will be the case if $\psi'' - C'' < 0$

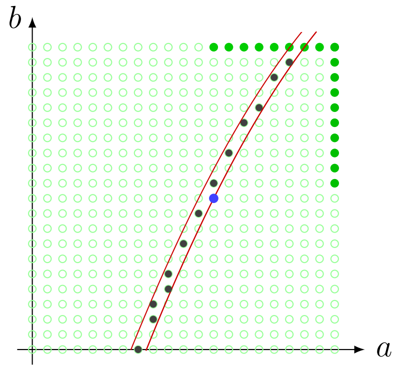


Figure 1: By discretization, IC constraints are not ensured with black points.

to zero with finer discretizations, and the sequence of optimal values converges to the optimal value of the continuous problem. These authors have considered a linear model including multiple agents and border constraints[5], which are not present in our setting. In contrast, it is considered a valuation function v that could be nonlinear.

Let (Q^n, V^n) be the solution of the discretized problem (9). Since these functions are defined on X_n, define the extensions $\widetilde{Q}^n, \widetilde{V}^n : [0,1]^2 \to \mathbb{R}$ as

$$\widetilde{Q}^n(x,y) := Q^n(a,b) \quad , \quad \widetilde{V}^n(x,y) := V^n(a,b)$$

where $(a,b) \in X_n$ is such that $a \le x < a + \frac{1}{n-1}$ and $b - \frac{1}{n-1} < y \le b$.

Let $\delta^*(\widetilde{Q}^n, \widetilde{V}^n)$ be the supremum over all IC constraint violations by the pair $(\widetilde{Q}^n, \widetilde{V}^n)$. That is, although some constraints are not fulfilled, we can be sure that for any $(a,b), (a',b') \in [0,1]^2$,

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(a',b') - (v(\widetilde{Q}^n(a',b'), a, b) - v(\widetilde{Q}^n(a',b'), a', b')) \ge -\delta^*(\widetilde{Q}^n, \widetilde{V}^n)$$

To guarantee the asymptotic feasibility of extensions $(\widetilde{Q}^n, \widetilde{V}^n)$, all IC constraint violations must uniformly converge to zero, as the next proposition shows.

Proposition 4.6. *We have $\delta^*(\widetilde{Q}^n, \widetilde{V}^n) \le O(\frac{1}{n-1})$.*

[5]These constraints are related to the allocation treated as a probability since, in their model, there are N buyers and J degrees of product quality.

The following proposition shows that optimality can be achieved in the limit.

Proposition 4.7. *Let OPT_n be the optimal value of the discretized problem, and let OPT^* be the optimal value of the continuous problem. Then, $\liminf_{n\to\infty} OPT_n \geq OPT^*$. Additionally, if $\lim_{n\to\infty} \widetilde{Q}^n(a,b)$ and $\lim_{n\to\infty} \widetilde{V}^n(a,b)$ exists for any $(a,b) \in [0,1]^2$, then $\lim_{n\to\infty} OPT_n = OPT^*$.*

5 Numerical Solution: Regulating a Monopolist Firm

Lewis and Sappington (1988b) studied the design of regulatory policy when the regulator is imperfectly informed about both the costs and the demand functions of the monopolist firm he is regulating. They considered that demand for the firm's product $q = Q(p,a)$ and the costs of producing output q, $C(q,b)$, involve firm's private information parameters (a,b) distributed over $\Theta = [\underline{a}, \overline{a}] \times [\underline{b}, \overline{b}]$ according to a strictly positive density function $f(a,b)$.

The regulator offers the firm a menu of unit prices p and corresponding subsidy t, according with firm's type revealed. The profit of the firm of type (a,b) is $p\,Q(p,a) - C(Q(p,a),b) + t$. Profit reservation level is type independent and normalized at zero. It is assumed that the regulator can ensure that the firm serves all demand at the established prices. The regulator's objective function is the expected consumer surplus net of the transfer to the firm

$$\int_{\underline{a}}^{\overline{a}} \int_{\underline{b}}^{\overline{b}} \{\Pi(Q(p(a,b),a),a) - p(a,b)Q(p(a,b),a) - t(a,b)\} f(a,b)db\,da \quad (10)$$

where $\Pi(Q,a) = \int_0^Q P(\xi,a)d\xi$, and $P(\cdot)$ denotes the inverse demand curve.

The regulator's problem is to design the menu of contracts $(p(a,b), t(a,b))$ in order to maximize (10) subject to individual rationality

$$p(a,b)Q(p(a,b),a) - C(Q(p(a,b),a),b) + t(a,b) \geq 0$$

and incentive compatibility constraints

$$p(a,b)Q(p(a,b),a) - C(Q(p(a,b),a),b) + t(a,b) \geq \\ p(\widehat{a},\widehat{b})Q(p(\widehat{a},\widehat{b}),a) - C(Q(p(\widehat{a},\widehat{b}),a),b) + t(\widehat{a},\widehat{b})$$

By setting

$$
\begin{aligned}
v(p,a,b) &= pQ(p,a) - C(Q(p,a),b)\\
H(p,a) &= pQ(p,a) - \Pi(Q(p,a),a)\\
V(a,b) &= v(p(a,b),a,b) + t(a,b)
\end{aligned}
$$

the regulator's problem can be written as

$$
\max_{p(\cdot),V(\cdot)} \int_{\underline{a}}^{\overline{a}} \int_{\underline{b}}^{\overline{b}} \{v(p(a,b),a,b) - H(p(a,b),a) - V(a,b)\} f(a,b) db\, da \tag{RP}
$$

subject to

(IR) $V(a,b) \geq 0 \qquad \forall\ (a,b) \in \Theta$

(IC) $V(a,b) - V(\widehat{a},\widehat{b}) \geq v(p(\widehat{a},\widehat{b}),a,b) - v(p(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) \qquad \forall\ (a,b),(\widehat{a},\widehat{b}) \in \Theta$

Note that this formulation fits the standard nonlinear pricing model. In this case, the new variable V is the firm's profit.

Lewis and Sappington have derived a solution for the particular example

$$
Q(p,a) = \alpha - p + a \quad , \quad C(q,b) = K + (c_0 + b)q \tag{11}
$$

with α, K and c_0 positive constants and a uniform distribution over $\Theta = [0,1]^2$. But, as Armstrong (1999) has noted, Lewis and Sappington's solution for this example cannot be right. Furthermore, in that paper, Armstrong argued that excluding a positive mass of types should be optimal, like in nonlinear pricing. However, because of the change in the variables he used, the type set is not convex, and his exclusion argument cannot strictly be applied. He also expressed the following:

- "*Nevertheless, I believe that the condition that the support be convex is strongly sufficient and that it will be the usual case that exclusion is optimal...*"

- "*I have not found it possible to solve this precise example correctly...*"

Therefore, we are facing a bidimensional adverse selection model with an unknown solution where a conjecture about the optimality of exclusion was made.

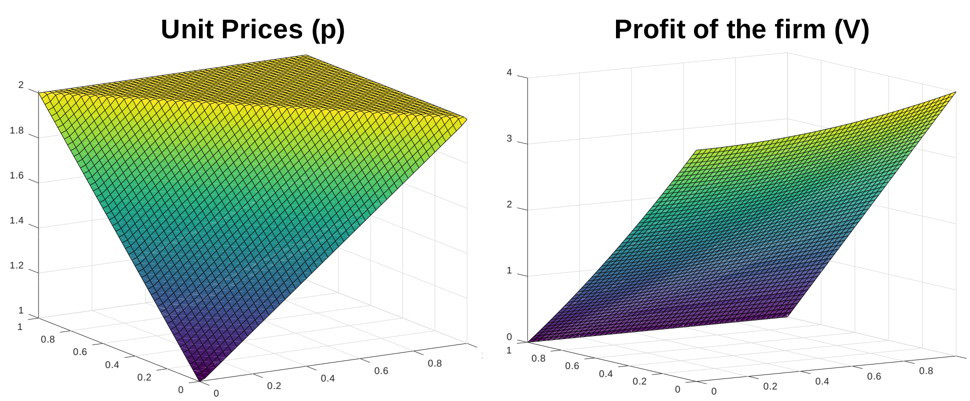
Given that $v(p,a,b) = (\alpha + a - p)(p - c_0 - b) - K$, then $v_{pa} = v_{pb} = 1$. In this case, $p(\cdot,\cdot)$ will be non-decreasing in a and b. Additionally, since $\frac{-v_{pb}}{v_{pa}} < 0$, the characteristic curves are strictly decreasing. Following the same considerations as in Section 4, it will be sufficient that each $(a,b)-$agent verifies the IC constraints with all the points over the set

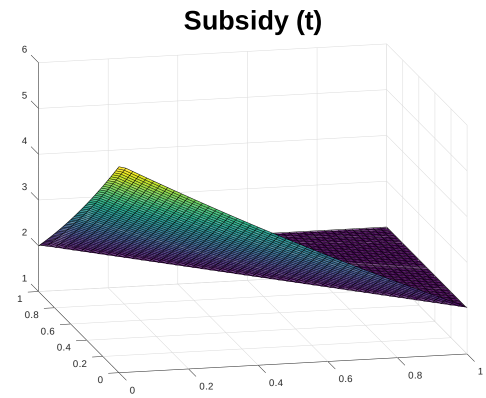
$$F^{(a,b)} := \{(0,s) \mid b \leq s \leq 1\} \cup \{(s,1) \mid 0 \leq s \leq a\}$$

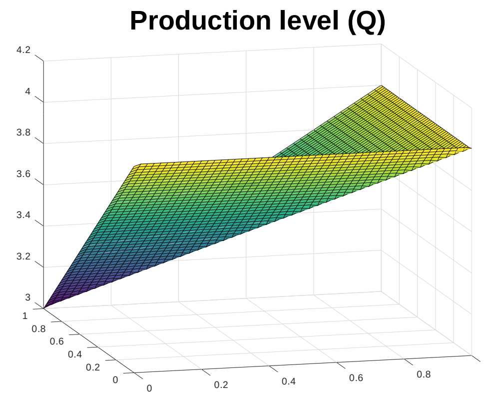
Note that the discretized problem has a unique solution in view of the linearity of IC constraints by the multiplicative separable form of v (see remark 4 on page 10) and the strict concavity of the objective function. Note also that the signs of v_a and v_b are endogenously determined, so condition (*) cannot be verified.

The problem was numerically solved for three different cases of c_0, α, and K. The type set was discretized into $n = 51$ points over each direction. The numerical solutions were obtained via Knitro/AMPL by using the active set algorithm. Next, it is shown the graphs of optimal prices, the firm's profit (solutions of (RP)), subsidies, and production level.

