PREPUBLICACIONES DEL DEPARTAMENTO DE MATEMÁTICA APLICADA
UNIVERSIDAD COMPLUTENSE DE MADRID
MA-UCM 2012-13

A numerical method to solve a duopolistic differential game in a closed-loop equilibrium

J. H. de la Cruz, B.Ivorra and A. M. Ramos

Octubre-2012

http://www.mat.ucm.es/deptos/ma
e-mail:matemática_aplicada@mat.ucm.es

A numerical method to solve a duopolistic differential game in a closed-loop equilibrium.

Jorge H. de la Cruz, Benjamin Ivorra, Angel M. Ramos

Dpt. de Matematica Aplicada,
Universidad Complutense de Madrid,
Pza. de Ciencias,3,28040 Madrid,Spain.

Abstract

In this work, we develope a numerical method to solve infinite time differential games in closed-loop equilibria. Differential games are thought to be run in dynamic decissions and competitive situations, such as marketing investments and pricing policies in a company. Closed-loop equilibria allow us to obtain strategies as a function of ourselves and our competitor. We apply our algorithm to a real data set of two competitive firms. We show how our algorithm is able to develop a different price-advertising strategy to get bigger benefits.

keywords: *differential games; closed-loop equilibrium; time series models; Lotka-Volterra models; Hamilton-Jacobi-Bellman equations; dynamic programming.*

1 Introduction

In this work, we will focus on how much a company should invest - during a fixed period - in television advertising and what should be the optimal price of its goods or services under a competitive environment, assuming a unique competitor, and in a dynamic context.

We are going to develop a rational methodology to optimize television advertising spending and the price of goods or services, taking into account that the other competitor is also rational. So, in the same way as us, he will optimize his spending and price, and it will be done with the aim of compete with us.

The existence of a time variable in this problem, for a given time period- such as a fiscal year-, drives us to a dynamical system. This system will have to capture the most relevant movements in sales, advertising, and price among competitors.

We will answer the question of this work by using both Game Theory and Optimal Control Theory. This approach is called “dynamic games”, concretely “differential games”.

According to [2] dynamic and differential games are based on "static games" and "Optimal Control Theory". Static games deal with decision process in which several players have to take a strategy in non-cooperative or cooperative situations. There are different examples in academic literature, such as the prisoner dilemma [2]. In this example, a static non-cooperative game, two thieves have to decide if they confess or not their crime. They do not know what is the decision of the other. But they know what they will win or lose if they confess or not. Finally, it is straightforward to demonstrate that the equilibrium (or final decision) is not the best for each player, but is not the worse. They try to think what could be the decision of the contrary and then, they try to maximize its decision. This kind of equilibrium will be discussed in the following pages, and it is called a Nash Equilibrium [4].

The birth of Game theory is usually considered around 1944, when Von Neumann and Morgensten published the book: "Theory of Games and Economic Behaviour". But this kind of problems worried lots of years ago to economists like Cournot or Stakelberg [2]. A game could be defined as a conjoint of well-known rules by all the players which determine what actions can be done and what are the pay-offs for taking a concrete action. The most interesting characteristic of a game is that our actions depend on those of the players. So, in order to define a strategy, all the players have to think of the other's decisions.

Optimal control theory could be considered as a decision process under the hypothesis that there is only one agent to take the decision. This decision is based on a maximization (or minimization) of a temporal function (or dynamic function) that reflects preferences, utility or benefit of the agent. If we relax the restriction of one player in Optimal Control theory, we are now in the Dynamic Games field.

The methodology we are going to explore to answer the main question is a field of dynamic games called differential games. It consists on a game in which each player has to optimize its benefits during a time interval, subject to a dynamical system -that reflects market movements- in which the player has to choose the different values of the control variables for all the exercise. Controls are the only variables the players can modify, and it will be the key of our problem. In this work we will explore a concrete kind of model solution, or equilibrium, called

"**closed-loop**"[1] [4],[8]. This solution has a clear interest in Marketing Science because it allows the player to react if the competitor changes its behavior surprisingly. We have worked in this kind of solution because, in general, marketing literature in differential games often consider open-loop equilibrium, for instance, [18]:

"Though the closed-loop differential game is more dynamic and realistic in nature, the solution is much harder to find than the open-loop games because it involves solving systems of partial differential equations [...]. On the contrary,

[1] The other type of equilibrio is called "open-loop". In the section 2 we will develop their definition and properties.

the open-loop games are much easier to solve because they only involve ordinary differential equations. In most cases, researches have to restrict themselves to open-loop games as a proxy to reality. Even with open-loop games, closed-form solutions are hard to find when the number of state variables involved is large".

In this practical case, we are going to use a (confidential) real data set from a company in Spain. We deal with a product with one clear competitor (but in other context of interest, for the sake of simplicity, all the competitors could be reduced to only one). Our company invests with regularity on TV advertising. A GRP, Gross Rating Point **(GRP)**, is the so called TV advertising pressure unit. Our competitor does not invest because it is a *white label* that only competes with promotions and price. Our data base is compound by several variables such as: the weekly sales, that represents the quantity of product units sold in the current week; the weekly competitor sales, which is also the quantity sold by our competitor in the current week; our GRPs and investment in euros per week; and the weekly prices of the good for us and our competitor.

So, in our problem, since our competitor does not advertise on TV, we will have to choose our advertising level and price, while our competitor only has to choose the price level. Those variables form part of the dynamical system we are going to develop in our work.

The main dynamic behavior is explained as follows: we have three levels of variables. First of all are the control variables. In our case, control variables are advertising spending and price. So, the value of those variables has to be decided by the user, in our context, after solving an optimizing problem. The second group of variables of the dynamical system is the state variables. State variables represent how our clients react to our movements in control variables. They can also affect each other in a bivariate sense, because it is reasonable to think that an increase in our sales provokes a decrease in the competitor sales, and vice-versa. Thirdly, the system needs to quantify our pay-offs[2] . They are measured with the benefit function. Benefit is a calculation that involves time, our preferences for having the money as soon as possible, how much we gain from our clients, and how much we spend in catching them.

Once we have defined our framework, we would remark that in our literature revision, we feel that closed-loop equilibrium in differential games is not as developed as open-loop equilibrium. The fact is the referred difficulty to obtain analytical solutions. Thus, in this work we have developed a numerical method, in line with dynamic programming literature, specific to solve marketing problems between two competitors. Then, we will provide an extended example with real data with the objective of being used in the decision making in companies.So, we are also focused not only in the numerical aspects but also in how to implement it.

In this work we will introduce two novel ideas:

1. Estimating state equations by using Lotka-Volterra models, versus the

[2]Game Theory needs to quantify pay-offs in order to have a criteria to choose among alternatives.

usual Lanchester models

2. Adapting dynamic programming algorithms to differential games in infinite periods.

Then, in order to validate our methodology, we consider a numerical benchmark example and compare the obtained results with those given by other methods found in the literature, as in [6],[12],[16]. And we will use real data provided by one important spanish company.

This work will be structured in 5 chapters. In the next chapter, the second one, we will deal with the theoretical model, specifying some concepts of dynamic systems involved and game theory to understand the rest of the work. The third chapter will be devoted to parameter estimation using statistical technics. The fourth section will explain the algorithm developed to answer our main question and its properties. We will show the obtained results in chapter 5, indicating the gain, in terms of benefit, by using our methodology. Finally, chapter 6 concludes.

2 Theoretical Model

2.1 variables and objectives

This section deals with the explanation of the theoretical model. Firstly, we will enounce the most important components of a differential game and the variables involved in the analysis. Then, we will explain three possible dynamical systems that could fit the data *a priori*. Finally, we will set the complete optimization model in a differential game form.

We start by defining a simple deterministic differential game on a time interval $[t_0, T]$. We will consider as in [4]:

1. A set of N players.
2. A vector u of controls with coordinates $u_i(t) \in U_i \subseteq \mathbb{R}^{m_i}$ for each player $i \in I = \{1, ...N\}$, and a vector of state variables$x(t) \in X \subseteq \mathbb{R}^n$. Here, X is called the state space and U_i is the set of admissible controls of player i.
3. The state equations describing the dynamic of the system:
   $\dot{x} = f(x(t), u_1(t), ..., u_N(t), t)$; $x(t_0) = x_0$.
4. A payoff functional for player i, $J_i(u(\cdot); t_0, x_0) = \int_{t_0}^{T} g_i(x(t), u(t), t)dt + S(x(T), T,)$, where function g_i is player i´s instantaneous payoff and function S is a terminal payoff.

In Table (1) , we list the notations and the variables we are going to use in the rest of the document[3]:

[3]Here we provide a short description. As the text advances, it is possible to understand more about notation and parameters meaning.

Table 1: list of variables

variable	description
$x_1(t)$	sales (in units) to our clients
$x_2(t)$	sales(in units) to our competitor's clients
$v_1(t)$	GRPs of our company at time t
$C_i(v_i)$	Cost function with $v_i$as argument
$c_{i,1}(t)$	set of estimated coefficients associated with the advertising cost in our company
$p_1(t)$	price of our good or service at time t
$p_2(t)$	price of our competitor's good or service at time t
ρ_1	Coefficient that measures the impact of GRPs on sales of our company
ϖ_1	Coefficient that measures the effect of our price on our demand
ϖ_2	Coefficient that measures the effect of competitor´s price on its demand
α_1	Our attraction coefficient. It quantifies the new customers that buy our product
α_2	Competitor attraction coefficient. It quantifies the new customers that buy competitor's product
β_{12}	Switching parameter. It quatifies the sensibility of customers of brand 1 (us) to switch to brand 2(competitor)
β_{21}	Switching parameter. It quatifies the sensibility of customers of brand 2(competitor) to switch to brand 1(us)
N_1	Our potential market
N_2	competitor's potential market
TM	$N_1 + N_2$
M	Market Share of one of the companies
m_i	unitary production cost associated with good or service x_i
x_i	sales in units of company i. It only will be used in one example

In order to model the possible dynamical system that could fit well our sales and the competitor´s ones as dependent variables, we have refreshed the baseline models applied in literature [8],[4],[5].[10],[11],[13],[14],[15] and[18]. Generally, the most common model in theoretical research is the "Lanchester model of combat " applied to marketing. As it is pointed in [10] the model could be re-written into marketing notions as:

$$\dot{M} = \rho_1 v_1^{1/2}(1 - M) - \rho_2 v_2^{1/2} M, \tag{1}$$

where M is the market share of one company, understanding Market share as the sales percentage of the company over the total sales of sector, and 1-M is the competitor´s share. v_1 and v_2[4] are the current investments in advertising and $\rho_1$and ρ_2 are two coefficients that measure the impact on market share of advertising efforts[5]. This model has been developed to compete on market share and it does not include prices as other control variables.

[10] introduces a similar model based on [15], but in one case adapted to a theoretical framework where competitors decide the advertising levels and the price competing on sales level (not in market shares). The model, for i=1,2, is defined as follows:

$$\dot{x_i} = \rho_i v_i(t)(TM - x_i(t) - x_j(t))^{1/2} D_i(p_i(t)) \tag{2}$$

Being ρ_i a coefficient to measure the impact of advertising, denoted by $v_i(t)$, on sales increase $\dot{x_i}$. TM is the market size for all the competitors, and $D_i(p_i(t))$ is a demand function that could be specified as a linear or non linear term.

This model could be a good candidate to fit our data. The problem, at first sight, is that we should have a good estimation of TM. Furthermore, this model, being completely multiplicative, does not contemplate the possibility of one variable (for instance $v_i(t)$) having a zero value for all the exercise (as it occurs in the problem considered during this work).

Taking these issues into consideration, we have decided to study another common model in applied research to analyze competitive situations. It is the so-called Lotka-volterra (LV) applied to marketing. We suggest some references such as [11] and [13] applied in telecommunications markets and in the internet market, respectively, with the aim to define strategic behavior. In a general sense, a LV model is used to manage population dynamics, as it is set in Equations (3), and (4):

$$\dot{x}_1 = \left[\alpha_1 \left(1 - \frac{\alpha_1 x_1}{N_1}\right) - \beta_{12} x_2\right] x_1 \tag{3}$$

$$\dot{x}_2 = \left[\alpha_2 \left(1 - \frac{\alpha_2 x_2}{N_2}\right) - \beta_{21} x_1\right] x_2 \tag{4}$$

In Biology literature, it is usual to consider $x_1$and $x_2$as different species of predators and preys. Understanding Marketing environment as a derivative of a problem of population dynamics, we can define, for instance, the sales of each product given by x_1 and x_2. In this context, where β_1 and β_2 appear as negative coefficients[6] in Equation (3) and Equation (4) the two variables

[4]We have included in this only case this variable, because is required by original model notation.

[5]the increase experimented by market share $\dot{M}$ is proportional to ρ_1 (and is nonlinear since it depends on the level of M at this time point).

[6]Could have any sense the increase in competitor sales causing an increase in our sales, for instance β_1 or β_2 have possitive coefficients? Perhaps this could be due to any exogenous factor that affects two variables to have a common trend, such as fashion or consumer propension.This factor should be isolated.

could be considered in populations dynamics as predators (and, of course, one is the prey of the other). In marketing context both variables will be called as *duopolistic competitives*. The way a "predator eats" in our model is nonlinear. For instance, the more the preys are in the market, the more the preys ''eat'' the predator. This is the key of the multiplicative functional terms between $x_1$and x_2 preceded by β_1 and β_2. Also, LV has a nonlinear (logistic concretely) growth director term: $\alpha_1 \left[1 - \frac{\alpha_1 x_1}{N_1}\right]$, with a market size, N_1, that reflects a natural growth with stronger increases in the begining and lower increases as far as we reach the N_1value.

When a LV model is applied with control purposes, in literature [17] it is common to include new variables (the controls) in a similar scheme than the latter model. In this context, we have decided to include the control variables in the traditional structure of LV models, preserving the non-linear relationship among state variables and control variables. In the following, this model will be referred as LV-1 :

LV-1

$$\dot{x}_1 = \left[\alpha_1 \left[1 - \tfrac{\alpha_1 x_1}{N_1}\right] - \beta_{12} x_2 + \rho_1 v_1 - \varpi_1 p_1\right] x_1$$
$$\dot{x}_2 = \left[\alpha_2 \left[1 - \tfrac{\alpha_2 x_2}{N_2}\right] - \beta_{21} x_1 - \varpi_2 p_2\right] x_2$$

This will be our baseline model. But, in order to discuss if there exist similar models appropriated to our problem, we will study two other models. In the one hand, we will allow this functional form to catch quadratic effects on the prices and advertising[7]:

LV-2

$$\dot{x}_1 = \left[\alpha_1 \left[1 - \tfrac{\alpha_1 x_1}{N_1}\right] - \beta_{12} x_2 + \rho_1 v_1 - \rho_{11} v_1^2 - \varpi_1 p_1 + \varpi_{11} p_1^2\right] x_1$$
$$\dot{x}_2 = \left[\alpha_2 \left[1 - \tfrac{\alpha_2 x_2}{N_2}\right] - \beta_{21} x_1 - \varpi_2 p_2 + \varpi_{22} p_2^2\right] x_2$$

A quadratic effect in this model represents the possibility of reach a point where an increase/decrease of the variable value has no more effect on sales. That uses to be called as *saturation effect.*Under this hypothesis, we will contrast in the next chapter if there are saturation effects in advertising and in price. The last model, proposed by [5] is an empirical Lotka Volterra model, based on a linear approximation, using Vectorial Autoregressive with exogenous variables models (VARX) imported from Econometric Theory. A VARX [7] is a stochastic system similar to previous model, but where the relationship among variables is linear and discrete:

VARX

[7]Under the hypothesis agents do not react to increases in price and advertising in the same way independent on the size of the increase.

