THE INSTITUTIONAL DETERMINANTS OF CO₂ EMISSIONS: 
A COMPUTATIONAL MODELLING APPROACH 
USING ARTIFICIAL NEURAL NETWORKS AND 
GENETIC PROGRAMMING

MARCOS ÁLVAREZ-DÍAZ 
GONZALO CABALLERO MIGUEZ 
MARIO SOLIÑO

FUNDACIÓN DE LAS CAJAS DE AHORROS 
DOCUMENTO DE TRABAJO 
N° 401/2008
De conformidad con la base quinta de la convocatoria del Programa de Estímulo a la Investigación, este trabajo ha sido sometido a evaluación externa anónima de especialistas cualificados a fin de contrastar su nivel técnico.

ISSN: 1988-8767

La serie DOCUMENTOS DE TRABAJO incluye avances y resultados de investigaciones dentro de los programas de la Fundación de las Cajas de Ahorros.
Las opiniones son responsabilidad de los autores.
The institutional determinants of CO\textsubscript{2} emissions: A computational modelling approach using Artificial Neural Networks and Genetic Programming

Marcos Álvarez-Díaz\textsuperscript{a}, Gonzalo Caballero Miguez\textsuperscript{b,\ast}, Mario Soliño\textsuperscript{b}

\textsuperscript{a}Economic Research Center (UIB-Sa Nostra), C/ del Ter 16 Son Fuster, 07009 Palma de Mallorca, Spain
\textsuperscript{b}ERENA and Department of Applied Economics, University of Vigo, Lagoas-Marcosende s/n, 36310 Vigo, Spain

Abstract

Understanding the complex process of climate change implies the knowledge of all possible determinants of CO\textsubscript{2} emissions. Most of the studies focused on the key role of economic growth on the emissions. Nevertheless, recent trends in economics have highlighted the relevance of the institutional structure of production on economic performance. This paper studies the influence of several institutional determinants on CO\textsubscript{2} emissions, clarifying which variables are relevant to explain this influence. For this aim, Genetic Programming and Artificial Neural Networks are used to find an optimal functional relationship between the CO\textsubscript{2} emissions and a set of historical, economic, geographical, religious and social variables, which are considered as a good approximation to the institutional quality of a country. Besides this, the paper compares the results using the computational methods with those employing a more traditional parametric perspective. Following the empirical results of the cross-country application, this paper generates new evidence on the binomial institutions and CO\textsubscript{2} emissions.

Keywords: Computational methods; CO\textsubscript{2} emissions; Institutional determinants

\textsuperscript{\ast} Corresponding author. Tel.: +34 986 812 442; fax: +34 986 812 401.

E-mail address: gcaballero@uvigo.es (G. Caballero)

Acknowledgements

Marcos Álvarez-Díaz wishes to thank European Commission for the financial support of the project “Climate Change and Impact Research: the Mediterranean Environment” (CIRCE project). Gonzalo Caballero was a visiting research fellow at the Donal Bren School of Environmental Science & Management (University of California, Santa Barbara); he thanks the financial support from Xunta de Galicia. Mario Soliño was visiting scholar at the Group of Environmental Economics of the Institute of Public Goods and Policies (CSIC, Centre for Social and Human Sciences).
1. Introduction

Climate change has emerged as one of the main issues in the political and scientific agenda. Nowadays it is widely accepted that this environmental problem represents a serious threat for the life conditions of hundred of million people; and it is considered one of the greatest challenges facing the world today. The principal consequence of climate change is associated generally with the increase of the temperature in the Earth. The global warming observed in recent decades has already led to negative and clearly perceptible effects for ecosystems, human settlements and economic activities such as tourism and agriculture, for example. It is directly responsible of more frequent and economically costly natural disasters (storms, hurricanes, floods and droughts), the melt of the ice caps, the rise of the sea level, and the irregular and unpredictable behaviour of the weather (IPCC, 2007).

The high concentration of greenhouse gases in the atmosphere is the main reason that explains the global warming and, therefore, the climate change. Among these gases, the anthropogenic carbon dioxide emissions (CO$_2$) are the main factor of the greenhouse effect. This fact justifies why CO$_2$ emissions are considered by the specialists as the best available indicator of climate change (Carlsson and Lundström, 2003; OECD, 2007; Quadrelli and Peterson, 2007). In spite of the international efforts to reduce the atmospheric level of CO$_2$, the emissions of this gas are still growing in many countries. Several papers in the specialized literature on energy economics and environmental economics have studied what possible factors contribute to the explanation of the increasing level of CO$_2$ emissions. Most of these efforts have focused on the key role of production and economic growth. For example, many studies provide empirical evidence of a monotonically increasing relation between CO$_2$ emissions and Gross National Product (GNP) (Shafik,
1994; Holtz-Eakin and Selden, 1995; Bertinelli and Strobl, 2005). However, as Bengoechea et al. (2001) point out, there may be other significant variables apart from the economic growth that explain the rising concentration of CO₂. Researchers are encouraged to make an effort to fill up this gap by discovering other possible determinants. In this sense, nowadays there is widespread agreement among economists that institutions matter and that they affect economic performance. Accordingly, it seems apparently reasonable to think that they must be included in the economic analysis on energy and environment. Thus, an effort to open the “black box” of the determinants of CO₂ emissions should focus on the role of organizational and institutional factors.

The new institutional economics (NIE) has constituted a program of research that has propelled the return of institutions into the agenda of mainstream economics. The NIE has already allowed significant advances in different areas such as economic history, economics of organization, law and economics, policy analysis, economic growth, development economics, ecological and environmental economics (Williamson, 2000; Ménard and Shirley, 2005; Vatn, 2005; Paavola and Adger, 2005; Culas, 2007). The coasean notion of transaction costs (Coase, 1937, 1960) and the northian notion of institutions (North, 1990) established the foundations for the theoretical framework of the NIE. Political rules, informal norms and enforcement mechanisms constitute the “rules of the game” of a society, and these rules establish a structure of incentives that affects the level of transaction costs and the efficiency in the economy. In a broad sense, institutions include social, historical, economic, legal and religious rules.

This paper studies the institutional determinants of CO₂ emissions. For this purpose, we analyze the relationship between the CO₂ emissions and a set of historical, geographical, religious, social and economic variables. This set of variables characterizes
some of the main institutional features of the countries in our sample (La Porta et al. 1999; Álvarez-Díaz and Caballero, 2008). Jointly to the conventional parametric approach, we make also use of computational modelling methods to specify the functional relationship between variables. Basically, one reason justifies the use of these novel techniques. The traditional modelling procedure has usually adopted a parametric perspective when the effects of institutions on diverse variables have been studied. Therefore, the functional form of these models is discretionally imposed by the researcher rather than observed in the data, and the unknown parameters are later estimated using some optimization procedure such as Ordinary Least Squares (OLS). Assuming this parametric point of view might cause a serious bias in the results due to a misspecification problem. It seems to be more suitable to consider a modelling perspective which does not assume any a priori and discretional hypothesis on the functional form of the model and allows obtaining models in which “data speak for themselves” without assuming any parametric restriction.

The great advances carried out in the field of Computer Science have allowed developing and applying new non-parametric techniques for the estimation and prediction of different scientific phenomena, and for modelling and solving engineering problems. One of these computational techniques, called Artificial Neural Networks (ANN), has been broadly applied to several economic and energy issues in recent years, such as the world green energy use (Ermis et al., 2007), the transport energy demand (Murat and Ceylan, 2006) and the energy consumption (Nizami and Al-Garni, 1995; Hamzaçebi, 2007; Sözen and Arcaklioglu, 2007). All these applications recognize the validity of this non-parametric approach in the study of energy issues. Another technique, Genetic Programming (GP), is inspired by Genetics and by the Darwinian theories of natural selection and survival (Holland, 1975; Koza, 1992; Mitchell 2001). The method has already been used
satisfactorily in different areas, including economics (Beenstock and Szpiro, 2002), finance (Álvarez-Díaz and Álvarez, 2003, 2005), natural resource economics (Álvarez-Díaz and Domínguez-Torreiro, 2006), energy (Azadeh and Tarverdian, 2007) and institutional economics (Álvarez-Díaz and Caballero, 2008). The increasing and intense spread of GP in the last years is mainly due to its advantages. Unlike other methods based on Computer Science, the GP offers explicitly a mathematical equation which allows a simple ad hoc interpretation of the results.

The main goal of this study is to apply these two computational methods to analyse how institutions influence the CO₂ emissions, and which variables are the most relevant to explain this influence across countries. We compare the results using these computational methods with those employing a more traditional parametric perspective (OLS). Besides this, a secondary objective is to detect possible misspecifications problems associated to the traditional parametric models. This verification is crucial in an empirical application and it should be always done in order to verify and corroborate the adequacy of the parametric results.

