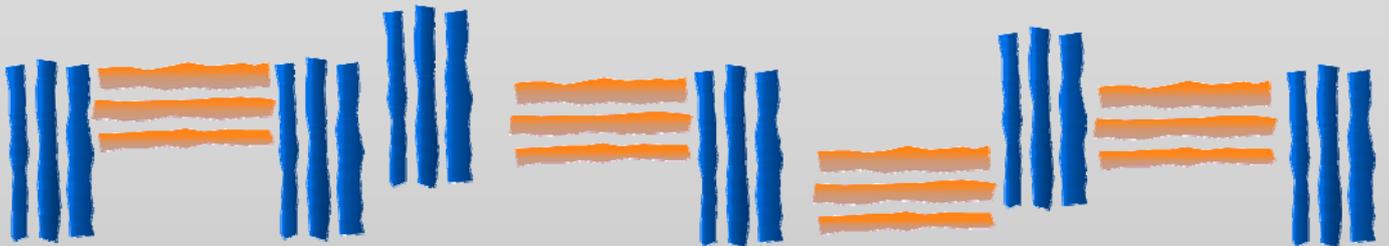


REVISTA DE
MÉTODOS CUANTITATIVOS
PARA LA ECONOMÍA Y LA EMPRESA



UNIVERSIDAD
**PABLO^D
OLAVIDE**
SEVILLA

Número 13
Junio de 2012
ISSN: 1886-516X
D.L: SE-2927-06

**REVISTA DE MÉTODOS CUANTITATIVOS
PARA LA ECONOMÍA Y LA EMPRESA**

**Journal of Quantitative Methods for
Economics and Business Administration**

Número 13. Junio de 2012.

ISSN: 1886-516X. D.L: SE-2927-06.

URL: <http://www.upo.es/RevMetCuant/>

Editores:

Dr. Eugenio M. Fedriani Martel
Universidad Pablo de Olavide
Ctra. de Utrera, km 1 - 41013
Sevilla (Spain).
Correo-e: efedmar@upo.es

Dr. Alfredo García Hernández-Díaz
Universidad Pablo de Olavide
Ctra. de Utrera, km 1 - 41013
Sevilla (Spain).
Correo-e: agarher@upo.es

Comité Editorial:

Dr. S. Ejaz Ahmed, University of Windsor, Ontario (Canadá)
Dr. Adam P. Balcerzak, Nicolaus Copernicus University, Toruń (Polonia)
Dr. Carlos A. Coello Coello, CINVESTAV-IPN, México D.F. (México)
Dr. Ignacio Contreras Rubio, Universidad Pablo de Olavide, Sevilla (España)
Dra. Manuela Coromaldi, University of Social Sciences UNISU, Roma (Italia)
Dr. Miguel Ángel Hinojosa Ramos, Universidad Pablo de Olavide, Sevilla (España)
Dr. Matías Irigoyen Testa, Universidad Nacional del Sur, Buenos Aires (Argentina)
Dr. M. Kazim Khan, Kent State University, Ohio (EE.UU.)
Dr. Manuel Laguna, University of Colorado at Boulder, Colorado (EE.UU.)
Dra. María Amparo León Sánchez, Universidad de Pinar del Río (Cuba)
Dr. Jesús López-Rodríguez, Universidad de A Coruña, (España)
Dr. Cecilio Mar Molinero, University of Kent, Canterbury (Reino Unido)
Dra. Ana M. Martín Caraballo, Universidad Pablo de Olavide, Sevilla (España)
Dra. M^a Carmen Melgar Hiraldo, Universidad Pablo de Olavide, Sevilla (España)
Dr. Gábor Nagy, University of Kent, Canterbury (Reino Unido)
Dr. José Antonio Ordaz Sanz, Universidad Pablo de Olavide, Sevilla (España)
Dr. Andrés Rodríguez-Pose, London School of Economics, Londres (Reino Unido)
Dr. José Manuel Rueda Cantuche, IPTS-DG J.R. Centre-European Commission
Dra. Mariagrazia Squicciarini, OECD, París (Francia)
Dra. Mariangela Zoli, Università degli Studi di Roma 'Tor Vergata', Roma (Italia)



Extreme Learning Machine to Analyze the Level of Default in Spanish Deposit Institutions

MONTERO-ROMERO, TERESA

Department of Management and Quantitative Methods, ETEA, Córdoba (Spain)

Correo electrónico: tmontero@etea.com

LÓPEZ-MARTÍN, MARÍA DEL CARMEN

Department of Economics, Legal Sciences and Sociology, ETEA, Córdoba (Spain)

Correo electrónico: mclopez@etea.com

BECERRA-ALONSO, DAVID

Department of Management and Quantitative Methods, ETEA, Córdoba (Spain)

Correo electrónico: dbecerra@etea.com

MARTÍNEZ-ESTUDILLO, FRANCISCO JOSÉ

Department of Management and Quantitative Methods, ETEA, Córdoba (Spain)

Correo electrónico: fjmestud@etea.com

ABSTRACT

The level of default in financial institutions is a key piece of information in the activity of these organizations and reveals their level of risk. This in turn explains the growing attention given to variables of this kind, during the crisis of these last years. This paper presents a method to estimate the default rate using the non-linear model defined by standard Multilayer Perceptron (MLP) neural networks trained with a novel methodology called Extreme Learning Machine (ELM). The experimental results are promising, and show a good performance when comparing the MLP model trained with the Levenberg-Marquard algorithm.

Keywords: level of default; financial institutions; neural networks; Extreme Learning Machine.

JEL classification: G21; G01; C45.

MSC2010: 62P05; 68T01; 82C32.

Análisis de la morosidad de las entidades financieras españolas mediante Extreme Learning Machine

RESUMEN

La morosidad en las entidades financieras es un dato muy importante de la actividad de estas instituciones pues permite conocer el nivel de riesgo asumido por éstas. Esto a su vez explica la creciente atención otorgada a esta variable, especialmente en los últimos años de crisis.

Este artículo presenta un método para estimar el nivel de la tasa de morosidad por medio de un modelo no lineal definido por la red neuronal *Multi-layer Perceptron* (MLP) entrenada con una nueva metodología llamada *Extreme Learning Machine* (ELM). Los resultados experimentales son prometedores, mostrando un buen resultado si se compara con el modelo MLP entrenado con el algoritmo de Levenberg-Marquard.

Palabras clave: nivel de morosidad; instituciones financieras; redes neuronales; Extreme Learning Machine.

Clasificación JEL: G21; G01; C45.

MSC2010: 62P05; 68T01; 82C32.



1. INTRODUCTION

The aim of this paper is to study the default rate in financial institutions through variables that define the entities' financial state. More specifically, we try to analyze how certain financial data (in the form of independent variables) is related to the level of default in each institution. The relevance of this study is justified by the importance that this last variable has in the context of today's economic crisis.

Linear economic models are not able to capture non-linear patterns. Moreover, linear techniques cannot distinguish effectively between random noise and non-linear relationships, and usually make assumptions about the independence and normality of data. In the last few years, artificial neural networks (ANNs) have emerged as a powerful statistical modeling technique (Bishop, 1995). ANNs are useful to detect the underlying functional relationships within a dataset and to perform several tasks like pattern recognition, classification, modeling, and prediction. The most popular neural network model may be the Multilayer Perceptron (MLP) (Bishop, 1995), trained with the well-known back-propagation algorithm, due to its simple architecture. Besides, a number of recent papers use ANNs to analyze traditional problems in Accounting and Finance (Coakley, 2000; Herbrick, 2000; McNelis, 2005; Parisi, 2006; Martínez-Estudillo *et al.*, 2007; Gutiérrez *et al.*, 2009; López-Martín *et al.*, 2011).

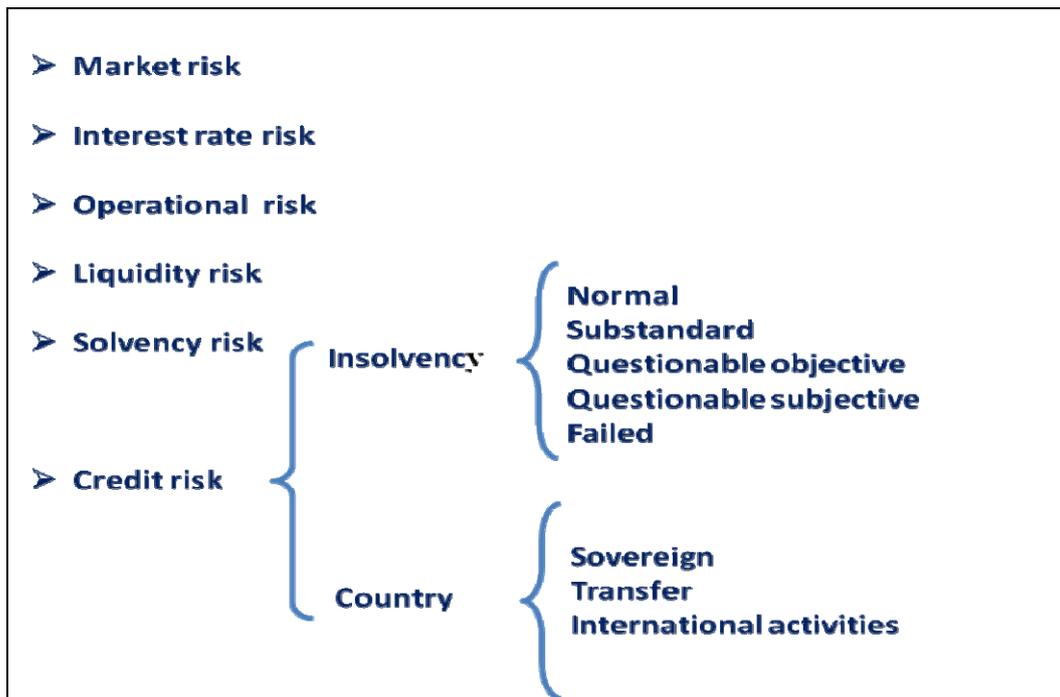
Taking all the above into account and as part of a greater research topic, let us introduce the main elements. The financial institutions studied in this study are Spanish deposit institutions with data since 2008. The methodology used is the non-linear model defined by MLP neural network trained with the Extreme Learning Machine (ELM) algorithm (Huang, 2004). ELM for MLP neural networks is a new algorithm that randomly chooses hidden nodes and analytically determines the output weights of the network. Theoretically, the ELM algorithm tends to provide a good generalization performance at an extremely fast learning speed. The experimental results based on artificial and real benchmarking problems show that ELM can result in a better generalization performance in many cases and can learn thousands of times faster than traditional learning algorithms for feed-forward neural networks (Huang *et al.*, 2011).

The structure of the rest of the paper is the following: Sections 2 and 3 analyze the level of default in Spanish deposit institutions and the importance of this variable in the current economic context. Section 4 shows the state of the art with respect to studies using the ELM algorithm in the field of Business and Finance. Section 5 describes the mathematical model defined by the MLP neural network used in the prediction task and the ELM algorithm. Then Section 6 presents the experimental application where the performance of the proposed approach is tested. Finally, the paper closes with some remarks and conclusions.

2. CREDIT RISK AND DEFAULT PAYMENT IN THE SPANISH FINANCIAL SYSTEM

The uncertainty characterising the present economic context has turned risk into a critical variable that must be kept under control. It is known that financial institutions face different kinds of risk, including market risk, credit risk, liquidity risk, and operational risk. Figure 1 shows various types of risk undergone by financial institutions. The main categories are defined in the following figure.

Figure 1. Financial institutions exposure risk



Source: elaborated by authors.

Market risk is the potential loss on an investment due to adverse changes in the factors affecting price and/or value. Interest rate risk is the potential loss from adverse changes affecting the interest rate. Operational risk shows potential loss as a result of failures in internal processes or potential loss when these internal processes are inadequate. The liquidity risk is the risk of not being able to turn an asset into cash. Solvency risk is the risk of debts not met at maturity. It refers to a position in the long term. And finally, credit risk is the potential loss assumed by the financial institution as a consequence of breach of contract obligations to counterparts.

One of the most important kinds of risk and the most studied in the field of financial institutions is credit risk; see some examples: Rodríguez Fernández, 1987; Cruz González, 1998; López and Saldenberg, 2000; Boal Velasco and González Sánchez, 2001; Soler and Miró, 2001; García Céspedes, 2005, among other studies.

This risk is the object of our study, i.e. the risk of not recovering money that has been lent, along with its corresponding interest. This work is focused on credit risk, not only because it can be the cause of great losses in financial institutions, but also because it has already caused interventions by *Banco de España* in the last two crisis-ridden years. It also indirectly influences the capital requirements demanded by Basel III regulations, as well as the recent Law 2/2011 for the reinforcement of financial institutions.

The tensions tackled by the national and international financial systems since the beginning of the crisis have obstructed its function as an essential driver of credit in the economy. In the Spanish scenario, credit institutions have undergone great difficulties in obtaining financial support. In turn, whatever financial support was granted came accompanied by deteriorated assets (mostly associated to the housing sector). Many loans were labeled as doubtful, and default rates reached surprising levels before the beginning of 2007. These problems have resulted in increasing difficulties for families and small companies to access credit, which is fundamental in order to generate jobs in the Spanish business network.

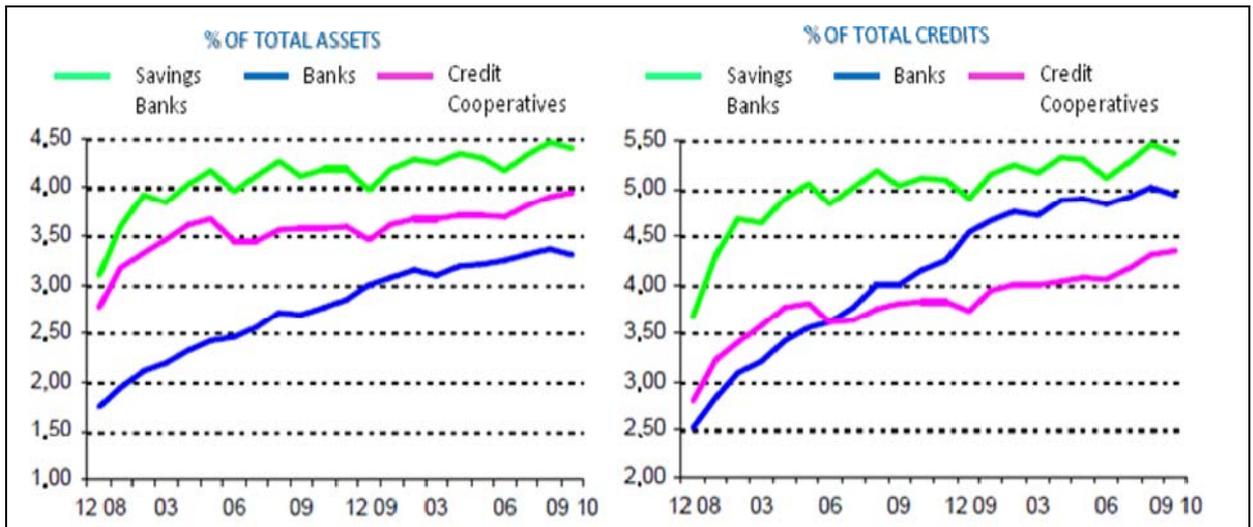
In this context, the crisis has emphasized the need for a clean, competitive and solid financial system that allows for mediation between those who own the financial resources and those who need them. A well-structured financial system is the main warranty for a country's productive economy to be able to access the financial support needed for investment projects that generate wealth and jobs. However, in order for this to take place, it is necessary to trust in the integrity of institutions and proper performance of the markets.