$x_1(t+1) = \alpha_{01} + \alpha_1 x_1(t) - \beta_{12} x_2(t) + \rho_1 v_1(t+1) - \varpi_1 p_1(t+1) + \varepsilon_1(t+1)$
$x_2(t+1) = \alpha_{02} + \alpha_2 x_2(t) - \beta_{21} x_1(t) - \varpi_1 p_2(t+1) + \varepsilon_2(t+1)$

Where $\begin{bmatrix} \varepsilon_1(t+1) \\ \varepsilon_2(t+1) \end{bmatrix}$ is a white noise vector with the following distribution:
$N\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} var(\varepsilon_1(t+1)) & cov(\varepsilon_1(t+1), \varepsilon_2(t+1)) \\ cov(\varepsilon_1(t+1), \varepsilon_2(t+1)) & var(\varepsilon_2(t+1)) \end{bmatrix}\right].$
In the following, in order to advance in the explaining of our methodology framework, we will simplify the dynamical system notation[8] as:

$\dot{x}_1 = g_1(x_1, x_2, v_1, p_1)$
$\dot{x}_2 = g_2(x_1, x_2, p_2)$

Leaving for the next section the correspondent discussion about what is the most adequate model to represent the system dynamics.

Each firm chooses its advertising and price level to maximize its discounted infinite-horizon profit, given by [10]:

$$J_1 = \int_0^\infty e^{-r_1 t}\left((p_1 - m)\, x_1(t) - C(v_1(t))\right) dt \tag{5}$$

$$J_2 = \int_0^\infty e^{-r_2 t}\left((p_2 - m)\, x_2(t)\right) dt \tag{6}$$

$$\dot{x}_1 = g_1(x_1, x_2, v_1, p_1) \tag{7}$$

$$\dot{x}_2 = g_2(x_1, x_2, p_2) \tag{8}$$

Player 1 has as control variables price and advertising ($p_1$and v_1 in Equation (5)). He has to decide, for all the time periods, the policy values of p_1, v_1 that will affect the sales as state variable ($x_1$in Equation (7)). But its sales will be affected also by the sales of player 2 (x_2 in Equation (7)). Every period, player 1 will receive a payoff that is registered in the benefit function (J_1 in Equation (5)). In the meanwhile, player 2 only can choose a pricing policy p_2, (in Equation (8)). This price will affect its sales x_2, as a state variable, which also will be affected by player 1 sales ($x_1$in Equation (8)). Finally, Player 2's payoffs will be computed in the benefit function (J_2in Equation (6)).

As explained previously, in this work we will focus in obtaining closed-loop strategies. If we allow the control function to be dependent on time and state variables (i.e., $u(x(t);t)$), as done here, we talk about a Markovian strategy, or also known as closed-loop strategy or feedback . Whereas in the case of a solely time-dependent control(i.e., $u(t)$) we consider an open-loop strategy. Next, following [4], we will introduce some concepts that we will use in the rest of this work.

[8]up to this moment, we don´t know what of the three systems is more adequate.

2.2 Nash equilibrium and Markov-Nash equilibrium, time consistency and subgame perfectness

As introduced in the first chapter, a Nash equilibrium is a solution of a game involving two or more players, in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only his own strategy unilaterally.

The solutions we are going to obtain in our framework can be included into the concept of Nash equilibrium, as we will see in this section. In a differential game if all opponents of player i use the so-called markovian strategies $u^j(t) = \theta^j\left(x(t), t\right), j \neq i$ where θ denotes a strategy, then player i faces a control problem of the form discussed previously. If we slightly modify our notation, the payer i's decision problem could be rewritten as:

$$\max J^i_{\theta^{-i}}\left(u^i\left(\cdot\right)\right) = \int_o^T e^{-r_i t} F^i_{\theta^{-i}}\left(x(t), u^i(t), t\right) dt \tag{9}$$

st:

$$\dot{x}(t) = f^i_{\theta^{-i}}\left(x(t), u^i(t), t\right) \tag{10}$$

$$x(0) = x_0 \tag{11}$$

$$u^i(t) \in U^i_{\theta^{-i}}\left(x\left(t\right), t\right) \tag{12}$$

where:

$$F^i_{\phi^{-i}}\left(x(t), u^i(t), t\right) = F^i(x, \theta^1(x,t), ...\theta^N(x,t), t)$$
$$f^i_{\phi^{-i}}\left(x(t), u^i(t), t\right) = f^i(x, \theta^1(x,t), ...\theta^N(x,t), t)$$
$$U^i_{\phi^{-i}}\left(x(t), t\right) = U^i(x, \theta^1(x,t), ...\theta^N(x,t), t)$$

Now it is possible to define a Markovian Nash equilibrium for the differential game.

Definition 2.1 We call a Markovian Nash equilibrium as the N-tuple $(\theta^1, \theta^2, ..., \theta^N)$ of functions $\theta^i : X \times [0, T]$ in $\mathbb{R}^m$, if for each i∈{1,2,. . . N} an optimal control path $u^i(\cdot)$ of the problem described by (Equations 9-12) exists and is given by the Markovian strategy $u^j(t) = \theta^j\left(x(t), t\right), j \neq i$.♣

This definition shows that finding a Markovian Nash equilibrium of an N-player differential game is equivalent to find Markovian strategies for the solutions of a system of N interdependent optimal control models.

In the discussion between open-loop solutions and closed-loop, [4] points out that every open loop strategy is, by definition, also a degenerated Markovian Strategy, which means that every open-loop equilibrium is also a Markovian Nash equilibrium. Thus, the set of open-loop equilibria of a particular game is a subset of the set of all Markovian Nash equilibria. In general, as said in this previous article, it is a proper subset.

However, some properties remarked in the literature of differential games, could help us to characterize better the properties of each equilibrium. In the following, we will study time consistency and subgame perfectness.

[4] also points that there is no reason to think that one of the equilibrium is better than others. It depends on the specific situation one want to depict. Nevertheless, it is adequate to study additional characteristics. To do so, we modify the previous model (eq 9-12) as follows:

We define a subgame $\Phi(x,t)$ by replacing the objective function for player i and the system dynamics by:

$$\int_t^T e^{-r_i(s-t)} F^i\left(x(s), u^i(s), s\right) ds \tag{13}$$

$$\dot{x}(s) = f^i_{\theta^{-i}}\left(x(s), u^i(s), s\right) \tag{14}$$

$$x(t) = x \tag{15}$$

Hence,$\Phi(x,t)$ is a differential game defined on the time interval $[t,T]$ with initial condition x(t)=x. So, we can next define time consistence.

Definition 2.2 Time consistency is defined if, for each $t \in [0,T\rangle$ the subgame $\Phi(x,t)$ admits a Markovian Nash equilibrium $\left(\phi^1,...,\phi^N\right)$ such that $\phi^i(y,s) = \theta^i(y,s)$ holds for all $i \in [1,2,...,N]$ and all $(y,s) \in X \times [t,T\rangle$.♣

In [4], it is demonstrated that any Markovian Nash equilibrium of a differential game is time consistent. Time consistency could be seen as a minimal requirement for the credibility of an equilibrium strategy. The idea is if player i has incentive to deviate from his strategy during the time interval, then the other players will not believe his announcement and will compute their own strategies by taking into account the expected future deviation of player i that, in general, will be different from the initial set.

As it is explained in [5] the closed-loop equilibrium we are looking for in this work has the properties of time consistency and subgame perfectness.

Definition 2.3 Let a Markovian Nash equilibrium $\left(\phi^1,...,\phi^N\right)$ be a Markovian Nash equilibrium for the game $\Phi(x_0,0)$. We call the equilibrium subgame perfect if, for each $(x,t) \in X \times [0,T]$, the subgame $\Phi(x,t)$ admits a Markovian Nash equilibrium $\left(\phi^1,...,\phi^N\right)$ such that $\phi^i(y,s) = \theta^i(y,s)$ holds for all $i \in [1,2,...,N]$and all $(y,s) \in X \times [t,T]$. A Markovian Equilibrium which is a subgame perfect is also called a Markov perfect Nash equilibrium.♣

As pointed out in [5], in contrast to time consistence, subgame perfectness not only requires that the restriction of a subgame to be a Markovian Nash equilibrium, but all the subgames are in a Markovian Nash equilibrium. It is possible to show[5] that open-loop equilibrium is not subgame pefect while closed-loop is. The other main difference between this kind of equilibria is how they describe the reality. As we have pointed, open-loop cannot observe how our competitors react to changes in the state of the game, but closed-loop can do it.

Finally, a way of solving a feedback equilibrium or closed-loop needs to specify a Hamilton Jacobi Bellman equation (HJB) [4],[8],[15]. In our case, this implies to solve for each player a differential equation in partial derivatives as follows:

$$r_i V_i = \max_{tv_i, p_i} \left[\left(\left(p_i - m_i \right) x_i - C(v_i) \right) + \nabla V_i \left(g_i(x_i, x_j, v_i, p_i) \right] \right. \tag{16}$$

where V_i is the value function (or objetive function) of player i $i \neq j$, and r_i is a time discount parameter of player i, as is defined in the literature. The way we are going to obtain the optimal strategies per player is based on numerical methods derived from Dynamic programming.

3 Parameter Estimation

In this section we will develop the parameter estimation of our problem. This section is divided in two main parts. First, we will be devoted to estimate the parameter values for each state equation.

3.1 Estimating State Equations

As we developed in Section 2, we will estimate three possible versions of these dynamic systems. The first of them is a LV model with two exogenous inputs (the control variables). The second version will allow a quadratic functional form for control variables. And the third model version is a possible linearization of the LV and is well-known in Econometric Literature [7] as a VARX model.

Due to our data is of discrete nature, we will set h=1 in the discretization scheme. So, we approximate $\dot{x} \approx \frac{x(t+1)-x(t)}{h}$. Thus, according to [5] we will consider this operation when it is necessary: $\tilde{x}(t+1) = \frac{x(t+1)-x(t)}{x(t)}$.

In the following, we present the different funtional forms we are going to estimate, LV1,LV2 and VARX models:

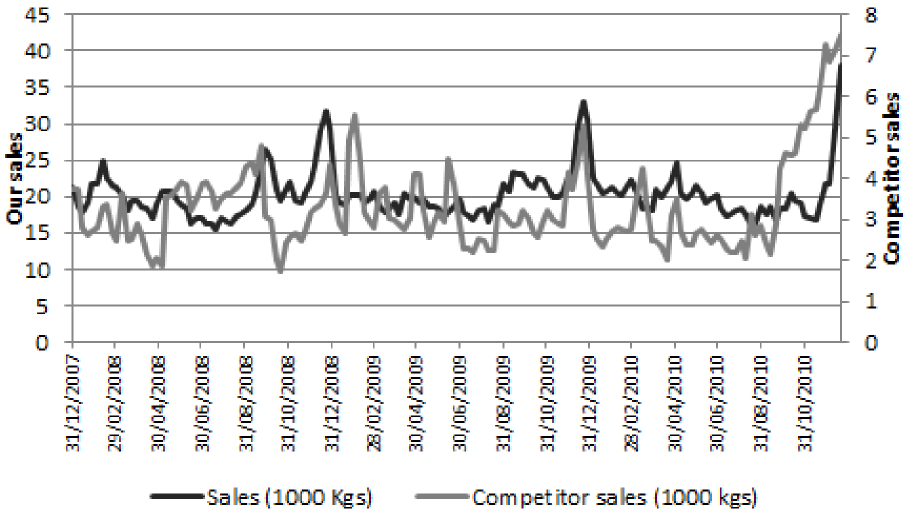
LV-1

$\tilde{x}_1(t+1) = \alpha_1 \left[1 - \frac{\alpha_1 x_1(t)}{N_1} \right] - \beta_{12} x_2(t) + \rho_1 v_1(t+1) - \varpi_1 p_1(t+1) + \varepsilon_1(t+1)$
$\tilde{x}_2(t+1) = \alpha_2 \left[1 - \frac{\alpha_2 x_2(t)}{N_2} \right] - \beta_{21} x_1(t) - \varpi_2 p_2(t+1) + \varepsilon_2(t+1)$

LV-2

$\tilde{x}_1(t+1) = \alpha_1 \left[1 - \frac{\alpha_1 x_1(t)}{N_1} \right] - \beta_{12} x_2(t) + \rho_1 v_1(t+1) - \rho_{11} v_1^2(t+1) - \varpi_1 p_1(t+1) + \varpi_{11} p_1^2(t+1) + \varepsilon_1(t+1)$
$\tilde{x}_2(t+1) = \alpha_2 \left[1 - \frac{\alpha_2 x_2(t)}{N_2} \right] - \beta_{21} x_1(t) - \varpi_2 p_2(t+1) + \varpi_2 p_2^2(t+1) + \varepsilon_2(t+1)$

Figure 1: Our sales and competitor's during the time period considered



VARX

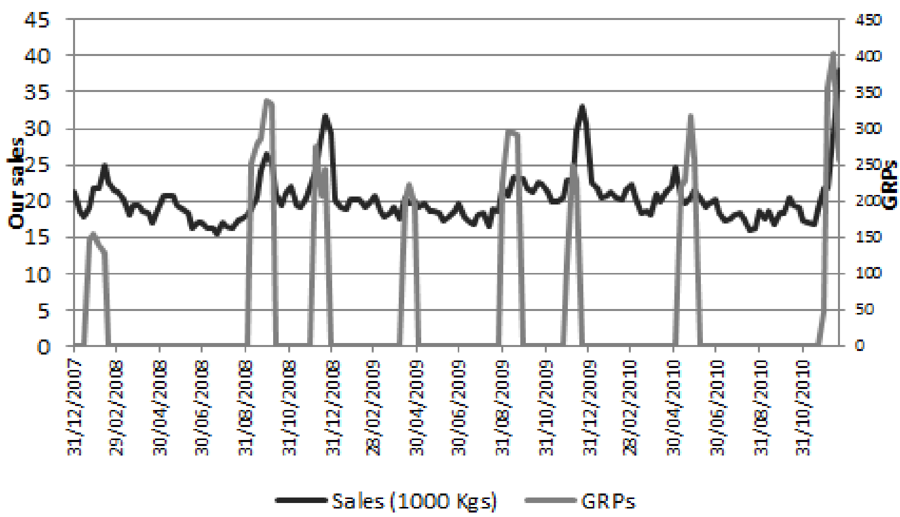
$$\tilde{x}_1(t+1) = \alpha_{01} + \alpha_1\tilde{x}_1(t) - \beta_{12}\tilde{x}_2(t) + \rho_1 v_1(t+1) - \varpi_1 p_1(t+1) + \varepsilon_1(t+1)$$
$$\tilde{x}_2(t+1) = \alpha_{02} + \alpha_2\tilde{x}_2(t) - \beta_{21}\tilde{x}_1(t) - \varpi_1 p_2(t+1) + \varepsilon_2(t+1)$$

where, in line with our methodology to estimate parameters (least squares estimation, LSE), we will impose, as an hypothesis for the stochastic terms, that $\begin{bmatrix} \varepsilon_1(t+1) \\ \varepsilon_2(t+1) \end{bmatrix}$ is a white noise vector with the following distribution: $\mathrm{N}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} var(\varepsilon_1(t+1)) & cov(\varepsilon_1(t+1), \varepsilon_2(t+1)) \\ cov(\varepsilon_1(t+1), \varepsilon_2(t+1)) & var(\varepsilon_2(t+1)) \end{bmatrix}\right]$. The reason of doing so is to ensure that the model is able to catch the most important movements of variables $\tilde{x}_1$ and $\tilde{x}_2$. If not, it will appear an autocorrelation structure [7] in the residual term $\begin{bmatrix} \varepsilon_1(t\hat{+}1) \\ \varepsilon_2(t\hat{+}1) \end{bmatrix}$, it is said, residual non independent and identically distributed as stated above.