The paper is presented as follows. After this introduction, Section 2 presents a brief explanation of the methods used in our study. In Section 3, the data are described. In Section 4, the results obtained for each method are presented. Finally, in Section 5, we draw our conclusions.
2. Computational Methods: Artificial Neural Network and Genetic Programming

Remarkable developments of computer hardware and software have allowed an improvement and generalized use of sophisticated non-parametric modelling methods. Among them, artificial neural networks and genetic programming are two of the most relevant methods. Next, we technically describe these computational methods.

2.1. Artificial Neural Networks

Artificial neural network is a modelling technique inspired by the findings of studies on how the brain and the nervous system work. It has been widely employed in numerous fields such as medicine, national defense and security, entertainment, robotics and physics, among many other. The use of ANN in economics and finance for forecasting and modelling purposes is noteworthy (Refenes, 1995). The literature on this topic distinguishes among different types of networks (Gately, 1996), although the feed-forward multi-layer network with a learning algorithm based on the back-propagation technique (Rumelhart et al., 1986) is certainly the most popular in economics and finance (Trippi and Turban, 1996). Other types of networks such as radial-basis function networks, recurrent neural networks or wavelets are also very useful, but much less used.

The main advantage offered by ANN is the great capability to detect and exploit any non-linearity that might exist in the data, even under conditions of incomplete data or where the presence of noise is important. Specifically, the technique is considered a universal functional approximator. Indeed, it has been demonstrated that a neural network correctly designed can approximate any continuous function to any desired level of

6
accuracy (Cybenko, 1989). In this manner, the technique is more suitable than traditional methods to model and predict phenomena characterized by a complex behavior (Bishop, 1995; Smith, 1995).

The great majority of empirical applications have showed that ANN scored as good as, or significantly better than, the traditional parametric methods. Nevertheless, the method is not exempt of some important drawbacks. For example, it is difficult to analyze the impact of the explanatory variables on the dependent variable and, moreover, it is difficult to perform traditional statistical inference to construct confidence intervals or check the statistical significance of the predictions. Moreover, the design of a neural network is a tedious and time-consuming procedure in which the user must specify a correct architecture. Finally, and more important, the great power of the neural networks to replicate data can be also a disadvantage. There is a risk that the network merely mimic data and it cannot generalize new observations. In this case, the model would fit irrelevant characteristics existing in the data rather than fitting the underlying function which links inputs and outputs. If this fact occurred, the network would lose its capacity to predict accurately untouched observations (overfitting problem).

Technically speaking, an ANN is composed by an input layer, an output layer and one or more hidden layers. Each layer has a group of process units called neurons or nodes. These nodes are connected to nodes at adjacent layer. The connections, called synapses, are weighted by a series of coefficients. The goal will be to find the values of these weights that minimize the forecast errors. Formally, the statistical formulation of a feed-forward network can be expressed as

$$y_0 = \Phi\left(\beta_0 + \sum_{k=1}^{H} \beta_k \cdot \Psi(t) \left(\theta_0 + \sum_{j=1}^{J} \alpha_j \cdot x_j\right)\right)$$
where \( \hat{y} \) is the output of the model, and the functions \( \Psi(\cdot) \) and \( \Phi(\cdot) \) are denoted as transfer functions of the hidden and output levels, respectively. The network has \( J \) inputs (explanatory variables or delays of the dependent variable), \( H \) process units (neurons) in the hidden level and one output. Initially, the weights \( \alpha_h \) and \( \beta_h \) are randomly determined within a given range. By means of an iterative learning process, based on the back-propagation technique, the values of these weights are modified such that the difference between the real value and the estimated value (i.e., the output of the neural network) is minimal. Several theoretical studies done on this sort of network, with just one hidden level and a high enough number of units, have demonstrated that it can approximate any non-linear function with a certain degree of accuracy (Cybenko, 1989; White, 1990).

In addition to the complexity of the data, the success of the prediction of a neural network depends, to a great extent, on the correct determination of its architecture. It is therefore necessary to accurately specify the number of inputs (\( J \)) and the number of process units in the hidden level (\( H \)), as well as to select the right structure of the transfer functions, (\( \Psi(\cdot) \) and \( \Phi(\cdot) \)). An excessive number of \( H \), for example, might create overfitting problems, thus eliminating any generalization. On the other hand, with an insufficient number of process units, the network could lose its forecasting capability because it would not fully exploit the non-linearity in the data. In the literature, one can find different rules for defining how many inputs and process units there must be in the hidden unit, but none of them is perfect nor have any of them ever been adopted as a general rule (Yao et al., 1999). A frequent recommendation is to determine \( J \) and \( H \) through a process of trial-and-error. Therefore, following this recommendation, it is necessary to consider
different architectures and choose that which produces the least errors in a sub-set reserved exclusively to this purpose, (i.e., the selection set).

With regard to the transfer functions \( \Psi(\cdot) \) and \( \Phi(\cdot) \), only a small number of “well-behaved” (bounded, monotonically increasing and differentiable) functions are used in practice (Zhang et al., 1998). Among them, we have chosen for our exercise the hyperbolic tangent function which is one of the most frequently employed.

Besides determining the number of inputs, hidden neurons and the form of the transform functions, an additional problem to be solved is that of the excessive variability in the results when considering different initial weights for the values of \( \alpha_j \) and \( \beta_h \) (Racine, 2000). The solution proposed in the specialized literature and followed in our application will be to run the networks considering different weights and choose that architecture which optimises the fit criterion in the selection sub-sample (Hu et al., 1999).

2.2. Genetic Programming

Genetic Algorithms, originally developed by Holland (1975), enclose a whole series of computing procedures inspired in biologic concepts based on the Theory of Evolution of Species: survival of the fittest individuals, reproduction and birth of offspring with a good genetic heritage. The basic characteristic of these procedures is to use some evolutionary rules observed in the Nature as inspiration for solving certain mathematical optimization process. Specifically, from the evolution of a random set of possible solutions and by means of applying operators based on natural selection concepts, these methods allow
finding an optimal approximation to the solution of different types of problems, including modelling issues.

In the specialized literature there is no a commonly accepted definition of genetic algorithms which allows us to distinguish them from other computational evolutionary methods. However, there exist many programs considered as genetic algorithms which present the following common elements: initial population of possible solutions to the problem, selection process using some fit criterion, and use of crossover and random mutation to generate new solutions (Mitchell, 2001). Different variations of genetic algorithms have been applied to model a large number of scientific and engineering problems. In this paper we have used a kind of genetic algorithm, called genetic programming (Koza, 1992; Álvarez et al., 2001), as a modelling tool. The evolutionary process was programmed in FORTRAN, and it can be explained by means of a series of stages. At a first stage, the genetic programming creates a random initial population of $N$ mathematical equations susceptible of representing accurately the relationship between the dependent variable $Y_i$ and a set of possible explanatory variables $X = \{X_{1i}, X_{2i}, \ldots, X_{0i}\}$. These mathematical equations are created by means of a random combination of operators and arguments in the following way:

$$S_j : \left( (A \otimes B) \otimes (C \otimes D) \right) \quad \forall 1 \leq j \leq N$$

where $A$, $B$, $C$, and $D$ are the arguments, the symbol $\otimes$ represents the mathematical operators and the subscript $j$ refers to each one of the $N$ equations belonging to the initial population. The arguments can be real numbers included in a certain interval (the coefficients of the model) or explanatory variables. Besides, the mathematical operators ($\otimes$) used will be the sum (+), subtraction (-), multiplication (·) and division (/), being the
latter ‘protected’ to prevent zero divisors. It is also possible to include other mathematical operators (such as logarithm or the trigonometric ones) but at the expense of increasing the complexity in the functional optimisation process.

At a second stage, after determining the initial population of candidates, the evolution process starts selecting those equations that fit best to the problem. For this purpose, the adjusted R-Square has been adopted as fitness criterion. This performance measure is defined as:

$$R^2_j = 1 - \frac{M - 1}{M - k} \cdot \frac{\sum_{i=1}^{M} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{M} (Y_i - \text{mean}(Y_i))^2} \quad \forall 1 \leq j \leq N$$

where $R^2_j$ is the adjusted R-Square obtained by equation $j$, $Y_i$ is the observed value, $\hat{Y}_i$ is the estimated value, $k$ is the number of explanatory variables and $M$ is the total number of observations in the sub-sample employed to train the genetic program. Later on, all equations of the initial population are classified in decreasing order according to the value of $R^2_j$. Those equations whose value of $R^2_j$ is very low are rejected, while those with a high value are more likely to survive, being the base for the next generation of equations.