The data gathered over the last few years confirm these problems. Figures 2 and 3 show how default rates in deposit institutions have risen over the last three years (2008-2010). These are seen to be higher in savings banks, than in banks and credit cooperatives. These latter two differ depending on whether rates are considered in terms of total assets or in terms of approved loans. Although the level of default has increased since 2007, it is also true that in the last two years there has been a dramatic decrease in the number of these three institutions' doubtful loans since their peak in late 2008 and early 2009. This phenomenon is more related to the greater number of requisites prior to credit approvals, than to the solvency of clients in general. This explains the constant and even the reduced outstanding balance of credits granted by institutions.

The data shown and the context in which we place this study prove how important it is for deposit institutions to have plenty of capital in order to be prepared for unexpected losses, such as those in times of decreased or moderately increased economic activity. This is one of the reasons that explain the capital regulation reforms that took place in the Basel Committee (Basel III), to increase demands to maintain resources, especially for those institutions that have

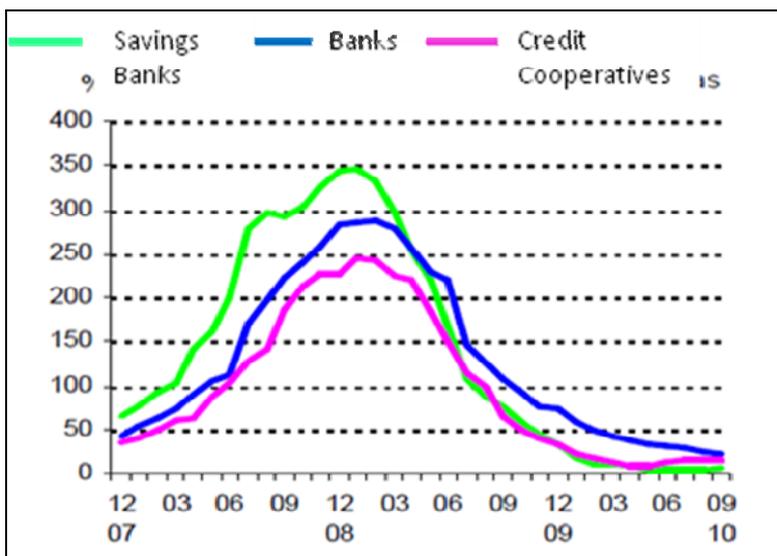
a greater capacity to absorb losses. Minimum solvency requisites in Spain were complemented at the beginning of this year with a new capital coefficient (that is regulated by 2/2011 Law, of Financial Institution Reinforcement), that in practice is even more demanding than the one imposed by Basel III, in order to dissipate any doubts that financial markets may have about the health of our financial system.

Figure 2. Level of default in institutions in terms of total assets and total credit



Source: CECA (2010).

Figure 3. Level of default variations from year to year in deposit entities



Source: CECA (2010).

3. INSTITUTIONS STUDIED, SOURCES, AND EXPERIMENTAL DESIGN

3.1. Institutions Analysed and Study Stages

As indicated in the introduction to this paper, the aim of this work is to estimate the level of default in Spanish deposit institutions using variables that define each entity's financial information. In order to do this, we start by studying data from 2008. This year is chosen since it is the first term in which, despite the already growing level of default, financial reports of deposit institutions still did not reflect the real magnitude of the level of default caused by the economic crisis. It is thus a term not too far from the transition to a recognizable state of crisis, and one where the level of default in financial institutions was still representative.

As we know, institutions that can approve credits or similar measures constitute what is called the group of credit institutions. Even so, this study has not taken this group into consideration, but instead that made up of deposit institutions, which account for more than 90% of the balance of all credit institutions. Thus, there is a deliberate elimination of some credit institutions that do not have the same kind of activity (such as specialized credit institutions) or those that have different features (such as the government's credit institution). With this criterion, the total number of deposit institutions available with annual accountings was 152. Out of these the level of default has been obtained for 111 of them (73.03% of the total). Table 1 shows distributions by types of institution, and the average level of default of each type.

Table 1. Sample distribution classified by the kind of institution, percentage of each type, and default payment average per institution in 2008

Institutions	# institutions	Percentage of each type	Default payment average
Banks	31	27.93%	2.86%
Savings banks	46	41.44%	3.06%
Credit cooperatives	34	30.30%	1.81%
TOTAL	111	100.00%	2.29%

Source: elaborated by authors.

3.2. The Dependent Variable: Financial Institutions' Level of Default

We can define credit risk as the risk of not recovering the money lent with its corresponding interest. Within this kind of risk, special attention is paid to insolvency risk, which refers to refund difficulties, either due to the level of default, or for subjective reasons (such as reasonable doubts). Then, within insolvency risk, there is normal risk, substandard, questionable objective, questionable subjective, and failed risk. The questionable objective risk has an objective component: nonpayment. This includes all of the debt instruments, regardless of the

client and the warranty, which are overdue either due to the main loan, interests or other expenses agreed upon in the contract, for more than 3 months and unless classified as failed. It also includes the risks (normally endorsements or third party involvements) in which the endorsed is in default payment. In questionable risk, operations are included for reasons other than default payment (whether overdue or not) whenever reasonable doubts over refund exist. This can be due to negative wealth, continued losses, delayed payments, or inadequate economic or financial structure.

Clearly, in strict terms every client in default is questionable, although not every questionable client is in default payment. This is because there is the possibility of classifying a risk as questionable for subjective reasons other than nonpayment. Still, in general, this distinction is not made when referring to “default payment”, and thus, in order to determine the default payment rate, the total of doubtful loans and default payers have been considered in all the loans. Thus, the level of default obtained in this way has been chosen as the identifying variable for credit risk for each institution. The information on each variable in 2008 has been obtained from reports on each of the institutions analysed.

3.3. Independent Variables in the Classification Model

The selection of independent variables has been carried out by carefully defining the elements in the financial statements, with prior identification of the operations generating default payment in deposit institutions (such as credits given by the institution), or those that indicate the existence of default payment (such as loans available for sale).

The information referring to these variables has been extracted from the balance sheets and the profit and losses accounts published by the institutions. Such financial statements have been obtained from *Confederación de Cajas de Ahorros* (Federation of Savings Banks), the *Asociación Española de Banca* (Spanish Banking Association), and the *Unión Nacional de Cooperativas de Crédito* (National Union of Credit Cooperatives).

The independent variables selected will be grouped into four categories: the kind of institution, assets, liabilities, and profit and loss, all expressed in terms of the percentage of the total balance.

The process of including variables in this model entails a previous selection based on discriminant analysis. From the preliminary reports of the financial states that could be related to the level of default, as mentioned above, the co-linearity between them was analysed using discriminant analysis. The independent variables chosen after that are included in Table 2.

Table 2. Independent variables

Aggrupation	Denomination	Variable type
Kind of institution	Institution credit cooperative	Binary
	Institution saving bank	Binary
Assets variables	Total holdings of securities (to deal)	Quantitative
	Holdings of securities (to deal): customers credits	Quantitative
	Total other financial assets at fair value (with changes in profits and loss)	Quantitative
	Other financial assets at fair value (with changes in profits and loss): credits to customers	Quantitative
	Financial assets on sale	Quantitative
	Loans (total)	Quantitative
	Loans: customers credits	Quantitative
	Loans: securities other than shares	Quantitative
	Loans: securities held in pawn	Quantitative
	Non-current assets available for sale	Quantitative
	Shares	Quantitative
	Real estate investment	Quantitative
Liability variables	Others liabilities at fair value (with changes in profits and loss)	Quantitative
	Derivatives	Quantitative
	Allowances	Quantitative
	Capital and reserves	Quantitative
	Contingent risks	Quantitative
	Contingent commitment	Quantitative
Profit and loss variables	Net interest income	Quantitative
	Return on equity instruments and non-interest income	Quantitative
	Provisioning expense (net)	Quantitative
	Operating profit	Quantitative
	Financial assets impairment losses (net)	Quantitative
	Profit (losses) in disposal of others assets not classified in non-current assets available for sale	Quantitative
	Losses in business combination	Quantitative
	Profit (losses) in non-current assets available for sale not classified as discontinued operations	Quantitative

Source: elaborated by authors.

3.4. Experimental Design

A database including 28 independent variables and 111 financial institutions is built, along with their corresponding level of default. Nominal variables have been turned into binary, one for each category, as shown in Table 2. In order to study the generalization of this model the database

was split into two subsets. One will be used to train our neural network, and the other will be used to test it. There are 73 institutions assigned to the training set, and 38 to test. In order to make the partition, subsets were chosen that included extreme cases in both training and test sets, in order to improve generalisation.

4. EXTREME LEARNING MACHINES IN FINANCIAL PROBLEMS

Over the last decade, computational intelligence techniques have been used in a wide range of applications (marketing and sales, risk assessment and accounting, manufacturing, finance, etc.). Of numerous computational intelligence techniques, artificial neural networks (ANNs) have been playing a dominant role. However, it is known that ANNs face some challenging issues such as slow learning speed and/or poor computational scalability. ELM, as an emergent technology that overcomes some challenges faced by other techniques, has recently attracted the attention of computational intelligence and machine learning communities, in both theory and applications.

This section shows “the state of the art” of studies that use the ELM algorithm to solve problems in the field of business and finance. It must be highlighted that the number of studies in this area today is limited. The following are the most relevant from our point of view:

- a) In Li *et al.*, 2009, the authors apply the ELM classifier to the field of credit scoring. They compare two ELM algorithms with two well-known classifiers, namely, K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), obtaining better results over two credit data sets from the University of California, Irvine (UCI).
- b) ELM has also been used in forecasting to overcome the drawbacks of existing neural network forecasting models. In Wong and Guo, 2010, a hybrid intelligent model combines a novel meta-heuristic optimisation technique, the harmony search (HS) algorithm (Mahdavi *et al.*, 2007), with ELM to construct a learning algorithm to obtain optimal neural network weights and achieve a better generalisation performance. The model is developed to tackle sales forecasting problems in the fashion retail supply chain.
- c) In Van Heeswijk *et al.*, 2009, the authors investigate the application of adaptive ensemble models of ELMs to the problem of one-step-ahead prediction in non-stationary time series. The proposed model achieves a test error comparable to the best methods with a low computational cost.
- d) Another work (Sorjamaa *et al.*, 2008) proposes a hybrid model that combines ELM with various intelligent heuristics such as the variable selection using the Partial Least Squares (PLS) and a projection based on Nonparametric Noise Estimation (NNE), to ensure proper results by the ELM method. Then, after the network is first created using the original ELM,

the selection of the most relevant nodes is carried out using a Least Angle Regression (LARS) ranking of the nodes and a Leave-One-Out estimation of the performances, leading to an Optimally-Pruned ELM (OP-ELM). Finally, the prediction accuracy of the global methodology is demonstrated using the ESTSP 2008 Competition and Poland Electricity Load datasets.

e) Finally, in a previous work (López-Martín *et al.*, 2011), we addressed the problem from the perspective of classification. This paper uses the ELM model for the classification of the default status of financial institutions. The technique proposed is shown to perform better than logistic regression for this problem.

This is, to the authors' knowledge, the first paper that applies the ELM algorithm to make a prediction about the level of default in deposit institutions from available financial information. We believe that the good performance obtained so far by the ELM algorithm in benchmark datasets will increase its use in applications on real problems in the field of finance and economics.

5. MATHEMATICAL MODELS

This section starts with a brief introduction to the MLP neural network model and regression with neural networks. Then the ELM algorithm used to estimate the neural network weights is presented in detail.

5.1. MLP Neural Network Model

Let us consider the training set given by N patterns $D = \{(\mathbf{x}_i, t_i) : \mathbf{x}_i \in \mathbb{R}^n, t_i \in \mathbb{R}, i = 1, 2, \dots, N\}$, where $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in}) \in \mathbb{R}^n$, represents the input and t_i is the output for each one of these input vectors. The model used in order to do regression here is given by the function:

$$f(\mathbf{x}) = \sum_{i=1}^M \beta_i g(\langle \mathbf{w}_i, \mathbf{x} \rangle + b_i),$$

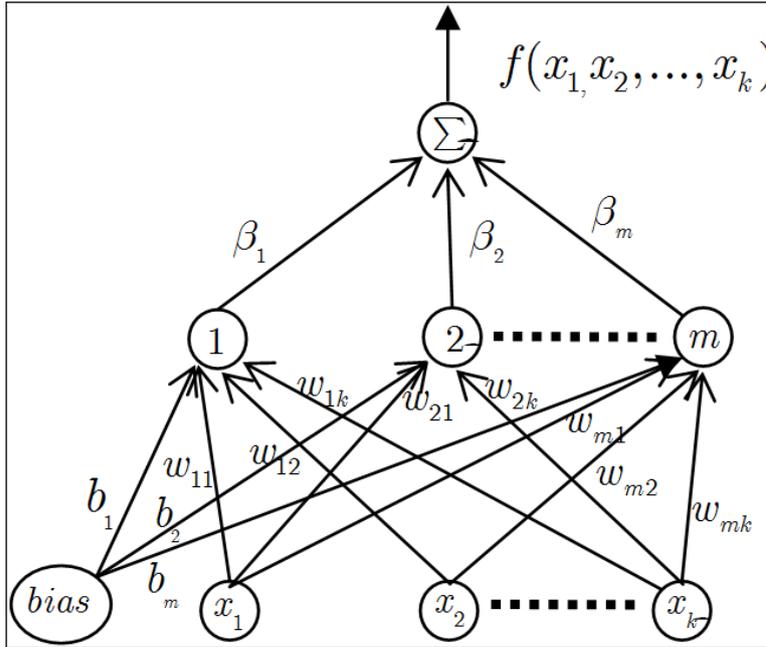
where the function g is the sigmoidal function given by:

$$g(t) = \frac{1}{1 + e^{-t}}.$$

This model is called the multilayer perceptron (MLP). In fact, this can be seen as a feed-forward neural network with n -nodes in the input layer (one for each independent variable), m nodes in the hidden layer, and one node in the output layer. $\mathbf{w}_i = (w_{i1}, \dots, w_{in})$ are the weight vectors that connect each i -th node in the hidden layer with the nodes in the input layer

corresponding to independent variables. β_i are the weights connecting each i -th node in the hidden layer with the output node. We define $\langle \mathbf{w}_i, \mathbf{x} \rangle$ as the Euclidean scalar product of input and weight vectors, and b_i as the independent bias associated to node i of the hidden layer. Figure 4 shows the structure of the feed-forward neural network at hand.

Figure 4. Structure of the MLP neural network proposed



Source: elaborated by authors.

The MLP model can be interpreted as a generalised linear model obtained as a combination of basis functions. Each one of these is a nonlinear function of a linear combination of the inputs, where the coefficients in the linear combination are adaptive parameters. For more details about the MLP model, the readers can see Bishop, 1995.