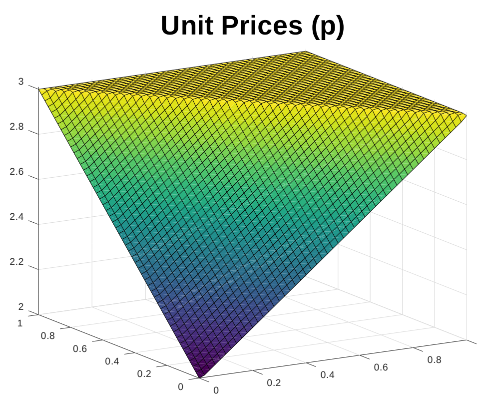
Case 1: $\boldsymbol{c_0 = 1}$, $\boldsymbol{\alpha = 5}$, $\boldsymbol{K = 2}$

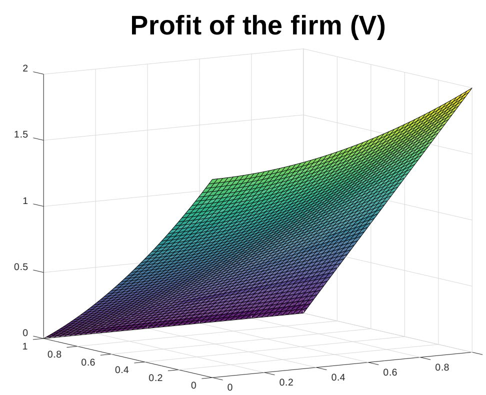


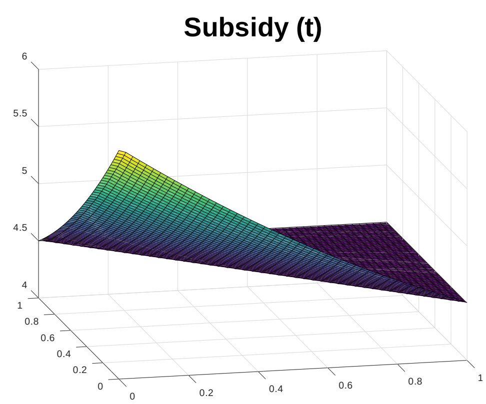


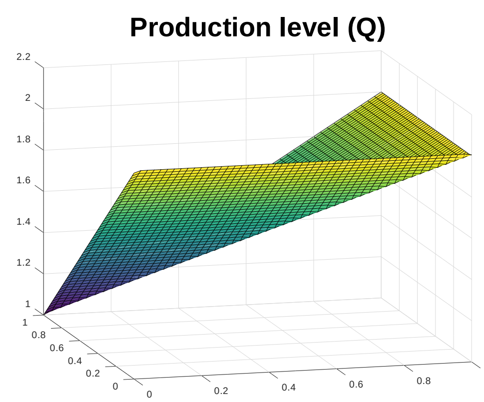


Case 2: $\boldsymbol{c_0 = 2\ ,\quad \alpha = 4\ ,\quad K = 4.5}$

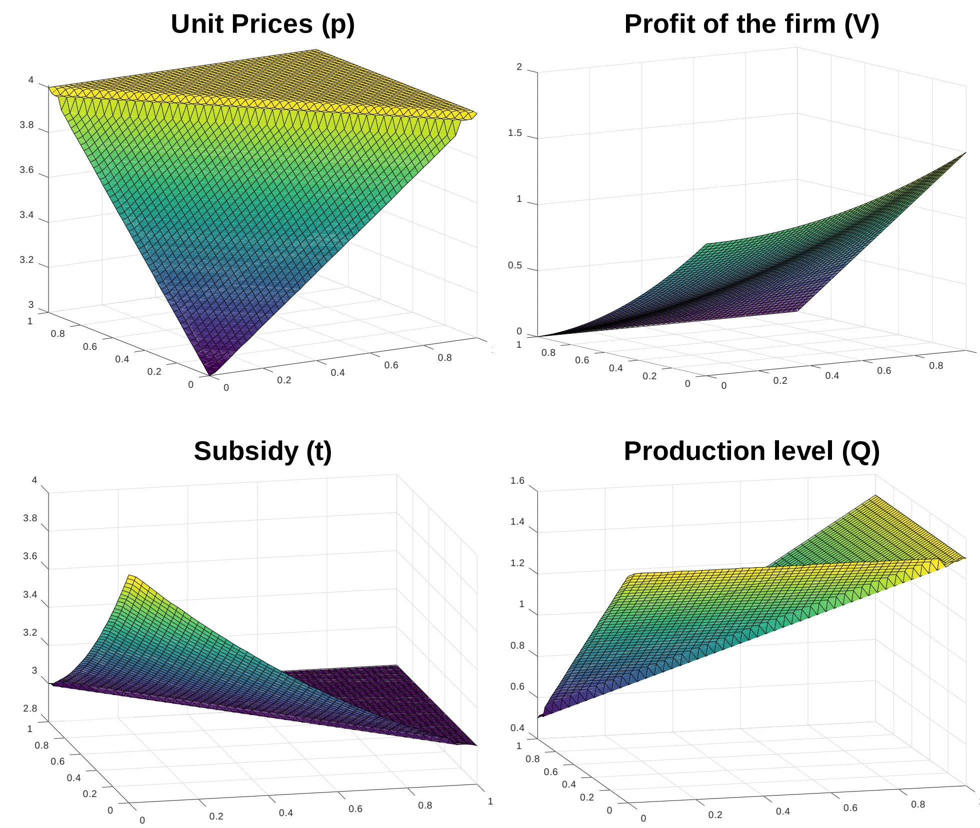




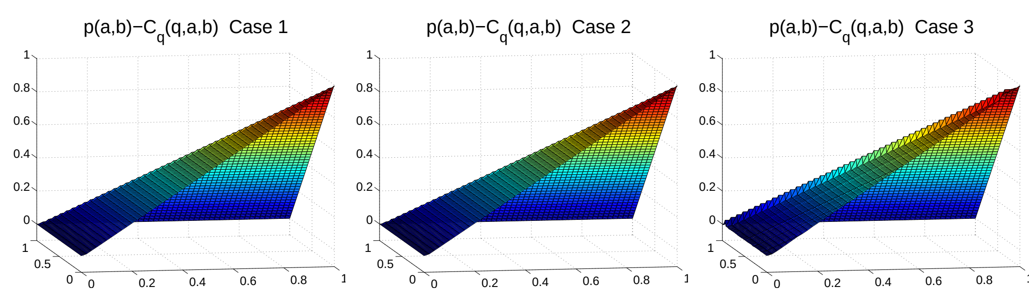




Case 3: $\boldsymbol{c_0 = 3}$, $\boldsymbol{\alpha = 4.5}$, $\boldsymbol{K = 3}$



It is also shown the numerical differences between unit prices and marginal costs.



Some insights from these solutions:

Due to this example derives into an optimization problem with linear constraints and unique solution, the numerical methods to solve it are efficient. Thus, the statements below are reliable:

1. It seems that at the optimum[6], all types (a, b) with $a + b \geq 1$ are bunching at unit price $c_0 + 1$, and the subsidy for them is the fixed cost K. Additionally, the unit price assigned to type $(0, 0)$ seems to be c_0.

2. In view of the numerical difference $p - C_q$, the regulator induces the firm to price above marginal costs for almost all (a, b) types rather than $a = 0$ or $b = 1$ (i.e., such types with the a priori lowest demand function or such types who obtain zero profit)[7].

3. The firm's numerical profit V suggests that there is no exclusion.

5.1 About the optimality of exclusion

Although the optimality of exclusion is a generic property, perhaps the most intriguing insight from the numerical solutions is the possibility that non-exclusion of a positive mass of types should be optimal, contrary to Armstrong's conjecture stated previously.

Furthermore, in Barelli et al. (2014), the authors relaxed Armstrong's strong conditions (strict convexity and homogeneity of degree one) and proved a more general result of the generic desirability of exclusion. For this example, they considered that prices belong to $[c_0 + 1, \alpha]$ to conclude that their result can be applied and hence confirm Armstrong's conjecture. However, the numerical results show that prices do not belong to $[c_0 + 1, \alpha]$. Therefore, their theorem should not be applied.

[6] I conjecture the optimum price p to be $p(a, b) = c_0 + a + b$ when $a + b \leq 1$ and $p(a, b) = c_0 + 1$ when $a + b > 1$

[7] In Baron and Myerson (1982), the authors analyzed a model in which the regulator is uncertain only about the firm's cost function. At the optimum, prices are above marginal costs for all cost realizations other than the lowest. In the model of Lewis and Sappington (1988a), the regulator is uncertain only about the position of the demand curves. In that model, if $C''(q) \geq 0$ (similar to here), setting prices at the level of marginal costs for the reported demand is optimal ($p = C_q$).

Next, I provide one technical argument explaining why Armstrong's Theorem about the desirability of exclusion formulated in the nonlinear pricing context could not be extended to this regulation model.

In nonlinear pricing, the customers' exit option is $q^{\text{out}} = 0$ and $t^{\text{out}} = 0$. Hence, the natural assumptions $v(0, a, b) = 0$ and $C(0) = 0$ imply that the monopolist's revenue $v(0, a, b) - C(0) - V(a, b)$ is zero when $V(a, b) = 0$ (that is, when type (a, b) is excluded). Then, the monopolist's penalty for causing some customers to exit the market is to not receive income from them.

On the other hand, in the regulation model, the exit option of the firm (a, b) is the unit price p^{out} and subsidy t^{out} at which profit $V(a, b)$ is zero. For such a firm, in the previous example given by (11), the regulator's benefit is $(\alpha + a - c_0 - b)Q(p^{\text{out}}, a) - (Q(p^{\text{out}}, a))^2/2 - K$ (in fact, if there is no production, this amount is $-K$). Thereby, in contrast with the monopolist, the regulator could have to assume a negative penalty by excluding a firm. Therefore, Armstrong's argument of comparing benefits (more income from agents still in the market) versus penalties (zero income from agents excluded) might not be applicable to this model.

Besides, by the model formulation, it can be understood that the regulator has no interest of excluding the firm (for example, if the firm is already operating and zero production is not desirable in the economy). Individual rationality constraints reflects that for all type of firm, whether producing or not, the subsidy must be such that the firm's profit is at least zero.

A model in which the regulator offers a menu of prices and subsidies at which when a firm cannot make a positive profit does not produce and regulator pays zero, should contain an extra variable (like the $r(\cdot)$ variable in Baron and Myerson (1982), Rochet (2009)) indicating the probability that the firm will be permitted to produce. The optimal value of this variable will indicate whether a type of firm is excluded or not.

Appendix A: Testing the method

In this appendix, the numerical solution of problem (9) is compared with the analytical solution of the following models in the literature: Laffont et al. (1987) have considered that monopolist faces customers with linear demand curves and is uncertain about both the slope and intercept of such linear demand, which yields on linear-quadratic customers' valuation $v(q,a,b) = aq - \frac{1+b}{2}q^2$. Basov (2001) proposed the Hamiltonian Approach and solved the generalization $v(q,a,b) = aq - \frac{1+b}{\gamma}q^\gamma$ with $\gamma \geq 2$, for which demand curves are concave. Araujo et al. (2019) have analized a case of convex demand curves, for which the customers' valuation is $v(q,a,b) = (c-b)\log(aq+1)$.

Two criteria to compare our approximations are presented. The first one is to compute the *average quadratic error (a.q.e.)* between analytic quantity Q^{real} and numerical quantity Q^{num} (the same calculation is shown for informational rent)

$$a.q.e.(Q^{\text{num}}, Q^{\text{real}}) = \frac{1}{n^2}\sum_{i,j=1}^{n}(Q^{\text{num}}_{i,j} - Q^{\text{real}}_{i,j})^2$$

The second criterion is merely a visual comparison. Despite no being formal, in practice the numerical approximations help us to formulate predictions about the functional form of the solution, such as the participation set or the contour levels (i.e., how types are bunching). Thus, there are provided graphs of quantity and informational rent for both numerical and analytical solutions, as well as their contour levels and cross-sections.