The equations that survived after the selection process are used to create the equations of a new solutions generation (i.e., reproduction process). In order to do that the so-called genetic operators will be applied: cloning, crossover and mutation. With the cloning operator, the fittest equations are replicated in the next generation. With the crossover operator pairs of equations with high values of $R^2_j$ are selected and they exchange part of their arguments and mathematical operators. Finally, mutation means that any operator or argument is randomly replaced in a small number of equations. The first top
ranked individuals are exempted from mutation, so that their information is not lost. Let us consider, for example, that the following equations belong to the initial population:

\[ S_1 : (A + B) / C \]
\[ S_2 : (D \cdot E) - G \]

where A, B, C, D, E and G are the equation arguments (coefficients and independent variables). Let us suppose that both expressions will survive the selection process and so they become the base equations for the next generation. The crossover operator means the random selection of a block of operators and arguments in each equation and their later exchange. For instance, let us suppose that the block (A+B) in expression \( S_1 \) and the argument \( G \) in expression \( S_2 \) have been selected. By means of an exchange of blocks two new equations appear as follows:

\[ S_3 : G / C \]
\[ S_4 : (D \cdot E) - (A + B) \]

As one can observe, the new equations inherit certain features from their parents.

Now let us suppose that the expression \( S_1 \) is selected again and the mutation operator is applied. So, the following equation can be obtained from \( S_1 \):

\[ S_5 : (A \cdot B) / C \]

where the mutation was the random alteration of a mathematical operator.

In short, the new population created from the initial population of equations is composed of cloned equations (such as \( S_2 \)), mutated expressions (such as \( S_5 \)), or crossed (such as \( S_3 \) and \( S_4 \)). From this moment, the process will repeat the selection and reproduction stages in an iterative way. After a given number of generations, determined by
the user, the iteration procedure ceases and an optimal mapping \( \hat{Y} = F(X_1, X_2, \ldots, X_k) \) is given by the strongest mathematical equation in the population.

3. Data

This paper analyses the functional relation between the CO\(_2\) emissions and a set of historical, geographical, economical, religious and social variables. Table 1 provides a brief description of the database used in this study. The CO\(_2\) emissions per capita (CO2) are the endogenous variable, while the other nine variables are considered as explanatory variables. In this way, the database constructed for our study contains complete information about 113 countries, whose emission levels are shown in Fig. 1.

*Fig. 1. Countries database and CO\(_2\) emissions per capita (metric tons, 2003)*
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETHF</td>
<td>Ethnolinguistic Fractionalization: Average value of five different indices of ethnolinguistic fractionalization. Its value ranges from 0 to 1 (Easterly and Levine (1997), as used in La Porta et al. (1999) and Acemoglu et al. (2001)).</td>
</tr>
<tr>
<td>PROT</td>
<td>Protestant Religion: Identifies the percentage of the population of each country that is protestant. The numbers are in percent (scale from 0 to 100). From La Porta et al. (1999).</td>
</tr>
<tr>
<td>CAT</td>
<td>Catholic Religion: Identifies the percentage of the population of each country that is catholic. The numbers are in percent (scale from 0 to 100). From La Porta et al. (1999).</td>
</tr>
<tr>
<td>OTHERS</td>
<td>Other Religion: Identifies the percentage of the population of each country that belongs to other religions (non-catholic and non protestant). The numbers are in percent (scale from 0 to 100). This data series was elaborated for this paper based on La Porta et al. (1999).</td>
</tr>
<tr>
<td>LATIT</td>
<td>Latitude: Absolute value of the latitude of the country (i.e., a measure of distance from the equator), scaled to take values between 0 and 1, where 0 is the equator. From the CIA, as used in La Porta et al. (1999).</td>
</tr>
</tbody>
</table>
Regarding to the dependent variable, the data come from the 2006 Human Development Report (UNDP, 2007). CO₂ emissions, measured in metric tons, include contributions to the carbon dioxide flux from the consumption of solid, liquid and gaseous fuels, gas flaring and the production of cement.

On the other hand, the explanatory variables include the ethnolinguistic fractionalization, the legal tradition (English Common Law, Socialist/Communist Law, French Commercial Code), the religion (Roman Catholic, Protestant, Others), the geographical (latitude) and economic (GNP) condition. We decided to consider these variables because they are considered relevant to explain the institutional quality across countries and, therefore, they are a good approximation to the “traditions and institutions by which authority in a country is exercised” (Kaufmann et al., 1999). The relevance of these variables as a good proxy of the quality of institutions has been tested by La Porta et al. (1999) and Álvarez-Díaz and Caballero (2008). Of course, several sets of institutional variables could be considered to study the institutional determinants of CO₂ emissions. Nevertheless, we had to select one of these possible sets, and we have decided to use that set that has been used in the previous relevant literature on the quality of institutions.

The nine explanatory variables may be characterized as follows. Ethnolinguistic fractionalization (ETHF) is the average value of five indices that measure the existence of different ethnolinguistic groups in a country and the percent of population not speaking the official language, among other elements. Legal tradition is a dummy explanatory variable that shows the legal origin of the company law or commercial code of each country, distinguishing the English tradition (ENG), the socialist tradition (SOCI) and the French

---

1 The data are originally from the Carbon Dioxide Information Analysis Center, Environmental Sciences Division, Oak Ridge National Laboratory, US Department of Energy.
tradition (FRENCH). In regard to religion, we consider the percentage of the population of each country that professes the protestant religion (PROT), the catholic religion (CAT) and other religions (OTHERS). Regarding economic development, the model incorporates the distance of the country from the equator (LATIT) and the logarithm of Gross National Product per capita (GNP). Table 2 shows descriptive statistics of the endogenous and explanatory variables.

Table 2. Descriptive statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Endogenous variable</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO2</td>
<td>4.5667</td>
<td>5.6345</td>
<td>0.10</td>
<td>33.60</td>
</tr>
<tr>
<td><strong>Continuous explanatory variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETHF</td>
<td>0.3341</td>
<td>0.2993</td>
<td>0.0</td>
<td>0.89</td>
</tr>
<tr>
<td>PROT</td>
<td>13.6423</td>
<td>22.3854</td>
<td>0.0</td>
<td>97.80</td>
</tr>
<tr>
<td>CAT</td>
<td>33.6685</td>
<td>36.6493</td>
<td>0.0</td>
<td>97.30</td>
</tr>
<tr>
<td>OTHERS</td>
<td>52.5196</td>
<td>39.5973</td>
<td>0.9</td>
<td>100.00</td>
</tr>
<tr>
<td>LATIT</td>
<td>0.2782</td>
<td>0.1881</td>
<td>0.01</td>
<td>0.72</td>
</tr>
<tr>
<td>GNP</td>
<td>7.3354</td>
<td>1.4555</td>
<td>4.72</td>
<td>10.15</td>
</tr>
<tr>
<td><strong>Dummy explanatory variables</strong></td>
<td></td>
<td></td>
<td></td>
<td>Frequency</td>
</tr>
<tr>
<td>ENG</td>
<td></td>
<td></td>
<td></td>
<td>34.23</td>
</tr>
<tr>
<td>SOCI</td>
<td></td>
<td></td>
<td></td>
<td>8.11</td>
</tr>
<tr>
<td>FRENCH</td>
<td></td>
<td></td>
<td></td>
<td>48.65</td>
</tr>
</tbody>
</table>
Once that the data have been described and before carrying out the modelling exercise, an important issue should be commented. Usually, in the majority of the empirical studies using computational methods, the total available data are divided into a training set (in-sample data) and a test set (out-of-sample or hold-out sample). In theory, the training set is used for the construction of the model while the test set is employed for measuring the adequacy and predictive ability of the method. Nevertheless, researchers using non-linear forecasting methods can be tempted to try different modeling structures in the training set and select that with the highest accuracy in the test set. The result would be a model that fits the data too closely and with a strong predictive ability in the test set. However, it would be not capable of generalizing and performing well with new data. The modelling procedure would suffer from an overfitting problem and the utility of the model would be practically null.