5.2. Regression and Neural Network Models

The fact that a neural network with sigmoidal transfer functions can approximate any given continuous function with the desired accuracy (Hornik, 1989) is a powerful basis for the application of neural networks to regression. Neural network regression is an instance of model-free or non-parametric regression. A model-free regression problem can be stated as follows. Given n pairs of vectors:

$$(\mathbf{x}_l, y_l) = (x_{l1}, x_{l2}, \dots, x_{lk}, y_l), \quad l = 1, 2, \dots, n,$$

that have been generated from unknown models

$$y_l = f(\mathbf{x}_l) + \varepsilon_l, \quad l = 1, 2, \dots, n,$$

where y is the response variable, \mathbf{x}_l is the independent variables vector, f is an unknown smooth non-parametric function that transforms a k -dimensional Euclidean space into \mathbb{R} :

$$f : \mathbb{R}^k \rightarrow \mathbb{R}$$

and ε_l are random variables with null mean, $E[\varepsilon_l] = 0$, and independent of \mathbf{x}_l . The aim of the regression is to construct an estimator \hat{f} , which is a function of the data (\mathbf{x}_l, y_l) , to approximate the unknown function f , and use this estimation to predict a new y given a new \mathbf{x} :

$$\hat{y} = \hat{f}(\mathbf{x}).$$

Let us consider a feed-forward neural network with k inputs and a hidden layer with m nodes. The hidden layer carries out a non-linear projection of the input vector \mathbf{x} to a vector \mathbf{h} where:

$$h_j = g\left(\sum_{i=1}^k w_{ji}x_i\right), \quad j = 1, \dots, m.$$

As previously stated, each node performs a non-linear projection of the input vector. $\mathbf{h} = g(\mathbf{x})$ tells how the output layer obtains its values from vector \mathbf{h} . Moreover, the projection performed by the hidden layer of a multilayer perceptron distorts data structure and inter-pattern distances in order to achieve a better approximation.

5.3. Extreme Learning Machine

Huang *et al.*, 2004, is the reference for the description of ELM. The regression problem can be formulated as an attempt to find solutions for $\mathbf{w}_i = (w_{i1}, \dots, w_{in})$ and β_i using the following system of equations:

$$f(\mathbf{x}_j) = \mathbf{t}_j, \quad j = 1, 2, \dots, N$$

where

$$f(\mathbf{x}_j) = \sum_{i=1}^m \beta_i g(\langle \mathbf{w}_i, \mathbf{x}_j \rangle + b_i), \quad j = 1, 2, \dots, N.$$

This system can also be expressed more concisely as $\mathbf{H}\boldsymbol{\beta} = \mathbf{T}$, where \mathbf{H} is the hidden layer's output matrix of the neural network given by:

$$\mathbf{H}(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m, b_1, \dots, b_m, \mathbf{x}_1, \dots, \mathbf{x}_N) = \begin{pmatrix} g(\langle \mathbf{w}_1, \mathbf{x}_1 \rangle + b_1) & \dots & g(\langle \mathbf{w}_m, \mathbf{x}_1 \rangle + b_m) \\ \vdots & \ddots & \vdots \\ g(\langle \mathbf{w}_1, \mathbf{x}_N \rangle + b_1) & \dots & g(\langle \mathbf{w}_m, \mathbf{x}_N \rangle + b_m) \end{pmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \mathbf{T} = \begin{bmatrix} t_1 \\ \vdots \\ t_m \end{bmatrix}.$$

Each column on matrix \mathbf{H} is made of the values of the corresponding hidden layer node, evaluated for each one of the patterns \mathbf{x}_i in the training set.

The ELM algorithm randomly selects the values for $\mathbf{w}_i = (w_{i1}, \dots, w_{in})$ and b_i , and then obtains corresponding values for $\beta_0, \beta_1, \dots, \beta_M$, from the generalized linear model. This is done by calculating the minimum quadratic solution of the linear system, given by:

$$\hat{\boldsymbol{\beta}} = \mathbf{H}^\dagger \mathbf{T}$$

where $\mathbf{H}^\dagger = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ is the generalized Moore-Penrose inverse matrix (Serre, 2002). The solution $\hat{\boldsymbol{\beta}} = \mathbf{H}^\dagger \mathbf{T}$ obtained has the following properties:

- a) It minimizes the training error:

$$\hat{\boldsymbol{\beta}} = \arg \min \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|$$

- b) It is the minimum Euclidean norm among all the possible solutions of the linear system:

$$\|\hat{\boldsymbol{\beta}}\| = \|\mathbf{H}^\dagger \mathbf{T}\| \leq \|\boldsymbol{\beta}\|$$

In short, the corresponding algorithm for this method is as follows (Figure 5).

Figure 5. ELM algorithm

Given a training set $D = \{(\mathbf{x}_i, t_i) : \mathbf{x}_i \in \mathbb{R}^n, t_i \in \mathbb{R}, i = 1, 2, \dots, N\}$, the activation function $g(t)$, and m neurons in the hidden layer:

Step 1: Assign arbitrary input weights for w and bias b .

Step 2: Calculate the hidden layer output matrix \mathbf{H} .

Step 3: Calculate the output weights $\hat{\boldsymbol{\beta}} = \mathbf{H}^\dagger \mathbf{T}$.

Source: elaborated by authors.

Geometrically speaking, the solution obtained corresponds to the orthogonal projection of vector \mathbf{T} , which determines the class corresponding to each pattern in the m -dimensional vector subspace (given by the number of nodes in the hidden layer) made by the column vectors of matrix \mathbf{H} .

Observe that when $N = m$ (i.e. there are as many nodes in the hidden layer as patterns in the training set), matrix \mathbf{H} is square and the corresponding system of equations has a unique solution, which is equivalent to saying that the training error is equal to zero. This happens because vector \mathbf{T} is in the subspace made by the m column vectors of matrix \mathbf{H} , and thus it can be expressed as the only possible linear combination. In this case, however, the generalisation error over the test set will be greater since overfitting takes place. It is therefore essential to determine the optimal number of nodes to avoid overfitting.

The ELM algorithm has been shown to have a good generalisation capability while it significantly reduces the time needed to train the neural network. For more details on this method the reader can see: Huang *et al.*, 2004; and Huang *et al.*, 2006.

5.4. The Levenberg-Marquardt Algorithm

The problem of learning in neural networks has been formulated in terms of the minimisation of an error function depending on the adaptive parameters (weights and biases) in the network. The Levenberg-Marquardt (LM) algorithm is a modification of the gradient descent method specifically designed to a sum-of-squares error function (Levenberg, 1944; Marquardt, 1963). The LM algorithm seeks to minimize the error function while at the same time trying to keep the step size small so as to ensure that the linear approximation of the error function remains valid. It has become a standard technique for nonlinear least-squares problems and can be thought of as a combination of steepest descent and the Gauss-Newton method. For this reason it has been chosen to optimize the MLP model in the comparison carried out in the next section. A description of the LM algorithm applied to neural network optimization can be seen in Bishop, 1995.

6. EXPERIMENTS

Several experiments are carried out in order to evaluate the performance of the regression method in predicting the default rate. The only preprocessing applied to the original database is the (-1,1) normalization of each one of the attributes, to facilitate the fitting processes of the neural networks.

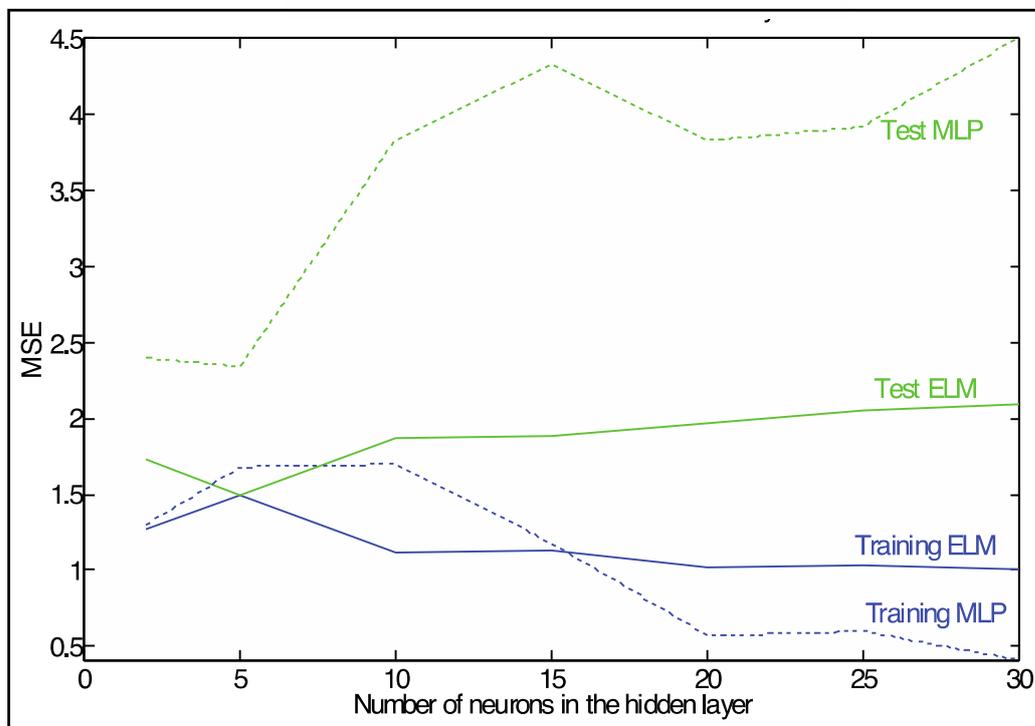
Firstly, the non-linear model defined by the ELM algorithm is compared to the standard MLP neural network training with the Levenberg-Marquardt algorithm (MLP-LM). The MLP-LM used in this experiment can be found in the MATLAB Toolbox for Neural Networks included in the 2010 release of MATLAB®. The code for ELM, also in MATLAB, can be downloaded from G. B. Huang's website (http://www.ntu.edu.sg/home/egbhuang/ELM_Codes.htm). For the comparison of the models the Mean Squared Error (MSE) has been used as given by:

$$MSE = \frac{1}{n} \sum_{l=1}^n (y_l - \hat{y}_l)^2$$

where y_i are the observed values and \hat{y}_i the predicted ones.

In order to give both ELM and MLP-LM a fair chance to perform accurately, we run both neural networks for a number of neurons in the hidden layer ranging from $m = 2$ to $m = 30$ neurons. The data set was divided into training (73 patterns) and testing (38 patterns) subsets to evaluate the generalization capability of the models. Both ELM and MLP-LM algorithms have been executed 30 times for the basic partition previously mentioned, and the average value for the regression results has been considered. Figure 6 shows the averaged MSE values for each method (after 30 runs), using the same training and test samples.

Figure 6. ELM vs. MLP-LM compared for different number of neurons



Source: elaborated by authors.

We can see how, as we increase the number of neurons in the hidden layer, the complexity of the model increases accordingly. This decreases the MSE in the training set, while it increases in the test set, giving the expected overfitting. The MLP-LM method has a more pronounced overfitting, while ELM works better in terms of generalisation.

Figure 6 shows how both methods perform best for $m = 5$ neurons in the hidden layer. A more detailed look at this particular instance can be seen in Table 3, which contains MSE average values in training and test subsets of the two models for 30 runs of each algorithm. Each model shows the mean, standard deviation, and best and worst results of MSE.

Table 3. MSE values in the training and test subsets of the two models for $m = 5$

Model	# runs	MSE training	Std. dev. MSE training	MSE testing	Std. dev. MSE testing	Best model (testing)	Worst model (testing)
ELM	30	1.501	0.053	1.493	0.235	1.211	2.494
MLP(L-M)	30	1.723	0.532	2.343	1.064	1.162	6.510

Source: elaborated by authors.

6.1. Statistical Comparison

In order to confirm that the differences observed both in the mean and the variance between ELM and MLP-LM (for $m = 5$) are significant, we have made the following verifications using the SPSS package. We have chosen the comparison for $m = 5$ since the best performance for both methods is obtained for this value. The comparison between the two procedures requires, as a first step, a test of normal fit for both distributions. The K-S test shows that both methods present normal distributions (p-values of 0.610 and 0.160 for MLP-LM and ELM, respectively).

Once the normality is verified, and since the partitions are the same for both methods in each repetition, it is now necessary to determine the possible relation between samples, which in turn depends on the method used for their comparison (t-test of either independent or related samples). We can assume the independence of the variables due to the low correlation coefficient (0.081 with p-value 0.670).

The comparison of the two measures, given the difference found in their variances, is significant (zero p-value in both contrasts). We can then conclude that the difference observed both in dispersion and in mean values for both methods is statistically significant, and thus the values of both parameters are greater for the MLP-LM experiment.

6.2. Prediction Capability

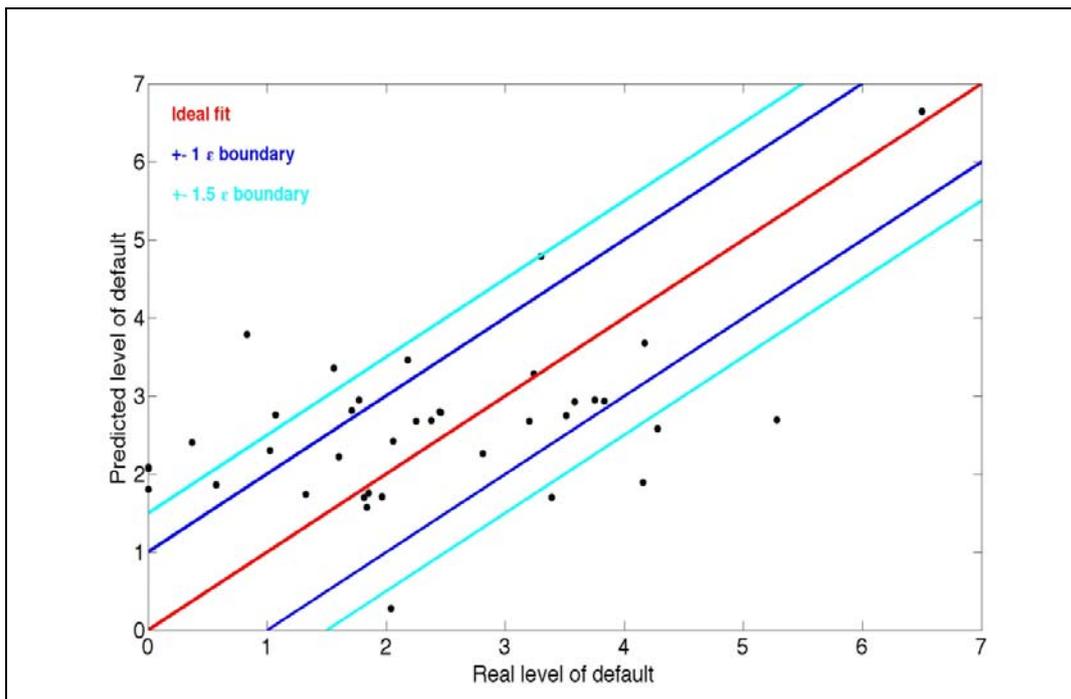
In order to learn more about the ability of this model to predict the default rate, the following measure has been considered. From the best model obtained with ELM (with $m = 5$ hidden nodes), the percentage of patterns is calculated for the training and testing sets whose predicted value belongs to an interval ε around the ideal fit. Table 4 shows the values of this measure for $\varepsilon=1$ and $\varepsilon=1.5$. Moreover, Figure 7 depicts the real and predicted values in our test set for the 30 runs of the ELM method.