Figure 1, shows the relationship between our sales and the competitor sales. Both show a stationary distribution around a mean value. They also have an important seasonal component, with great picks in December. In the estimation procedure, we will have to isolate seasonality, in order to obtain white noise residuals and efficient estimation of parameters.

It is also possible to see some relationship between our sales and the GRPs investment. The next Figure 2 shows a kind of non linear effect between GRP

Figure 2: our sales and advertising effort in GRPs



and sales, as predicted by the LV theory. We can see that when we invest on advertising, we obtain additional sales. However, there is not a clear direct relationship (in which lower levels of advertising report lower levels of sales), so this situation could be an evidence of a non linear relationship.

Before estimating any model, we analyze in detail the data set. First, we relate our sales in time t and the competitor sales in time t-1. We make this comparison because discrete state equations in LV model are Markovians equations, it is said, as we can see in models LV1 and LV2, the future state depends only of the current state. Here we will show two different kind of relationship. Figure 3 (left) represents the level of both variables. This relationship is positive, and it represents a counterintuitive correlation. We should expect a negative relationship between our sales and competitor sales, as our theoretical models set. But since our hipothetized market model LV is non-linear, our sales (in percentage change) will be affected by the level of competitor sales in the previous moment. In this case, the relationship showed by Figure 3 (right) shows some evidences in favor of our theoretical model

Regarding the price effects, we can see in Figure 4 a negative effect between our percentual changes in sales and the current price, perhaps with a little evidence of a non linear behavior.

Similar results could be found in the relationship between percentage increase in competitor sales in time t versus our sales in time t-1, and perhaps a non linear relationship with their price (figure 5).

Least squares estimation of parameters [7] will be obtained next. We will

Figure 3: our sales (in level and in % change)versus competitor's sales

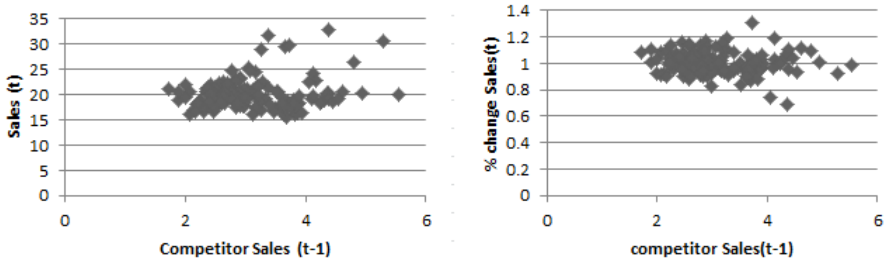


Figure 4: our sales (in % change) versus our price

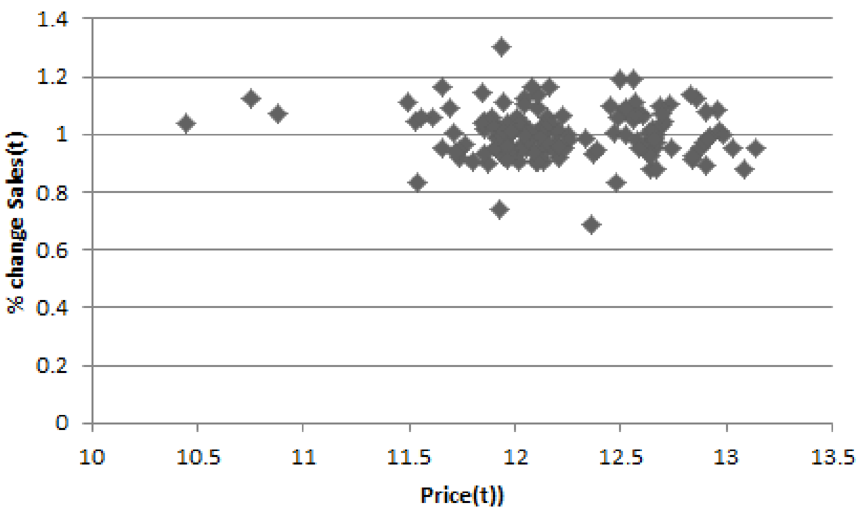
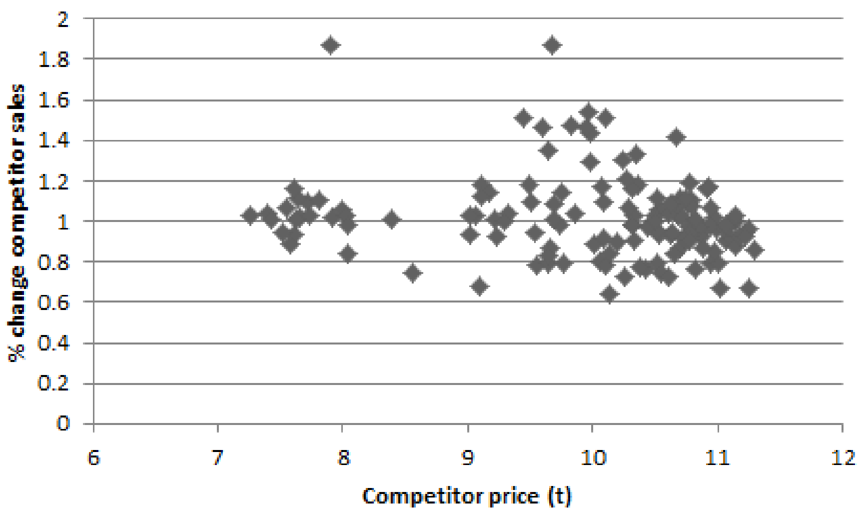


Figure 5: competitor sales (in % change) versus competitor price



estimate parameters and make some validation test to discuss about how likely this estimated model is in line with the theoretical approach. Concretely, this validation process will be[9] :

- Structural test. To answer the question: Are the model parameters constant over time? We will use Qusum test . [7]
- Autocorrelation test: It will be used to see in what level the estimated model provides i.i.d[10]. residuals. [7],[5]
- And we will discuss about how the data fit the theory.

In the Table (2) we provide all the coefficient estimations with their associated standard deviation and some fit statistics such as R squared, Akaike information Criteria and Schwartz Criteria. We also provide Durbin-Watson autocorrelation test.

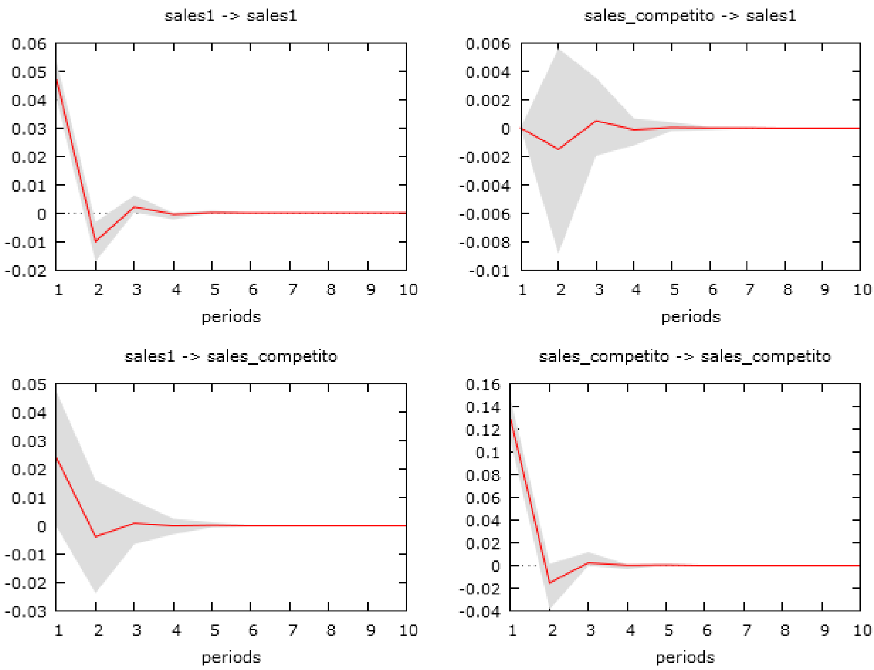
We observe that LV2 model appears to fit better the data . Equation 1 of LV2 allows quadratic effects in price but not in advertising (it was not statistically significative). Furthermore, equation 2 in LV2 also incorporates a quadratic effect in price. Moreover, parameters do not change significantly across equations, excepting intercept parameter in LV-2 (due to the inclusion of a quadratic value for the price). We would like to point out that we have estimated this model filtering the seasonality of the data. That means that we have included

[9]In Appendix A we present this statistic test.
[10]Identical and independent distributed

Table 2: Estimated coefficientes for LV-1,LV-2 and VARX models

	LV1 model		LV2 model		VAR MODEL	
	Equation 1	Equation 2	Equation 1	Equation 2	Equation 1	Equation 2
Dependent Variable	$\frac{x_1(t+1)-x_1(t)}{x_1(t)}$	$\frac{x_2(t+1)-x_2(t)}{x_2(t)}$	$\frac{x_1(t+1)-x_1(t)}{x_1(t)}$	$\frac{x_2(t+1)-x_2(t)}{x_2(t)}$	$\frac{x_1(t+1)-x_1(t)}{x_1(t)}$	$\frac{x_2(t+1)-x_2(t)}{x_2(t)}$
α_{01}					0.695	1.285
std deviation					0.201	0.557
α_{02}						
std deviation						
α_1	0.690		5.410		-0.200	-0.042
std deviation	0.146		2.070		0.096	0.267
α_2		1.540		3.732		
std deviation		0.348		1.320		
$\frac{\alpha_1}{N}$	-0.022		-0.022			
std deviation	0.003		0.003			
$\frac{\alpha_2}{M}$		-0.083		-0.046		
std deviation		0.017		0.017		
β_1	-0.032		-0.034		-0.016	-0.115
std deviation	0.008		0.007		0.037	0.102
β_2		-0.029		-0.040		
std deviation		0.009		0.009		
ρ_1	0.0004		0.00035		0.0002	-0.0003
std deviation	0.0001		0.00001		0.0001	0.0002
ρ_{11}						
std deviation						
ω_1	-0.023		-0.800		-0.032	-0.061
std deviation	0.010		0.342		0.013	0.037
ω_2		-0.047		-0.534	-0.014	-0.037
std deviation		0.014		0.260	0.006	0.016
ω_{11}			0.032			
std deviation			0.014			
ω_{21}				0.029		
std deviation				0.014		
R-squared	0.775	0.538	0.790	0.553	0.720	0.480
Dwatson	1.899	1.834	1.900	1.902	1.900	1.830
AIC	-423.048	-73.220	-435.760	-75.213	-420.760	-72.000
SC	-252.617	61.490	-262.289	60.120	-262.289	62.150

Figure 6: Impulse response graphics of estimated VARX. (left -down:response of our sales to a competitor possitive shock. Rigth-up:response of competitor´s sales to a possitive shock of us).



51 dummy variables to catch the seasonal cycle that the rest of the explicative variables could not explain. Doing so, we increase the precision of the models and we ensure no autocorrelation in residuals, as we will see in the following.

In addition, we will depict in Figure (6) an impulse-responde graph obtained by VAR modeling. This graphs are habitual in Econometrics Empirical research [5] because they help us to understand estimated parameters. Don´t forget that in LV1 and LV2 as in VARX model the explained variables also appear as explicative variables in the opposite equations. So, impulse responses help us to understand how one impact in the disturbance term of one variable can affect the endogenous variables (i.e., one change in one variable also affects, due to the dynamic system, the other variables). We can interpret this result as if the impact is slight the LV models can be approximated by a linear model (as a VARX):

As we can see in Figure (6) -left, the impulse response is the effect of our sales to a positive shock in competitor sales. The effect is negative but confidence bounds seems to about not a negligible effect. Figure (6)-rigth shows the impulse response of competitor sales due to a positive shock in our sales. The effect is clearly positive. That reflects our previous descriptive work, where we found that the negative effect is only found in a non linear relationship. So, this could be an evidence that the linear model is not the most accurate to reflect a duopolistic competition.

To conclude, we will present in Figure (7) and Figure (8) two fit graphs. In all them, we show in red the real variable, and in the other colors, are represented the fit ability of each model. For the sake of readability, we present some interesting sub-sample zoom.

Regarding to the correlation among different fittings and the real observation, In equation 1, LV1 and LV2 have similar correlation coefficient among reality and fit (near 80%) the worst model is VARX (with less than 70%). Is the same case in equation 2, due to correlation coefficient among models and reality is near 75% in LV but aroung a 70% in VARX model.

In the next chapters we will use LV2 model results[11] due to its non-linear price effect on state equations. This property will allow us to obtain a more rich optimization model.

3.2 Estimating cost functions

As in [1],[10], the advertising market benefit function is of the form :

$$\Pi_1(x_1(t), u_1(t)) = (p_1 - m_1)x_1(t) - C(v_1(t)) \tag{17}$$

Since p_1 is well-known and we have decided to restrict m=0, we focus on $C(v_1(t))$. As is refered in some references [1],[10],[14] in order to operate in the model, theoretic approach uses to work with a quadratic function, such as:$C(v_1(t)) = c_1 v_1 + c_2 v_1^2$ or $C(v_1(t) = c_1 v_1^2$.

The main problem with quadratic functional forms is the fact that the more the GRPs I invest up to the inflexion point, the more the cheaper could become (in this case, we could have the same cost for 800 GRPs than for 200). And our data only shows the behavior of advertising cost till 400 GRPs. So, logarithmic functional form could be more interesting, due to its functional form, representing a saturation point. Least squares estimation [1] provides this equation :

$$C(tv_1) = 26917 log(v_1 + 1) \tag{18}$$

with R-squared of 95%[12].