In order to detect overfitting problems and develop a useful and fair modelling exercise, researchers should follow the technical and practical recommendations and guidelines proposed in the literature on computer science (Bishop, 1995). These recommendations advise to divide the sample into three sub-sets. For this reason, in this paper we assume this advice and consider three sub-sets when applying computational methods. The first one is composed by 67 observations randomly chosen, and it is reserved to train the ANN and develop the evolutionary process of the genetic programming. The second sub-set covers 23 observations randomly chosen. It constitutes the selection sub-sample and is used to adjust the technical parameters of the genetic programming and to find an optimal architecture of the neural network. Finally, the last 23 observations are not employed in the modelling process. These untouched data conform the out-of-sample set and the value of the accuracy measure obtained in this sub-sample is employed to check the adequacy of the model. Specifically, it will be necessary that the considered accuracy
measure shows a similar and relatively high value in the selection and out-of-sample sets. If this condition was verified, it would be proved the ability of the non-parametric method set-up to get a model that generalizes new observations and, therefore, without showing overfitting problems.

4. Results

4.1. OLS Results

In order to analyze the institutional determinants of CO₂ emissions, we model firstly the relationship between variables assuming the parametric point of view. Specifically, we estimate a linear model using OLS regression in the in-sample. The explanatory variables of the OLS model were chosen following the backwards stepwise procedure with a significance level at 10%. As we can observe in Table 3, OLS results reveal that the relevant variables to explain the CO₂ emissions are GNP, ENG, ETHF, PROT and SOCI. The sign of the estimated coefficients is positive for GNP, ENG and SOCI, while it is negative for ETHF and PROT. The positive relationship between CO₂ emissions and GNP is in accordance with the a priori expectative and the economic logic: more economic growth implies more CO₂ emissions. Moreover, the results also reveal that those countries with a Socialist or English Law tradition tend to produce higher levels of CO₂ emissions, whereas those countries with a high level of ethnolinguistic fractionalization and a high percentage of protestant population have a more controlled level of emissions. Other variables such as CAT, OTHERS or LATIT do not have a significant effect on CO₂ emissions.
### Table 3. OLS results

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coefficients (p-values)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>-5.583 (0.00)</td>
</tr>
<tr>
<td>GNP</td>
<td>8.851 (0.00)</td>
</tr>
<tr>
<td>ENG</td>
<td>0.413 (0.01)</td>
</tr>
<tr>
<td>ETHF</td>
<td>-1.174 (0.00)</td>
</tr>
<tr>
<td>PROT</td>
<td>-0.007 (0.03)</td>
</tr>
<tr>
<td>SOCI</td>
<td>0.666 (0.00)</td>
</tr>
</tbody>
</table>

**In-Sample** 0.8348  
**Adjusted R-Square**  
**Out-of-Sample** 0.7825

Table 3 also shows the Adjusted R-Square for the first 90 observations (the in-sample), and for the last 23 observations (the out-of-sample). This fit criterion exhibits a relatively high value for the in-sample (0.8473), while that for the out-of-sample the value decreases up to 0.7925. This reduction, even small, seems to reveal a possible lack of generalization using the OLS model. The model gets a high value in the sample where the coefficients are estimated (in-sample) but the performance is damaged when new data are considered (out-of-sample). Thus, the model could suffer from a problem of misspecification, biasing the results and, therefore, misunderstanding our conclusions. For
example, are really the selected variables the most important to explain the CO₂ emissions? We could question as well if the effect of the selected variables are real or spurious because of assuming a specific and rigid functional form. In order to validate and investigate the possible existence of a bias in the OLS results, we compare them with those obtained employing computational non-parametric methods.

4.2. GP Results

Regarding the GP results, the optimal model survival to the evolutionary process is given by the expression shown in Table 4.

Table 4. GP results

<table>
<thead>
<tr>
<th>Equation</th>
<th>Survival Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ CO₂_i = 3 \cdot GNP_i - ETHF_i - \frac{1.38}{3.0170 \cdot GNP_i \cdot (0.07 + GNP_i)} ]</td>
<td>GNP (+) ETHF (-)</td>
</tr>
</tbody>
</table>

| Training | 0.8143 |
| Selection | 0.8141 |
| Out-of-sample | 0.8256 |

As we can see analyzing the equation provided by the GP, there are only two variables that are relevant to explain the CO₂ emissions: GNP and ETHF. Jointly to the positive and foreseen effect of GNP, ETHF shows again a negative and linear effect on CO₂. This implies that for a same level of economic growth, those countries with a higher
ethnolinguistic fractionalization show lower levels of CO\textsubscript{2} emissions. On the other hand, the GP did not discover any clear effect of the religious, legal and geographical variables.

The adjusted R-Square is relatively high (over 0.80) and nearly constant when Training, Selection and Out-of-sample are considered (0.8143, 0.8141 and 0.8256, respectively). This stability reveals the absence of a possible lack of generalization. It seems that the model has discovered a good approximation to the general pattern existing in the data rather than memorizing some specific features of the individual observations (overfitting problem).

4.3. ANN Results

Before analyzing and discussing the results obtained by the ANN, it is necessary to explain how the inputs were finally selected. Given the complexity of the modelling process using ANN, the literature usually recommends to select as inputs those variables that were already chosen using a simpler method such as the linear model (Bishop, 1995). Therefore, we should insert as inputs of our ANN the explanatory variables of the OLS model: GNP, ENG, ETHF, PROT and SOCI. Table 5 depicts the optimum number of neurons and the Adjusted R-Square for an ANN with those variables used in the linear model. Observing the results, we can underline, first, the stability of the adjusted R-Square in the different sub-samples and, second, the fact that the network has considerably improved the out-of-sample performance of the linear model (0.8497 versus 0.7825). It seems, therefore, that there exists an important non-linear relationship between CO\textsubscript{2} emissions and the explanatory variables considered in this study. Unlike the OLS model, the network is capable of exploiting the nonlinearity existing in the data.
In spite of the good results following the recommendation suggested in the literature, it can be possible to improve our results considering those inputs that already were selected by another non-parametric method. This justification lies in the fact that a set of explanatory variables can be “linearly” appropriate, but from a non-linear point of view we could have found a better set of variables. For example, we can use the survival variables obtained by the GP (GNP and ETHF) as inputs of a network. Table 5 also shows the results for this case. The new specification of the network does not improve the previous result. However, it is still better than the OLS model in terms of fitness and stability.

Table 5. ANN results

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Neurons</th>
<th>Adjusted R-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>In-Sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Training Selection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Out-of-sample</td>
</tr>
<tr>
<td>GNP</td>
<td>3</td>
<td>0.8528 0.8378 0.8497</td>
</tr>
<tr>
<td>ENG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ETHF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROT</td>
<td>5</td>
<td>0.8417 0.8355 0.8420</td>
</tr>
<tr>
<td>SOCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNP</td>
<td>4</td>
<td>0.8734 0.8745 0.8631</td>
</tr>
<tr>
<td>ETHF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNP</td>
<td>OTHERS</td>
<td></td>
</tr>
<tr>
<td>CAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS</td>
<td>PROT</td>
<td></td>
</tr>
<tr>
<td>LATIT</td>
<td>ETHF</td>
<td></td>
</tr>
</tbody>
</table>
Finally, another way of selecting a set of variables as inputs of our ANN is to assume a trial-and-error approach. Different sets of inputs are considered, and that combination of variables with the highest fitness in the selection sub-sample will be finally chosen. The computational requirement is too high but, however, it can be a useful way of, first, improving even more the fitness and, second, knowing what variables are the most relevant to explain the CO₂ emissions. Analyzing all possible subsets of variables is not feasible from a practical point of view. For this reason, in our study we have restricted the number of combinations to one hundred sets of inputs. As we can see in Table 5, the inputs chosen using this approach are GNP, CAT, OTHERS, PROT, LATIT and ETHF. The Adjusted R-Square shows a high value and a great stability in the training, selection and out-of-sample sets (0.8734, 0.8745 and 0.8631, respectively). In this sense, this input set provides the best fitness in comparison with other possible combinations.

5. Conclusions

In the traditional economic literature, the CO₂ emissions have been explained by the level of production and economic growth. Nevertheless, there may be other variables that explain the rising concentration of CO₂. In this paper we consider a set of institutional factors in order to examine their functional relationship with CO₂ emissions around the world. This analysis helps to open the “black box” of the determinants of CO₂ emissions. For this aim, we employ two computational and non-parametric methods: Genetic Programming and Artificial Neural Networks. These sophisticated techniques imply a solid improvement when the relationship between variables is non-linear. The computational methods allow us to approach accurately to the underlying relationship which links CO₂
emissions with institutional and economic variables. Moreover, we have compared the results using the computational methods with those employing a more traditional parametric perspective.