Table 4. Values of measure for different ε

Sets	$\varepsilon=1$	$\varepsilon=1.5$
Training	60.3%	75.4%
Testing	55.2%	71.0%

Source: elaborated by authors.

Figure 7. Predicted vs. real levels of default in a sample test set



Source: elaborated by authors.

7. CONCLUSION

Today the Spanish financial system is in a process of rearrangement due to the problems caused by the current economic crisis. This is particularly relevant in terms of the relationship between financial institutions and the real estate and construction industries. A critical element for analysis is the level of default, which in turn indicates the quality of the assets held by the entities and the subsequent credit risk they are exposed to. We must bear in mind that, from the point of view of financial institutions, an adequate management of this risk is critical to guarantee the future of the entities themselves. It is also fundamental for the overall strength and reliability of the entire financial system is the fact that the institutions need to improve their performance within the economic system.

Financial authorities also need to be permanently aware of each institution's credit risk situation, and thus demanding greater controls and restrictions in the case where any of them drift towards more problematic scenarios. The demand for solvency needs to be founded on clear and reliable methods of evaluation. These are the motivations that explain the interest in methods that permit the level of default (especially for institutions that anticipate a greater credit risk) to be estimated, given certain magnitudes related to the accounting balance of each entity.

In order to carry out this estimate, a neural network has been used that is based on a non-linear and non-parametric model trained by the ELM algorithm. There are two benefits to be derived from the use of this approach:

- a) From a financial point of view, this algorithm provides an acceptable estimate of the level of default. This can be useful for both the entities and the authorities involved.
- b) The method allows for a good generalisation and a fast learning adaptation of the data in use. The results obtained by the ELM method provide better performance when compared to those of the MLP-LM method.

ACKNOWLEDGEMENTS

This work was supported in part by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the *Junta de Andalucía*, Spain, under Project P2011-TIC-7508. We would also like to thank the board of reviewers for their careful study and their ideas for improving both the writing and the experiments presented in this paper.

REFERENCES

- Bishop, C.M. (1995) *Neural networks for pattern recognition*, Oxford, Oxford University Press.
- Boal Velasco N. and González Sánchez, M. (2001) “Estimación del riesgo de crédito mediante modelos internos”, *Banca & Finanzas*, n. 66, pp. 40–45.
- Coakley, J.R. and Brown, C.E. (2000) “Artificial Neural Networks in Accounting and Finance: Modelling Issues”, *International Journal of Intelligent Systems in Accounting, Finance & Management*, n. 9, pp. 119–144.
- Confederación Española de Cajas de Ahorro (CECA) (2010) “Informe sobre la morosidad en las cajas de ahorros”, Mimeo.
- Cruz González, F.J. de la (1998) “Enfoques cuantitativos para la predicción del riesgo de crédito”, en *Predicción de la insolvencia empresarial*, Madrid, Monografías AECA.
- García Céspedes, J.C., (2005) “Nuevas técnicas de medición del riesgo de crédito”, *Revista Economía Financiera*, n. 5, abril, pp. 86–114.
- Gutiérrez, P.A., Segovia-Vargas, M.J., Salcedo-Sanz, S., Hervás-Martínez, C., Sanchis, A., Portilla-Figueras, J.A., Fernández-Navarro, F. (2009) “Hybridizing logistic regression with product unit and RBF networks for accurate detection and prediction of banking crises”, *Omega*, doi:10.1016/j.omega.2009.11.001.
- Herbrich, D., Keilbach, M., Graepel, T., Bollmann-Sdorra, P., and Obermayer, K. (2000) “Neural Networks in Economics: Background, applications and new developments”, in

Advances in Computational Economics: Computational techniques for Modelling Learning Economics, T. Brenner, Editor, Kluwer Academics. pp.169–196 .

- Hornik, K. (1989). “Multilayer feedforward neural networks are universal approximators”, *Neural Networks*, 2 (5), pp. 359–366.
- Huang, G.B., Wang, D.H., and Lan, Y. (2011) “Extreme learning machines: a survey”, *International Journal of Machine Learning and Cybernetics*, pp. 1–16.
- Huang, G., Zhu, Q., and Siew, C. (2004). “Extreme learning machine: a new learning scheme of feedforward neural networks”, *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, 70, pp. 985–990.
- Huang, G., Zhu, Q., and Siew, C. (2006). “Extreme learning machine: Theory and applications”, *Neurocomputing*, 70 (1-3), pp. 489–501.
- Levenberg, K. (1944) “A Method for the Solution of Certain Non-linear Problems in Least Squares”, *Quarterly of Applied Mathematics*, 2(2) , Jul, pp. 164–168.
- Li, F.C. Wang, P.K., and Wang, G.E. (2009) “Comparison of the primitive classifiers with extreme learning machine in credit scoring”, en *Industrial Engineering and Engineering Management, 2009. IEEM 2009. IEEE International Conference*, pp. 685–688.
- López, J.A. and Saidenberg, M.R. (2000) “Evaluating credit risk models”, *Journal of Banking & Finance*, vol. 24, n. 1-2, pp. 151–165.
- López-Martín, M.C., Montero-Romero, M.T., Becerra-Alonso, D., and Martínez Estudillo, F.J. (2011), “Clasificación por nivel de morosidad de las entidades de depósito españolas mediante redes neuronales”, *Anales de Economía aplicada 2011*, p. 373.
- Mahdavi, M. Fesanghary, M., and Damangir, E. (2007) “An improved harmony search algorithm for solving optimization problems”, *Applied Mathematics and Computation* 188, n°. 2, pp. 1567–1579.
- Marquardt, D.W. (1963). “An Algorithm for the Least-Squares Estimation of Nonlinear Parameters”, *SIAM Journal of Applied Mathematics*, 11(2), pp. 431–441, Jun.
- Martínez Estudillo, F.J., Hervás Martínez, C., Torres Jiménez, M., and Martínez Estudillo, A.C. (2007) “Modelo no lineal basado en redes neuronales de unidades producto para la clasificación. Una aplicación a la determinación del riesgo en tarjetas de crédito”, *Revista de Métodos Cuantitativos para la Economía y la Empresa*, n. 3, junio, pp. 40–62.
- Mcnelis, P.D. (2005) *Neural Networks in Finance: Gaining Predictive Edge in the Market*, Advanced Finance Series, Elsevier Academic Press

- Parisi, A., Parisi, F., and Díaz, D. (2006), “Modelos de Algoritmos Genéticos y Redes Neuronales en la Predicción de Índices Bursátiles Asiáticos”, *Cuadernos de Economía*, n. 43, pp. 251–284.
- Rodríguez Fernández, J.M. (1987) “Crisis en los bancos privados españoles: un modelo logit”, *Investigaciones Económicas*, suplemento, pp. 59–64.
- Serre, D. (2002) *Matrices: theory and applications*, New York, Springer.
- Soler, M. and Miró, A. (2001) “Enfoques cuantitativos para riesgo de crédito de particulares y su aplicación a realidades nacionales diferentes”, *Perspectivas del sistema financiero*, n. 72, pp. 43–56.
- Sorjamaa, A., Miche, Y., Weiss, R., and Lendasse, A. (2008) “Long-term prediction of time series using NNE-based projection and OP-ELM”, en *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference*, pp. 2674–2680.
- Van Heeswijk, M., Miche, Y., Lindh-Knuutila, T., Hilbers, P., Honkela, T., Oja, E., and Lendasse, A. (2009) “Adaptive ensemble models of extreme learning machines for time series prediction”, *Artificial Neural Networks–ICANN 2009*, pp. 305–314.
- Wong, W.K. and Guo, Z.X. (2010) “A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm”, *International Journal of Production Economics*, pp. 614–624.

APPENDIX: DATA SOURCES

Asociación Española de Banca: “Estados financieros públicos individuales 2008” (balance y cuenta de pérdidas y ganancias):

<http://www.aebanca.es/es/EstadosFinancieros/index.htm?pAnio=2008>

Confederación Española de Cajas de Ahorro (CECA): “Estados financieros públicos individuales 2008” (balance y cuenta de pérdidas y ganancias):

<http://www.cajasdeahorros.es/balance.htm>

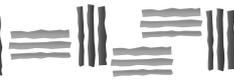
Unión Nacional de Cooperativas de crédito (UNACC): “Estados financieros públicos individuales 2008” (balance y cuenta de pérdidas y ganancias):

http://www.ruralvia.com/rsi_data/downloadPDF?p_report=4310_unacc_1.rdf&p_modelo=4310&p_periodo=20081231,

http://www.ruralvia.com/rsi_data/downloadPDF?p_report=2300_unacc_1.rdf&p_modelo=2300&p_periodo=20081231



UNIVERSIDAD
PABLO DE OLAVIDE
SEVILLA



REVISTA DE MÉTODOS CUANTITATIVOS PARA
LA ECONOMÍA Y LA EMPRESA (13). Páginas 24–53.
Junio de 2012. ISSN: 1886-516X. D.L: SE-2927-06.
URL: <http://www.upo.es/RevMetCuant/art.php?id=57>

Determinantes de la temporalidad en el mercado laboral ecuatoriano

CHÁVEZ, YANNIRA

Instituto Nacional de Estadística y Censos, Quito (Ecuador)

Correo electrónico: yannira_chavez@inec.gob.ec

MEDINA, PAÚL

Instituto Gregorio Millán, Universidad Carlos III de Madrid (España)

Correo electrónico: plmedina@espe.edu.ec

RESUMEN

Las diferencias que existen dentro del mercado laboral ecuatoriano, en las diferentes ramas de actividad, evidencian los niveles de formación que los trabajadores deben tener para evitar efectos negativos. Por tal motivo, en este estudio se analiza el papel que las características personales, laborales y geográficas desempeñan en la probabilidad de obtener un contrato temporal frente a uno indefinido. El análisis es realizado por rama de actividad, para determinar qué características posibilitarían la existencia del contrato temporal en cada una de ellas. Para lograr este objetivo, se estiman modelos de regresión logística utilizando los datos de la Encuesta de Empleo, Desempleo y Subempleo desde el 2º trimestre del año 2007 al 2º trimestre del año 2010, elaborada por el Instituto Nacional de Estadística y Censos (INEC).

Palabras clave: contrato temporal; contrato indefinido; regresión logística; rama de actividad.

Clasificación JEL: J23; J24; J41.

MSC2010: 00A71; 03C98; 93E10.

Determinants of Temporality in Ecuadorian Labor Market

ABSTRACT

The differences that exist inside the labor Ecuadorian market, in the different branches of activity, demonstrate the training levels that the workers must have to avoid negative effects. However, in this study, the role played by individual, jobs and residence characteristics are analyzed on the probability of having a fixed-term employment versus permanent employment. It is analyzed concretely by branch of activity, to check which are the characteristics that would make it possible the existence of the fixed-term employment in each of them. To achieve this aim, there are estimated models of logistic regression using the information of the Survey of Employment, Unemployment and Underemployment from 2nd quarter of 2007 to the 2nd quarter of 2010, elaborated by the National Institute of Statistics and Censuses (INEC).

Keywords: fixed-term employment; permanent employment; logistic regression; branch of activity.

JEL classification: J23; J24; J41.

MSC2010: 00A71; 03C98; 93E10.



1. Introducción

El mercado laboral ecuatoriano ha vivido cambios estructurales en las décadas de los 80 y los 90 del siglo XX, así como en el primer lustro del nuevo siglo, cuando se empezó con las propuestas de la flexibilidad laboral, una mayor apertura comercial, las propuestas privatizadoras, la apertura financiera, las reformas de la política cambiaria y la reforma tributaria; fueron los puntales básicos de este cambio [1]. Las tasas de desempleo han estado en un rango del 8 al 15 %, con su máximo en la crisis de 1999 y su mínimo en el año de 1980. El subempleo ha tenido unas tasas igualmente altas, llegando hasta el 54 % en 1991 y el nivel más bajo en 1980, con el 30 %. En el período de 1999 a 2006, los niveles de desempleo se estabilizaron a niveles similares a los de la pre-crisis (1994-1998). En este período se tuvo un pico del 17 % de desempleo total en febrero del 2000; posteriormente, el desempleo más alto que se tuvo fue del 12,05 % en abril del 2004. El mes con menos desempleo fue diciembre del 2002 con el 7,74 % [5]. En el período del 2007 al 2010¹ se presentó una mejoría sensible a la registrada en las últimas décadas, manteniéndose en promedio una tasa del 7,5 %; la tasa más alta se registró en marzo del 2010 con un 9,1 % y la más baja en diciembre del 2010 con un 6,1 % [15].

La flexibilidad laboral en Ecuador empezó desde 1990 con la Ley de Maquilas y la contratación a tiempo parcial [23], proporcionando otras formas de contratación como: contratos por tiempo parcial y contratos eventuales. Dentro del concepto de contrato temporal se distinguen dos grupos. El primero considera a los contratos eventuales, es decir, contratos por obra, a destajo, por horas y por jornal. Estos tipos de contratos son por tiempo determinado o por necesidades eventuales de las empresas. Generalmente, los trabajadores que se encuentran bajo esta relación laboral pueden ser despedidos sin que el empleador incurra en grandes pagos, de modo que los trabajadores terminen usualmente sin derecho a ninguna indemnización. Esta última característica podría explicar el dramático crecimiento de esta modalidad laboral. El segundo grupo considera a los contratos de temporada, es decir, aquellos que en razón de la costumbre o de la contratación colectiva se celebran entre una empresa o empleador y un trabajador o grupo de trabajadores para que realicen trabajos cíclicos o periódicos, en razón de la naturaleza discontinua de sus labores, gozando estos contratos de estabilidad, entendida como el derecho de los trabajadores a ser llamados a prestar sus servicios en cada temporada que se requieran. Este tipo de contrato, a diferencia de los anteriores, implica la obligación por parte del empleador de renovar el contrato por las siguientes temporadas; caso contrario, el empleador debe incurrir en el pago de indemnización [10].

Por otro lado, se debe tener presente que en el año 2000 se dictó la Ley para la Transformación Económica de Ecuador “Trole 1” [24], que incorporó la contratación por horas, además de implementar al dólar como moneda de curso legal. Otra de las reformas laborales importantes fue la tercerización, introducida en 1988 con posteriores reformas en 2006, que no es más que una intermediación de empleo, para formalizar un contrato de trabajo entre empleadores y empleados, para brindar servicios de personal a terceras personas.

Centrándose en el aspecto de la tercerización, no ha sido una reforma positiva para la creación y la calidad del empleo², debido a que los contratos temporales han presentado una evolución progresiva a lo largo de los tres últimos años de análisis. Sin embargo, se presenta un notable decrecimiento a partir

¹En el año 2007, el INEC introdujo un cambio metodológico en la forma de calcular la tasa de desempleo [12].

²De acuerdo al Ec. Rafael Correa, presidente de la República del Ecuador [5].

de que la Asamblea Constituyente del año 2008, mediante el Mandato Constituyente N° 8 [20]³, en su parte esencial, elimina y prohíbe la tercerización laboral. Así, a partir de este mandato, se produce un decrecimiento en la temporalidad laboral. Pero no se evidencia un efecto progresivo en el largo plazo, pues se presenta una tendencia creciente que se mantiene a lo largo del año 2009. Esto podría deberse a que las empresas ecuatorianas no están en la capacidad de brindar estabilidad a sus trabajadores, ni brindar posibilidades de formación y promoción interna, generándose altas tasas de rotación, menor antigüedad laboral y salarios más bajos para este grupo de asalariados; es decir, existe un desajuste educativo y de movilidad laboral [18].