After omitted the IC constraints on the southeast for each type, it is required to verify a posteriori if the omitted constraints are indeed strictly satisfied. Such verification can only be numerical. For this reason, it is provided some graphs showing whether a fixed type (a_i, b_j) (in blue) is IC with all the others in X_n, drawing a green point if such IC is satisfied and a red point if it is not. Note that all the types on the southeast of the points considered are satisfying IC.

Because of numerical optimization, as well as the limitations by the discretization pointed out in the remarks of subsection (4.2), it does not surprise the existence of red points in some graphs; however, the violations may be considered small. The value of $\delta = \delta_{i,j}$ in each one of the graphs indicates the minimum violation of IC constraints between (a_i, b_j) and the red points. That is, if we allow some tolerance $tol_{i,j} > -\delta_{i,j}$, all the IC constraints will be verified for such (a_i, b_j) blue point. Furthermore, by defining $\delta^* = \min_{i,j}\{\delta_{i,j}\}$ if $tol > -\delta^*$ then all the IC constraints will be satisfied for any type in X_n. Such δ^* value is provided in each example.

The numerical solutions were performed via Knitro/AMPL using the Active Set Algorithm. Otimization process stopped if one of the following tolerances were achieved: `maxit`$= 10^4$, `feastol`$= 10^{-15}$, `xtol`$= 10^{-15}$, `opttol`$= 10^{-15}$, where `maxit` is the maximum number of iterations, `feastol` refers to feasibility tolerance, `xtol` is the relative change of decision variables and `opttol` is the optimality KKT sttoping tolerance. In all examples, `xtol` were achieved first.

5.2 Example 1. Linear Demand

In Laffont et al. (1987) the authors have solved the monopolist's problem for the data

$$v(q,a,b) = aq - \frac{(1+b)}{2}q^2 \quad , \quad C(q) = 0 \quad , \quad f(a,b) = 1$$

The solutions q and T they have found are:

$$q(a,b) = \begin{cases} 0 & , \quad a \leq \frac{1}{2} \\[2ex] \dfrac{4a-2}{4b+1} & , \quad \dfrac{1}{2} \leq \dfrac{a+2b}{4b+1} \leq \dfrac{3}{5} \\[2ex] \dfrac{3a-1}{2+3b} & , \quad \dfrac{3}{5} \leq \dfrac{2a+b}{2+3b} \leq 1 \end{cases}$$

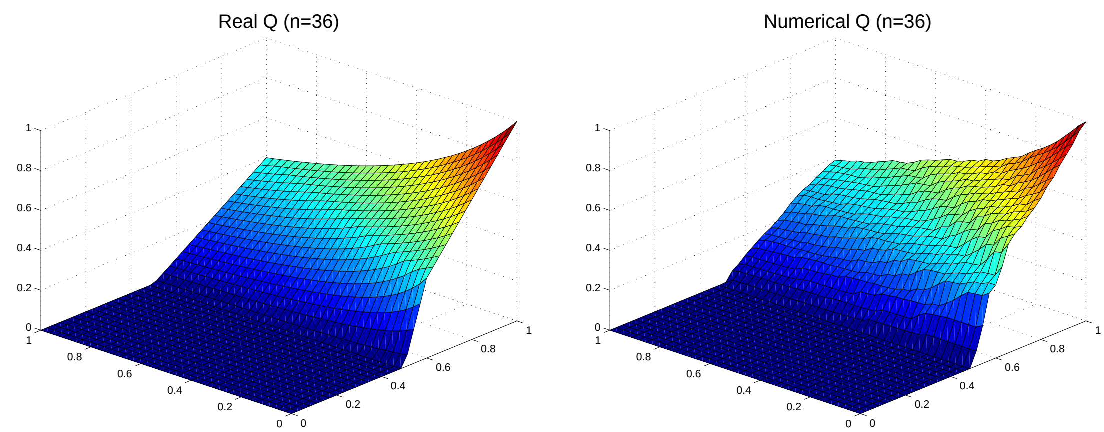
$$T(q) = \begin{cases} \dfrac{q}{2} - \dfrac{3q^2}{8} & , \quad q \leq \dfrac{2}{5} \\[2ex] \dfrac{q}{3} - \dfrac{q^2}{6} + \dfrac{1}{30} & , \quad \dfrac{2}{5} \leq q \leq 1 \end{cases}$$

Note that v_q is linear in each a and b variables. Then, by Proposition 4.5, it can be reduced even more the number of constraints. The exact number of IC constraints is $\frac{1}{2}(3n^3 - 3n^2 - 4n + 4)$ instead of $n^4 - n^2$ as in the original problem. The problem was solved with $n = 36$. For this value, 67 970 IC constraints were considered whereas 1 610 350 were ruled out.

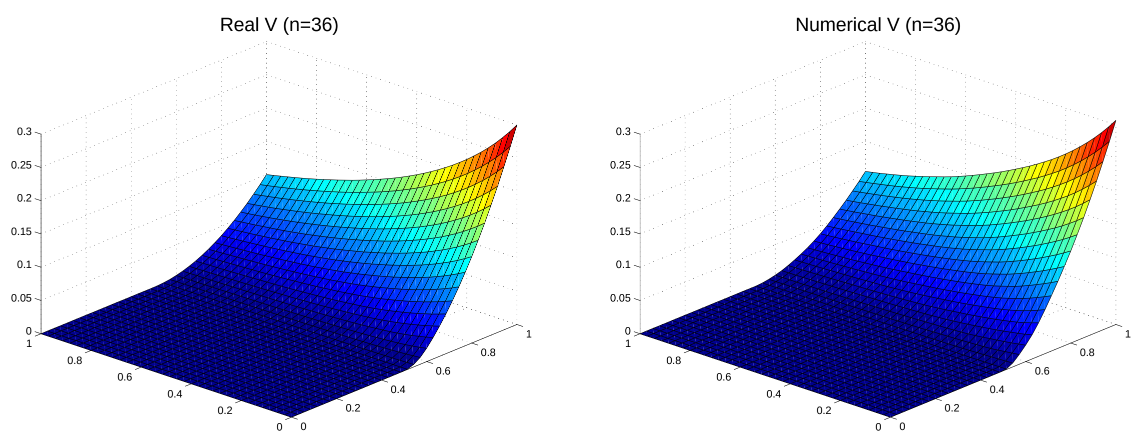
Next, the numerical result is compared with the analytical (real) one.

$$
\begin{aligned}
a.q.e.(Q^{num}, Q^{real}) &= 3.6442{\times}10^{-4} \\
a.q.e.(V^{num}, V^{real}) &= 0.1149{\times}10^{-4} \\
\left| \text{ profit}^{num} - \text{profit}^{real} \right| &= 9.6668{\times}10^{-4} \\
\delta^* &= -1.70804{\times}10^{-4}
\end{aligned}
$$