According to the specialized literature and the economic insights, our results show that the level of GNP is the main determinant of CO₂ emissions. Nevertheless, the empirical results of this paper prove too that the institutional variables matter in determining the level of CO₂ emissions. In this sense, independently of the level of production, the institutional structure of society affects the economic organization of production, by which the institutional rules are useful for the understanding of the levels of CO₂ emissions across countries. Specifically, the application of different methods has allowed us to conclude the relevant influence of the ethnolinguistic fractionalization on CO₂ emissions: for a fixed level of GNP, the higher the ethnolinguistic fractionalization is, the lower the emissions of CO₂ are. The empirical results point out that those countries with a homogeneous ethnolinguistic society have an institutional structure of production that does not favor the control of the emission levels.

In summary, the institutional rules allow us a better understanding of the level of CO₂ emissions across countries. Thus, when we consider the institutional factors as explanatory variables of CO₂ emissions, the fitness of the models is around the 85%. The major results are associated with the computational methods. In this sense, the paper presents empirical evidence about the explanatory ability of the binomial GNP (positive effect) and ethnolinguistic fractionalization (negative effect) on CO₂ emissions. Moreover, an artificial neural network points out that other social, cultural and geographical variables are useful to explain CO₂ emissions across countries. This paper has opened a research avenue that will require new efforts in the future to advance in the knowledge of the
institutional determinants of CO₂ emissions. In this sense, this paper has allowed us a first advance, but it is only a small piece of empirical work with a set of variables. We recognize that more effort should be taken into account for future endeavours and developments of this research, including the study of the role of other institutional and non-institutional variables.

References


Bishop, C.M., 1995. Neural Networks for pattern Recognition, Oxford University Press.


<table>
<thead>
<tr>
<th>Número</th>
<th>Título</th>
<th>Autor/Autores</th>
</tr>
</thead>
<tbody>
<tr>
<td>159/2000</td>
<td>Participación privada en la construcción y explotación de carreteras de peaje</td>
<td>Ginés de Rus, Manuel Romero y Lourdes Trujillo</td>
</tr>
<tr>
<td>160/2000</td>
<td>Errores y posibles soluciones en la aplicación del <em>Value at Risk</em></td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>161/2000</td>
<td>Tax neutrality on saving assets. The spahish case before and after the tax reform</td>
<td>Cristina Ruza y de Paz-Curbera</td>
</tr>
<tr>
<td>163/2000</td>
<td>El control interno del riesgo. Una propuesta de sistema de límites riesgo neutral</td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>164/2001</td>
<td>La evolución de las políticas de gasto de las Administraciones Públicas en los años 90</td>
<td>Alfonso Utrilla de la Hoz y Carmen Pérez Esparrells</td>
</tr>
<tr>
<td>165/2001</td>
<td>Bank cost efficiency and output specification</td>
<td>Emili Tortosa-Ausina</td>
</tr>
<tr>
<td>166/2001</td>
<td>Recent trends in Spanish income distribution: A robust picture of falling income inequality</td>
<td>Josep Oliver-Alonso, Xavier Ramos y José Luis Raymond-Bara</td>
</tr>
<tr>
<td>167/2001</td>
<td>Efectos redistributivos y sobre el bienestar social del tratamiento de las cargas familiares en el nuevo IRPF</td>
<td>Nuria Badenes Plá, Julio López Laborda, Jorge Onrubia Fernández</td>
</tr>
<tr>
<td>168/2001</td>
<td>The Effects of Bank Debt on Financial Structure of Small and Medium Firms in some European Countries</td>
<td>Mónica Melle-Hernández</td>
</tr>
<tr>
<td>169/2001</td>
<td>La política de cohesión de la UE ampliada: la perspectiva de España</td>
<td>Ismael Sanz Labrador</td>
</tr>
<tr>
<td>170/2002</td>
<td>Riesgo de liquidez de Mercado</td>
<td>Mariano González Sánchez</td>
</tr>
<tr>
<td>171/2002</td>
<td>Los costes de administración para el afiliado en los sistemas de pensiones basados en cuentas de capitalización individual: medida y comparación internacional.</td>
<td>José Enrique Devesa Carpio, Rosa Rodriguez Barrera, Carlos Vidal Meliá</td>
</tr>
<tr>
<td>172/2002</td>
<td>La encuesta continua de presupuestos familiares (1985-1996): descripción, representatividad y propuestas de metodología para la explotación de la información de los ingresos y el gasto.</td>
<td>Llorenç Pou, Joaquín Alegre</td>
</tr>
<tr>
<td>173/2002</td>
<td>Modelos paramétricos y no paramétricos en problemas de concesión de tarjetas de credito.</td>
<td>Rosa Puertas, María Bonilla, Ignacio Olmeda</td>
</tr>
</tbody>
</table>
174/2002 Mercado único, comercio intra-industrial y costes de ajuste en las manufacturas españolas. José Vicente Blanes Cristóbal

175/2003 La Administración tributaria en España. Un análisis de la gestión a través de los ingresos y de los gastos. Juan de Dios Jiménez Aguilera, Pedro Enrique Barrilao González


177/2003 Effects of ATMs and Electronic Payments on Banking Costs: The Spanish Case. Santiago Carbó Valverde, Rafael López del Paso, David B. Humphrey

178/2003 Factors explaining the interest margin in the banking sectors of the European Union. Joaquín Maudos y Juan Fernández Guevara

179/2003 Los planes de stock options para directivos y consejeros y su valoración por el mercado de valores en España. Mónica Melle Hernández


181/2003 The Euro effect on the integration of the European stock markets. Mónica Melle Hernández

182/2004 In search of complementarity in the innovation strategy: international R&D and external knowledge acquisition. Bruno Cassiman, Reinhilde Veugelers

183/2004 Fijación de precios en el sector público: una aplicación para el servicio municipal de suministro de agua. Mª Ángeles García Valiñas

184/2004 Estimación de la economía sumergida en España: un modelo estructural de variables latentes. Ángel Alañón Pardo, Miguel Gómez de Antonio

185/2004 Causas políticas y consecuencias sociales de la corrupción. Joan Oriol Prats Cabrera

186/2004 Loan bankers’ decisions and sensitivity to the audit report using the belief revision model. Andrés Guiral Contreras and José A. Gonzalo Angulo

187/2004 El modelo de Black, Derman y Toy en la práctica. Aplicación al mercado español. Marta Tolentino García-Abadillo y Antonio Díaz Pérez

188/2004 Does market competition make banks perform well?. Mónica Melle

189/2004 Efficiency differences among banks: external, technical, internal, and managerial. Santiago Carbó Valverde, David B. Humphrey y Rafael López del Paso
190/2004 Una aproximación al análisis de los costes de la esquizofrenia en españa: los modelos jerárquicos bayesianos
F. J. Vázquez-Polo, M. A. Negrín, J. M. Cavasés, E. Sánchez y grupo RIRAG

191/2004 Environmental proactivity and business performance: an empirical analysis
Javier González-Benito y Óscar González-Benito

192/2004 Economic risk to beneficiaries in notional defined contribution accounts (NDCs)
Carlos Vidal-Meliá, Inmaculada Domínguez-Fabian y José Enrique Devesa-Carpio

193/2004 Sources of efficiency gains in port reform: non parametric malmquist decomposition tfp index for Mexico
Antonio Estache, Beatriz Tovar de la Fé y Lourdes Trujillo

194/2004 Persistencia de resultados en los fondos de inversión españoles
Alfredo Ciriaco Fernández y Rafael Santamaría Aquilué

195/2005 El modelo de revisión de creencias como aproximación psicológica a la formación del juicio del auditor sobre la gestión continuada
Andrés Guiral Contreras y Francisco Esteso Sánchez

196/2005 La nueva financiación sanitaria en España: descentralización y prospectiva
David Cantarero Prieto

197/2005 A cointegration analysis of the Long-Run supply response of Spanish agriculture to the common agricultural policy
José A. Mendez, Ricardo Mora y Carlos San Juan

198/2005 ¿Refleja la estructura temporal de los tipos de interés del mercado español preferencia por la liquidez?
Magdalena Massot Perelló y Juan M. Nave

199/2005 Análisis de impacto de los Fondos Estructurales Europeos recibidos por una economía regional: Un enfoque a través de Matrices de Contabilidad Social
M. Carmen Lima y M. Alejandro Cardene

200/2005 Does the development of non-cash payments affect monetary policy transmission?
Santiago Carbó Valverde y Rafael López del Paso

201/2005 Firm and time varying technical and allocative efficiency: an application for port cargo handling firms
Ana Rodríguez-Álvarez, Beatriz Tovar de la Fé y Lourdes Trujillo

202/2005 Contractual complexity in strategic alliances
Jeffrey J. Reuer y Africa Ariño

203/2005 Factores determinantes de la evolución del empleo en las empresas adquiridas por opa
Nuria Alcalde Fradejas y Inés Pérez-Soba Aguilar