De acuerdo a la realidad ecuatoriana señalada, el crecimiento dentro del mercado laboral se ha basado en la precarización al trabajador, olvidando que la productividad de los empleados depende no solo de su aptitud y de la inversión que se realiza en ellos, tanto dentro como fuera del puesto de trabajo, sino también de su motivación y de la intensidad de su esfuerzo [7]. La teoría del capital humano señala que la inversión en la formación de los trabajadores hace que este se convierta en un factor para disminuir el desempleo; es por esto que se pretende verificar si, en Ecuador, la formación es un factor influyente al momento de tener un contrato indefinido. Además, es preciso recalcar la hipótesis de que los trabajadores que se encuentran en período de prueba son más eficientes, de que existe una auto preparación, de que los empleadores tienen un conocimiento previo de sus trabajadores al conocer las capacidades y habilidades del trabajador frente a las funciones que debe desempeñar [19].

Por lo expuesto, este trabajo tiene dos objetivos. El primero consiste en identificar las características primordiales que afectan o determinan que una persona tenga un contrato temporal y, de esta manera, poder influir en dichas características. El segundo, consiste en determinar en qué medida la temporalidad es un fenómeno que depende tanto de las características de los trabajadores como de la rama de actividad de las empresas.

El artículo se estructura como sigue. En la sección 2 se expone la metodología y las especificaciones econométricas que servirán de base para las estimaciones que se realizarán. En la sección 3 se presenta el modelo desarrollado que permite el contraste entre la técnica de estimación empleada, los datos y variables utilizadas. En la sección 4 se señalan las conclusiones más relevantes. Al final, se presenta un anexo en el que se detalla el análisis descriptivo de los datos utilizados.

2. Metodología

En esta sección se plantean las especificaciones econométricas necesarias para analizar los factores determinantes al momento de tener un contrato temporal, en contraposición con el contrato indefinido.

La modelización estadística con datos binarios tiene como fin adaptar las herramientas de regresión convencionales para relacionar la variable respuesta (Y) con variables explicativas (X_1, X_2, \dots, X_n) [14]. Dentro de este contexto, los modelos más habituales son: el modelo Logit y el modelo Probit, con la diferencia de que el uno utiliza la función de distribución logística y el otro la normal tipificada [6]. En este estudio se utilizará el modelo Logit por ser el modelo que más se ha potenciado en la aplicación de los estudios empíricos [21].

³Eliminación y prohibición de la tercerización, intermediación laboral, contratación laboral por horas y cualquier forma de precarización de las relaciones de trabajo.

Para cuantificar la influencia que ejercen las variables explicativas, en la probabilidad de tener una relación laboral temporal frente a una indefinida, se especifica el siguiente modelo Logit:

$$\Pr(Y = y_j) = \frac{\exp(\eta_i)}{1 + \exp(\eta_i)}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, m, \quad (1)$$

donde

$$Y = \begin{cases} 1 & \text{contrato temporal,} \\ 0 & \text{contrato indefinido} \end{cases}$$

y

$$\eta_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_n X_{in}. \quad (2)$$

La ecuación (1) representa una función de distribución de probabilidad; en consecuencia, toma sus valores entre 0 y 1. Para estimar los parámetros del modelo, se utiliza el método de máxima verosimilitud.

Tras estimar el modelo, en el siguiente paso se comprobará la significación estadística de cada uno de los coeficientes de la regresión; para ello, se emplea básicamente el estadístico de Wald⁴, ya que nos permite evaluar de forma individual la significancia de la covariable introducida.

Para la interpretación de los parámetros β , sin conocer la magnitud de los mismos ya que no es un modelo lineal, el signo indica que, si el estimador es positivo, incrementos en la variable asociada causarían incrementos en la probabilidad de tener un contrato temporal ($\Pr(Y=1)$). Y si el estimador es negativo, significará que incrementos en la variable asociada causarían disminuciones en la probabilidad de tener un contrato temporal ($\Pr(Y=1)$).

3. Modelo

En esta sección se presenta un modelo que permite contrastar la técnica de estimación empleada con los datos y las variables utilizadas, para probar las hipótesis de esta investigación:

- i Determinar las características que afectan el tener un contrato temporal.
- ii Determinar en qué medida la temporalidad depende de la rama de actividad.

3.1. Datos

La información estadística utilizada en el análisis proviene de la Encuesta de Empleo, Desempleo y Subempleo (ENEMDU) desde el II trimestre del año 2007 al II trimestre del año 2010 del área urbana [16], con las siguientes restricciones en la base:

1. Se considera a las personas de 15 a 65 años de edad, debido a que los Artículos 82 y 86 del Código de la Niñez y Adolescencia [22] señalan que la edad mínima para realizar un contrato laboral se fija en 15 años para todo tipo de trabajo; y, por otro lado, se establece hasta los 65 años de edad, pues es la edad mínima en la que los empleados pueden jubilarse.

⁴El estadístico de Wald contrasta la hipótesis de que un coeficiente aislado es distinto de 0 y sigue una distribución normal de media 0 y varianza 1.

2. Se seleccionaron a los individuos que están trabajando y tienen un contrato laboral. Además, se considera a individuos únicos en los diferentes períodos de tiempo; es decir, si un individuo Z fue encuestado en el segundo trimestre del año 2007 y el mismo individuo Z es encuestado nuevamente en el segundo trimestre del año 2008, se tomará la información más actual, es decir, la del segundo trimestre del año 2008.
3. Inicialmente, se contó con una muestra conformada por 586 018 observaciones a nivel nacional. Una vez eliminados los valores de las variables que no se encuentran en el dominio del estudio especificado en los numerales anteriores, la muestra final con la que se realizó el estudio es de 66 655 trabajadores del área urbana con contrato laboral.

Dentro de esta muestra, la composición o estructura de las variables de control que se incorporarán en el modelo toman en cuenta características personales, laborales y geográficas, ya que se pretende verificar si estas son determinantes al momento de tener un contrato temporal. La selección de las variables independientes se basa, en primera instancia, en estudios precedentes que se han realizado en España [8, 9, 13], pues a nivel de Ecuador no se ha encontrado un estudio previo sobre el tema referente.

En segunda instancia, el fundamento para determinar qué variables independientes son las que mejor explican la variable dependiente (1 tener un contrato temporal, 0 tener un contrato fijo), se llevó a cabo una serie de submodelos, evaluando la confusión e iteración [4], evitando de esta manera tener resultados estadísticos sesgados; de esta manera cada variable que integra el modelo multivariado es relevante en el estudio analizado, ya que tiene un aporte significativo.

3.1.1. Características personales

Dentro de las características personales que se analizan, tenemos: el *Sexo*, la *Edad*, el *Nivel de Instrucción* y el *Estado Civil*, con el objeto de analizar diferencias existentes entre las mismas.

1. El porcentaje de individuos que son hombres es el 60,9% y las mujeres el 39,1%; es así que la mayoría de los datos corresponden a individuos con un contrato laboral entre 15 y 65 años de edad que son hombres, reflejándose una diferencia de género en la determinación de contratos. Estos datos se muestran en la Tabla 1.

Sexo	Número	Porcentaje
Hombre	40 599	60,90
Mujer	26 056	39,09
Total	66 655	100

Tabla 1: Frecuencias y porcentajes para la variable *Sexo*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

2. De los individuos con un contrato laboral se puede observar, en la Tabla 2, que un 76,74% se encuentra entre personas de 15 a 45 años de edad y que tan solo un 7,65% corresponden a personas entre 56 y 65 años de edad, evidenciándose que, en Ecuador, la población económicamente activa se encuentra concentrada en personas jóvenes.

Rango de edad	Número	Porcentaje
15-25	18 586	27,88
26-35	18 357	27,54
36-45	14 213	21,32
46-55	10 400	15,60
56-65	5 099	7,65
Total	66 655	100

Tabla 2: Frecuencias y porcentajes para la variable *Rango de Edad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

- En la Tabla 3 se puede observar la estructura de la muestra con respecto al *Nivel de Instrucción* que se maneja de manera original, concentrándose los porcentajes más altos en personas con educación secundaria y personas con educación superior universitaria en un 33,54 % y un 30,38 %, respectivamente, indicando que, en Ecuador, existen dos grupos de concentración en los que existe, por un lado, personas que únicamente tienen acceso a una educación básica, por el otro, personas que tienen una educación superior y a la vez el nivel de individuos que tienen post-grado es mínimo.

Nivel de instrucción	Número	Porcentaje
Ninguno	819	1,23
Centro de alfabetización	171	0,26
Primaria	14 695	22,05
Educación básica	2 962	4,44
Secundaria	22 359	33,54
Educación media	3 047	4,57
Superior no universitaria	974	1,46
Superior universitaria	20 248	30,38
Post-grado	1 380	2,07
Total	66 655	100

Tabla 3: Frecuencias y porcentajes para la variable *Nivel de Instrucción*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

- La clasificación que se presenta en la variable *Estado Civil* es amplia, determinándose seis estados civiles. Las categorías de Separado, Divorciado y Viudo representan, en conjunto, un 11,36 %, siendo así más bajo que las otras categorías, como se puede observar en la Tabla 4. Se evidencia que en Ecuador la cantidad de personas casadas es mayor que los otros estados civiles.

3.1.2. Características laborales

Las características laborales que se analizan son: la *Ocupación*, *Antigüedad*, *Jornada Laboral*, *Tamaño de Establecimiento* y *Capacitación*, con el objeto de analizar si el tipo de empleo en el cual se encuentra la persona es un factor que tiene influencia en el tipo de contrato.

- En la muestra se observa que la mayor parte de los trabajadores se encuentran en el sector privado (60,55 %), mientras que se refleja un nivel menor al 1 % para empleados tercerizados, como se

Estado civil	Número	Porcentaje
Casado(a)	24 035	36,06
Soltero(a)	23 605	35,41
Unión libre	11 440	17,16
Separado(a)	4 800	7,20
Divorciado(a)	1 768	2,65
Viudo(a)	1 007	1,51
Total	66 655	100

Tabla 4: Frecuencias y porcentajes para la variable *Estado Civil*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

observa en la Tabla 5. Esto nos indica que el sector privado es el que ofrece mayor empleo. Cabe recalcar que la ocupación de Jornalero o Peón hace referencia a personas que se dedican a tareas sencillas y rudimentarias; se los identifica como aquellas personas que pactan su remuneración por día o jornada independientemente de la periodicidad de pago [17]. La clasificación que se toma en cuenta dentro de la variable *Ocupación* es la definida por el INEC en la ENEMDU, que permite abarcar la realidad del mercado de trabajo dentro de las ocupaciones que se desempeñan en Ecuador. Se observa en la Tabla 36 (ver página 52), que existen algunos contratos a jornal (2%) que no son contratos temporales; hecho que puede ser explicado, por las características culturales de la región Costa, pues es muy común que trabajadores de las haciendas bananeras laboren de manera continua sin tener establecido de forma legal un contrato indefinido. Sin embargo, por el tiempo que llevan realizando esta actividad, en muchos casos años, en la encuesta señalan tener un contrato indefinido.

Ocupación	Número	Porcentaje
Empleado privado	40 359	60,55
Empleado de gobierno	13 487	20,23
Jornalero o peón	7 933	11,90
Empleado doméstico	4 413	6,62
Empleado terciarizado	463	0,69
Total	66 655	100

Tabla 5: Frecuencias y porcentajes para la variable *Ocupación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

- De los individuos con un contrato laboral, cuya *Antigüedad* es mayor a un año, se tienen que estos representan el 69,27%. Evidenciando que la mayoría de población en Ecuador se mantiene trabajando al menos más de un año, como se puede observar en la Tabla 6.

Antigüedad	Número	Porcentaje
Mayor que un año	46 169	69,27
Menor o igual a un año	20 486	30,73
Total	66 655	100

Tabla 6: Frecuencias y porcentajes para la variable *Antigüedad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

3. En la Tabla 7 se puede observar que el 83,87% de los individuos con un contrato laboral tienen una *Jornada Laboral* completa, reflejándose, en este caso, la eliminación de contratos temporales vía mandato [10], por lo cual las personas presenten una mayor propensión en contratos con jornada completa.

Jornada laboral	Número	Porcentaje
Jornada completa	55 901	83,87
Jornada parcial	10 754	16,13
Total	66 655	100

Tabla 7: Frecuencias y porcentajes para la variable *Jornada Laboral*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

4. Es interesante observar en la Tabla 8 que el 64,59% de los establecimientos tienen menos de 100 trabajadores, reflejándose así que la mayoría de los centros de trabajo que existen en Ecuador son pequeños, medianos, o que se encuentran en crecimiento.

Tamaño del establecimiento	Número	Porcentaje
Menos de 100	43 052	64,59
100 y más	23 603	35,41
Total	66 655	100

Tabla 8: Frecuencias y porcentajes para la variable *Tamaño del establecimiento*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

5. En la Tabla 9 se puede observar que tan solo el 23,57% de las personas con un contrato reciben cursos de capacitación en el lugar de trabajo, evidenciando que los centros de trabajo ecuatorianos que están en la capacidad económica o física de brindar capacitación a su empleado son limitados, ya que la mayoría de los centros de trabajo del país son pequeños, medianos o se encuentran en expansión.

Capacitación	Número	Porcentaje
Sí	15 717	23,57
No	50 938	76,42
Total	66 655	100

Tabla 9: Frecuencias y porcentajes para la variable *Capacitación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

3.1.3. Características geográficas

En la Tabla 10 se puede observar que entre Sierra y Costa se encuentra el 94,07% de la muestra, en la Amazonía el 5,43% y en las Zonas no delimitadas, tan solo se encuentra el 0,50%. Esta variable es importante para reflejar si existe o no diferencia en relación al lugar de residencia, en la posibilidad de contar con un contrato laboral estable. Regiones geográficas o ciudades autorepresentadas como son: Quito, Cuenca, Ambato pertenecen a la Región Sierra y Machala, Guayaquil pertenecen a la Región Costa.

Ciudades autorepresentadas y regiones	Número	Porcentaje
Resto de la Sierra	17 072	25,61
Resto de la Costa	15 547	23,32
Guayaquil	8 830	13,25
Quito	6 857	10,29
Cuenca	5 057	7,59
Machala	4 872	7,31
Ambato	4 465	6,70
Amazonía	3 619	5,43
Zonas no delimitadas*	336	0,50
Total	66 655	100

Tabla 10: Frecuencias y porcentajes para la variable *Ciudades Autorepresentadas y Regiones*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

* Las *Zonas no delimitadas* se refiere a las zonas de las Golondrinas, Manga del Cura y el Piedrero. Se definen como *no delimitadas* ya que no pertenecen a ninguna región o provincia de manera específica.