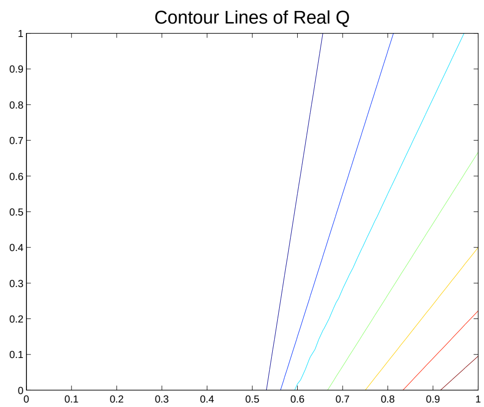
Comparing Quantity

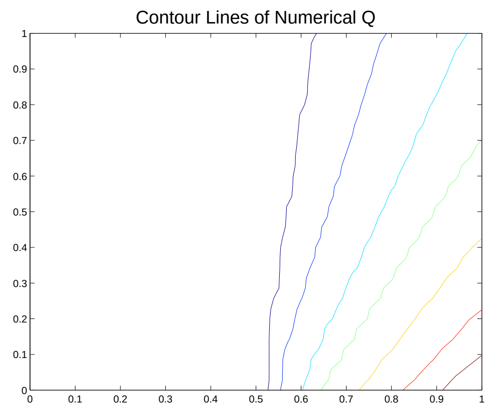


Comparing Informational Rent

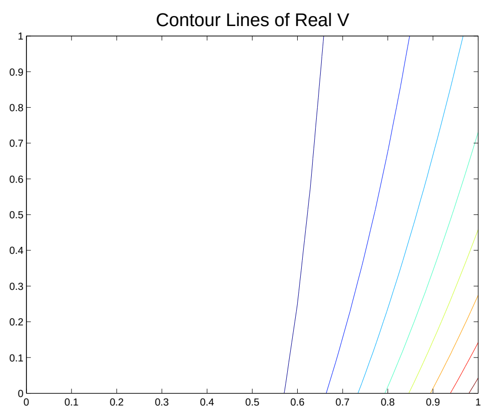


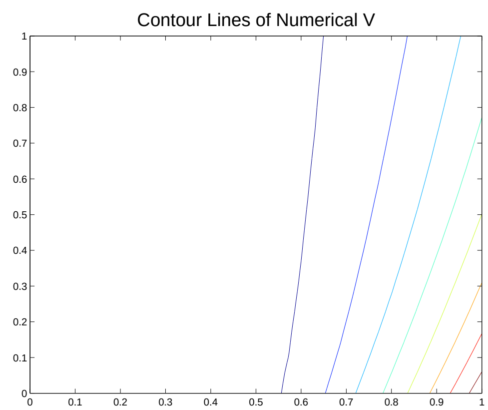
Contour Lines of Quantity



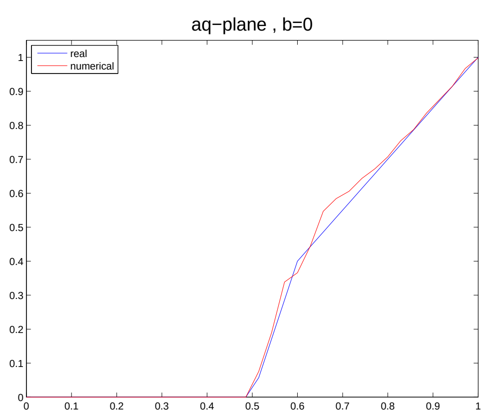


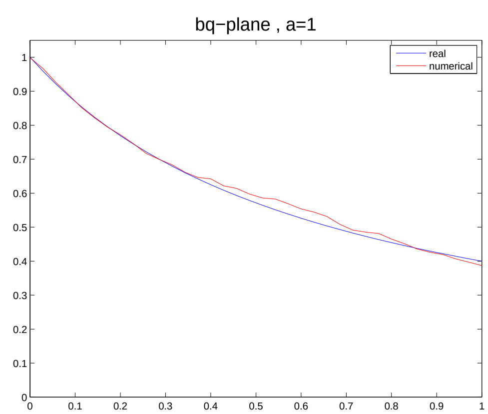
Contour Lines of Informational Rent



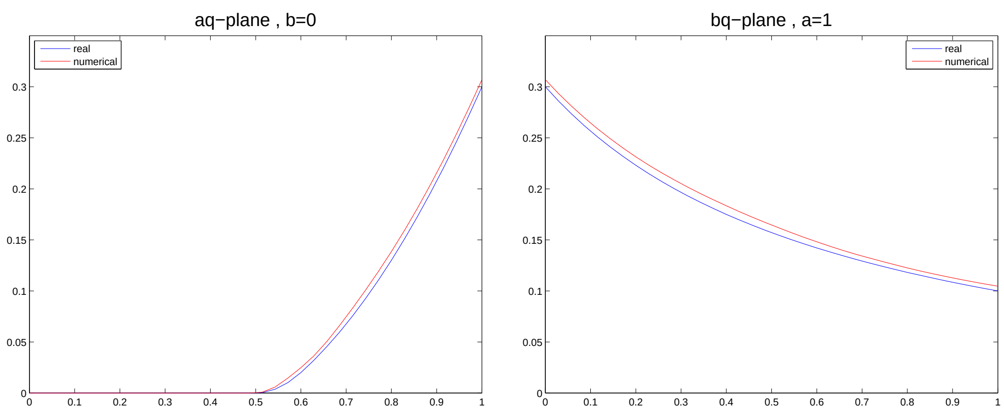


Cross-sections of Quantity

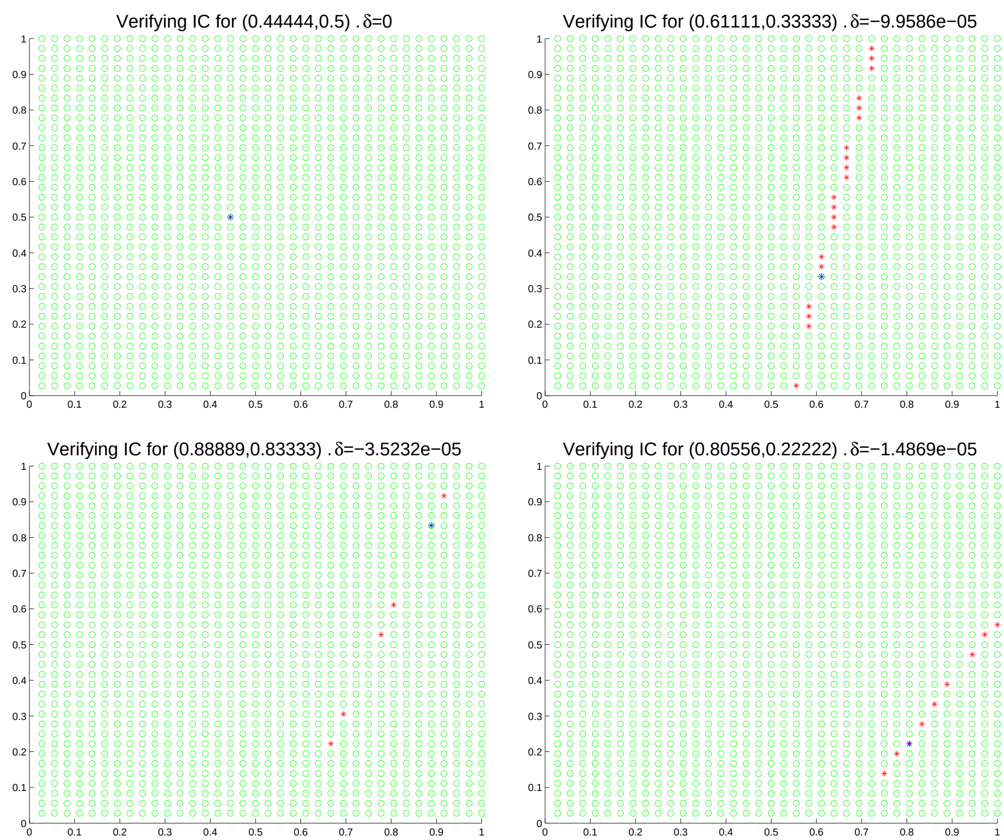




Cross-sections of Informational Rent



Verifying IC constraints



5.3 Example 2. Concave Demand

In Basov (2001) the author solved the original problem for the data

$$v(q,a,b) = aq - \frac{(c+b)}{\gamma} q^{\gamma} \quad , \quad C(q) = 0 \quad , \quad f(a,b) = 1$$

where $c > \frac{1}{2}$ and $\gamma \geq 2$ are constants. The solutions q and T he have found are:

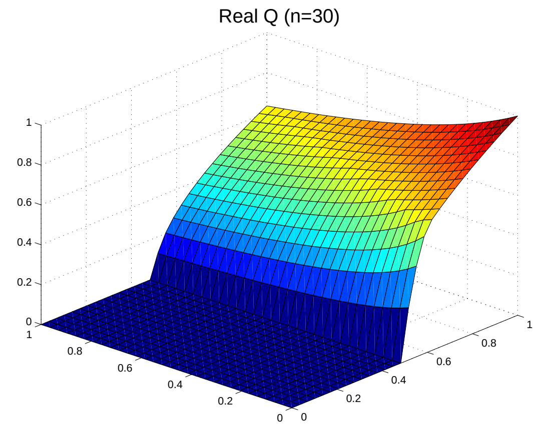
$$q(a,b) = \begin{cases} 0 & , \quad a \leq \frac{1}{2} \\ \left(\frac{4a-2}{4b+2c-1}\right)^{\frac{1}{\gamma-1}} & , \quad (3+2c)a - 2b \leq 2c+1 \\ \left(\frac{3a-1}{3b+2c}\right)^{\frac{1}{\gamma-1}} & , \quad (3+2c)a - 2b > 2c+1 \end{cases}$$

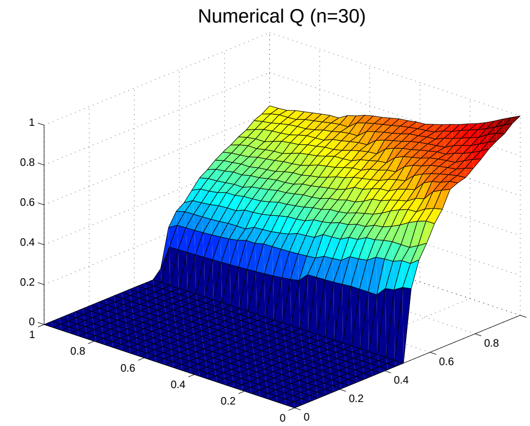
$$T(q) = \begin{cases} \frac{q}{2} - \frac{(\frac{c}{2}+\frac{1}{4})}{\gamma} q^{\gamma} & , \quad q \leq \left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} \\ \frac{1}{6}\left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} - \frac{(\frac{c}{6}+\frac{1}{4})}{\gamma}\left(\frac{2}{3+2c}\right)^{\frac{\gamma}{\gamma-1}} + \frac{q}{3} - \frac{c}{3\gamma} q^{\gamma} & , \quad q > \left(\frac{2}{3+2c}\right)^{\frac{1}{\gamma-1}} \end{cases}$$

The discretized problem was solved for the case $\gamma = 3$ with $n = 30$ points over each axis. For this value, 39 092 incentive compatibility constraints were considered, and 770 008 were eliminated.

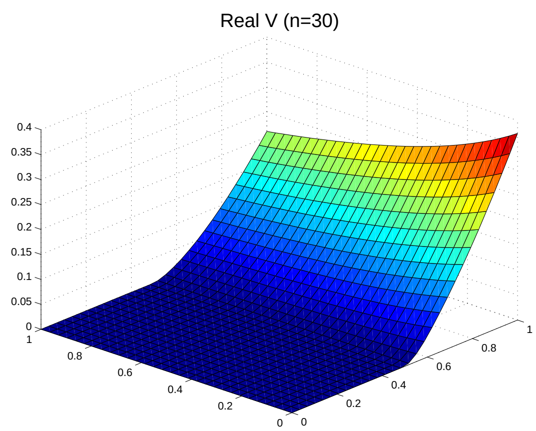
$$\begin{aligned} a.q.e.(Q^{num}, Q^{real}) &= 4.5853 \times 10^{-4} \\ a.q.e.(V^{num}, V^{real}) &= 0.0384 \times 10^{-4} \\ \left| \text{ profit}^{num} - \text{profit}^{real} \right| &= 2.5717 \times 10^{-3} \\ \delta^* &= -7.82371 \times 10^{-4} \end{aligned}$$