Elena Olmedo, Juan M. Valderas, Ricardo Gimeno and Lorenzo Escot
<table>
<thead>
<tr>
<th>Year</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>205/05</td>
<td>Precio de la tierra con presión urbana: un modelo para España</td>
<td>Esther Decimavilla, Carlos San Juan y Stefan Sperlich</td>
</tr>
<tr>
<td>206/05</td>
<td>Interregional migration in Spain: a semiparametric analysis</td>
<td>Adolfo Maza y José Villaverde</td>
</tr>
<tr>
<td>207/05</td>
<td>Productivity growth in European banking</td>
<td>Carmen Murillo-Melchior, José Manuel Pastor y Emili Tortosa-Ausina</td>
</tr>
<tr>
<td>208/05</td>
<td>Explaining Bank Cost Efficiency in Europe: Environmental and Productivity Influences.</td>
<td>Santiago Carbó Valverde, David B. Humphrey y Rafael López del Paso</td>
</tr>
<tr>
<td>209/05</td>
<td>La elasticidad de sustitución intertemporal con preferencias no separables intratemporalmente: los casos de Alemania, España y Francia.</td>
<td>Elena Márquez de la Cruz, Ana R. Martínez Cañete y Inés Pérez-Soba Aguilar</td>
</tr>
<tr>
<td>210/05</td>
<td>Contribución de los efectos tamaño, book-to-market y momentum a la valoración de activos: el caso español.</td>
<td>Begoña Font-Belaire y Alfredo Juan Grau-Grau</td>
</tr>
<tr>
<td>211/05</td>
<td>Permanent income, convergence and inequality among countries</td>
<td>José M. Pastor and Lorenzo Serrano</td>
</tr>
<tr>
<td>212/05</td>
<td>The Latin Model of Welfare: Do ‘Insertion Contracts’ Reduce Long-Term Dependence?</td>
<td>Luis Ayala and Magdalena Rodríguez</td>
</tr>
<tr>
<td>213/05</td>
<td>The effect of geographic expansion on the productivity of Spanish savings banks</td>
<td>Manuel Illueca, José M. Pastor and Emili Tortosa-Ausina</td>
</tr>
<tr>
<td>214/05</td>
<td>Dynamic network interconnection under consumer switching costs</td>
<td>Ángel Luis López Rodríguez</td>
</tr>
<tr>
<td>215/05</td>
<td>La influencia del entorno socioeconómico en la realización de estudios universitarios: una aproximación al caso español en la década de los noventa</td>
<td>Marta Rahona López</td>
</tr>
<tr>
<td>216/05</td>
<td>The valuation of spanish ipos: efficiency analysis</td>
<td>Susana Álvarez Otero</td>
</tr>
<tr>
<td>217/05</td>
<td>On the generation of a regular multi-input multi-output technology using parametric output distance functions</td>
<td>Sergio Perelman and Daniel Santín</td>
</tr>
<tr>
<td>218/05</td>
<td>La gobernanza de los procesos parlamentarios: la organización industrial del congreso de los diputados en España</td>
<td>Gonzalo Caballero Miguez</td>
</tr>
<tr>
<td>219/05</td>
<td>Determinants of bank market structure: Efficiency and political economy variables</td>
<td>Francisco González</td>
</tr>
<tr>
<td>220/05</td>
<td>Agresividad de las órdenes introducidas en el mercado español: estrategias, determinantes y medidas de performance</td>
<td>David Abad Diaz</td>
</tr>
</tbody>
</table>
221/2005  Tendencia post-anuncio de resultados contables: evidencia para el mercado español
Carlos Forner Rodríguez, Joaquín Marhuenda Fructuoso y Sonia Sanabria García

222/2005  Human capital accumulation and geography: empirical evidence in the European Union
Jesús López-Rodríguez, J. Andrés Faíña y Jose Lopez Rodriguez

223/2005  Auditors' Forecasting in Going Concern Decisions: Framing, Confidence and Information Processing
Waymond Rodgers and Andrés Guiral

José Ramón Canelo de la Torre, J. Andrés Faíña and Jesús López-Rodríguez

225/2005  The effects of ownership structure and board composition on the audit committee activity: Spanish evidence
Carlos Fernández Méndez and Rubén Arrondo García

226/2005  Cross-country determinants of bank income smoothing by managing loan loss provisions
Ana Rosa Fonseca and Francisco González

Alejandro Estellér Moré

228/2005  Region versus Industry effects: volatility transmission
Pilar Soriano Felipe and Francisco J. Climent Diranzo

Daniel Vázquez-Bustelo and Sandra Valle

Alfonso Palacio-Vera

231/2005  Reconciling Sustainability and Discounting in Cost Benefit Analysis: a methodological proposal
M. Carmen Almansa Sáez and Javier Calatrava Requena

232/2005  Can The Excess Of Liquidity Affect The Effectiveness Of The European Monetary Policy?
Santiago Carbó Valverde and Rafael López del Paso

Miguel Angel Barberán Lahuerta

Víctor M. González

Waymond Rodgers, Paul Pavlou and Andres Guiral.

Francisco J. André, M. Alejandro Cardenete y Carlos Romero.
Santiago Carbó-Valverde, Francisco Rodríguez-Fernández y Gregory F. Udell.

238/2006  Trade Effects Of Monetary Agreements: Evidence For Oecd Countries.  
Salvador Gil-Pareja, Rafael Llorca-Vivero y José Antonio Martínez-Serrano.

Marcos Álvarez-Díaz y Gonzalo Caballero Miguez.

240/2006  La interacción entre el éxito competitivo y las condiciones del mercado doméstico como deter-
minantes de la decisión de exportación en las Pymes.  
Francisco García Pérez.

241/2006  Una estimación de la depreciación del capital humano por sectores, por ocupación y en el 
tiempo.  
Inés P. Murillo.

Manuel A. Gómez.

243/2006  Measuring efficiency in education: an analysis of different approaches for incorporating 
non-discretionary inputs.  
Jose Manuel Cordero-Ferrera, Francisco Pedraja-Chaparro y Javier Salinas-Jiménez

244/2006  Did The European Exchange-Rate Mechanism Contribute To The Integration Of Peripheral 
Countries?.  
Salvador Gil-Pareja, Rafael Llorca-Vivero y José Antonio Martínez-Serrano

Marta Pascual and David Cantarero

246/2006  Measurement and analysis of the Spanish Stock Exchange using the Lyapunov exponent with 
digital technology.  
Salvador Rojí Ferrari and Ana Gonzalez Marcos

247/2006  Testing For Structural Breaks In Variance Withadditive Outliers And Measurement Errors.  
Paulo M.M. Rodrigues and Antonio Rubia

Joaquín Maudos and Juan Fernández de Guevara

Desiderio Romero Jordán, José Félix Sanz Sanz y César Pérez López

250/2006  Regional Income Disparities in Europe: What role for location?.  
Jesús López-Rodriguez and J. Andrés Faíña

251/2006  Funciones abreviadas de bienestar social: Una forma sencilla de simultanear la medición de la 
eficiencia y la equidad de las políticas de gasto público.  
Nuria Badenes Plá y Daniel Santín González

252/2006  “The momentum effect in the Spanish stock market: Omitted risk factors or investor behaviour?”.  
Luis Muga and Rafael Santamaria

253/2006  Dinámica de precios en el mercado español de gasolina: un equilibrio de colusión tácita.  
Jordi Perdiguero García
José M. Pastor, Empar Pons y Lorenzo Serrano

255/2006 Environmental implications of organic food preferences: an application of the impure public 
goods model. 
Ana María Aldanondo-Ochoa y Carmen Almansa-Sáez

José Félix Sanz-Sanz, Desiderio Romero-Jordán y Santiago Álvarez-García

257/2006 La internacionalización de la empresa manufacturera española: efectos del capital humano 
genérico y específico. 
José López Rodríguez

María Martínez Torres

259/2006 Efficiency and market power in Spanish banking. 
Rolf Färe, Shawna Grosskopf y Emili Tortosa-Ausina.

260/2006 Asimetrías en volatilidad, beta y contagios entre las empresas grandes y pequeñas cotizadas en la 
bolsa española. 
Helena Chuliá y Hipòlit Torró.

José Antonio Ortega.

262/2006 Accidentes de tráfico, víctimas mortales y consumo de alcohol. 
José Mª Arranz y Ana I. Gil.

263/2006 Análisis de la Presencia de la Mujer en los Consejos de Administración de las Mil Mayores Empresas Españolas. 
Ruth Mateos de Cabo, Lorenzo Escot Mangas y Ricardo Gimeno Nogués.

Ignacio Álvarez Peralta.

Jaime Vallés-Giménez y Anabel Zárate-Marco.