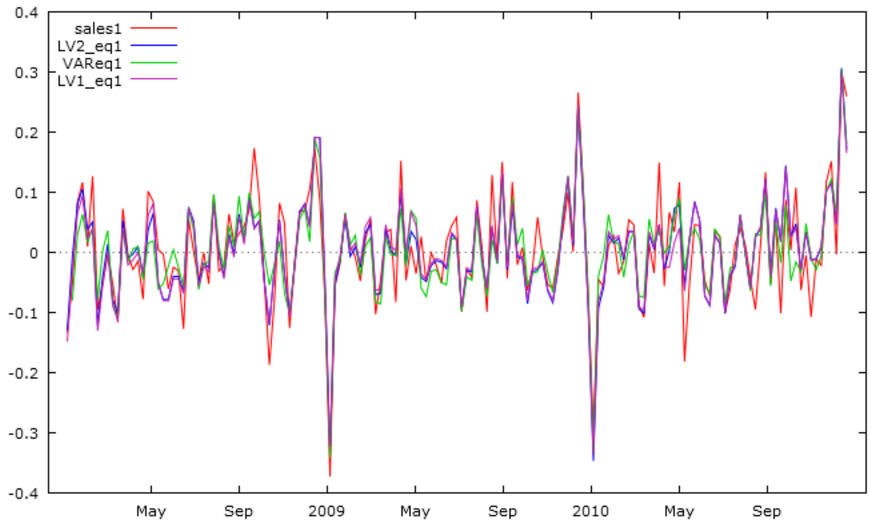
In Figure (10) we can see the fit ability of the model to catch the nonlinearity between GRPs invested and its cost. Following previous estimation, our cost-benefit function will be:

$$\Pi_1 = p_1 x_1(t) - 26917 log(v_1 + 1) \tag{19}$$

[11] LV1 performs in a similar way than LV2, but we preffer the latter for the optimization purposes.

[12] we use ln(x+1) due to the existence of zero values in the variable

Figure 7: Graphs of model performance, model for our sales



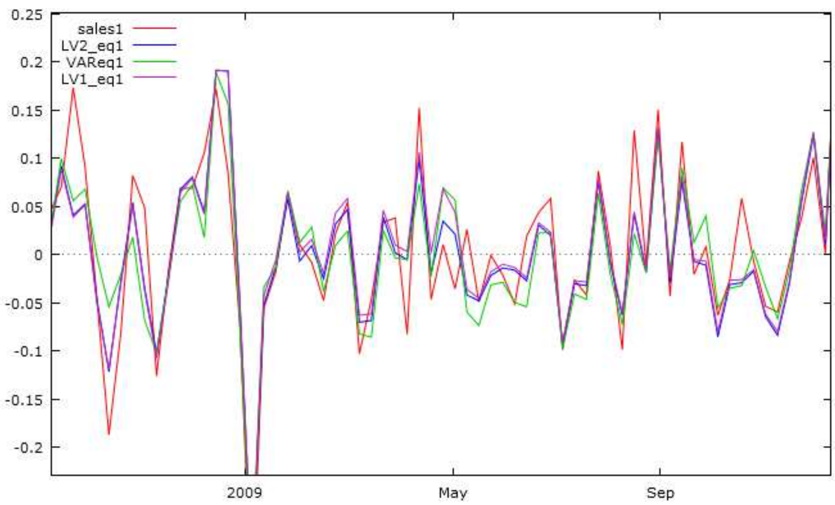
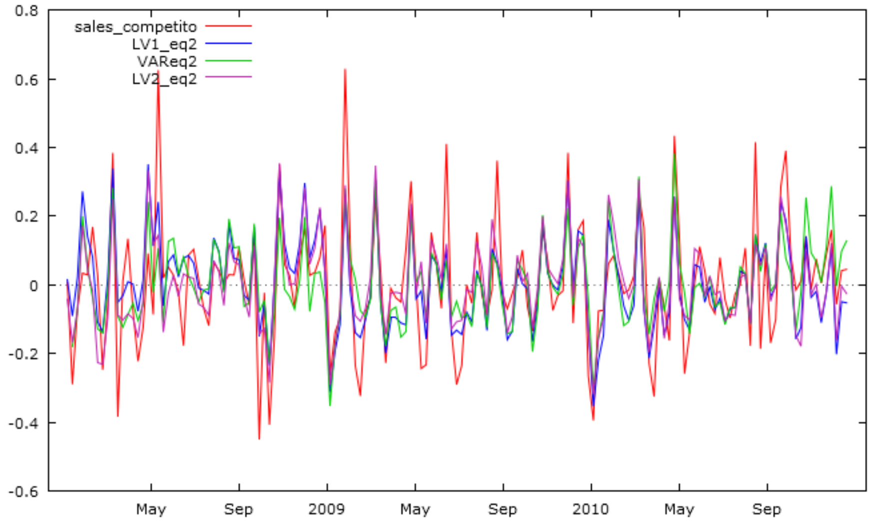


Figure 8: Graphs of model performance, model for competitor´s sales



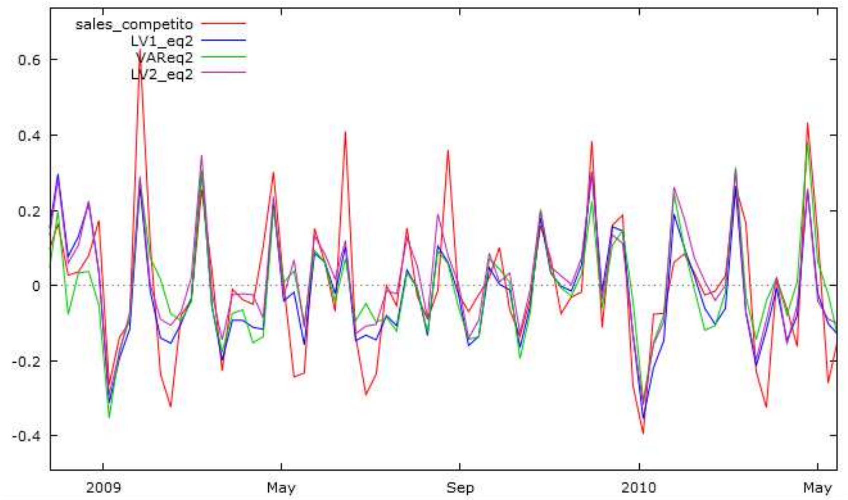


Figure 9: Relationship between cost(real investment of our company) and GRPs

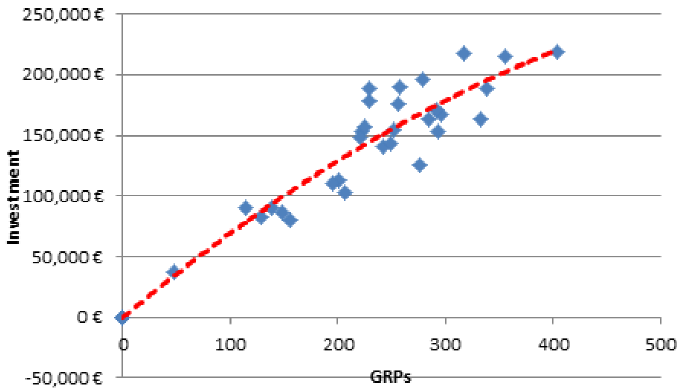
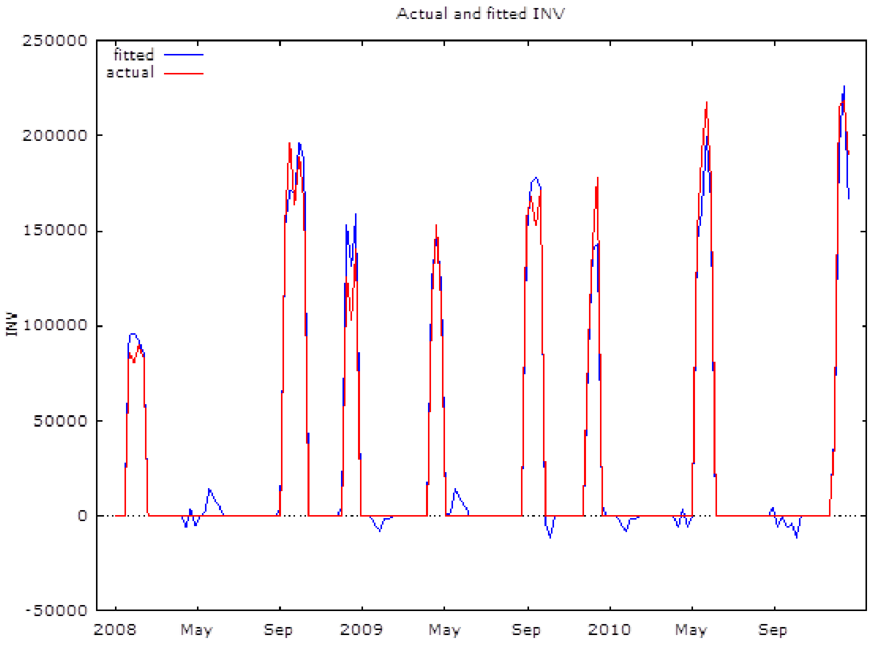


Figure 10: Fitting of cost function



In the case of competitor company, since it has not advertising costs, this should be its benefit function:

$$\Pi_2 = p_2 x_2(t) \tag{20}$$

In this section we have obtained parameter values for all the necessary functions to apply our optimization algortithm. We remark that in the next chapters we will rely on a state equation type Lotka-Volterra with quadratic effects on price. The cost function in our company will relate GRPs with its real cost by a logartihmic function representing a kind of saturation point as the empirical evidence shows in this case.

4 Dynamic Programming algorithm

In this section, we provide a bibliographic revision about dynamic programming to solve optimal control problems. We recall that our objective is to solve a problem of the form:

$$\max_{v_1,p_1} J_1 = \int_0^{\infty} e^{-r_1 t} \left((p_1 - m)\, x_1(t) - C(v_1(t)) \right) dt \tag{21}$$

$$\max_{p_2} J_2 = \int_0^{\infty} e^{-r_2 t} \left((p_2 - m)\, x_2(t) \right) dt \tag{22}$$

$$\dot{x}_1 = g_1 \left(x_1, x_2, v_1, p_1 \right), \ x_1(0) = x_{1,0} \tag{23}$$

$$\dot{x}_2 = g_{12} \left(x_1, x_2, p_2 \right) \ x_2(0) = x_{2,0} \tag{24}$$

Using a compact notation, our objective is to evaluate:

$$\max_{u_i(t)} J_i = \int\limits_0^{\infty} f(s, x(s), u(s); \alpha) ds \tag{25}$$

st:

$$\dot{x}(s) = g(s, x(s), u(s); \alpha), \ x(0) = x_0 \tag{26}$$

where $u(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}^{\boldsymbol{m}}$ is the control function (in our case is componed by the variables (v_1, p_1, p_2), concretely $u_1 = u(v_1, p_1)$ and $u_2 = u(p_2)$, $x(\cdot) : \mathbb{R} \longrightarrow \mathbb{R}^{\boldsymbol{n}}$is the state function (in our case, formed by x_1, x_2), $\alpha \in\mathbb{R}^{A}$ is the vector of exogenous and constant parameters,$x_0 \in\mathbb{R}^{n}$ is the initial state. We can define the optimal value function $V(\cdot)$as the maximum value objective function that can be obtained starting at any time $t \in [0, T]$ and in any admissible state $x(t)$ given the parameter vector α. More precisely, we have the following definition of $V(\cdot)$:

$$V_i(\alpha, t, x_t) = \max_{u_i(t)} \int_0^\infty f(s, x(s), u(s); \alpha) ds \tag{27}$$

st:

$$\dot{x}(s) = g(s, x(s), u(s); \alpha),\ x(0) = x_0 \tag{28}$$

Once we have defined the value function, we can solve the considered problem. As said in Chapter 2, we should obtain an analytical solution by solving the associated Hamilton-Jacobi-Bellman equation (see Equation (16)). But it is necessary to point out that we have to optimize two different HJB equations (one per player). Since in this game we have to optimize taking into account that our competitor also optimizes, we have to solve a system of two connected HJB equations.

Since HJB equations are only solvable in few cases (e.g.,as in linear-quadratic models [4] we will present some particular numerical methods to do it.

The method we propose has been developed using viscosity solutions in dynamic programming technics. In a few words, we will have to discretize the problem in time and in space, and then iterate a value and a policy function until convergence. This solution is stationary in time, and is adapted to infinite time horizon games.

Next, we are going to explain all those concepts. Following [6],[12] we will define h>0 as the discretization time step. The continuous time optimal control problem is approximated by a first order discrete time scheme given by:

$$V_{i,h}(\alpha, t, x_t) = \max_{u_i(t)} J_{i,h}(x_i, u_i, \alpha) \tag{29}$$

for i=1,2 and defining:

$$J_{i,h}(x_i, u_i, \alpha) = h \sum_{t=0}^{\infty} \beta^t f_i(x_{i,h}(t), u_{i,t}; \alpha) \tag{30}$$

where $\beta = 1 - r_i h$ and $x_{i,h}$ is defined by the discrete dynamics:

$$x_{i,h}(t+1) = \varphi_{i,h}(x_{i,h}(t), u_{i,t}) := x_{i,h}(t) + h g_i(x_{i,h}(t), u_{i,t}) \tag{31}$$

The optimal value function $V_{i,h}$ is the unique solution of the discrete Hamilton-Jacobi-Bellman equation:

$$V_{i,h}(\alpha, t, x_t) = \max_{u_i(t)} \{h f_i(x_{i,h}(t), u_{i,t}; \alpha) + \beta V_{i,h}(\varphi_{i,h}(x_{i,h}(t), u_{i,t}))\} \tag{32}$$

If we define the dynamic programming operator T_h by:

$$T_h(V_{i,h})(\alpha, t, x_t) = \max_{u_i(t)} \{h f_i(x_{i,h}(t), u_{i,t}; \alpha) + \beta V_{i,h}(\varphi_{i,h}(x_{i,h}(t), u_{i,t}))\} \tag{33}$$

Then, $V_{i,h}$ can be characterized as the unique solution of the fixed point equation:

$$(V_{i,h})(\alpha,t,x_t) = T_h(V_{i,h})(\alpha,t,x_t) \tag{34}$$

The next step in the algorithm is to approximate the solution on a grid denoted by Γ covering a compact subset Ω of the state space. So, denoting the nodes of the grid Γ by x^n,we are looking for an approximation $V_{i,h}^{\Gamma}$ satisfying:

$$(V_{i,h}^{\Gamma})(\alpha,t,x_t) = T_h(V_{i,h}^{\Gamma})(\alpha,t,x_t) \tag{35}$$

The value of $V_{i,h}^{\Gamma}$ for points which are not in the grid Gamma are determined by interpolation. This iteration used to reach a fixed point will be called as "Value iterations". In [16] the author proposes an algorithm (programmed in MATLAB) that does the value iterations but also include a kind of policy iterations. They point out that in 2D problems it could be more efficient and quicker not to use a discretization of possible control variables to evaluate it in each iteration. However they propose obtain optimal candidates by operating explicitly in the problem due to functions are explicit and derivables. To do so, they approximate the value function in the so-called "policy iterations" using Taylor theorem as follows:

$$V_{i,h}(\varphi_{i,h}(x_{i,h}(t),u_{i,t})) \cong V_{i,h}(x_{i,h}(t)) + \frac{\partial V_{i,h}}{\partial x_i} g_i(x_{i,h}(t),u_{i,t}), \tag{36}$$

Substituting Equation (36) into Equation (32), we have:

$$(V_{i,h})(\alpha,t,x_t) = \frac{h}{1-\beta} \max_{u_i(t)} \left\{ f(x_{i,h}(t),u_{i,t};\alpha) + \beta \frac{\partial V_{i,h}}{\partial x_i} g_i(x_{i,h}(t),u_{i,t}) \right\} \tag{37}$$

So, in order to obtain from Equation (37) the optimal value for u_i,one can operate algebraically to obtain an explicit expression:

$$\frac{\partial f(x_{i,h}(t),u_{i,t};\alpha)}{\partial u_{i,t}} + \beta \frac{\partial V_{i,h}}{\partial x_i} \frac{\partial g_i(x_{i,h}(t),u_{i,t})}{\partial u_{i,t}} = 0 \tag{38}$$

and, as [16] proposes, we can derivate $V(\cdot)$by discretizing for steps k>1:

$$\frac{\partial V_{i,h}^{k}(x_{i,h}(t))}{\partial x_{i,h}} \cong \frac{V_{i,h}^{k-1}(x_{i,h}(t)+\triangle x_{i,h}) - V_{i,h}^{k-1}(x_{i,h}(t))}{\triangle x_{i,h}} \tag{39}$$

we leave for next steps to research if using the exact gradient will cause problems and if the maximum is located at the boundary of an element on which we do the interpolation, since at these points the gradient will not be defined .