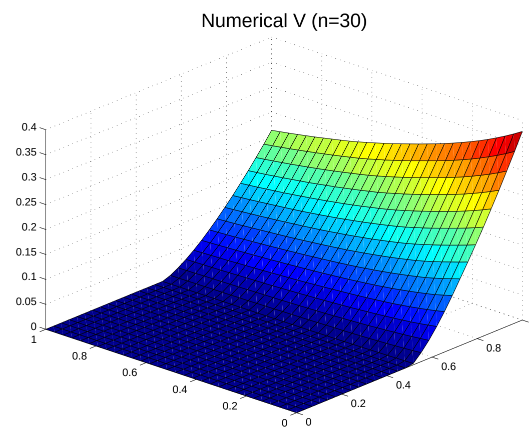
Comparing Quantity



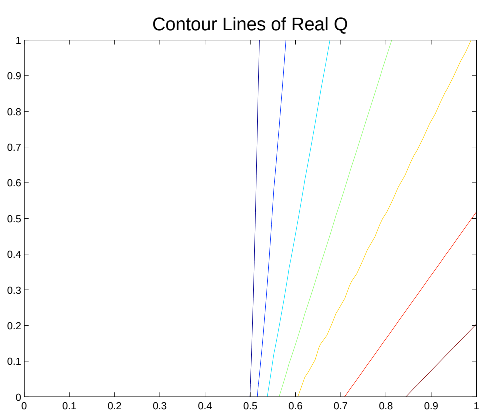


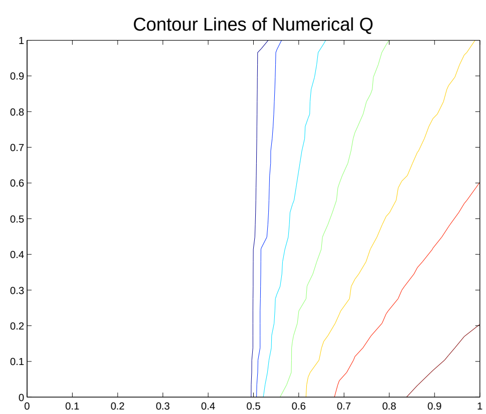
Comparing Informational Rent



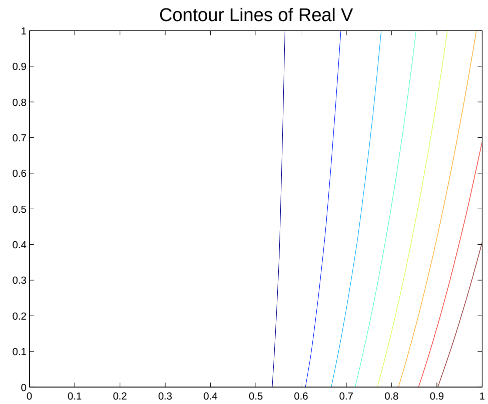


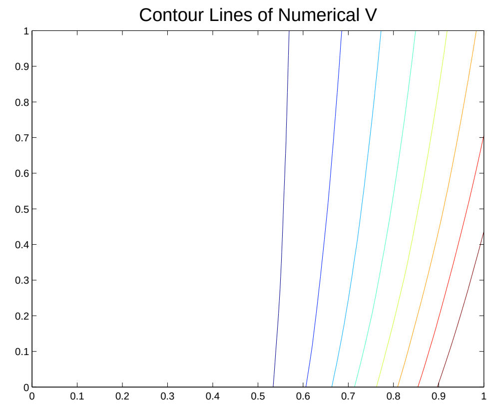
Contour Lines of Quantity



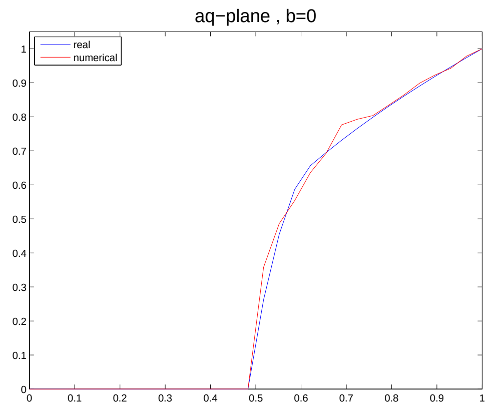


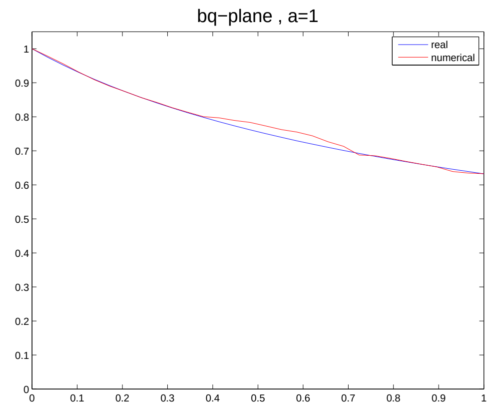
Contour Lines of Informational Rent



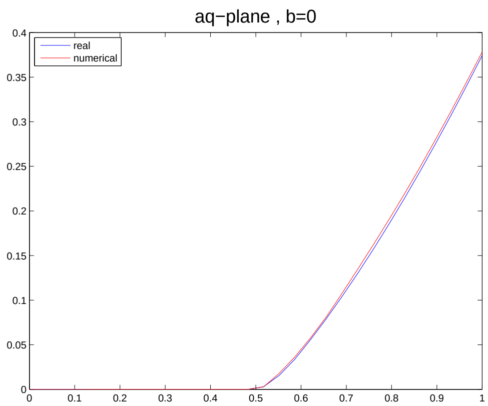


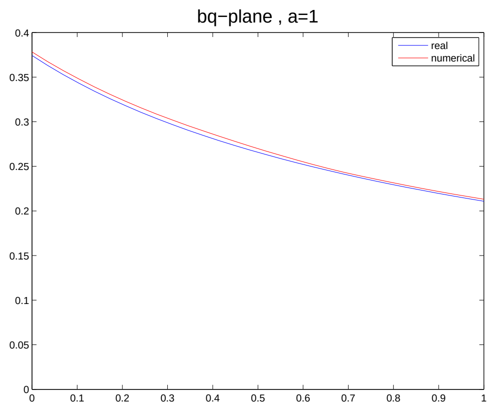
Cross-sections of Quantity



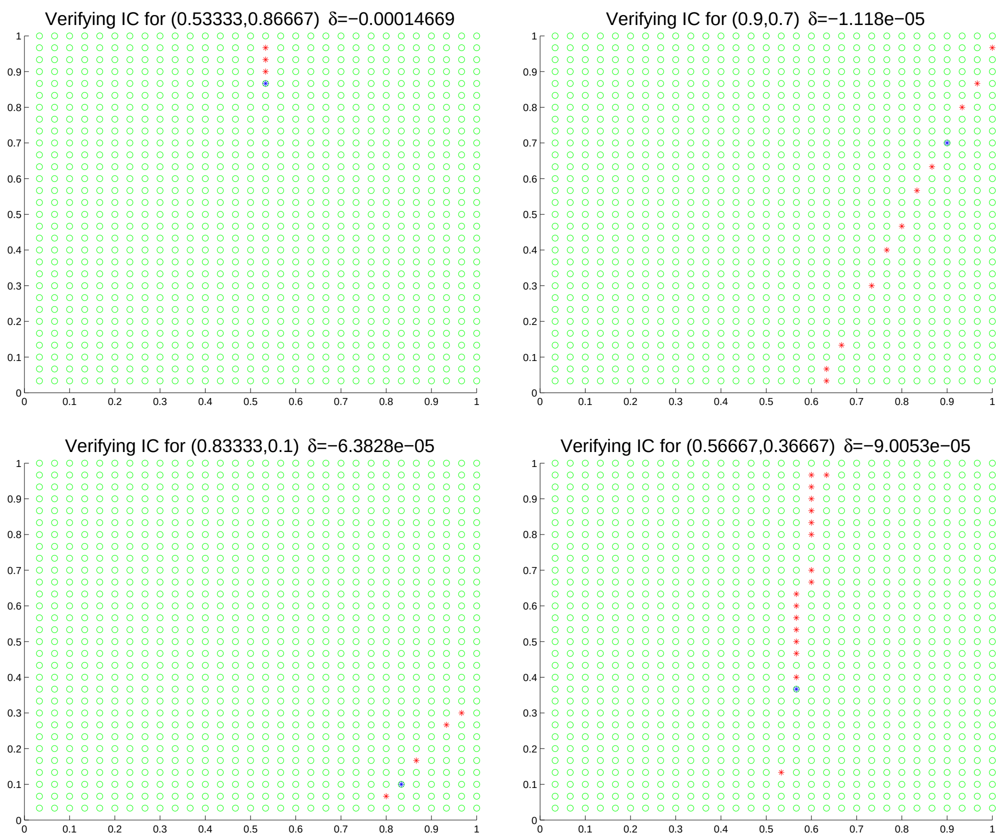


Cross-sections of Informational Rent





Verifying IC constraints



5.4 Example 3. Convex Demand Curves

In Araujo et al. (2019) the authors have analized the monopolist's problem for the case

$$v(q,a,b) = (c-b)\log(aq+1) \quad , \quad C(q) = \lambda q \quad , \quad f(a,b) = 1$$

where $c \geq 1$ and $\lambda \in (0,1)$ are given. In this case, the solution proposed is not given in a closed-form. To express the analytical solution, define

$$\begin{aligned}
D(r) &= \lambda r(1-r) \\
E(r) &= \lambda(1-r) - \lambda r\log(r) - c\,r(1-r) \\
F(r) &= 2c\,r\log(r) + c(1-r) - \lambda\log(r)
\end{aligned}$$

and let $\underline{r} \in]0,1[$ be the solution of

$$(2c\,r - \lambda)\log(r) + c(1-r) = 0$$

Also define

$$\phi(r) = \frac{-E(r) + \sqrt{E(r)^2 - 4D(r)F(r)}}{2D(r)} \quad , \quad \forall\, r \in]\underline{r}, 1[$$

Finally, given $(a,b) \in [0,1]^2$ define $q(a,b)$ as follows

- If $b \geq c - (c\underline{r})/a$, let $q(a,b) := 0$.
- If $b < c - (c\underline{r})/a$, let $r(a,b) \in]\underline{r}, 1[$ be the solution of

$$\frac{c-b}{br} - \frac{c}{ab} = \frac{-E(r) + \sqrt{E(r)^2 - 4D(r)F(r)}}{2D(r)}$$

  such that $\phi(r(a,b)) > 0$ and $\phi'(r(a,b)) > 0$, and let $q(a,b) := \phi(r(a,b))$.

Furthermore, the tariff as a function of r over $]\underline{r}, 1[$ can be expressed as

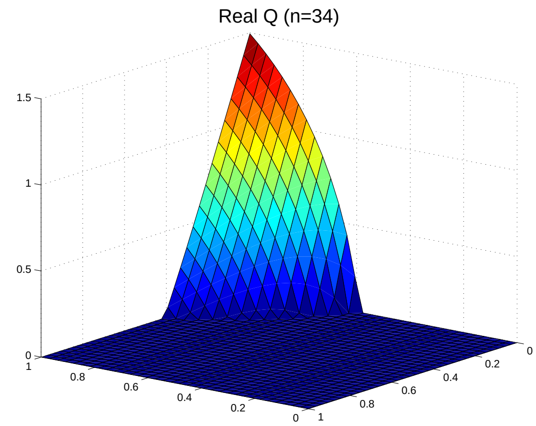
$$T(r) = \int_{\underline{r}}^{r} v_q(\phi(\tilde{r}), \tilde{r}, 0)\phi'(\tilde{r})d\tilde{r}$$

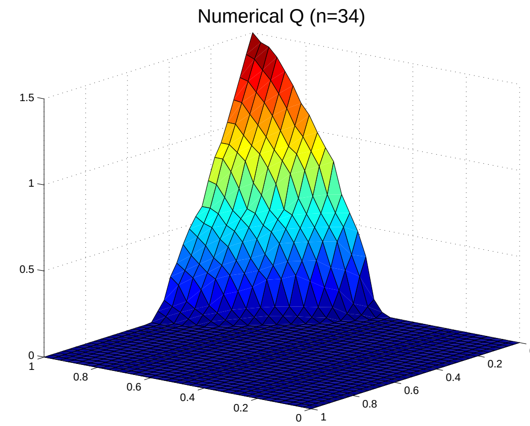
Agent type (a,b) has to transfer $t(a,b) = T(r(a,b))$ to the monopolist, which determines $V(a,b)$. In this way the variables q and V are defined over $[0,1]^2$.

On the other hand, the discretized problem was solved for the case $c = 1$ and $\lambda = 0.4$ with $n = 34$ points over each axis.