266/2006 Health Human Capital And The Shift From Foraging To Farming. 
Paolo Rungo.

Juan Luis Jiménez y Jordi Perdiguero.

Desiderio Romero-Jordán y José Félix Sanz-Sanz.

269/2006 Banking competition, financial dependence and economic growth 
Joaquín Maudos y Juan Fernández de Guevara

270/2006 Efficiency, subsidies and environmental adaptation of animal farming under CAP 
Werner Kleinhans, Carmen Murillo, Carlos San Juan y Stefan Sperlich
A. García-Lorenzo y Jesús López-Rodríguez

272/2006 Riesgo asimétrico y estrategias de momentum en el mercado de valores español
Luis Muga y Rafael Santamaría

273/2006 Valoración de capital-riesgo en proyectos de base tecnológica e innovadora a través de la teoría de opciones reales
Gracia Rubio Martín

274/2006 Capital stock and unemployment: searching for the missing link
Ana Rosa Martínez-Cañete, Elena Márquez de la Cruz, Alfonso Palacio-Vera and Inés Pérez-Soba Aguilar

275/2006 Study of the influence of the voters’ political culture on vote decision through the simulation of a political competition problem in Spain
Sagrario Lantarón, Isabel Lillo, Mª Dolores López and Javier Rodrigo

276/2006 Investment and growth in Europe during the Golden Age
Antonio Cubel and Mª Teresa Sanchis

277/2006 Efectos de vincular la pensión pública a la inversión en cantidad y calidad de hijos en un modelo de equilibrio general
Robert Meneu Gaya

278/2006 El consumo y la valoración de activos
Elena Márquez y Belén Nieto

279/2006 Economic growth and currency crisis: A real exchange rate entropic approach
David Matesanz Gómez y Guillermo J. Ortega

280/2006 Three measures of returns to education: An illustration for the case of Spain
Maria Arrazola y José de Hevia

281/2006 Composition of Firms versus Composition of Jobs
Antoni Cunyat

282/2006 La vocación internacional de un holding tranviario belga: la Compagnie Mutuelle de Tramways, 1895-1918
Alberte Martinez López

283/2006 Una visión panorámica de las entidades de crédito en España en la última década.
Constantino García Ramos

Alberte Martinez López

285/2006 Los intereses belgas en la red ferroviaria catalana, 1890-1936
Alberte Martinez López

286/2006 The Governance of Quality: The Case of the Agrifood Brand Names
Marta Fernández Barcala, Manuel González-Díaz y Emmanuel Raynaud

287/2006 Modelling the role of health status in the transition out of malthusian equilibrium
Paolo Rungo, Luis Currais and Berta Rivera

288/2006 Industrial Effects of Climate Change Policies through the EU Emissions Trading Scheme
Xavier Labandeira and Miguel Rodríguez
<table>
<thead>
<tr>
<th>Issue</th>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>290/2006</td>
<td>La producción de energía eléctrica en España: Análisis económico de la actividad tras la liberalización del Sector Eléctrico</td>
<td>Fernando Hernández Martínez</td>
</tr>
<tr>
<td>291/2006</td>
<td>Further considerations on the link between adjustment costs and the productivity of R&amp;D investment: evidence for Spain</td>
<td>Desiderio Romero-Jordán, José Félix Sanz-Sanz and Inmaculada Álvarez-Ayuso</td>
</tr>
<tr>
<td>292/2006</td>
<td>Una teoría sobre la contribución de la función de compras al rendimiento empresarial</td>
<td>Javier González Benito</td>
</tr>
<tr>
<td>294/2006</td>
<td>Testing the parametric vs the semiparametric generalized mixed effects models</td>
<td>María José Lombardía and Stefan Sperlich</td>
</tr>
<tr>
<td>295/2006</td>
<td>Nonlinear dynamics in energy futures</td>
<td>Mariano Matilla-García</td>
</tr>
<tr>
<td>296/2006</td>
<td>Estimating Spatial Models By Generalized Maximum Entropy Or How To Get Rid Of W</td>
<td>Esteban Fernández Vázquez, Matías Mayor Fernández and Jorge Rodríguez-Valez</td>
</tr>
<tr>
<td>297/2006</td>
<td>Optimización fiscal en las transmisiones lucrativas: análisis metodológico</td>
<td>Félix Domínguez Barrero</td>
</tr>
<tr>
<td>298/2006</td>
<td>La situación actual de la banca online en España</td>
<td>Francisco José Climent Diranzo y Alexandre Momparler Pechuán</td>
</tr>
<tr>
<td>299/2006</td>
<td>Estrategia competitiva y rendimiento del negocio: el papel mediador de la estrategia y las capacidades productivas</td>
<td>Javier González Benito y Isabel Suárez González</td>
</tr>
<tr>
<td>300/2006</td>
<td>A Parametric Model to Estimate Risk in a Fixed Income Portfolio</td>
<td>Pilar Abad and Sonia Benito</td>
</tr>
<tr>
<td>301/2007</td>
<td>Análisis Empírico de las Preferencias Sociales Respecto del Gasto en Obra Social de las Cajas de Ahorros</td>
<td>Alejandro Esteller-Moré, Jonathan Jorba Jiménez y Albert Solé-Ollé</td>
</tr>
<tr>
<td>302/2007</td>
<td>Assessing the enlargement and deepening of regional trading blocs: The European Union case</td>
<td>Salvador Gil-Pareja, Rafael Llorca-Vivero y José Antonio Martínez-Serrano</td>
</tr>
<tr>
<td>303/2007</td>
<td>¿Es la Franquicia un Medio de Financiación?: Evidencia para el Caso Español</td>
<td>Vanesa Solís Rodríguez y Manuel González Díaz</td>
</tr>
<tr>
<td>305/2007</td>
<td>Spain is Different: Relative Wages 1989-98</td>
<td>José Antonio Carrasco Gallego</td>
</tr>
</tbody>
</table>
323/2007 Wood and industrialization. evidence and hypotheses from the case of Spain, 1860-1935. Iñaki Iriarte-Goñi and María Isabel Ayuda Bosque


325/2007 Monetary policy and structural changes in the volatility of us interest rates. Juncal Cuñado, Javier Gomez Biscarri and Fernando Perez de Gracia

326/2007 The productivity effects of intrafirm diffusion. Lucio Fuentelsaz, Jaime Gómez and Sergio Palomas


328/2007 El grado de cobertura del gasto público en España respecto a la UE-15 Nuria Rueda, Begoña Barruso, Carmen Calderón y Mª del Mar Herrador

329/2007 The Impact of Direct Subsidies in Spain before and after the CAP'92 Reform Carmen Murillo, Carlos San Juan and Stefan Sperlich

330/2007 Determinants of post-privatization performance of Spanish divested firms Laura Cabeza García and Silvia Gómez Ansón

331/2007 ¿Por qué deciden diversificar las empresas españolas? Razones oportunistas versus razonas económicas Almudena Martínez Campillo

332/2007 Dynamical Hierarchical Tree in Currency Markets Juan Gabriel Brida, David Matesanz Gómez and Wiston Adrián Risso

333/2007 Los determinantes sociodemográficos del gasto sanitario. Análisis con microdatos individuales Ana María Angulo, Ramón Barberán, Pilar Egea y Jesús Mur

334/2007 Why do companies go private? The Spanish case Inés Pérez-Soba Aguilar

335/2007 The use of gis to study transport for disabled people Verónica Cañal Fernández

336/2007 The long run consequences of M&A: An empirical application Cristina Bernad, Lucio Fuentelsaz and Jaime Gómez

337/2007 Las clasificaciones de materias en economía: principios para el desarrollo de una nueva clasificación Valentín Edo Hernández


339/2007 Impacts of an iron and steel plant on residential property values Celia Bilbao-Terol

340/2007 Firm size and capital structure: Evidence using dynamic panel data Víctor M. González and Francisco González
341/2007 ¿Cómo organizar una cadena hotelera? La elección de la forma de gobierno
Marta Fernández Barcala y Manuel González Díaz

342/2007 Análisis de los efectos de la decisión de diversificar: un contraste del marco teórico “Agencia-Stewardship”
Almudena Martínez Campillo y Roberto Fernández Gago

343/2007 Selecting portfolios given multiple eurostoxx-based uncertainty scenarios: a stochastic goal programming approach from fuzzy betas
Enrique Ballestero, Blanca Pérez-Gladish, Mar Arenas-Parra and Amelia Bilbao-Terol

344/2007 “El bienestar de los inmigrantes y los factores implicados en la decisión de emigrar”
Anastasia Hernández Alemán y Carmelo J. León