The next three algorithms define our working proccess to obtaing optimal policies. As it has been depicted in this chapter, we have a main algorith and two different sub-rutines to obtain value iterations and policy iterations as follows:

Algorithm 1 2D algorithm to solve differential games by dynamic programming

1. Choose $\triangle x_1, \triangle x_2$, tolerance,$r_i, r_j$,$h$,$x_{1min}, x_{1max}, x_{2min}, x_{2max}$,calculante $m = \frac{x_{1max}-x_{1min}}{\triangle x_1}, n = \frac{x_{2max}-x_{2min}}{\triangle x_2}$.Generate the 2D grid ($\Gamma$) using x_1^m, x_2^n
2. Initialize the following variables (all are matrix with dimension $m \times n$):
   (a) value function player 1 and 2$\rightarrow V01 = 0, V02 = 0$
   (b) controls for player 1 and 2 (p=price, tv=advertising)$\rightarrow p01 = 0, p02 = 0, v1 = 0$
   (c) Gradients for player 1 and 2, $\frac{\partial V_{i,h}^k(x_{i,h}(t))}{\partial x_{i,h}} \cong \frac{V_{i,h}^{k-1}(x_{i,h}(t)+\triangle x_{i,h})-V_{i,h}^{k-1}(x_{i,h}(t))}{\triangle x_{i,h}}$, $\rightarrow grad01 = 0, grad02 = 0$
   (d) obtain the first values for state equations for players 1 and 2$\rightarrow$ $x_{i,h}(t+1) = x_{i,h}(t) + hg_i(x_{i,h}(t), p01(t), p02(t), v1(t))$
   (e) interpolate Value function outside Γ for new state values $x_{i,h}(t+1)$
3. initialize error values (for value function and policy function):
   (a) WHILE it<itmax or *Error*>tolerance
      i. it=it+1
      ii. VALUE ITERATION$\rightarrow$ *obtain Error* value per player
      iii. POLICY ITERATION$\rightarrow$ *obtain Error* policy per player
   (b) update error values per player (value function and policy function). *Error* is the maximmum of then.
4. END when convergence is achieved (Error is less than tolerance, or we have reached the maximmum of iterations)

Algorithm 2 VALUE ITERATION

Note that here we take as given al the policy values (p01,p02,v1=0) and iterate value function. In each iteration, the two value functions are connected as we have explained before.

1. Initialize *Value error* per player. Initialize utility function per player→ $f_i(x_{i,h}(t), u_{i,t}; \alpha)$, initialize $V_{i,h}(\varphi_{i,h}(x_{i,h}(t), u_{i,t}) = 0$
2. WHILE it<itmax or *Error* value>tolerance
   (a) it=it+1
   (b) calculate for i=1,2→ $V0i^{it} = hf_i(x_{i,h}(t), u_{i,t}; \alpha) + \beta V_{i,h}^{it}(\varphi_{i,h}(x_{i,h}(t), u_{i,t}))$
   (c) actualize by interpolation $V_{i,h}^{it+1}(\varphi_{i,h}(x_{i,h}(t), u_{i,t})$ using $x_{i,h}(t+1)$
   (d) calculate$V0i^{it+1} = hf_i(x_{i,h}(t), u_{i,t}; \alpha) + \beta V_{i,h}^{it+1}(\varphi_{i,h}(x_{i,h}(t), u_{i,t}))$
   (e) $Error\ value_i^{it+1} = V0i^{it+1} - V0i^{it}$
   (f) error is the max $Error\ value_i^{it+1}$
3. END when convergence is achieved

Next, we explain some of the properties of the algorithms and operations presented in this chapter. Following [13], by using the contraction mapping theorem in the algorithm (we have dropped the h and the i for the sake of simplicity) $V^{n+1} = TV^n$, V^n will converge to the infinite value horizon value function for any initial guess V^0. The sequence of control which rules u^n will also converge to the optimal control rule. This demonstration is explained with detail in [9]. But there are two practical problems when implementing a discretization, as is suggested in [9]:

1. The limit can never be achieved. Numerical methods rely on finite iterations
2. T is a functional, a mapping which takes a function and creates a new function.

We have to approximate this last idea by using interpolation, as it has been explained before. In the following, we will derivate the error bounds in a programming dynamic algorithm and then we will examine the properties of linear interpolation (the method is used in our algorithm, as suggested by [6],[16]).

Assuming the continuity and the derivability of the involved functions, is shown in [9] and [12], that error bounds of T when we have V^∞ (it is said, value solved in infinite time) : $\|V - V^\infty\| \leq \|TV - V\|/(1-\beta)$.

Linear interpolation satisfies particularly desirable properties for approximate dynamic programming. Assuming we know that V is monotonically in-

Algorithm 3 POLICY ITERATION

Here we have to obtain a candidate to policies (p01,p02*,v1*). We make this operations in each stage taking into account that the two equations are connected as we pointed out before*:

1. Initialize *Error policy* per player.

2. WHILE it<itmax or *Error* policy>tolerance

   (a) calculate gradient for i=1,2$\rightarrow$ $\frac{\partial V_{i,h}^{k}(x_{i,h}(t))}{\partial x_{i,h}} \cong \frac{V_{i,h}^{k-1}(x_{i,h}(t)+\triangle x_{i,h})-V_{i,h}^{k-1}(x_{i,h}(t))}{\triangle x_{i,h}} \rightarrow grad01, grad02$

   (b) obtain optimal control candidates$\rightarrow$$\frac{\partial f(x_{i,h}(t),u_{i,t};\alpha)}{\partial u_{i,t}} + \beta\frac{\partial V_{i,h}}{\partial x_{i}}\frac{\partial g_{i}(x_{i,h}(t),u_{i,t})}{\partial u_{i,t}} = 0 \rightarrow u_{i,h}* = p01*, p02*, v1*$

   (c) actualize state equations$\rightarrow$ $x_{i,h}(t+1) = x_{i,h}(t) + hg_{i}(x_{i,h}(t), p01*(t), p02*(t), v1*(t))$

   (d) actualize value function $\rightarrow V0i^{it+1} = hf_{i}(x_{i,h}(t), u*_{i,t};\alpha) + \beta V_{i,h}^{it+1}(\varphi_{i,h}(x_{i,h}(t), u*_{i,t}))$

   (e) $Error\, policy_{i}^{it+1} = V0i^{it+1} - V0i^{it}$

   (f) error is the max $Error\, policy_{i}^{it+1}$

3. END when convergence is achieved

creasing, we consider the interval $I_i \equiv [x_i, x_{i+1})]$. The monotonicity implies, for $x \in I$:$V_i \leq TV(x) \leq V_{i+1}$, where we have relaxed notation to remark the use of interpolation function $\hat{T}$. So, having an approximation error on I_i at most: $\triangle_i \equiv V_{i+1} - V_i$ the linear interpolation method, implies:

$$\|TV - \hat{T}V\| \leq \max_i \triangle_i \tag{40}$$

Where this expression for error of interpolation is demonstrated in [9], this ensures that linear interpolation satisfies particularly desirable properties for approximate dynamic programming.

4.1 An application in 1D and 2D problems

We now apply the algorithms developed in this chapter to two base-line problems.

EXAMPLE 1

Firstly, we propose the Brock-Mirman growth deterministic model [6],[12]. This model is simple and has one analytic solution (being c=consum, k=capital):

$$V(k_0) = \max_k \sum_{t=0}^{\infty} \beta^t log(c_t) \tag{41}$$

st:

$$c_t = Ak_t^{\alpha} - k_{t+1} \tag{42}$$

where $k_0$is fixed, $0 < \beta < 1, A > 0, 0 < \alpha < 1$.

The analytic solution is of the form: $V(k_0) = B + Clog(k_0)$ where , $B = \frac{log((1-\alpha\beta)A) + \frac{\alpha\beta}{1-\alpha\beta} log(A\alpha\beta)}{1-\beta}$ and $C = \frac{\alpha}{1-\beta}$. We will set, as in [6], $A = 5, \alpha = 0.34, \beta = 0.95$. We define a grid of possible capitals (k_t)in the interval [0.1,10]. We use our algorith adapted to 1D, taking as initial value function as zero, we get this solution in Figure(11):

In this first example, we have chosen as tolerance value h^2. Where h is the kapital increase in the defined grid. If h=0.01, tolerance will have to be $tol = 0.0001$. Following [12], this tolerance is fixed such as:

$$\|V_{n+1}^h - V_n^h\| \leq h^2 \tag{43}$$

Due to the fact that T is a contractive operator with modulues, the fixed point should be bounded, taking Equation (43) , by:

$$\|V_{n+1}^h - V_n^h\| \leq \frac{h^2}{1-\beta} \tag{44}$$

So, in this EX1, $\|V_{n+1}^h - V_n^h\| \leq \frac{0.0001}{1-0.95} = 0.002$. As we can see, taking L2, our results are in line with we expected a priori. If we make a fine grid, we

Figure 11: Numerical solution of EX1 and error estimation. h=0.01

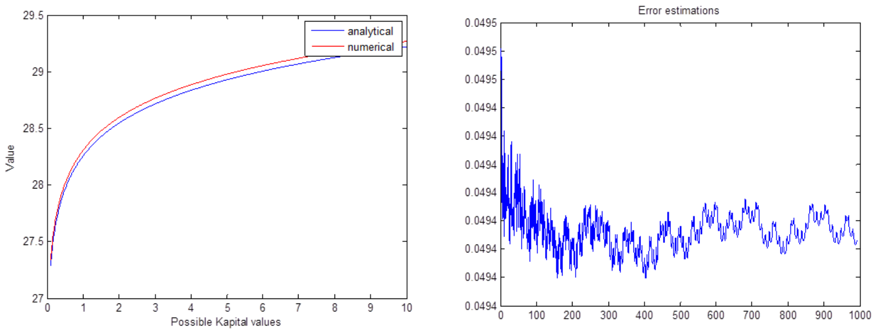
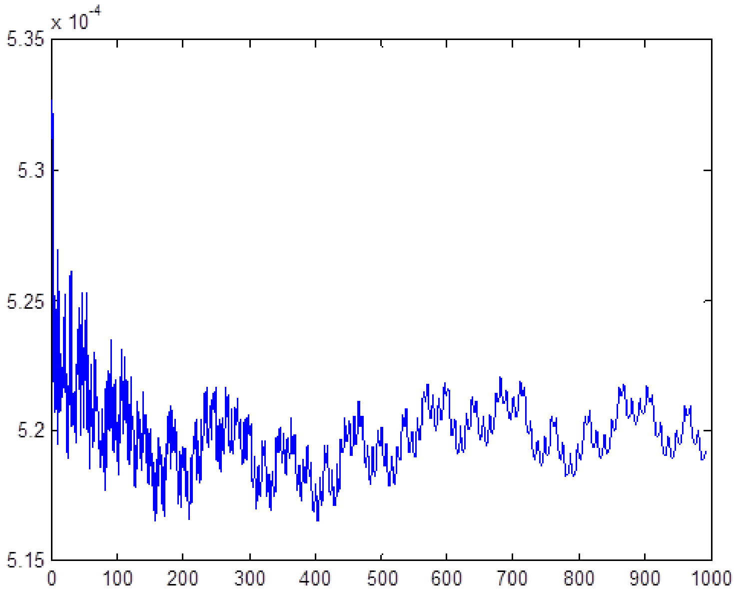


Figure 12: Error estimation when h=0.001



can reduce the error. Following that idea, we will use a space discretization of $h = 0.001$.

EXAMPLE 2

The second problem is a 2D growth model without analytic solution, but it has been solved with the same results than in [6] and [16]. Since it is used as a baseline model to prove that the method is reliable, we show that the results of our algorithm (which is a variation of [16] for solving a dynamic game with more than one control) are comparable with the ones published in [6] and [16]. Let:

$$g(x,u) = k_1 x_1^{1/2} - \frac{x_1}{1 + k_2 x_1^4} - c_1 x_2 - \frac{c_2 x_2^2}{2} - \frac{a u^2}{2} \tag{45}$$

with:

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_2(t) - \sigma x_1(t) \\ u(t) \end{bmatrix} \tag{46}$$

and, taking the same values than in [6] and [16], we obtain the same solutions shown in Figure (13) As it is docummented in bibliografy, $u(x_1, x_2)$, obtained in Figure (13), shows a special characteristic of optimal control. It is the discontinuity in the politics (u) showing a line wich is called $skiba - line$.[6],[16].

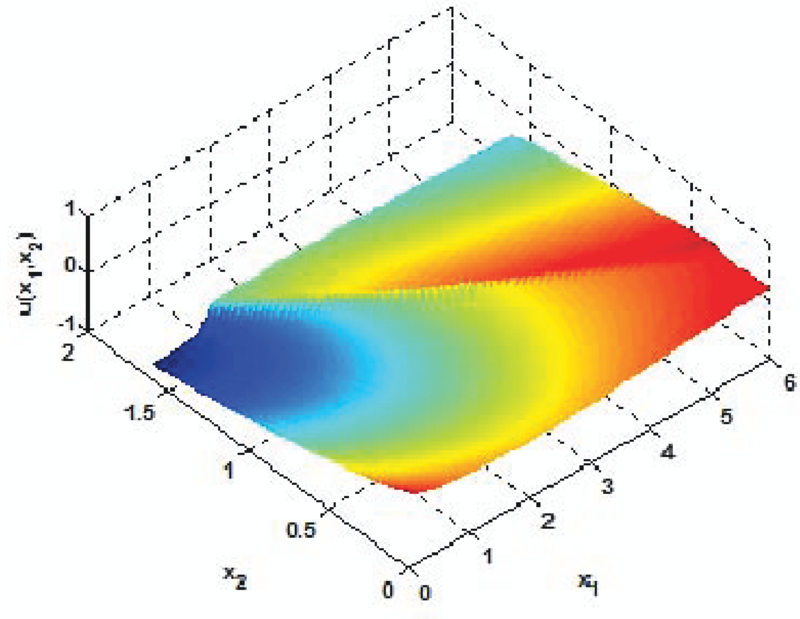
In this chapter, we have show the main characteristics of our algorithm to solve differential games by using dynamic programming. We have shown two benchmark examples based on bibliography to illustrate the ability of our method to solve this kind of problems.