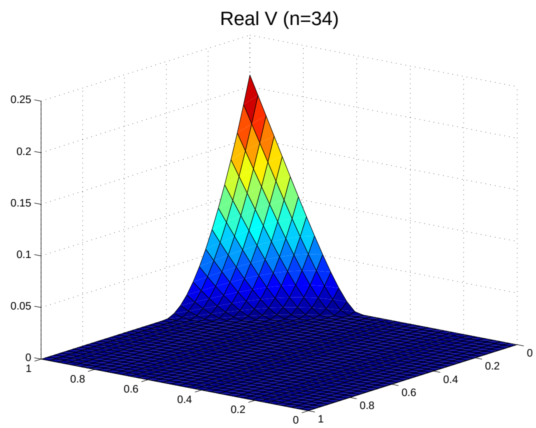
$$\begin{aligned}
a.q.e.(Q^{num}, Q^{real}) &= 2.6300{\times}10^{-3} \\
a.q.e.(V^{num}, V^{real}) &= 2.6064{\times}10^{-5} \\
\left|\ \text{profit}^{num} - \text{profit}^{real}\right| &= 2.3191{\times}10^{-2} \\
\delta^* &= -5.77989{\times}10^{-4}
\end{aligned}$$

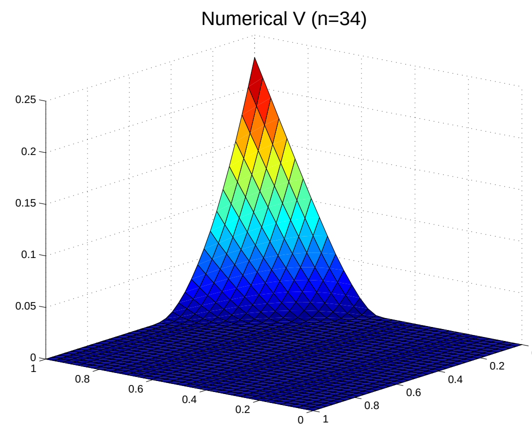
Comparing Quantity



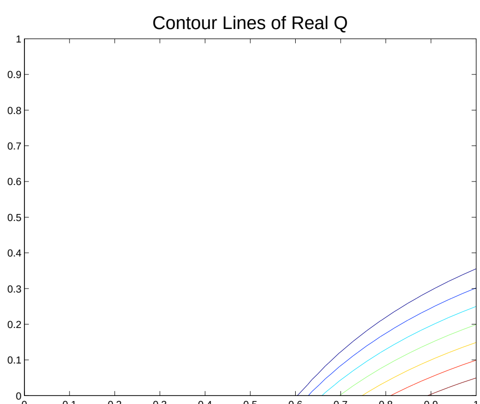


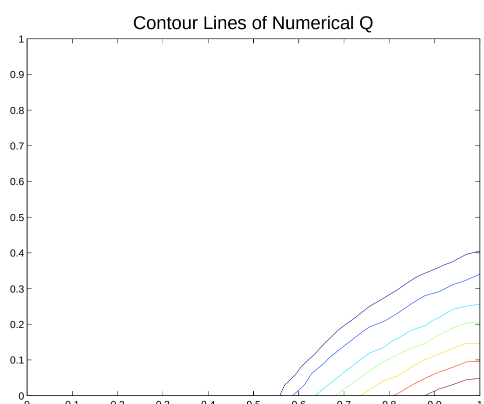
Comparing Informational Rent



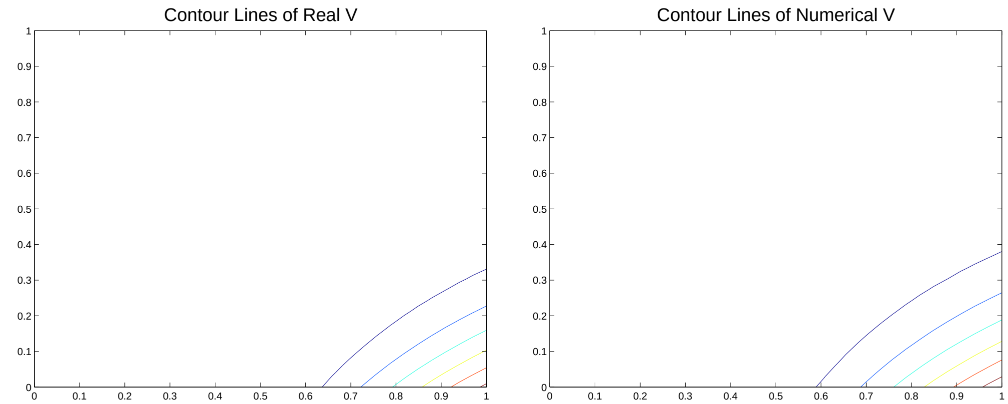


Contour Lines of Quantity

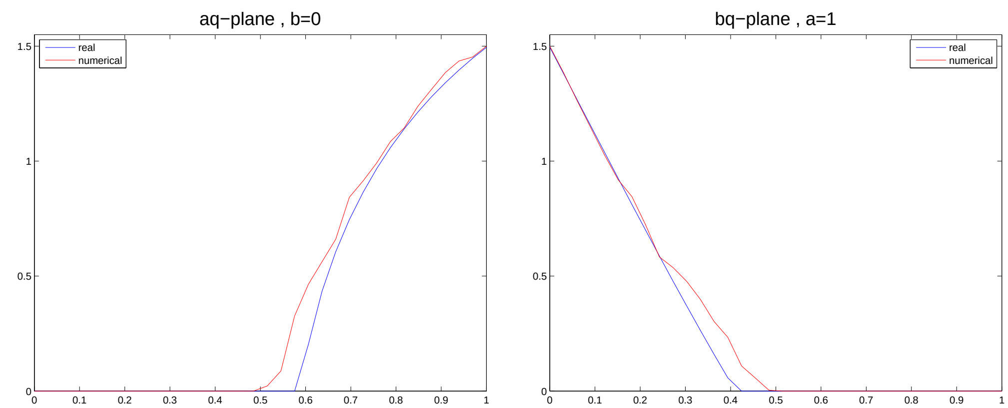




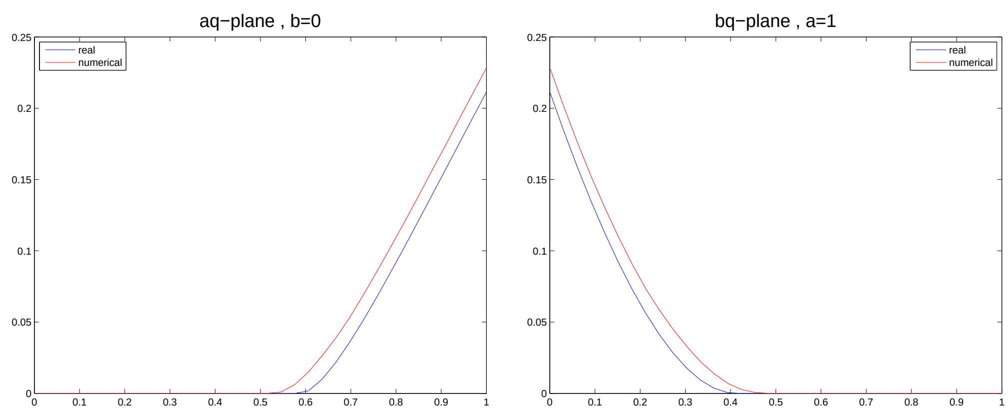
Contour Lines of Informational Rent



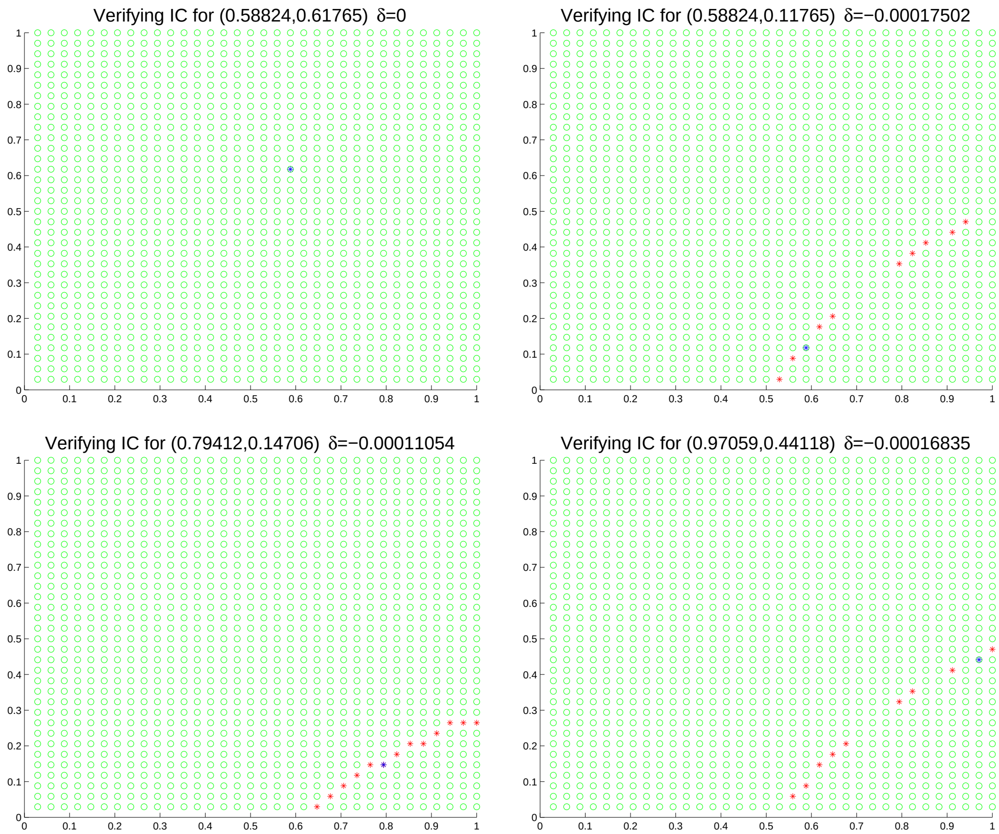
Cross-sections of Quantity



Cross-sections of Informational Rent



Verifying IC constraints



Appendix B: Mathematical Proofs

Proof of Proposition 4.1. Fix $q \in Q$; by Assumption 2, $v_q(q, \cdot, b)$ is strictly increasing and $v_q(q, \widehat{a}, \cdot)$ is strictly decreasing, so $\widehat{a} > a$ and $\widehat{b} < b$ imply $v_q(q, a, b) < v_q(q, \widehat{a}, b)$ and $v_q(q, \widehat{a}, b) < v_q(q, \widehat{a}, \widehat{b})$, respectively. Thus, $v_q(q, a, b) < v_q(q, \widehat{a}, \widehat{b})$. □

Proof of Proposition 4.2. Fix $(\widehat{a}, \widehat{b})$ such that $a < \widehat{a}$ and $b > \widehat{b}$. Define $F(x, y) := V(x, y) - v(q(\widehat{a}, \widehat{b}), x, y) \;\; \forall \; (x, y) \in [0, \widehat{a}] \times [\widehat{b}, 1]$. Then,

$$F_x = V_a(x, y) - v_a(q(\widehat{a}, \widehat{b}), x, y) = v_a(q(x, y), x, y) - v_a(q(\widehat{a}, \widehat{b}), x, y)$$

$$F_y = V_b(x, y) - v_b(q(\widehat{a}, \widehat{b}), x, y) = v_b(q(x, y), x, y) - v_b(q(\widehat{a}, \widehat{b}), x, y)$$

Conditions $q_a \geq 0$ and $q_b \leq 0$ imply that $q(x, y) \leq q(\widehat{a}, \widehat{b})$. From Assumption 2, we obtain $F_x \leq 0$ and $F_y \geq 0$. Then, since $a < \widehat{a}$ and $b > \widehat{b}$, we have $F(a, b) \geq F(\widehat{a}, \widehat{b})$. That is, (a, b) is IC with $(\widehat{a}, \widehat{b})$. □

Proof of Proposition 4.3. If $(x, y) \in CC(\widehat{a}, \widehat{b})$, then $q(\widehat{a}, \widehat{b}) = q(x, y)$. Therefore, by the taxation principle, $t(\widehat{a}, \widehat{b}) = T(q(\widehat{a}, \widehat{b})) = T(q(x, y)) = t(x, y)$. Because (a, b) is IC with $(\widehat{a}, \widehat{b})$, we have

$$v(q(a, b), a, b) - t(a, b) \geq v(q(\widehat{a}, \widehat{b}), a, b) - t(\widehat{a}, \widehat{b}) = v(q(x, y), a, b) - t(x, y)$$

that is, (a, b) is IC with (x, y). □

Proof of Proposition 4.4. Since (a, b) is IC with $(\widehat{a}, \widehat{b})$ and $(\widehat{a}, \widehat{b})$ is IC with (x, y), we have

$$\begin{aligned} V(a, b) - V(x, y) + v(q(x, y), x, y) \geq& \\ & v(q(\widehat{a}, \widehat{b}), a, b) - v(q(\widehat{a}, \widehat{b}), \widehat{a}, \widehat{b}) + v(q(x, y), \widehat{a}, \widehat{b}) \end{aligned} \tag{12}$$

Additionally, because $v_q(q, \widehat{a}, \widehat{b}) \leq v_q(q, a, b) \; \forall \; q \in Q$ and $q(x, y) \leq q(\widehat{a}, \widehat{b})$, $\int_{q(x,y)}^{q(\widehat{a},\widehat{b})} v_q(q, \widehat{a}, \widehat{b}) dq \;\leq\; \int_{q(x,y)}^{q(\widehat{a},\widehat{b})} v_q(q, a, b) dq$. Then,

$$v(q(\widehat{a}, \widehat{b}), a, b) - v(q(\widehat{a}, \widehat{b}), \widehat{a}, \widehat{b}) + v(q(x, y), \widehat{a}, \widehat{b}) \geq v(q(x, y), a, b) \tag{13}$$

Therefore, from (12) and (13), (a, b) is IC with (x, y). □

Proof of Theorem 4.1. Fix any $(a,b), (\widehat{a},\widehat{b}) \in [0,1]^2$. Let us prove that (a,b) is IC with $(\widehat{a},\widehat{b})$.