Andrea Martínez-Noya and Esteban García-Canal

346/2007 Diferencias salariales entre empresas públicas y privadas. El caso español
Begoña Cueto y Nuria Sánchez- Sánchez

347/2007 Effects of Fiscal Treatments of Second Home Ownership on Renting Supply
Celia Bilbao Terol and Juan Prieto Rodriguez

348/2007 Auditors’ ethical dilemmas in the going concern evaluation
Andres Guiral, Waymond Rodgers, Emiliano Ruiz and Jose A. Gonzalo

Susana Morales Sequera y Carmen Pérez Esparrells

350/2007 Socially responsible investment: mutual funds portfolio selection using fuzzy multiobjective programming
Blanca Mª Pérez-Gladish, Mar Arenas-Parra , Amelia Bilbao-Terol and Mª Victoria Rodriguez-Uría

351/2007 Persistencia del resultado contable y sus componentes: implicaciones de la medida de ajustes por devengo
Raúl Iñiguez Sánchez y Francisco Poveda Fuentes

352/2007 Wage Inequality and Globalisation: What can we Learn from the Past? A General Equilibrium Approach
Concha Betrán, Javier Ferri and Maria A. Pons

353/2007 Eficacia de los incentivos fiscales a la inversión en I+D en España en los años noventa
Desiderio Romero Jordán y José Félix Sanz Sanz

354/2007 Convergencia regional en renta y bienestar en España
Robert Meneu Gaya

355/2007 Tributación ambiental: Estado de la Cuestión y Experiencia en España
Ana Carrera Poncela

356/2007 Salient features of dependence in daily us stock market indices
Luis A. Gil-Alana, Juncal Cuñado and Fernando Pérez de Gracia

357/2007 La educación superior: ¿un gasto o una inversión rentable para el sector público?
Inés P. Murillo y Francisco Pedraja
358/2007 Effects of a reduction of working hours on a model with job creation and job destruction
Emilio Domínguez, Miren Ullibarri y Idoya Zabaleta

359/2007 Stock split size, signaling and earnings management: Evidence from the Spanish market
José Yagüe, J. Carlos Gómez-Sala and Francisco Poveda-Fuentes

360/2007 Modelización de las expectativas y estrategias de inversión en mercados de derivados
Begoña Font-Belaire

361/2008 Trade in capital goods during the golden age, 1953-1973
Mª Teresa Sanchis and Antonio Cubel

362/2008 El capital económico por riesgo operacional: una aplicación del modelo de distribución de pérdidas
Enrique José Jiménez Rodríguez y José Manuel Feria Domínguez

363/2008 The drivers of effectiveness in competition policy
Joan-Ramon Borrell and Juan-Luis Jiménez

364/2008 Corporate governance structure and board of directors remuneration policies: evidence from Spain
Carlos Fernández Méndez, Rubén Arrondo García and Enrique Fernández Rodríguez

365/2008 Beyond the disciplinary role of governance: how boards and donors add value to Spanish foundations
Pablo De Andrés Alonso, Valentín Azofra Palenzuela y M. Elena Romero Merino

366/2008 Complejidad y perfeccionamiento contractual para la contención del oportunismo en los acuerdos de franquicia
Vanessa Solís Rodríguez y Manuel González Díaz

367/2008 Inestabilidad y convergencia entre las regiones europeas
Jesús Mur, Fernando López y Ana Angulo

368/2008 Análisis espacial del cierre de explotaciones agrarias
Ana Aldanondo Ochoa, Carmen Almansa Sáez y Valero Casanovas Oliva

369/2008 Cross-Country Efficiency Comparison between Italian and Spanish Public Universities in the period 2000-2005
Tommaso Agasisti and Carmen Pérez Esparrells

370/2008 El desarrollo de la sociedad de la informacion en España: un análisis por comunidades autónomas
María Concepción García Jiménez y José Luis Gómez Barroso

371/2008 El medioambiente y los objetivos de fabricación: un análisis de los modelos estratégicos para su consecución
Lucía Avella Camarero, Esteban Fernández Sánchez y Daniel Vázquez-Bustelo

372/2008 Influence of bank concentration and institutions on capital structure: New international evidence
Víctor M. González and Francisco González

373/2008 Generalización del concepto de equilibrio en juegos de competición política
Mª Dolores López González y Javier Rodrigo Hitos

374/2008 Smooth Transition from Fixed Effects to Mixed Effects Models in Multi-level regression Models
Maria José Lombardía and Stefan Sperlich
375/2008 A Revenue-Neutral Tax Reform to Increase Demand for Public Transport Services
Carlos Pestana Barros and Juan Prieto-Rodriguez

376/2008 Measurement of intra-distribution dynamics: An application of different approaches to the European regions
Adolfo Maza, María Hierro and José Villaverde

377/2008 Migración interna de extranjeros y ¿nueva fase en la convergencia?
María Hierro y Adolfo Maza

378/2008 Efeto de la Reforma del Sector Eléctrico: Modelización Teórica y Experiencia Internacional
Ciro Eduardo Bazán Navarro

379/2008 A Non-Parametric Independence Test Using Permutation Entropy
Mariano Matilla-García and Manuel Ruiz Marín

380/2008 Testing for the General Fractional Unit Root Hypothesis in the Time Domain
Uwe Hassler, Paulo M.M. Rodrigues and Antonio Rubia

381/2008 Multivariate gram-charlier densities
Esther B. Del Brio, Trino-Manuel Ñíguez and Javier Perote

382/2008 Analyzing Semiparametrically the Trends in the Gender Pay Gap - The Example of Spain
Ignacio Moral-Arce, Stefan Sperlich, Ana I. Fernández-Sainz and Maria J. Roca

383/2008 A Cost-Benefit Analysis of a Two-Sided Card Market
Santiago Carbó Valverde, David B. Humphrey, José Manuel Liñares Zegarra and Francisco Rodríguez Fernández

384/2008 A Fuzzy Bicriteria Approach for Journal Deselection in a Hospital Library
M. L. López-Avello, M. V. Rodríguez-Uría, B. Pérez-Gladish, A. Bilbao-Terol, M. Arenas-Parra

385/2008 Valoración de las grandes corporaciones farmaceúticas, a través del análisis de sus principales intangibles, con el método de opciones reales
Gracia Rubio Martín y Prosper Lamothe Fernández

386/2008 El marketing interno como impulsor de las habilidades comerciales de las pymes españolas: efectos en los resultados empresariales
Mª Leticia Santos Vijande, Mª José Sanzo Pérez, Nuria García Rodríguez y Juan A. Trespalacios Gutiérrez

387/2008 Understanding Warrants Pricing: A case study of the financial market in Spain
David Abad y Belén Nieto

388/2008 Aglomeración espacial, Potencial de Mercado y Geografía Económica: Una revisión de la literatura
Jesús López-Rodríguez y J. Andrés Faiña

389/2008 An empirical assessment of the impact of switching costs and first mover advantages on firm performance
Jaime Gómez, Juan Pablo Maícas

390/2008 Tender offers in Spain: testing the wave
Ana R. Martínez-Cañete y Inés Pérez-Soba Aguilar
La integración del mercado español a finales del siglo XIX: los precios del trigo entre 1891 y 1905
Mariano Matilla García, Pedro Pérez Pascual y Basilio Sanz Carnero

Cuando el tamaño importa: estudio sobre la influencia de los sujetos políticos en la balanza de bienes y servicios
Alfonso Echazarra de Gregorio

Una visión cooperativa de las medidas ante el posible daño ambiental de la desalación
Borja Montaño Sanz

Efectos externos del endeudamiento sobre la calificación crediticia de las Comunidades Autónomas
Andrés Leal Marcos y Julio López Laborda

Technical efficiency and productivity changes in Spanish airports: A parametric distance functions approach
Beatriz Tovar & Roberto Rendeiro Martín-Cejas

Network analysis of exchange data: Interdependence drives crisis contagion
David Matesanz Gómez & Guillermo J. Ortega

Explaining the performance of Spanish privatised firms: a panel data approach
Laura Cabeza García and Silvia Gomez Anson

Technological capabilities and the decision to outsource R&D services
Andrea Martínez-Noya and Esteban García-Canal

Hybrid Risk Adjustment for Pharmaceutical Benefits
Manuel García-Goñi, Pere Ibern & José María Inoriza

The Team Consensus–Performance Relationship and the Moderating Role of Team Diversity
José Henrique Dieguez, Javier González-Benito and Jesús Galende

The institutional determinants of CO₂ emissions: A computational modelling approach using Artificial Neural Networks and Genetic Programming
Marcos Álvarez-Díaz, Gonzalo Caballero Miguez and Mario Soliño