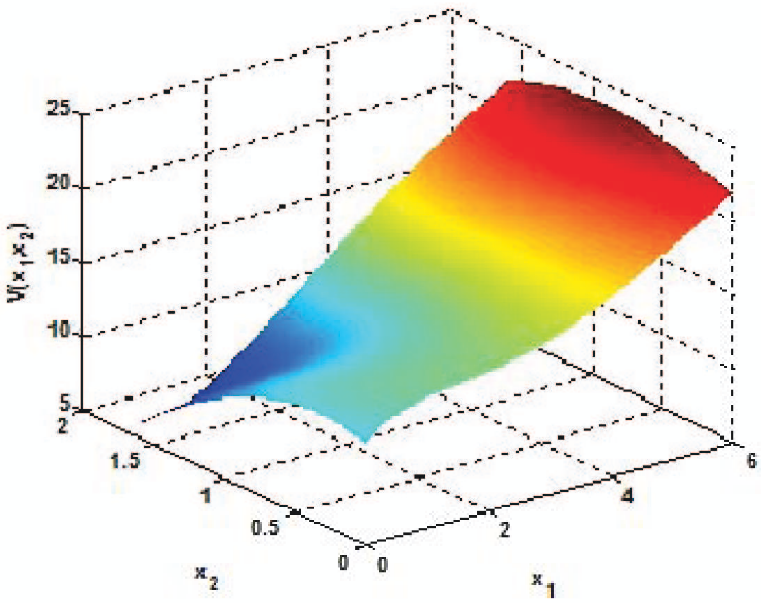
5 The 'How much' war. A numerical solution

In this final chapter, we will apply the previous algorithms introduced in Section 3 to obtain optimal solutions to our primary problem. We will deliver the following results: Value function, optimal price and optimal advertising investments for our company and the competitor in a closed-loop form (which depending of our sales value and the competitor sales in the state space). We compare the obtained results with real policies considered by a particular company (anonymous due to legal issues) to check which method provide the best results.

We remember that the considered problem is:

Figure 13: solution to EX-2 by our algorithm





obtain a closed-loop approximation to the following differential game:

$$\max_{tv_1,p_1} J_1 = \int_0^{\infty} e^{-r_1 t} \left(p_1 x_1(t) - c_1 log(v_1(t))\right) dt \tag{47}$$

$$\max_{p_2} J_2 = \int_0^{\infty} e^{-r_2 t} \left(p_2 x_2(t)\right) dt \tag{48}$$

with dynamics:

$\dot{x}_1 = \left[\alpha_1 \left[1 - \frac{\alpha_1 x_1}{N_1}\right] - \beta_{12} x_2 + \rho_1 v_1 - \varpi_1 p_1 + \varpi_{11} p_1^2\right] x_1$

$\dot{x}_2 = \left[\alpha_2 \left[1 - \frac{\alpha_2 x_2}{N_2}\right] - \beta_{21} x_1 - \varpi_2 p_2 + \varpi_{22} p_2^2\right] x_2$

and with parameter values:

	r_1	r_2	c_1	α_1	$\frac{\alpha_1}{N_1}$	β_{12}
player 1	0.05		26917	5.41	0.022	0.034
player 2		0.05				

	ρ_1	ω_1	ω_{11}	α_2	$\frac{\alpha_2}{N_2}$	β_{21}	ω_2	ω_{22}
player 1	0.00035	0.80	0.032					
player 2				3.73	0.046	0.040	0.534	0.029

In Section 3, we showed that the two versions of the LV state equations could be considered as similar to fit the real data. As we can see, the only difference between models, is the quadratic price variable in the two equations of LV2. This is a slight difference, but it allows us to obtain in our algorithm an explicit expression for the price variable in the Step 2-d of the *policy iteration* (see algorithm 3).

Using LV1 as the state equation, to obtain optimal price, we realize that the price variable does not appear after derivation. This happens when the control variables are linear in benefit and in state equations [13]. The optimal control obtained is called “bang –bang”. The result of a bang-bang control is the maximum or the minimum value for the control, depending if the value of variables are greater than zero or not. We have decided to not derive bang-bang controls following the LV1 model.

Solving the closed-loop equilibria for two players, we obtain the results shown in Figure(13), considering a tolerance of 0.01. As it is shown in this Figure, the results are intuitive:

1. The value Function increases as our sales increase. But it has an intense growth when our competitor is in low sales levels and viceversa.

2. The price function has a dominant value (around 12.5 %). But is interesting to comment that this price function has a growing part as our sales increase from 0 to 15 (approx.). This could be represent a leader mechanism, in which up to some sales level (in this case, around 15), the company is a leader and can up the price, taking profit of this advantage.

3. The advertising function has a decreasing value as our sales increase. It is plausible due to, the more we are known, the less we need to advertise. Note that if our competitor has more sales than us, the advertising value is bigger than in other cases.

4. Similar conclusions can be taken from competitor's results.

Once we have results, it is possible to compare the performance of the model versus reality. In figure (16), we compare real GRPs invested by a particular compnay and the one recommended by the model. As it is shown in the chart, our model does not use to recommend more than 150 GRPs per week. Due to the costs structure and the dynamics, it could be possible to interpret the real GRPs as oversized. On the other hand, pricing policy recommended by our model is more aggresive. It recommends a constant price around 12,5 % In reality, price has moved from 13% to 12%.

Regarding the Economic performance of the model versus reality, we have calculated a benefit using Equation(17). The most important losses are incurred in the time periods with TV advertising. As is depicted in Figure (17), up to 2009 there is a slight gain, every week, by using our model due to pricing advantage. In fact,we have obtained an improvement in the benefit with our model of a 7.7% in comparison with the real strategy.

In this section, we have shown the main advantages of using a differential game to make decissions. It is important to note that this model, once it is estimated, is able to obtain all the strategy for the fixed time period by the company. And, following competitor evolution, it is possible to actualize the model to adapt its response to the shocks that could affect both companies.

6 Conclusions and Further work

In this work, we have developed a numerical method to solve infinite time differential games in closed-loop equilibria. Differential games are thought to be run in dynamic decisions and competitive situations, such as marketing investments and pricing policies in a company. Closed-loop equilibria allow us to obtain strategies as a function of ourselves and our competitor. We have applied our algorithm to a real data set of two spanish competitive firms. We show that with our algorithm it is possible to develop different price-advertising strategy to get around more than 7,7% of benefits.

We remark that the main contribution of this work has been the numerical method, adapting the algorithm from dynamic programming, and the parameter estimation in real-life cases, using Lotka-Volterra family of differential equations adapted to our context. We propose for next steps this research lines:

Figure 14: Our approximate solution

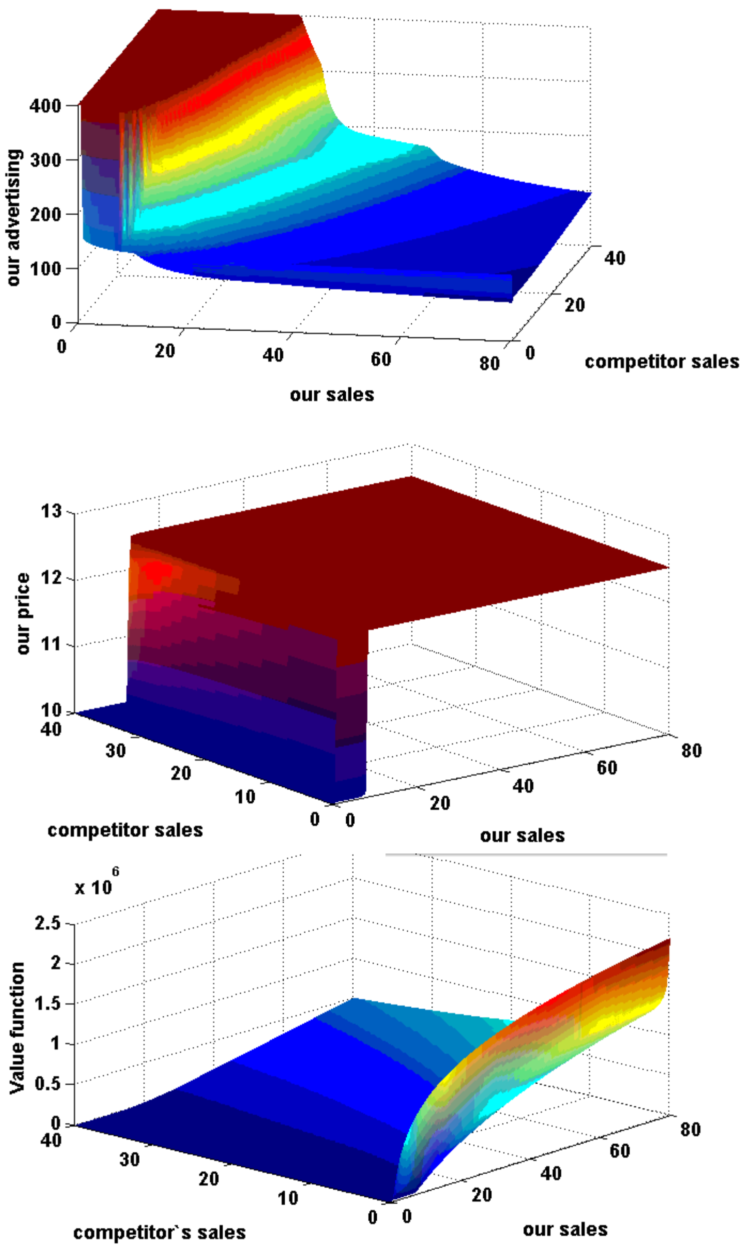


Figure 15: Competitor approximate solution.

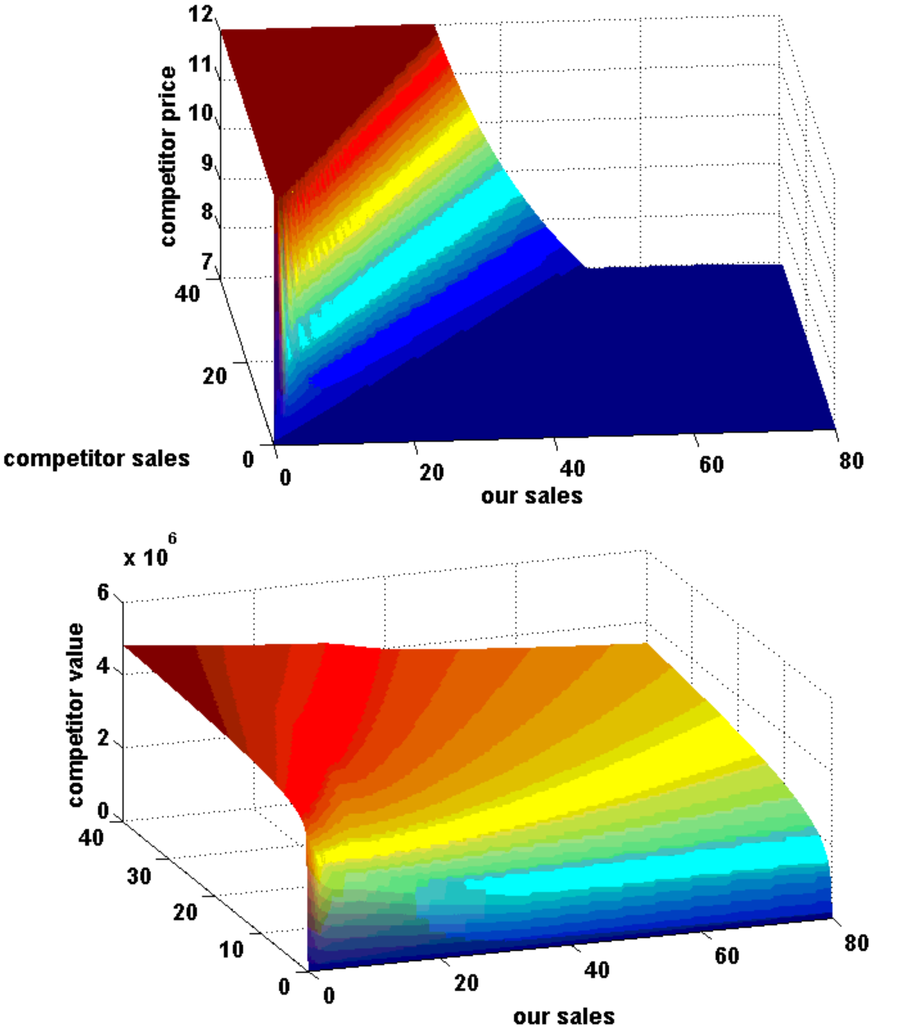


Figure 16: Recommended policies versus reality.