3.1.4. Observaciones para el análisis

- El conocer cómo se encuentran distribuidas las características personales, laborales y geográficas de los individuos, en que tengan un contrato temporal o un contrato indefinido, nos permite conocer las variables con sus respectivas categorías disponibles y, principalmente, determinar las que serán de interés en el análisis.
- Para las variables *Nivel de Instrucción* y *Estado Civil* se realizará una agrupación de las categorías que se presentaron en la sección anterior, para facilidad en la interpretación de resultados. Es así que para la variable *Nivel de Instrucción* se unifica la información de los individuos que se registra con el sistema tradicional al sistema actual de educación [17], para tener un nivel de información homogéneo. En la variable *Estado Civil* se realizó una unificación entre Casado y Unión Libre, Divorciado y Separado, pues son categorías que, en principio, por su condición se las puede agrupar, dejando a las categorías de Soltero y Viudo por separado, para tener un nivel de información más agregado.
- A las variables *Edad* y *Número de Trabajadores* que, en principio, se presentaban como variables discretas, con el fin de tener un nivel de información que tenga un mayor aporte al análisis desarrollado, se las consideró categóricas. A la variable *Edad* se la dividió en rangos, es así que se empezó probando con una amplitud de 5, resultando no significativa en el modelo, por lo que se realizaron pruebas hasta obtener una amplitud que se ajustase al análisis, siendo esta de 10. En lo que respecta a la variable *Número de Trabajadores*, se la representa en forma dicotómica, de tal manera que cuando el número de trabajadores es menor que 100 se le denomina con 1 y a los de 100 y más trabajadores se le denomina con 0. En el Anexo A se puede observar cómo queda la distribución final de los datos.

3.2. Resultados del modelo

Un análisis preliminar indica diferencias marcadas en grupos de individuos destacándose, principalmente, características como el nivel de instrucción, la edad, el estado civil, la antigüedad, la capacitación que reciben los empleados; con respecto a tener un contrato temporal frente a un contrato indefinido. No cabe duda que, mediante el análisis descriptivo detallado en el Anexo A, se evidencia que un bajo nivel de instrucción, estar entre los 15 y 35 años de edad, ser soltero, tener una antigüedad inferior a un año y no recibir ningún tipo de capacitación por parte del empleador, es el perfil característico de las personas que mantienen un contrato laboral temporal. La evidencia empírica mostrada hace evidente la posibilidad de formular un modelo econométrico que permita conocer cuáles son las características de los asalariados que influyen en tener un contrato temporal o un contrato indefinido.

Los resultados obtenidos a través del modelo econométrico planteado, para averiguar el grado de influencia de cada una de las variables incluidas en el mismo, se muestran en la Tabla 11. En la citada tabla se explican los coeficientes estimados (β), su error estándar (E.T.), su significación estadística (Sig.) con la prueba de Wald, el Exp (β) y sus intervalos de confianza.

Sobre las características principales, evidenciadas en el análisis descriptivo sobre los individuos con contratos temporales y contratos indefinidos, se seleccionó un individuo de referencia, el cual posee características de protección (por ejemplo, Nivel de Instrucción Superior Universitaria) para realizar el análisis del modelo. El objeto de esta elección es la de poder tener comparabilidad en el análisis.

3.2.1. Análisis

En primer lugar, se puede observar que los coeficientes son significativos con un nivel de confianza del 95 % para todas las variables incluidas en el modelo. Cabe recalcar que para determinadas categorías de las siguientes variables: *Nivel de Instrucción* y *Ocupación*, en las categorías Post-grado, Superior no Universitaria y Empleado Doméstico, respectivamente, no son estadísticamente significativas al 95 %, ya que el valor de Sig. es mayor que 0,05. Sin embargo, no se eliminan las categorías que resultaron no significativas de estas variables, ya que en forma global las variables son significativas; es decir, en el caso de las variables categóricas, si alguna resulta significativa, la variable queda incluida en el modelo final, aun cuando el resto no lo sea [11]. Además, debido a que las mismas variables de forma univariante son significativas, se decide dejarlas en el modelo multivariado, ya que aportan información extra dentro del objetivo de la investigación [8] .

En segundo lugar, para comprobar si el modelo es globalmente significativo, se evalúa con la prueba del logaritmo del cociente de verosimilitudes [2, 3], en la que se obtiene una significación estadística de (0,00); indicando así que el modelo con las variables incluidas mejora el ajuste de forma significativa.

En tercer lugar, se evalúa el ajuste del modelo mediante una “valoración de prueba diagnosticada”. Esto quiere decir que se comprueba cómo clasifica el modelo obtenido a nuevos individuos de la muestra, en comparación con la realidad observada. Con el objetivo de evidenciar la clasificación que realiza el modelo con las diez variables predictoras, se presenta la Tabla 12 de clasificación, en la que se puede apreciar cómo el modelo obtenido clasifica correctamente a 27 934 (de los 35 545) trabajadores temporales ($Y=1$), por lo que su sensibilidad es del 78,60 %. Por otra parte, el modelo clasifica correctamente a 23 403 (de los 31 110) trabajadores con contrato indefinido, por lo que la especificidad del modelo es del 75,20 %. De forma global se puede decir que se ha clasificado correctamente al 77,0 % de los

Variables en la ecuación ^a								
	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95 % para Exp(B)	
							Inferior	Superior
Constante	-3,440*	0,053	4201,741	1	0,000	0,032		
Sexo								
Hombre	0,175*	0,022	63,286	1	0,000	1,192	1,141	1,244
Edad			469,147	4	0,000			
15-25	0,271*	0,032	72,571	1	0,000	1,312	1,232	1,396
26-35	0,220*	0,028	61,595	1	0,000	1,246	1,180	1,317
46-55	-0,261*	0,033	61,479	1	0,000	0,771	0,722	0,823
56-65	-0,555*	0,044	155,625	1	0,000	0,574	0,526	0,627
Nivel de instrucción			511,308	6	0,000			
Post-grado	0,047***	0,079	0,359	1	0,549	1,048	0,898	1,224
Superior no universitaria	-0,116***	0,084	1,886	1	0,170	0,891	0,755	1,051
Educación media	0,196*	0,026	56,528	1	0,000	1,216	1,156	1,280
Educación básica	0,575*	0,028	430,459	1	0,000	1,778	1,684	1,877
Centro de alfabetización	0,854*	0,211	16,323	1	0,000	2,349	1,552	3,556
Ninguna	1,035*	0,107	93,065	1	0,000	2,816	2,282	3,475
Estado civil			65,787	3	0,000			
Viudo	0,223*	0,081	7,604	1	0,006	1,250	1,067	1,464
Soltero	0,129*	0,024	28,153	1	0,000	1,137	1,084	1,193
Separado/Divorciado	0,234*	0,034	47,148	1	0,000	1,264	1,182	1,351
Antigüedad								
Menor igual a un año	1,501*	0,023	4178,587	1	0,000	4,485	4,285	4,693
Ocupación			1923,849	4	0,000			
Empleado doméstico	0,092**	0,054	2,875	1	0,090	1,097	0,986	1,220
Empleado/Obrero privado	0,098*	0,036	7,623	1	0,006	1,103	1,029	1,183
Empleado tercerizado	1,692*	0,124	187,192	1	0,000	5,429	4,260	6,918
Jornalero o peón	3,693*	0,093	1588,866	1	0,000	40,150	33,483	48,143
Jornada laboral								
Jornada parcial	0,767*	0,029	711,831	1	0,000	2,153	2,035	2,278
Tamaño del establecimiento								
Menos de 100	0,614*	0,029	457,254	1	0,000	1,847	1,746	1,954
Capacitación								
No recibe	0,884*	0,026	1112,924	1	0,000	2,420	2,298	2,549
Ciudades autorepresentadas y regiones			2235,333	8	0,000			
Zonas no delimitadas	1,762*	0,163	117,162	1	0,000	5,824	4,233	8,012
Machala	1,743*	0,052	1123,408	1	0,000	5,714	5,160	6,327
Guayaquil	1,529*	0,045	1149,38	1	0,000	4,614	4,224	5,040
Quito	0,663*	0,047	196,496	1	0,000	1,940	1,768	2,128
Ambato	0,604*	0,052	138,850	1	0,000	1,830	1,654	2,025
Resto de la Sierra	1,054*	0,042	638,166	1	0,000	2,869	2,644	3,114
Resto de la Costa	1,521*	0,043	1276,544	1	0,000	4,578	4,211	4,976
Amazonía	1,216*	0,056	467,668	1	0,000	3,372	3,020	3,765

Tabla 11: Resultados de las variables incluidas en el modelo. Fuente: elaboración propia a partir de la encuesta ENEMDU.

^a El individuo de referencia es mujer, en una edad entre 36 y 45 años, con un nivel de estudios superior universitaria, casado/unión libre, con una antigüedad superior a un año, empleada de gobierno, con jornada laboral completa, en una empresa de más de 100 trabajadores, recibe cursos de capacitación y reside en Cuenca.

(*) Coeficiente significativo al 5 %, (**) al 10 %, (***) no significativos [9].

individuos, siendo considerada una predicción aceptable [3]. Además, mediante la curva característica operativa del receptor (curva de COR), la cual discrimina la puntuación óptima de corte para una prueba de detección selectiva, la probabilidad de clasificar correctamente un par de individuos con contrato temporal y contrato indefinido es de 0,85; es decir, el poder de discriminación del modelo construido es de, aproximadamente, el 85 %.

Tabla de clasificación				
Observado		Pronosticado		
		Y		Porcentaje correcto
		Contrato indefinido	Contrato temporal	
Y	Contrato indefinido	23403	7707	75,2
	Contrato temporal	7611	27934	78,6
Porcentaje global				77,0

Tabla 12: Tabla de clasificación. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Finalmente, se presenta el resumen de la validación del modelo.

R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
0,351	0,469

Tabla 13: Resumen de la validación del modelo. Fuente: elaboración propia a partir de la encuesta ENEMDU.

El **R cuadrado de Cox y Snell** y el **R cuadrado de Nagelkerke** presentados en la Tabla 13 tienen un valor de 0,351 y 0,469, respectivamente, indicando que el 35,10 % y el 46,9 % de la variación de la variable dependiente es explicada por las variables incluidas en el modelo. Cabe indicar que el R cuadrado de Nagelkerke es una versión corregida del R cuadrado de Cox y Snell. Además, según [3, 13], se puede considerar que los valores obtenidos representan un buen ajuste en la estimación del modelo.

3.2.2. Interpretación de resultados

La interpretación de los resultados obtenidos, mostrados en la Tabla 11, se basarán en las características personales, laborales y geográficas consideradas en este estudio.

Las características personales que afectan positivamente a la probabilidad de aumentar la obtención de una relación laboral temporal son: que un individuo tenga nivel de educación media, o educación básica, o centro de alfabetización, o ninguna; que sea hombre; que tenga una edad entre 15 y 35 años. Por otro lado, las características que posibilitan una disminución en la probabilidad de tener un contrato temporal son: que tenga un nivel de instrucción superior, sea mujer, casada y con una edad comprendida entre los 46 y 65 años de edad.

Las características laborales que conducen a aumentar asalariados temporales son: que tenga una jornada laboral parcial, que tenga una experiencia menor o igual a un año de trabajo, que esté trabajando en una empresa de menos de 100 trabajadores, que sea jornalero o peón y que no reciba cursos de capacitación. Por el contrario, las características que influyen para disminuir la probabilidad del contrato temporal son: que tenga una jornada laboral completa, que tenga una experiencia mayor que un

año, que se encuentre en una empresa de más de 100 trabajadores, que reciba cursos de capacitación y que desarrolle su trabajo dentro del sector público.

De acuerdo a las ciudades autorepresentadas y regiones que se incorporaron en el modelo, se puede notar que en Guayaquil, Machala, resto de la Costa, Amazonía y Zonas no delimitadas, el contrato temporal tienen mayor probabilidad de ocurrir. Por el contrario, en Quito, Ambato, Cuenca y resto de la Sierra esta probabilidad disminuye.

3.2.3. Escenarios del modelo

En esta sección, a partir de la base de estudio, se analizan tres diferentes escenarios; considerando los factores de protección y de riesgo, que presentan tres individuos seleccionados. Cabe indicar que el valor $Exp(\beta)$, obtenido para cada una de las variables utilizadas en el modelo, será el indicador de si ésta es o no una variable de riesgo.

El estudio y planteamiento de estos escenarios tiene por objeto observar la bondad de las estimaciones realizadas. A continuación, se presentan en la Tabla 14 los resultados para cada escenario. Los escenarios propuestos son: Pesimista, Moderado y Optimista.

- **Escenario Pesimista:** si se consideran las características menos favorables que puede tener un asalariado, las mismas serían las siguientes: un individuo hombre, entre 15 y 25 años, ningún nivel de instrucción, separado o divorciado, con una antigüedad menor igual que un año, jornalero o peón, con jornada parcial, en un establecimiento de menos de 100 trabajadores, no recibe cursos de capacitación y reside en las zonas no delimitadas de Ecuador. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,998 de tener un contrato temporal (véase la Tabla 14).
- **Escenario Moderado:** si se consideran condiciones aceptables de un asalariado en el mercado laboral, las mismas serían las siguientes: un individuo hombre de entre 26 y 35 años de edad, con un nivel de instrucción superior universitaria, soltero, empleado privado, con una antigüedad mayor que un año, con jornada completa, en un establecimiento de más de 100 trabajadores, que recibe capacitación y que habita en la ciudad de Guayaquil. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,216 de tener un contrato temporal (véase la Tabla 14).
- **Escenario Optimista:** si se consideran condiciones relativamente buenas y estables dentro del mercado laboral, el escenario optimista debería ser el siguiente: un individuo mujer, entre 56 y 65 años de edad, con post-grado, casada, empleada pública, con una antigüedad mayor a un año, con jornada completa, en un establecimiento de más de 100 trabajadores, que recibe cursos de capacitación y habita en la ciudad de Cuenca. Bajo las características señaladas, la persona seleccionada tendrá una probabilidad de 0,018 de tener un contrato temporal (véase la Tabla 14).

En los escenarios desarrollados, la distinción del sexo marca una diferencia; es decir, un hombre con las características que se señalan en los tres diferentes escenarios presenta una probabilidad de tener un contrato temporal mayor que la de mujeres.

Características del individuo		Escenarios		
		Pesimista	Moderado	Optimista
Sexo	Hombre	1	1	0
	Mujer	0	0	1
Edad	15-25	1	0	0
	26-35	0	1	0
	36-45	0	0	0
	46-55	0	0	0
	56-65	0	0	1
Nivel de instrucción	Post-grado	0	0	1
	Superior no universitaria	0	0	0
	Superior universitaria	0	1	0
	Educación media	0	0	0
	Educación básica	0	0	0
	Centro de alfabetización	0	0	0
	Ninguna	1	0	0
Estado civil	Viudo	0	1	0
	Soltero	0	1	0
	Separado/Divorciado	1	0	0
	Casado/Unión libre	0	0	1
Antigüedad	Mayor a un año	0	1	1
	Menor igual a 1 año	1	0	0
Ocupación	Empleado doméstico	0	0	1
	Empleado privado	0	1	0
	Empleado público	0	0	1
	Empleado tercerizado	0	0	0
	Jornalero o peón	1	0	0
Jornada laboral	Jornada completa	0	1	1
	Jornada parcial	1	0	0
Tamaño del establecimiento	Menos de 100	1	0	0
	Más de 100	0	1	1
Capacitación	Sí recibe	0	1	1
	No recibe	1	0	0
Ciudades autorepresentadas y regiones	Zonas no delimitadas	1	0	0
	Machala	0	0	0
	Guayaquil	0	1	0
	Quito	0	0	0
	Ambato	0	0	0
	Cuenca	0	0	1
	Resto de la Sierra	0	0	0
	Resto de la Costa	0	0	0
Amazonía	0	0	0	
$Pr(Y = 1)$		0,998	0,216	0,018

Tabla 14: Escenarios para el modelo. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Comprobación de la hipótesis (i): *Determinar las características que afectan a tener un contrato temporal.*

Las características como: el *Nivel de Educación*, el tipo de *Ocupación*, la *Antigüedad* y la *Jornada Laboral* son variables que, al modificar su estado, afectan en mayor medida a la probabilidad de obtener un contrato temporal.