If $q(\widehat{a},\widehat{b}) = q^{\text{out}}$ (that is, if type $(\widehat{a},\widehat{b})$ is excluded), we have $V(\widehat{a},\widehat{b}) = 0$, so from the IR constraint $V(a,b) \geq 0$, we can write

$$V(a,b) - V(\widehat{a},\widehat{b}) \geq v(q^{\text{out}},a,b) - v(q^{\text{out}},\widehat{a},\widehat{b})$$

in view of $v(q^{\text{out}},a,b) = v(q^{\text{out}},\widehat{a},\widehat{b})$ by Assumption 1.

If $q(\widehat{a},\widehat{b}) \neq q^{\text{out}}$, since $CC(\widehat{a},\widehat{b})$ is strictly increasing, there are three possible cases:

Case 1 $CC(\widehat{a},\widehat{b})$ intersects $F^{(a,b)}$:
Let (x,y) the point of intersection. Because (a,b) is IC with (x,y) and $(x,y) \in CC(\widehat{a},\widehat{b})$, by Proposition 4.3, (a,b) is IC with $(\widehat{a},\widehat{b})$.

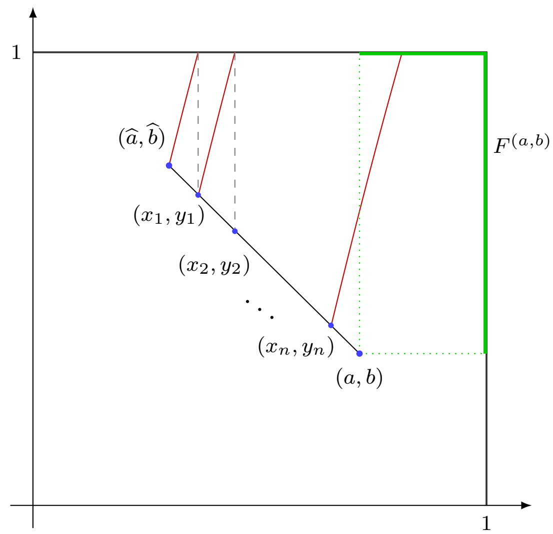
Case 2 $CC(\widehat{a},\widehat{b})$ intersects $\{(1,s) : 0 \leq s \leq b\}$:
Since $CC(\widehat{a},\widehat{b})$ is strictly increasing, then $\widehat{b} < b$. If $\widehat{a} > a$, by Proposition 4.1, we have that $(a,b) \preceq (\widehat{a},\widehat{b})$. Then, (a,b) is IC with $(\widehat{a},\widehat{b})$. If $\widehat{a} \leq a$, consider $(x,y) \in CC(\widehat{a},\widehat{b}) \cap \text{conv}\{(a,b),(1,0)\}$ [8]. Then, (x,y) is such that $x > a$ and $y < b$, and we are in the previous case. That is, (a,b) is IC with (x,y), and by Proposition 4.3, (a,b) is IC with $(\widehat{a},\widehat{b})$.

Case 3 $CC(\widehat{a},\widehat{b})$ intersects $\{(s,1) : 0 \leq s \leq a\}$ (Figure 2 illustrates for this case):
Since $CC(\widehat{a},\widehat{b})$ is strictly increasing, $\widehat{a} < a$. Without the loss of generality, we consider that $\widehat{b} > b$ [9]. Let $(x_1,1) \in CC(\widehat{a},\widehat{b}) \cap \{(s,1) : 0 \leq s \leq a\}$, and $(x_1,y_1) \in \{(x_1,y) : y \in \mathbb{R}\} \cap \text{conv}\{(\widehat{a},\widehat{b}),(a,b)\}$. Note that $q(\widehat{a},\widehat{b}) < q(x_1,y_1)$ and, by Proposition 4.1, $(x_1,y_1) \preceq (a,b)$. Since (x_1,y_1) is IC with $(x_1,1)$ (due to $(x_1,1) \in F^{(x_1,y_1)}$), by Proposition 4.3, (x_1,y_1) is IC with $(\widehat{a},\widehat{b})$. Then, by Proposition 4.4, it will be sufficient that $CC(x_1,y_1) \cap F^{(a,b)} \neq \emptyset$ to conclude that (a,b) is IC with $(\widehat{a},\widehat{b})$. If $CC(x_1,y_1) \cap F^{(a,b)} = \emptyset$, repeat the procedure taking $(x_2,1) \in CC(x_1,y_1) \cap \{(s,1) : 0 \leq s \leq a\}$ and $(x_2,y_2) \in \{(x_2,y) : y \in \mathbb{R}\} \cap \text{conv}\{(x_1,y_1),(a,b)\}$. Similarly to the above, we have $q(x_1,y_1) < q(x_2,y_2)$, $(x_2,y_2) \preceq (a,b)$ and that (x_2,y_2) is IC with (x_1,y_1). Then, by Proposition 4.4, it will be sufficient that $CC(x_2,y_2) \cap F^{(a,b)} \neq \emptyset$ to conclude that (a,b) is IC with (x_1,y_1), and therefore, by Proposition 4.4, that (a,b) is IC with $(\widehat{a},\widehat{b})$. If $CC(x_2,y_2) \cap F^{(a,b)} = \emptyset$, we set up the point (x_3,y_3), and so

[8] $\text{conv}\{(a,b),(1,0)\}$ is the convex hull of these points.

[9] Otherwise, replace $(\widehat{a},\widehat{b})$ for any point in $CC(\widehat{a},\widehat{b})$ on the northwest of (a,b).

Figure 2: Illustration of Theorem 4.1 proof.



on. Note that $\frac{d}{dq}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ and $\frac{d}{da}\left(\frac{v_{qa}}{v_{qb}}\right) \geq 0$ implies $\frac{d}{dr}a_s(r,1) \geq 0$ in view of

$$\frac{d}{dr}a_s(r,1) = \frac{d}{dr}\Big(-\frac{v_{qb}}{v_{qa}}(q(r,1),r,1)\Big) = \Big(\frac{v_{qb}}{v_{qa}}\Big)^2\Big[\frac{d}{dq}\Big(\frac{v_{qa}}{v_{qb}}\Big) \times q_a(r,1) + \frac{d}{da}\Big(\frac{v_{qa}}{v_{qb}}\Big)\Big]$$

That is, the slope of the characteristic curves at the border $(r,1)$ is non-decreasing which guarantees that for large enough n, $CC(x_n, y_n) \cap F^{(a,b)} \neq \emptyset$ because (x_n, y_n) will be close to (a,b) and $CC(x_n, y_n)$ is strictly increasing. Thus, applying Proposition 4.4 n times, we have that (a,b) is IC with $(\widehat{a}, \widehat{b})$. □

Proof of Proposition 4.5.

1. Since $v_q(q,\cdot,\widehat{b})$ is concave and $v_q(q,a,\cdot)$ is convex:

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,\widehat{b}) \leq v_{qa}(q,a,\widehat{b})(\widehat{a}-a)$$

$$v_q(q,a,b) - v_q(q,a,\widehat{b}) \geq v_{qb}(q,a,\widehat{b})(b-\widehat{b})$$

then

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b) \leq v_{qa}(q,a,\widehat{b})(\widehat{a}-a) + v_{qb}(q,a,\widehat{b})(\widehat{b}-b)$$

Besides, if $(a,b) \preceq (\widehat{a},\widehat{b})$ then $0 \leq v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b)$. Therefore

$$\frac{\widehat{b}-b}{\widehat{a}-a} \leq \frac{-v_{qa}(q,a,\widehat{b})}{v_{qb}(q,a,\widehat{b})}$$

in view of $\widehat{a} > a$ and $-v_{qb} > 0$.

2. Since $v_q(q,\cdot,b)$ is concave and $v_q(q,\widehat{a},\cdot)$ is convex:

$$v_q(q,\widehat{a},b) - v_q(q,a,b) \geq v_{qa}(q,\widehat{a},b)(\widehat{a}-a)$$

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,\widehat{a},b) \geq v_{qb}(q,\widehat{a},b)(\widehat{b}-b)$$

then

$$v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b) \geq v_{qa}(q,\widehat{a},b)(\widehat{a}-a) + v_{qb}(q,\widehat{a},b)(\widehat{b}-b)$$

Besides, if $\frac{\widehat{b}-b}{\widehat{a}-a} \leq \frac{-v_{qa}(q,\widehat{a},b)}{v_{qb}(q,\widehat{a},b)}$ with $\widehat{a} > a$ and $v_{qb} < 0$ then $v_{qa}(q,\widehat{a},b)(\widehat{a}-a) + v_{qb}(q,\widehat{a},b)(\widehat{b}-b) \geq 0$ thus $v_q(q,\widehat{a},\widehat{b}) - v_q(q,a,b) \geq 0$ for any $q \in Q$, that is $(a,b) \preceq (\widehat{a},\widehat{b})$

□

Proof of Proposition 4.6. The proof is based on the two following lemmas.

Lemma 1. *Given $(a,b) \in X_n$, $\forall$ $(x,y) \in F^{(a,b)}$, we have*

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y),a,b) - v(\widetilde{Q}^n(x,y),x,y) - O(\tfrac{1}{n-1})$$

That is, since $(a,b) \in X_n$ verifies IC with all points in $\widetilde{F}^{(a,b)} = F^{(a,b)} \cap X_n$, it satisfies a relaxed IC version with all points in the continuous set $F^{(a,b)}$ with some tolerance that is asymptotically zero. The next lemma shows that between any two points on the grid X_n, the same relaxed IC version holds.