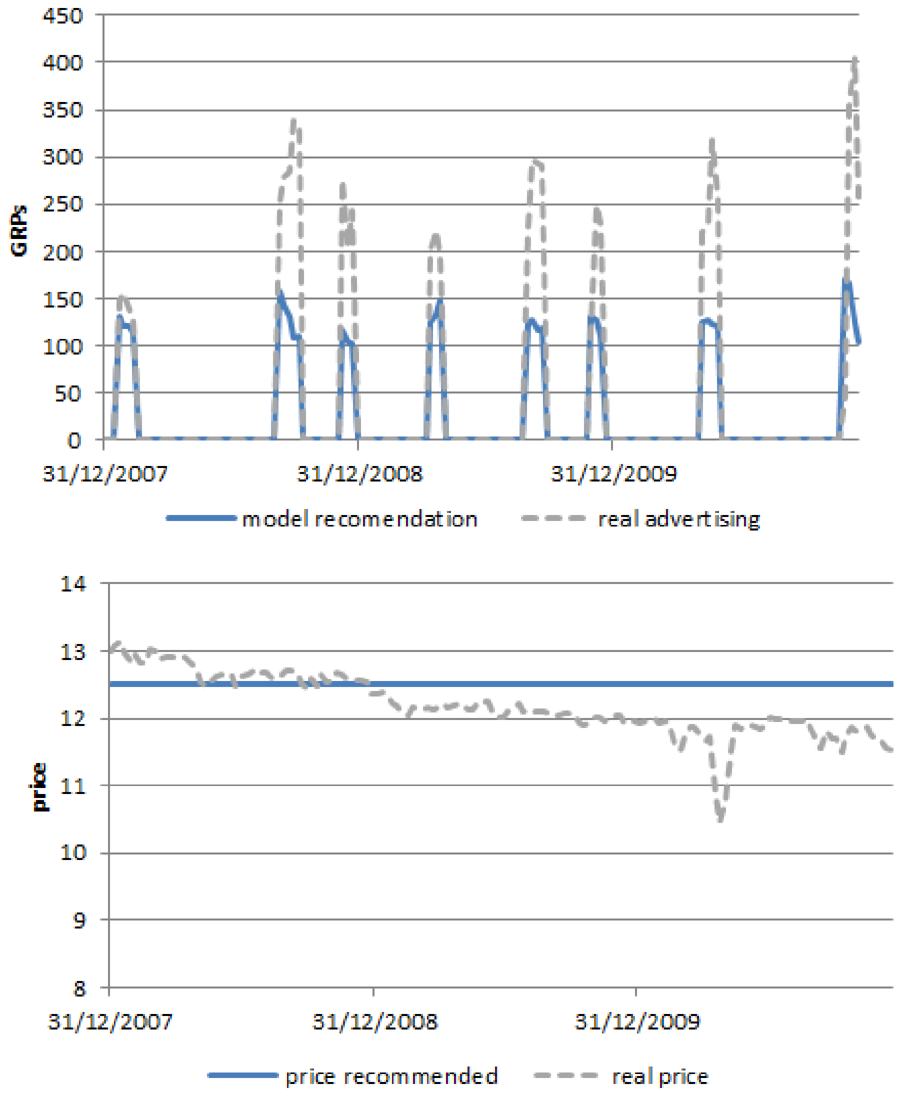
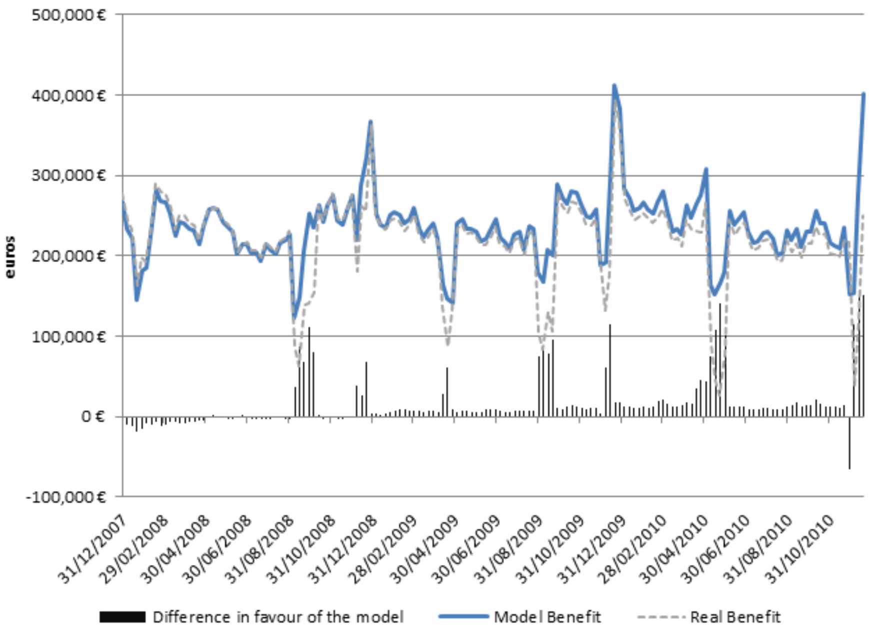


Figure 17: A study of the benefit



1. Adapting the algorithm for stochastic games. Since we deal with estimated parameters and models, we have a measure of error that could be incorporated in the method.
2. Research of new ways to obtain maximal points in a more efficient way taking into account that usually there are more than one control in our functions.
3. Using more different methodologies to estimate parameters, and compare which is more adaptable to data for this purposes. As can be seen in the Apendix, it could be necessary to include more lags in the variables, so state equations should be re-formulated in a multi-step approach.
4. Trying different functional forms such as competing in price differential or a typical Stackelberg game (in which one of the company is a leader and the rest is a follower).

7 Bibliography

1. Arnason R, Steinsham S, Sandal L, Vestergaard N (2004). Optimal Feedback Controls: Comparative evaluation of the cod fisheries in Denmark, Iceland, and Norway. *American Journal of Agricultural Economics*.86(2) (May 2004): 531-542.

2. Basar T,Olsder G.J. (1987). Dynamic Non-Cooperative Game Theory. *Society for Industrial and Applied Mathematics*

3. Cerdo, E. (2000). Optimizacion Dinamica, 1 Edicion. *Prentice Hall.*

4. Dockner E, Jorgensen S, Van Long N, Sorger G. (2000). Differential games in economics and management science. *Cambridge University press.*

5. Ewing B.T, R. K. (2006). Time series analysis of a predator-prey system: Application of VAR and generalized impulse response function. *Journal of Ecological Economics.* 60(Jan 2007): 605-612.

6. Grïene L, Semmpler W. (2004). Using dynamic programming with adaptive grids to solve control problems in economics. *Journal of Economic Dynamics and Control.* 28(Dec 2004): 2427-2456.

7. Hamilton,J.D. (1994). Time Series Analysis. *Princeton University Press.*

8. Jorgensen S, Z. G. (2006). Developments in differential game theory and numerical methods:economic and management applications. *Journal of Computational Management Science.* 4(Apr 2007): 159-181.

9. Kenneth L. Judd, A. S. (2004). Numerical Dynamic Programming with Shape Preserving Splines. *Working paper*

10. Krishnamoorthy A, P. A. (2010). Optimal pricing and Advertising in a Durable - good duopoly. *European Journal of Operational Research,* 200(2)(Jan 2010): 486-497.

11. Lopez L, S. M. (2001). Defining strategies to win in the Internet market. *Physica A.* 301 (jan 2001) 512-534.

12. Manuel S Santos, J. V.-A. (1998). Analysis of a Numerical Dynamic Programming Algorithm Applied to Economic Models. *Econometrica.* 66 (2) (1998) 409-426.

13. Michalakelis, C. (2009). Modeling Competition in the Telecommunications Market Based on Concepts of Population Biology. *IEEE transactions on systems, man, and cybernetics.*41(2) (2011) 200-210.

14. Wan,Q. (1996). A duopolistic model of dynamic Competitive advertising and empirical validation. *Nanayang Technological University, Working paper series.*

15. Sethi S, Prasad A, He X. (2008). Optimal advertising and pricing in a new product adoption model. *Journal of Optimization Theory and Applications.*139 (2) (2008)351-360

16. Sigurd, A. (2009). Essais on explotation on natural resources. Optimal control theory applied on multispecies fisheries and fossil fuel extraction. *Doctoral Thesis.*

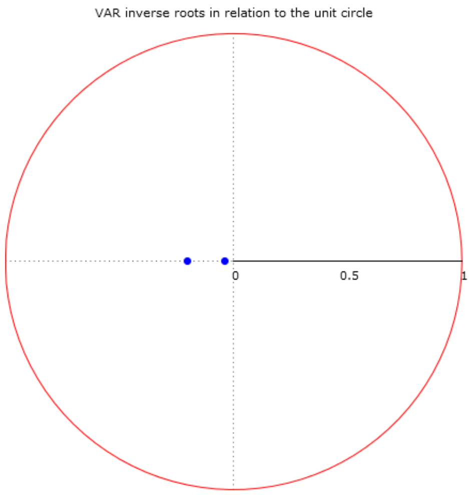
17. Stukalin D, Schmidt. W. (2011). Some Applications of Optimal Control in Sustainable Fishing in the Baltic Sea. *Applied Mathematics.*7(2) (2011) 854-865.

18. Sun, M. (2006). Dynamic Platform Competition in two-sided markets. *ProQuest Dissertations and Theses.*

Appendix. Estimated Models

In this Appendix we include the most important test of specification in Econometric Models of Section 3.

Figure 18: Specificacion test VARX



Equation 1: Ljung-Box Q' = 65.9137 with p-value = P(Chi-square(60) > 65.9137) = 0.28
Equation 2: Ljung-Box Q' = 117.868 with p-value = P(Chi-square(60) > 117.868) = 1.21e-005

Residual correlation matrix, C (2 x 2)
1.0000 0.16625 0.16625 1.0000
Eigenvalues of C
0.83375 1.16625
Doornik-Hansen test Chi-square(4) = 21.1133 [0.0003]

Null hypothesis: the regression parameters are zero for the variables
GRPs, preciox, precioy
Test statistic: Chi-square(6) = 35.5627, with p-value = 3.35204e-006

Figure 19: Specification Test LV-1 Eq1

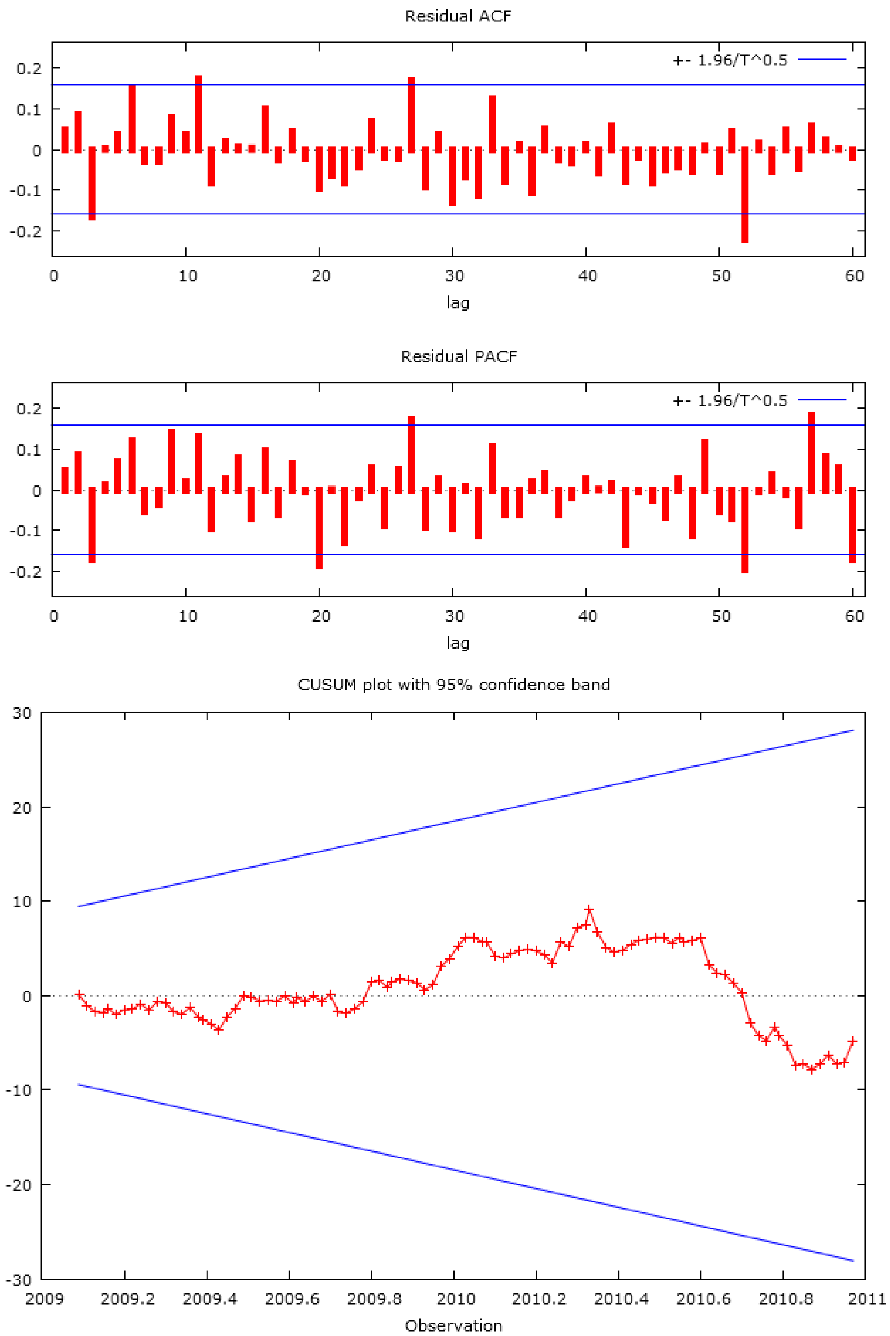


Figure 20: Specification test LV-2 eq2

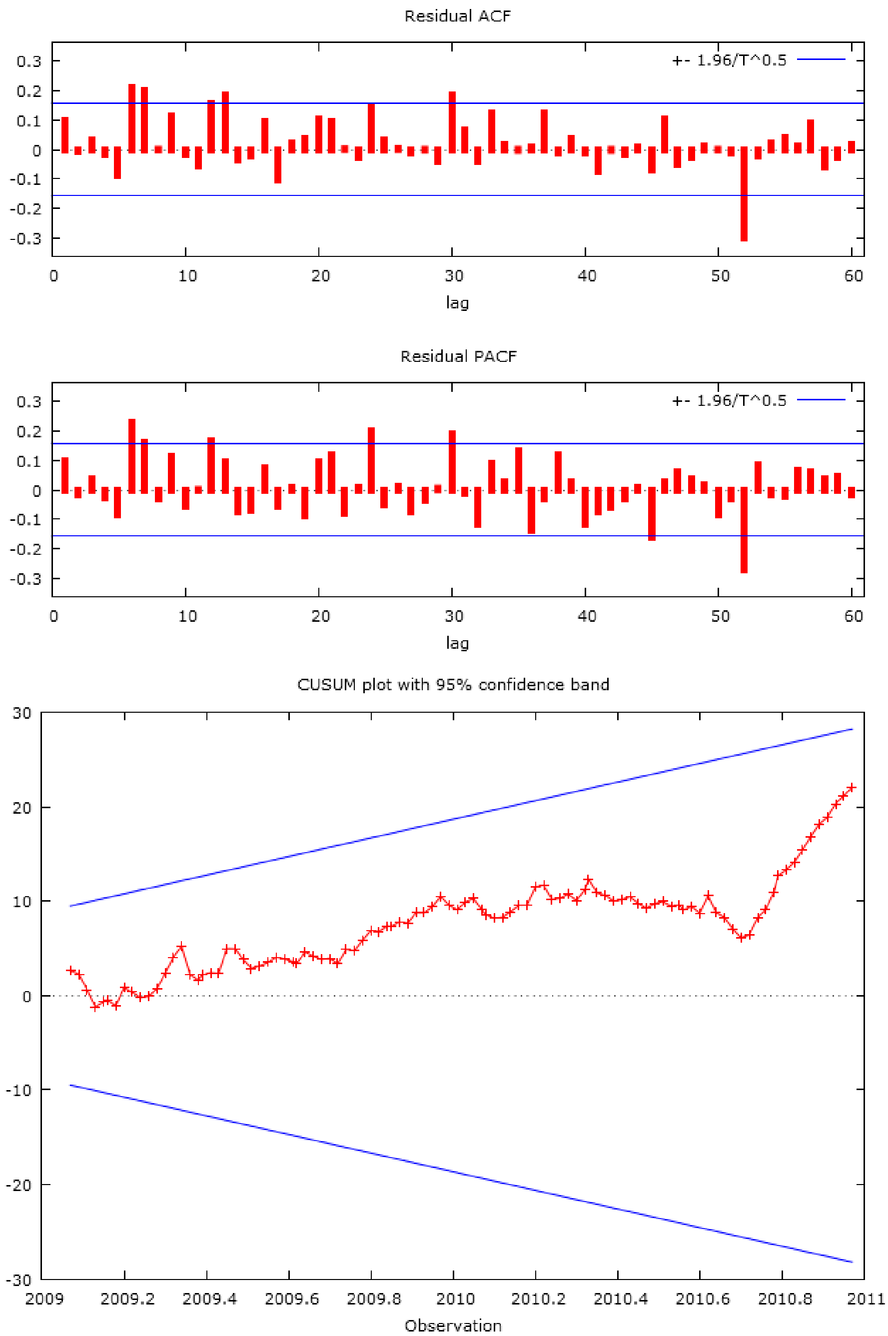
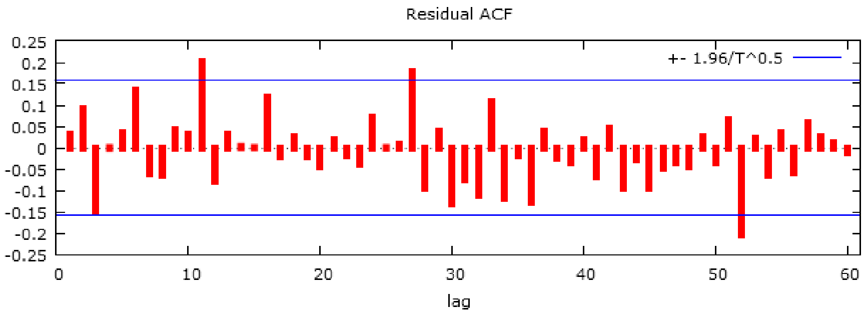
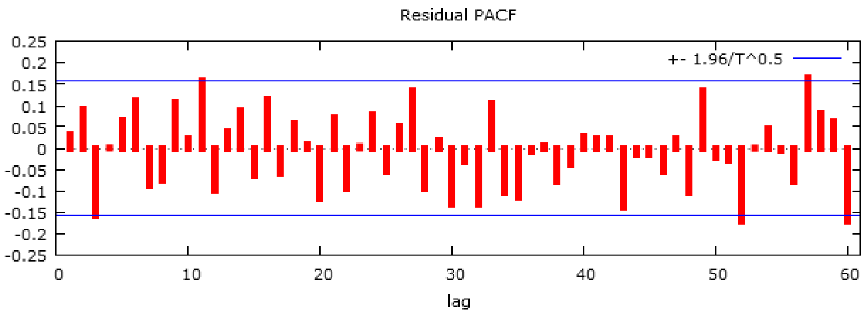