En particular, cuando un individuo es jornalero o peón ése es el factor de riesgo que más influye en la obtención de tener un contrato temporal. Esto se observa en el escenario pesimista, con las condiciones establecidas, se obtuvo una probabilidad alta en la predicción del modelo. Las Zonas no delimitadas son las que presentan una mayor probabilidad de temporalidad y Cuenca es la ciudad de menor probabilidad.

3.3. Análisis por rama de actividad

El objetivo de este análisis es identificar las características relevantes de cada rama de actividad para, de esta manera, conocer cuál es el grupo de individuos predominante en cada rama y así tener un aporte que permita evidenciar y constatar las razones de la existencia del contrato temporal en cada una de ellas.

Para lograr el objetivo señalado, se realizó un análisis desagregado de acuerdo a la Clasificación Internacional Industrial Uniforme (CIIU) a un dígito, en la cual se consideran 17 Ramas de Actividad que permiten clasificar al establecimiento donde se encuentra el empleado. Cabe señalar que los datos existentes para el sector “Q. Organizaciones y órganos extraterritoriales” no son suficientes para realizar el análisis, pues la única información que se posee es la de 7 individuos a lo largo del período de análisis, y su representación porcentual dentro del conjunto de ramas de actividad consideradas es apenas el 0,01 % (véase la Figura 1). En la Figura 1 se puede observar la distribución de las distintas Ramas de Actividad, dentro del universo considerado para el estudio.

A continuación, se presentan los resultados obtenidos, enfocando principalmente en los *Rangos de Edad*, en el *Nivel de Instrucción*, en el tipo de *Ocupación*, y en las *Ciudades Autorepresentadas y Regiones*. Sin embargo, no se prestará mayor atención a las variables, *Sexo*, *Antigüedad*, *Jornada Laboral*, *Tamaño del Establecimiento* y *Capacitación*, pues son variables que, al aplicarlas en el modelo por Rama de Actividad, en general, no presentan variaciones sobre cada una de ellas; es decir, estas variables tienen el mismo efecto que se tiene en el modelo presentado en la Tabla 11.

Es necesario tomar en cuenta que los resultados que se presentan son realizados en base a una muestra, por lo que la agrupación de los individuos en determinadas ciudades representarían el comportamiento general de la región. Por ejemplo, los resultados de la ciudad de Machala explicaría el comportamiento en la región Costa, pues se encuentra en ella.

▪ A. Agricultura, ganadería, caza y silvicultura

En la Tabla 15 se observa que el factor de protección determinante es el tipo de *Ocupación* que desempeñen; es así, individuos con ocupación de empleado privado dentro de esta rama de actividad tienen menos propensión de tener un contrato temporal. Por otro lado, los factores de riesgo son: la *Edad*, el *Nivel de Instrucción* y el lugar donde residan; por ejemplo, un individuo

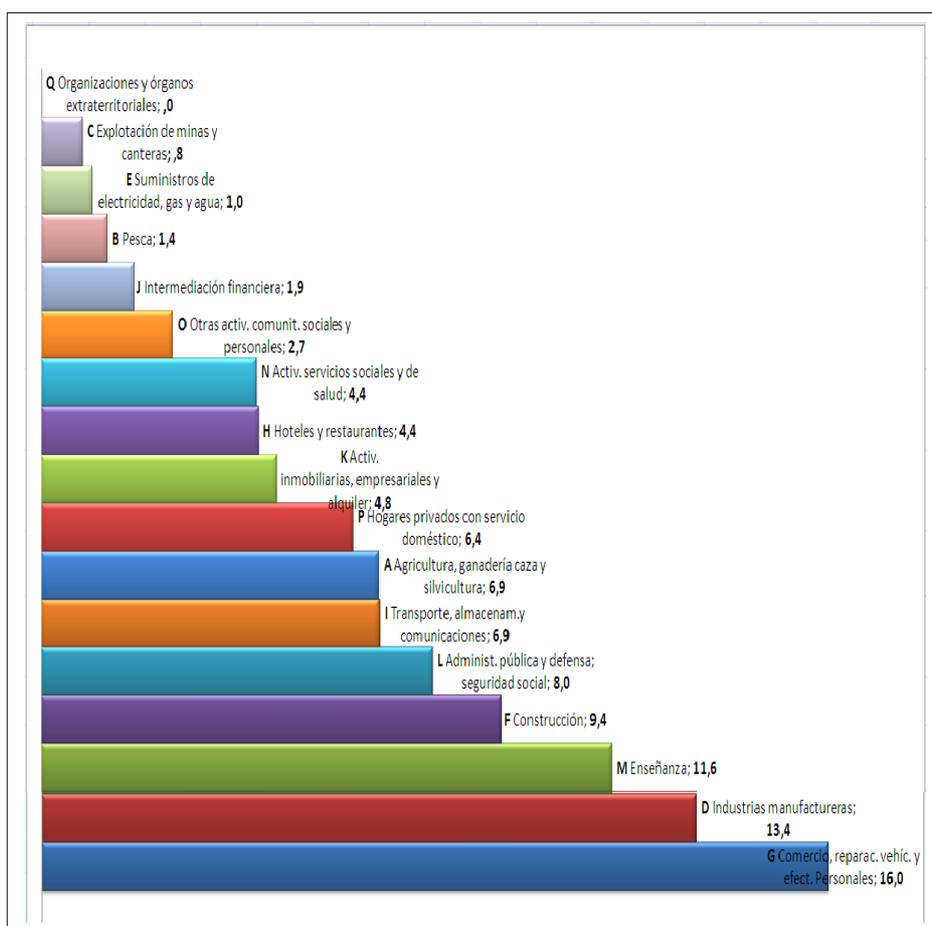


Figura 1: Distribución porcentual de individuos por Rama de Actividad. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Rama de Actividad	Características		Influencia
A. Agricultura, ganadería, caza y silvicultura	Rango de edad	15-25	Positiva
		26-35	Positiva
	Nivel de instrucción	Educación media	Positiva
		Educación básica	Positiva
		Ninguna	Positiva
	Ocupación	Empleado/Obrero privado	Negativa
	Ciudades autorepresentadas y regiones	Machala	Positiva
		Amazonía	Positiva
Zonas no delimitadas		Positiva	
Guayaquil		Positiva	
Resto de la Sierra		Positiva	
	Resto de la Costa	Positiva	

Tabla 15: Caracterización de los individuos con contrato temporal en la Rama A. Fuente: elaboración propia a partir de la encuesta ENEMDU.

entre 15 y 35 años de edad, con una educación media, o básica, o ningún tipo de educación, y que resida en Machala, o Guayaquil, o Resto de la Costa, o Resto de la Sierra, o Amazonía, o Zonas no delimitadas, es más propenso a tener un contrato temporal.

■ **B. Pesca**

Rama de actividad	Características		Influencia
B.Pesca	Rango de edad	15-25	Positiva
	Nivel de instrucción	Educación básica	Positiva
	Ocupación	Empleado/Obrero privado	Negativa

Tabla 16: Caracterización de los individuos con contrato temporal en la Rama B. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 16 se puede notar que los factores de riesgo son: el *Nivel de Instrucción* y la *Edad*; es así que individuos de 15 a 25 años con educación básica presentan una mayor propensión a tener un contrato temporal. Por otro lado, el factor de protección determinante es el tipo de *Ocupación*, pues individuos con empleo privado dentro de esta rama de actividad fomentan la disminución de la probabilidad de tener un contrato temporal. En esta rama de actividad no se puede determinar en qué región o ciudad existe una mayor probabilidad de una relación laboral temporal, debido a que este sector económico no se desarrolla en las ciudades que abarca la representación del universo considerado.

■ **C. Explotación de minas y canteras**

Rama de actividad	Características		Influencia
C. Explotación de minas y canteras	Rango de edad	15-25 46-55	Positiva Negativa
	Nivel de instrucción	Educación básica	Positiva
	Ocupación	Empleado/Obrero privado	Negativa

Tabla 17: Caracterización de los individuos con contrato temporal en la Rama C. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 17 se observa que entre los factores de riesgo se tiene la *Edad* y el *Nivel de Instrucción*; es así que un individuo de 15 a 25 años de edad, con un nivel de educación básica, presenta una mayor propensión de tener un contrato temporal. Por otro lado, como factor de protección se encuentra la edad, pues un individuo entre 46 y 55 años de edad tiene menos propensión de tener un contrato temporal. Al igual que en la rama de actividad Pesca, en el universo considerado no es posible establecer en que región o ciudad aumentaría el contrato temporal.

■ **D. Industrias manufactureras**

En la Tabla 18 se puede notar que el factor de protección determinante es la *Edad*: personas entre 46 y 65 años de edad dentro de esta rama de actividad tienen menos propensión de tener un contrato temporal. Por otro lado, los factores de riesgo que se presentan son el bajo nivel de instrucción, ser soltero, separado o divorciado, ser empleado tercerizado o jornalero peón y que resida en Machala, o Guayaquil, o Resto de la Costa, o Quito, o Ambato, o Resto de la Sierra, o Amazonía, siendo Machala y el Resto de la Costa los que presentan una mayor propensión de tener un contrato temporal.

Rama de actividad	Características		Influencia
D. Industrias manufactureras	Rango de edad	46-55	Negativa
		56-65	Negativa
	Nivel de instrucción	Ninguna	Positiva
		Centro de alfabetización	Positiva
		Educación básica	Positiva
Estado civil	Soltero	Positivo	
Ocupación	Separado/Divorciado	Positivo	
Ciudades autorepresentadas y regiones	Ocupación	Empleado tercerizado	Positiva
		Jornalero o peón	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Quito	Positiva
Ambato	Positiva		
Resto de la Sierra	Positiva		
Resto de la Costa	Positiva		
Amazonía	Positiva		

Tabla 18: Caracterización de los individuos con contrato temporal en la Rama D. Fuente: elaboración propia a partir de la encuesta ENEMDU.

▪ E. Suministro de electricidad, gas y agua

Rama de actividad	Características		Influencia
E. Suministro de electricidad, gas y agua	Rango de edad	15-25	Positiva
		26-35	Positiva
	Nivel de instrucción	Educación media	Positiva
	Ocupación	Empleado tercerizado	Positiva

Tabla 19: Caracterización de los individuos con contrato temporal en la Rama E. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 19 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se incrementará para los individuos de entre 15 y 35 años de edad, con educación media y que sea un empleado tercerizado. Por otra parte, en esta rama de actividad no se evidencia un factor de protección.

▪ F. Construcción

Rama de actividad	Características		Influencia
F. Construcción	Nivel de instrucción	Centro de alfabetización	Positiva
		Educación básica	Positiva
		Educación media	Positiva
	Estado civil	Soltero	Positiva
Ocupación	Empleado/Obrero público	Negativa	
	Empleado tercerizado	Positiva	
Ciudades autorepresentadas y regiones	Ocupación	Jornalero o peón	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Resto de la Sierra	Positiva
		Resto de la Costa	Positiva
Amazonía	Positiva		

Tabla 20: Caracterización de los individuos con contrato temporal en la Rama F. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 20 se puede apreciar que la probabilidad de tener un contrato temporal, en esta

rama de actividad, aumenta para los individuos con bajo nivel de instrucción, solteros, con un empleo tercerizado o jornalero y que vivan en Machala, o Guayaquil, o Resto de la Sierra, o Resto de la Costa, o Amazonía. Por otra parte, la característica de los individuos que posibilitan una disminución en la probabilidad de obtener una relación laboral temporal es que sea empleado público.

▪ **G. Comercio al por mayor y menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos**

Rama de actividad	Características		Influencia
G. Comercio al por mayor y menor; reparación de vehículos automotores, motocicletas, efectos personales y enseres domésticos	Rango de edad	15-25	Positiva
		46-55	Negativa
	Nivel de instrucción	Ninguna	Positiva
		Centro de alfabetización	Positiva
		Educación básica	Positiva
		Educación media	Positiva
	Estado civil	Soltero	Positiva
		Separado/Divorciado	Positiva
	Ocupación	Empleado tercerizado	Positiva
		Jornalero o peón	Positiva
	Ciudades autorepresentadas y regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
Guayaquil		Positiva	
Quito		Negativa	
Ambato		Negativa	
Resto de la Sierra		Negativa	
Resto de la Costa		Positiva	
Amazonía	Positiva		

Tabla 21: Caracterización de los individuos con contrato temporal en la Rama G. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 21 se puede apreciar que entre los factores de protección para tener un contrato temporal tenemos: la *Edad* entre 46 a 55 años y vivir en Quito, o Ambato, o en Resto de la Sierra. Por otra parte, como factores de riesgo para tener un contrato temporal tenemos a individuos entre los 15 y 25 años de edad, con un nivel de educación bajo, que sean solteros, o separados, o divorciados, que se desempeñen como empleado tercerizado, o jornalero o peón y que vivan en Machala, o Guayaquil, o Resto de la Costa, o Amazonía, o Zonas no delimitadas.

▪ **H. Hoteles y restaurantes**

Rama de actividad	Características		Influencia
H. Hoteles y restaurantes	Rango de edad	15-25	Positiva
	Nivel de instrucción	Educación básica	Positiva
	Ciudades autorepresentadas y regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Quito	Negativa
		Ambato	Negativa
		Resto de la Sierra	Negativa
		Resto de la Costa	Positiva
		Amazonía	Negativa

Tabla 22: Caracterización de los individuos con contrato temporal en la Rama H. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 22 se puede notar que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es por la agrupación de los individuos entre 15 y 25 años de edad, con

un nivel de educación básica, y que vivan en Machala, o Guayaquil, o Resto de la Costa, o Zonas no delimitadas. Por otra parte, las características que posibilitan una disminución en la probabilidad de tener un contrato temporal es que vivan en Quito, o Ambato, o Resto de la Sierra, o Amazonía.

■ **I. Transporte, almacenamiento y comunicaciones**

Rama de actividad	Características		Influencia
I. Transporte, almacenamiento y comunicaciones	Rango de edad	15-25	Positiva
		26-35	Positiva
	Nivel de instrucción	Educación básica	Positiva
		Educación media	Positiva
Ocupación	Empleado tercerizado	Positiva	
	Jornalero o peón	Positiva	
Ciudades autorepresentadas y regiones	Zonas no delimitadas	Positiva	
	Machala	Positiva	
	Guayaquil	Negativa	
	Quito	Negativa	
	Resto de la Sierra	Negativa	
	Resto de la Costa	Positiva	
Amazonía	Negativa		

Tabla 23: Caracterización de los individuos con contrato temporal en la Rama I. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 23 se puede apreciar que la probabilidad para que incremente el contrato temporal, en esta rama de actividad, es en los individuos entre los 15 y 35 años de edad, con un nivel de educación básica o media, desenvolviéndose en un empleo tercerizado o de jornalero peón y que vivan en Machala, o resto de la Costa, o Zonas no delimitadas. Por otra parte, las características que posibilitan la disminución de esta probabilidad, de tener un contrato temporal, es que residan en Guayaquil, o Quito, o Resto de la Sierra, o Amazonía.