Lemma 2. *Given $(a,b),(\widehat{a},\widehat{b}) \in X_n$, we have*

$$V^n(a,b) - V^n(\widehat{a},\widehat{b}) \geq v(Q^n(\widehat{a},\widehat{b}),a,b) - v(Q^n(\widehat{a},\widehat{b}),\widehat{a},\widehat{b}) - O(\tfrac{1}{n-1})$$

We return to the proof of Proposition 4.6. Given $(a,b),(a',b')\in[0,1]^2$, it will be sufficient to prove that

$$\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(\widetilde{Q}^n(a',b'),a,b)-v(\widetilde{Q}^n(a',b'),a',b'))\geq -O(\tfrac{1}{n-1})$$

Let $(\widehat{a},\widehat{b}),(\widehat{a}',\widehat{b}')\in X_n$ be such that $\widehat{a}\leq a<\widehat{a}+\frac{1}{n-1}$, $\widehat{b}-\frac{1}{n-1}<b\leq\widehat{b}$ and $\widehat{a}'\leq a'<\widehat{a}'+\frac{1}{n-1}$, $\widehat{b}'-\frac{1}{n-1}<b'\leq\widehat{b}'$. Let $q=\widetilde{Q}^n(a',b')=Q^n(\widehat{a}',\widehat{b}')$. Since $\widetilde{V}^n(a,b)=V^n(\widehat{a},\widehat{b})$, $\widetilde{V}^n(a',b')=V^n(\widehat{a}',\widehat{b}')$ we have

$$\begin{aligned}\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(q,a,b)-v(q,a',b'))=&\\ &V^n(\widehat{a},\widehat{b})-V^n(\widehat{a}',\widehat{b}')-(v(q,\widehat{a},\widehat{b})-v(q,\widehat{a}',\widehat{b}'))\\ &+v(q,\widehat{a},\widehat{b})-v(q,a,b)+v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\end{aligned}\tag{14}$$

Since $(\widehat{a},\widehat{b}),(\widehat{a}',\widehat{b}')\in X_n$, by Lemma 2,

$$V^n(\widehat{a},\widehat{b})-V^n(\widehat{a}',\widehat{b}')-(v(q,\widehat{a},\widehat{b})-v(q,\widehat{a}',\widehat{b}'))\geq -O(\tfrac{1}{n-1})\tag{15}$$

Besides, since v is differentiable, $v(q,\cdot,\cdot)$ is Lipschitz over $[0,1]^2$ (with constant L), then

$$\left|v(q,\widehat{a},\widehat{b})-v(q,a,b)\right|\leq L||(\widehat{a},\widehat{b})-(a,b)||\leq O(\tfrac{1}{n-1})$$

which implies

$$v(q,\widehat{a},\widehat{b})-v(q,a,b)\geq -O(\tfrac{1}{n-1})\tag{16}$$

Similarly

$$\left|v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\right|\leq L||(a',b')-(\widehat{a}',\widehat{b}')||\leq O(\tfrac{1}{n-1})$$

which implies

$$v(q,a',b')-v(q,\widehat{a}',\widehat{b}')\geq -O(\tfrac{1}{n-1})\tag{17}$$

Therefore, from (14) using (15), (16) and (17) we obtain

$$\widetilde{V}^n(a,b)-\widetilde{V}^n(a',b')-(v(q,a,b)-v(q,a',b'))\geq -O(\tfrac{1}{n-1})$$

□

Proof of Lemma 1. Let $(x,y) \in F^{(a,b)}$ be such that $x = 1$ (case $y = 1$ is analogous), and let $\widehat{b}$ be such that $\widehat{b} - \frac{1}{n-1} < y \leq \widehat{b}$. Since (Q^n, V^n) are the solutions of problem (9), (a,b) satisfies IC with $(1, \widehat{b})$

$$V^n(a,b) - V^n(1,\widehat{b}) \geq v(Q^n(1,\widehat{b}), a, b) - v(Q^n(1,\widehat{b}), 1, \widehat{b})$$

By definition, $\widetilde{Q}^n(x,y) = Q^n(1,\widehat{b})$ and $\widetilde{V}^n(x,y) = V^n(1,\widehat{b})$. Additionally, in view of $(a,b) \in X_n$, we have $\widetilde{V}^n(a,b) = V^n(a,b)$. Then,

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y), a, b) - v(\widetilde{Q}^n(x,y), 1, \widehat{b}) \tag{18}$$

On the other hand, since v is Lipschitz,

$$\left| v(\widetilde{Q}^n(x,y), 1, \widehat{b}) - v(\widetilde{Q}^n(x,y), x, y) \right| \leq L||(1,\widehat{b}) - (x,y)|| = O(\tfrac{1}{n-1})$$

Then,

$$-v(\widetilde{Q}^n(x,y), 1, \widehat{b}) \geq -v(\widetilde{Q}^n(x,y), x, y) - O(\tfrac{1}{n-1}) \tag{19}$$

Therefore, from (18) and (19),

$$\widetilde{V}^n(a,b) - \widetilde{V}^n(x,y) \geq v(\widetilde{Q}^n(x,y), a, b) - v(\widetilde{Q}^n(x,y), x, y) - O(\tfrac{1}{n-1})$$

□

Proof of Lemma 2. If $CC(\widehat{a},\widehat{b}) \cap F^{(a,b)} = (x,y)$, we apply Lemma 1 for (a,b) with (x,y), and considering that $Q^n(\widehat{a},\widehat{b}) = Q^n(x,y)$ and $t(x,y) = t(\widehat{a},\widehat{b})$, we conclude. Other cases are treated analogously as in the proof of Theorem 4.1. □

Proof of Proposition 4.7. Let $(\overline{Q}, \overline{V})$ denote the solution for the continuous problem, and let $(\overline{Q}^n, \overline{V}^n)$ be their restriction on the grid X_n. If (Q^n, V^n) are the solutions of the discretized problem and OPT_n is the optimal value, we have

$$\begin{aligned} OPT_n &\geq \sum_{i=1}^{n}\sum_{j=1}^{n} w(i,j)(v(\overline{Q}^n_{i,j}, a_i, b_j) - \overline{V}^n_{i,j} - C(\overline{Q}^n_{i,j}))f(a_i,b_j) \\ &= \int_0^1\int_0^1 (v(\overline{Q}(a,b),a,b) - \overline{V}(a,b) - C(\overline{Q}(a,b)))f(a,b)\,da\,db - O(\frac{1}{n}) \\ &= OPT^* - O(\frac{1}{n}) \end{aligned}$$

Then, $\liminf_{n\to\infty} OPT_n \geq OPT^*$.

On the other hand, if $\exists$ $\lim_{n\to\infty} \widetilde{Q}^n(a,b)$ and $\lim_{n\to\infty} \widetilde{V}^n(a,b)$ for any $(a,b) \in [0,1]^2$, define

$$\widehat{Q}(a,b) := \lim_{n\to\infty} \widetilde{Q}^n(a,b) \quad , \quad \widehat{V}(a,b) := \lim_{n\to\infty} \widetilde{V}^n(a,b)$$

By Proposition 4.6, $(\widehat{Q}, \widehat{V})$ is feasible. Hence

$$\begin{aligned}
OPT^* &\geq \int_0^1 \int_0^1 (v(\widehat{Q}(a,b),a,b) - \widehat{V}(a,b) - C(\widehat{Q}(a,b)))f(a,b)\,da\,db \\
&= \lim_{n\to\infty} \Big(\int_0^1 \int_0^1 (v(\widetilde{Q}^n(a,b),a,b) - \widetilde{V}^n(a,b) - C(\widetilde{Q}^n(a,b)))f(a,b)\,da\,db \Big) \\
&= \lim_{n\to\infty} \Big(\sum_{i=1}^n \sum_{j=1}^n w(i,j)(v(\widetilde{Q}^n_{i,j},a_i,b_j) - \widetilde{V}^n_{i,j} - C(\widetilde{Q}^n_{i,j}))f(a_i,b_j) + O(\tfrac{1}{n-1}) \Big) \\
&= \lim_{n\to\infty} \Big(\sum_{i=1}^n \sum_{j=1}^n w(i,j)(v(Q^n_{i,j},a_i,b_j) - V^n_{i,j} - C(Q^n_{i,j}))f(a_i,b_j) + O(\tfrac{1}{n-1}) \Big) \\
&= \lim_{n\to\infty} \Big(OPT_n + O(\tfrac{1}{n-1}) \Big)
\end{aligned}$$

where equalities are true by the dominated convergence theorem (each $\widetilde{Q}^n$ and $\widetilde{V}^n$ are bounded), by the finite approximation of the integral, by the definition of $\widetilde{Q}^n$ and $\widetilde{V}^n$, and because (Q^n, V^n) is the solution of the discretized problem. $\square$

References

Araujo, A., Vieira, S., and Calagua, B. (2019). A necessary optimality condition in two-dimensional screening. Working Paper.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica: Journal of the Econometric Society*, pages 51–75.

Armstrong, M. (1999). Optimal regulation with unknown demand and cost functions. *Journal of Economic Theory*, 84(2):196–215.

Barelli, P., Basov, S., Bugarin, M., and King, I. (2014). On the optimality of exclusion in multi-dimensional screening. *Journal of Mathematical Economics*, 54:74–83.

Baron, D. P. and Myerson, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica: Journal of the Econometric Society*, pages 911–930.

Basov, S. (2001). Hamiltonian approach to multi-dimensional screening. *Journal of Mathematical Economics*, 36(1):77–94.

Belloni, A., Lopomo, G., and Wang, S. (2010). Multidimensional mechanism design: Finite-dimensional approximations and efficient computation. *Operations Research*, 58(4-part-2):1079–1089.

Berg, K. and Ehtamo, H. (2009). Learning in nonlinear pricing with unknown utility functions. *Annals of Operations Research*, 172(1):375.

Evans, L. (1998). *Partial Differential Equations*, volume 19 of *Graduate studies in mathematics*. American Mathematical Society.

Judd, K., Ma, D., Saunders, M. A., and Su, C.-L. (2018). Optimal income taxation with multidimensional taxpayer types. Working Paper.

Laffont, J.-J. and Martimort, D. (2001). *The theory of incentives: the principal-agent model*. Princeton university press.

Laffont, J.-J., Maskin, E., and Rochet, J.-C. (1987). Optimal nonlinear pricing with two-dimensional characteristics. *Information, Incentives and Economic Mechanisms*, pages 256–266.

Lewis, T. R. and Sappington, D. E. (1988a). Regulating a monopolist with unknown demand. *The American Economic Review*, pages 986–998.

Lewis, T. R. and Sappington, D. E. (1988b). Regulating a monopolist with unknown demand and cost functions. *The RAND Journal of Economics*, pages 438–457.

McAfee, R. P. and McMillan, J. (1988). Multidimensional incentive compatibility and mechanism design. *Journal of Economic theory*, 46(2):335–354.

Mussa, M. and Rosen, S. (1978). Monopoly and product quality. *Journal of Economic theory*, 18(2):301–317.

Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica: Journal of the Econometric Society*, pages 61–73.

Rochet, J.-C. (2009). Monopoly regulation without the spence–mirrlees assumption. *Journal of Mathematical Economics*, 45(9-10):693–700.

Tarkiainen, R. and Tuomala, M. (1999). Optimal nonlinear income taxation with a two-dimensional population; a computational approach. *Computational Economics*, 13(1):1–16.

Wilson, R. (1996). Nonlinear pricing and mechanism design. *Handbook of computational economics*, 1:253–293.