Figure 21: Specification test LV-2 eq1





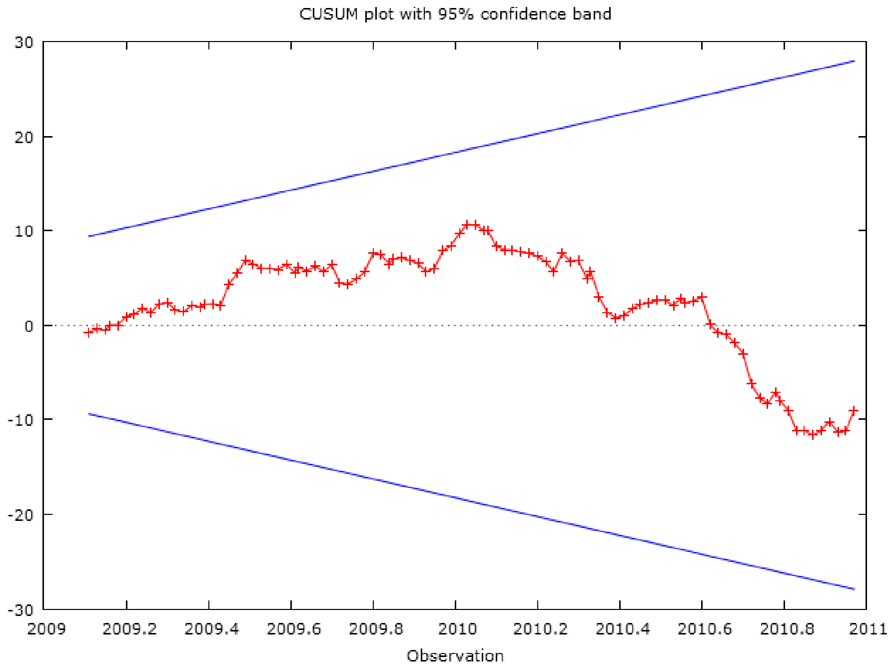
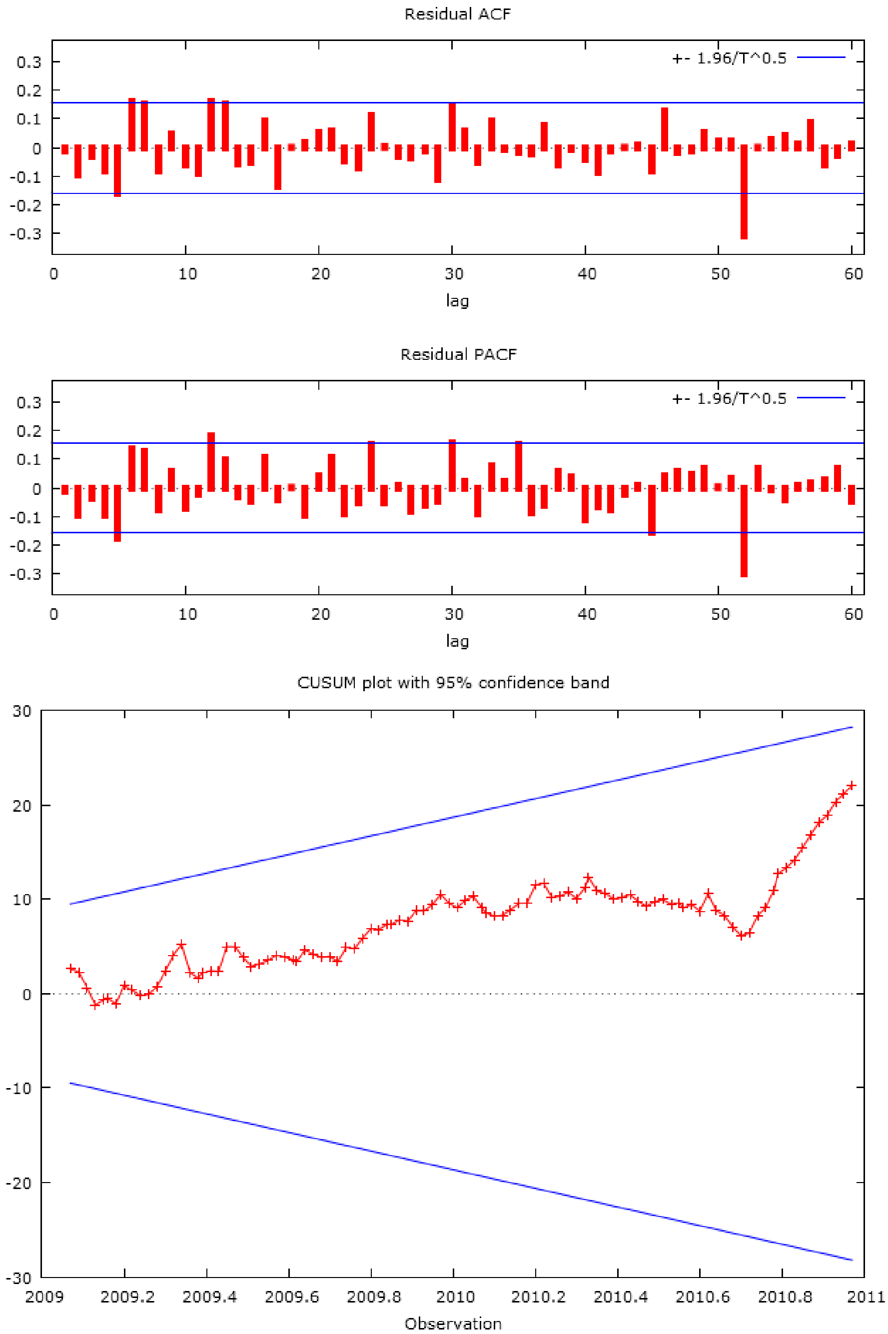


Figure 22: Specification test LV-2 Eq2



PREPUBLICACIONES DEL DEPARTAMENTO
DE MATEMÁTICA APLICADA
UNIVERSIDAD COMPLUTENSE DE MADRID
MA-UCM 2011

1. APPROXIMATING TRAVELLING WAVES BY EQUILIBRIA OF NON LOCAL EQUATIONS, J. M. Arrieta, M. López-Fernández and E. Zuazua.

2. INFINITELY MANY STABILITY SWITCHES IN A PROBLEM WITH SUBLINEAR OSCILLATORY BOUNDARY CONDITIONS, A. Castro and R. Pardo

3. THIN DOMAINS WITH EXTREMELY HIGH OSCILLATORY BOUNDARIES, J́. M. Arrieta and M. C. Pereira

4. FROM NEWTON EQUATION TO FRACTIONAL DIFFUSION AND WAVE EQUATIONS, L. Váquez

5. EL CÁLCULO FRACCIONARIO COMO INSTRUMENTO DE MODELIZACIÓN, L. Váquez and M. P. Velasco

6. THE TANGENTIAL VARIATION OF A LOCALIZED FLUX-TYPE EIGENVALUE PROBLEM, R. Pardo, A. L. Pereira and J. C. Sabina de Lis

7. IDENTIFICATION OF A HEAT TRANSFER COEFFICIENT DEPENDING ON PRESSURE AND TEMPERATURE, A. Fraguela, J. A. Infante, Á. M. Ramos and J. M. Rey

8. A NOTE ON THE LIOUVILLE METHOD APPLIED TO ELLIPTIC EVENTUALLY DEGENERATE FULLY NONLINEAR EQUATIONS GOVERNED BY THE PUCCI OPERATORS AND THE KELLER–OSSERMAN CONDITION, G. Díaz

9. RESONANT SOLUTIONS AND TURNING POINTS IN AN ELLIPTIC PROBLEM WITH OSCILLATORY BOUNDARY CONDITIONS, A. Castro and R. Pardo

10. BE-FAST: A SPATIAL MODEL FOR STUDYING CLASSICAL SWINE FEVER VIRUS SPREAD BETWEEN AND WITHIN FARMS. DESCRIPTION AND VALIDATION. B. Ivorra B. Martínez-López, A. M. Ramos and J.M. Sánchez-Vizcaíno.

11. FRACTIONAL HEAT EQUATION AND THE SECOND LAW OF THERMODYNAMICS, L. Vázquez , J. J. Trujillo and M. P. Velasco

12. LINEAR AND SEMILINEAR HIGHER ORDER PARABOLIC EQUATIONS IN R^N, J. Cholewa and A. Rodríguez Bernal

13. DISSIPATIVE MECHANISM OF A SEMILINEAR HIGHER ORDER PARABOLIC EQUATION IN R^N, J. Cholewa and A. Rodríguez Bernal

14. DYNAMIC BOUNDARY CONDITIONS AS A SINGULAR LIMIT OF PARABOLIC PROBLEMS WITH TERMS CONCENTRATING AT THE BOUNDARY, A. Jimémez-Casas and A. Rodríguez Bernal

15. DISEÑO DE UN MODELO ECONÓMICO Y DE PLANES DE CONTROL PARA UNA EPIDEMIA DE PESTE PORCINA CLÁSICA, E. Fernández Carrión, B. Ivorra, A. M. Ramos, B. Martínez-López, Sánchez-Vizcaíno.

16. BIOREACTOR SHAPE OPTIMIZATION. MODELING, SIMULATION, AND SHAPE OPTIMIZATION OF A SIMPLE BIOREACTOR FOR WATER TREATMENT, J. M. Bello Rivas, B. Ivorra, A. M. Ramos, J. Harmand and A. Rapaport

17. THE PROBABILISTIC BROSAMLER FORMULA FOR SOME NONLINEAR NEUMANN BOUNDARY VALUE PROBLEMS GOVERNED BY ELLIPTIC POSSIBLY DEGENERATE OPERATORS, G. Díaz

PREPUBLICACIONES DEL DEPARTAMENTO
DE MATEMÁTICA APLICADA
UNIVERSIDAD COMPLUTENSE DE MADRID
MA-UCM 2012

1. ON THE CAHN-HILLIARD EQUATION IN H^1(R^N), J. Cholewa and A. Rodríguez Bernal

2. GENERALIZED ENTHALPY MODEL OF A HIGH PRESSURE SHIFT FREEZING PROCESS, N. A. S. Smith, S. S. L. Peppin and A. M. Ramos

3. 2D AND 3D MODELING AND OPTIMIZATION FOR THE DESIGN OF A FAST HYDRODYNAMIC FOCUSING MICROFLUIDIC MIXER, B. Ivorra, J. L. Redondo, J. G. Santiago, P.M. Ortigosa and A. M. Ramos

4. SMOOTHING AND PERTURBATION FOR SOME FOURTH ORDER LINEAR PARABOLIC EQUATIONS IN R^N, C. Quesada and A. Rodríguez-Bernal

5. NONLINEAR BALANCE AND ASYMPTOTIC BEHAVIOR OF SUPERCRITICAL REACTION-DIFFUSION EQUATIONS WITH NONLINEAR BOUNDARY CONDITIONS, A. Rodríguez-Bernal and A. Vidal-López

6. NAVIGATION IN TIME-EVOLVING ENVIRONMENTS BASED ON COMPACT INTERNAL REPRESENTATION: EXPERIMENTAL MODEL, J. A. Villacorta-Atienza and V.A. Makarov

7. ARBITRAGE CONDITIONS WITH NO SHORT SELLING, G. E. Oleaga

8. THEORY OF INTERMITTENCY APPLIED TO CLASSICAL PATHOLOGICAL CASES, E. del Rio, S. Elaskar, and V. A. Makarov

9. ANALYSIS AND SIMPLIFICATION OF A MATHEMATICAL MODEL FOR HIGH-PRESSURE FOOD PROCESSES, N. A. S. Smith, S. L. Mitchell and A. M. Ramos

10. THE INFLUENCE OF SOURCES TERMS ON THE BOUNDARY BEHAVIOR OF THE LARGE SOLUTIONS OF QUASILINEAR ELLIPTIC EQUATIONS. THE POWER LIKE CASE, S.Alarcón, G.Díaz and J.M.Rey

11. SUSTAINED INCREASE OF SPONTANEOUS INPUT AND SPIKE TRANSFER IN THE CA3-CA1 PATHWAY FOLLOWING LONG TERM POTENTIATION IN VIVO, O. Herreras, V. Makarov and A. Fernández--Ruiz

12. ELLIPTIC EQUATIONS IN WEIGHTED BESOV SPACES ON ASYMPTOTICALLY FLAT RIEMANNIAN MANIFOLDS, U. Brauer and L. Karp

13. A NUMERICAL METHOD TO SOLVE A DUOPOLISTIC DIFFERENTIAL GAME IN A CLOSED-LOOP EQUILIBRIUM, J. H. de la Cruz, B.Ivorra and A. M. Ramos