■ **J. Intermediación financiera**

Rama de actividad	Características		Influencia
J. Intermediación financiera	Nivel de instrucción	Educación básica	Positiva
		Educación media	Positiva
	Ocupación	Empleado privado	Negativa
		Empleado tercerizado	Positiva
Ciudades autorep. y regiones	Quito	Negativa	
	Machala	Positiva	

Tabla 24: Caracterización de los individuos con contrato temporal en la Rama J. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 24 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se ve afectada positivamente por las personas que tenga un nivel de educación básica o media, que sea empleado de tercerizado y que vivan en Machala. Por otra parte, las características que hacen que disminuya la probabilidad de tener un contrato temporal son: ser empleado privado y que vivan en Quito.

■ **K. Actividades inmobiliarias, empresariales y de alquiler**

En la Tabla 25 se puede apreciar que la probabilidad de incrementar el contrato temporal, en esta rama de actividad, es para los individuos con un nivel de educación básica o media, que

Rama de actividad	Características		Influencia
K. Actividades inmobiliarias, empresariales y de alquiler	Nivel de instrucción	Educación básica	Positiva
		Educación media	Positiva
	Estado civil	Viudo	Positiva
		Separado/Divorciado	Positiva
		Soltero	Positiva
Ocupación	Empleado tercerizado	Positiva	
Ciudades autorepresentadas y regiones	Machala	Positiva	
	Guayaquil	Positiva	
	Quito	Negativa	
	Resto de la Sierra	Negativa	
	Resto de la Costa	Positiva	
	Amazonía	Positiva	

Tabla 25: Caracterización de los individuos con contrato temporal en la Rama K. Fuente: elaboración propia a partir de la encuesta ENEMDU.

sean viudos, o separados, o divorciados, o solteros, que tengan un empleo tercerizado y que vivan en Machala, o Guayaquil, o Resto de la Costa. Por otra parte, las características que hacen que disminuya esta probabilidad es que vivan en Quito, o Resto de la Sierra.

▪ **L. Administración pública y defensa; planes de seguridad social de afiliación obligatoria**

Rama de actividad	Características		Influencia
L. Administración pública y defensa; planes de seguridad social de afiliación obligatoria	Rango de edad	15-25	Positiva
		26-35	Positiva
		56-65	Negativa
	Nivel de instrucción	Educación básica	Positiva
		Separado/Divorciado	Positiva
	Estado civil	Soltero	Positiva
Machala		Positiva	
Ciudades autorepresentadas y regiones	Ambato	Negativa	
	Resto de la Sierra	Negativa	

Tabla 26: Caracterización de los individuos con contrato temporal en la Rama L. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 26 se observa que la probabilidad de obtener un contrato temporal, en esta rama de actividad, se ve afectada positivamente para las personas entre 15 y 35 años de edad, con un nivel de educación básica, que sea separado, o divorciado, o soltero y, que vivan en Machala. Por otra parte, las características que posibilitan la disminución de la probabilidad de una relación laboral temporal son: que se encuentre entre los 56 y 65 años de edad, y que vivan en Ambato, o el resto de la Sierra.

▪ **M. Enseñanza**

En la Tabla 27 se puede notar que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es por la agrupación de los individuos entre 15 y 35 años de edad, con ningún nivel de instrucción o centro de alfabetización y que vivan en Machala, o Guayaquil, o el resto de la Costa. Por otra parte, las características que hacen que la probabilidad de obtener un contrato temporal disminuya son que los individuos tengan una edad entre 46 y 65 años y que vivan en Ambato o Quito.

Rama de actividad	Características		Influencia
M. Enseñanza	Rango de edad	15-25	Positiva
		26-35	Positiva
		46-55	Negativa
		56-65	Negativa
M. Enseñanza	Nivel de instrucción	Ninguna	Positiva
		Centro de alfabetización	Positiva
	Ciudades autorepresentadas y regiones	Machala	Positiva
		Guayaquil	Positiva
		Ambato	Negativa
		Resto de la Costa	Positiva
		Quito	Negativa

Tabla 27: Caracterización de los individuos con contrato temporal en la Rama M. Fuente: elaboración propia a partir de la encuesta ENEMDU.

■ N. Actividades de servicios sociales y de salud

Rama de actividad	Características		Influencia
N. Actividades de servicios sociales y de salud	Rango de edad	15-25	Positiva
		26-35	Positiva
		46-55	Negativa
		56-65	Negativa
	Nivel de instrucción	Superior no universitaria	Negativa
		Educación básica	Positiva
	Ocupación	Empleado privado	Negativa
	Ciudades autorepresentadas y regiones	Machala	Positiva

Tabla 28: Caracterización de los individuos con contrato temporal en la Rama N. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 28 se aprecia que la probabilidad de aumentar la contratación temporal, en esta rama de actividad, es para los individuos entre 15 y 35 años de edad, con un nivel de educación básica y que vivan en Machala. Por otra parte, las características que hacen que la probabilidad de obtener un contrato temporal disminuya son: que tengan de 46 a 65 años de edad, con un nivel de instrucción superior no universitario y un empleo privado.

■ O. Otras actividades comunitarias sociales y personales de tipo servicios

Rama de actividad	Características		Influencia
O. Otras actividades comunitarias sociales y personales de tipo servicios	Rango de edad	56-65	Negativa
	Estado civil	Soltero	Positiva
		Separado/Divorciado	Positiva
	Ocupación	Empleado tercerizado	Positiva
	Ciudades autorepresentadas y regiones	Quito	Negativa
		Amazonía	Negativa
		Resto de la Sierra	Negativa
Machala		Positiva	
	Guayaquil	Positiva	
	Resto de la Costa	Positiva	

Tabla 29: Caracterización de los individuos con contrato temporal en la Rama O. Fuente: elaboración propia a partir de la encuesta ENEMDU.

En la Tabla 29 se observa que la probabilidad de obtener un contrato temporal, en esta rama de

actividad, se incrementa para las personas solteras, o separadas, o divorciadas, y que vivan en Machala, o Guayaquil, o el resto de la Costa. Por otra parte, las características que incentivan la disminución de esta probabilidad son: que tenga una edad entre 56 y 65 años y que vivan en Quito, o en Amazonía, o resto de la Sierra.

▪ **P. Hogares privados con servicio doméstico**

Rama de actividad	Características		Influencia
P. Hogares privados con servicio doméstico	Rango de edad	15-25	Positiva
		56-65	Negativa
	Estado civil	Soltero	Negativa
	Ciudades autorepresentadas y regiones	Zonas no delimitadas	Positiva
		Machala	Positiva
		Guayaquil	Positiva
		Quito	Positiva
Ambato		Positiva	
Resto de la Sierra	Positiva		
Resto de la Costa	Positiva		
Amazonía	Positiva		

Tabla 30: Caracterización de los individuos con contrato temporal en la Rama P. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Características		Rama de actividad															
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
Edad	15-25	+	+	+		+		+	+	+			+	+	+		+
	26-35	+				+			+				+	+	+		
	36-45																
	46-55			-	-			-						-	-	-	-
	56-65				-								-	-	-	-	-
Nivel de instrucción	Post-grado																
	Superior no universitaria														-		
	Educación media	+			+	+	+	+		+	+	+					
	Educación básica	+	+	+	+		+	+	+	+	+	+	+		+		
	Centro de alfabetización				+		+	+						+			
Estado civil	Ninguna	+			+		+							+			
	Viudo											+					
	Soltero				+		+	+				+	+			+	-
	Separado/Divorciado				+			+				+	+			+	
	Casado/Unión libre																
Ocupación	Empleado doméstico																
	Empleado privado	-	-	-							-				-		
	Empleado público						-										
	Empleado tercerizado				+	+	+	+		+	+	+				+	
	Jornalero o peón				+		+	+		+							
Ciudades autorepresentadas y regiones	Resto de la Sierra	+			+		+	-	-	-		-	-			-	+
	Resto de la Costa	+			+		+	+	+	+		+		+		+	+
	Guayaquil	+			+		+	+	+	-		+		+		+	+
	Quito				+			-	-	-	-	-	-	-	-	-	+
	Cuenca																
	Machala	+			+		+	+	+	+	+	+	+	+	+	+	+
	Ambato				+			-	-				-	-			+
	Amazonía	+			+		+	+	-	-		+				-	+
Zonas no delimitadas	+						+	+	+							+	

Tabla 31: Factores de riesgo y protección por rama de actividad. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Por ejemplo, para la la Rama de Actividad B tenemos que los factores de riesgo serían: ser una persona entre 15 y 25 años de edad y tener educación básica; por el contrario, el factor de protección sería: ser empleado privado. El análisis es similar para el resto de Ramas de Actividad.

En la Tabla 30 se puede apreciar que la probabilidad de disminuir la contratación temporal, en esta rama de actividad, es para los individuos entre 56 y 65 años de edad, solteros. Por otra parte, las características que permiten aumentar la probabilidad de tener un contrato temporal son: que tenga una edad entre 15 y 25 años, y que vivan en Quito, o Ambato, o Resto de la Sierra, o Machala, o Guayaquil, o Resto de la Costa, o Amazonía, o Zonas no delimitadas. Además, es necesario conocer que en esta rama se agrupan las empleadas domésticas y los empleados domésticos, y que por tal razón el establecimiento no sobrepasa en ninguno de los casos a 100 trabajadores.

Comprobación de la hipótesis (ii): *Determinar en qué medida la temporalidad depende de la rama de actividad.*

De acuerdo a la Rama de Actividad, los factores de riesgo y protección cambian, como se puede observar en la Tabla 31, denotando a (+) como el factor de riesgo y a (-) como el factor de protección.

4. Conclusiones

A partir de datos de la Encuesta de Empleo, Desempleo y Subempleo del Instituto Nacional de Estadísticas y Censos del Ecuador y de la aplicación de un modelo de regresión logística, usando la desagregación de las características personales, laborales y geográficas de los individuos, este trabajo ayuda a comprender la realidad laboral ecuatoriana, pues permite detectar las diferencias en el acceso al empleo temporal según la rama de actividad económica, usando la desagregación de las características personales, laborales y geográficas de los individuos.

Los resultados obtenidos han mostrado que la temporalidad en Ecuador no solo depende de la estructura productiva existente, sino que es un fenómeno asociado a las características de los trabajadores y a nuestra cultura empresarial. Resaltan la influencia que tiene el nivel de instrucción, el tiempo de trabajo, el tipo de jornada laboral, la capacitación al momento de obtener un contrato temporal, y el hecho de ser hombre o mujer: los hombres presentan una mayor propensión de tener un contrato temporal; esto podría explicarse por el hecho de que existen tareas sencillas y rudimentarias, como las que exigen los cultivos agrícolas, la cría de ganado, la pesca, la caza, la industria y construcción que requieran la utilización de herramientas manuales y, a menudo, un esfuerzo considerable.

Se observa que el tipo de ocupación es una variable influyente sobre la determinación del tipo de contrato que la persona tenga o pueda llegar a obtener, evidenciando que si una persona tiene una ocupación de jornalero o peón, es una condición que afecta fuertemente a la probabilidad de tener un contrato temporal. Esta situación implica un estancamiento de la situación laboral de las personas que se desenvuelven dentro de esta ocupación, logrando que no puedan tener un nivel salarial ascendente, afectando con esto en el nivel económico de las familias.

En referencia a las ciudades autorepresentadas y regiones, se obtiene que en la Región Costa, particularmente en las ciudades de Machala y Guayaquil se concentran los contratos temporales. Una posible explicación sería el desarrollo en las actividades de: agricultura, ganadería, caza y silvicultura; industrias manufactureras; hoteles y restaurantes; transporte; almacenamiento y comunicaciones. En

las actividades mencionadas se pueden presentar contratos eventuales, por obra o por horas. Generalmente son actividades que se desarrollan por temporadas, es decir, por períodos del año en los cuales tienen mayor desarrollo y surge la necesidad de incrementar el contrato temporal.

En definitiva, se ha comprobado a través de un modelo de regresión logística el conocimiento empírico del comportamiento social, dentro del mercado laboral ecuatoriano, poniendo en evidencia cómo se encuentra concentrado el contrato temporal.

Referencias

- [1] A. Acosta (2005) “Breve historia económica del Ecuador”, Corporación Editorial Nacional, Quito, pp. 202–230.
- [2] M. Aguayo (2007) “Cómo hacer una regresión logística con SPSS paso a paso (I)”, Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702012, pp. 1–16.
- [3] M. Aguayo y E. Lore (2007) “Cómo hacer una regresión logística binaria paso a paso (II): análisis multivariante”, Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702013, pp. 1–35.
- [4] M. Aguayo y L. Monge (2007) “Confusión e iteración (2): su abordaje en el análisis multivariante”, Fundación Andalucía Beturia para la Investigación en Salud, Dot. N° 0702011, p. 5.
- [5] V. Aguiar (2007) “El mercado laboral ecuatoriano: propuesta de una reforma”, edición electrónica. Texto completo en www.eumed.net/libros/2007a/240/, ISBN: 978-84-690-5504-5, pp. 5–25.
- [6] A. Alejandro, J. Pérez, R. Kizys y L. Manzanedo (2002) “Regresión Logística Binaria”, Universidad Oberta de Catalunya, Barcelona, pp. 5–16.
- [7] G. Becker (1962) “Investment in human capital: A theoretical analysis”, *Journal of Political Economy*, Volumen 70, pp. 9–49.
- [8] A. Caparrós y L. Navarro (2008) “Temporalidad, segmentación laboral y actividad productiva: ¿existen diferencias regionales?”, *Estadística Española*, Vol. 50, N° 168, pp. 206–226.
- [9] A. Caparrós, L. Navarro y M. Rueda (2004) “Efectos de la temporalidad sobre la formación recibida durante el empleo”, *Universidad de Málaga, Cuadernos de Economía*. Vol. 27, 051-074, pp. 52–73.
- [10] Codificación 17 (2005) “Codificación del Código del Trabajo,” Registro Oficial Suplemento 167, diciembre, título I, capítulo I.
- [11] E. Domínguez y D. Aldana (2001) “Regresión Logística. Un ejemplo de su uso en endocrinología”, *Rev. Cubana Endocrinol.*, Vol. 12, pp. 58–64.
- [12] C. Gallardo y H. Capa (2008) “Nueva metodología y cálculo de indicadores de la Encuesta de Empleo, Desempleo y Subempleo”, *Revista Ecuatoriana de Estadística*, Instituto Nacional de Estadísticas y Censos del Ecuador, pp. 51–53.
- [13] C. Gamero (1999) “Satisfacción laboral y tipo de contrato en España”, *Investigative Radiology* 34, Vol. 10, 636-642, pp. 421–429.

- [14] L. Hachuel, G. Boggio, D. Wojdyla y E. Servy (2005) “Interpretación y comparación de modelos de regresión logística para el estudio de la desocupación”, Décimas Jornadas Investigaciones en la Facultad de Ciencias Económicas y Estadística, Instituto de Investigaciones Teóricas y Aplicadas, pp. 1–12.
- [15] Instituto Nacional de Estadísticas y Censos del Ecuador (2010) “Estadísticas de Empleo, Mercado Laboral”. Disponible en: www.inec.gob.ec/web/guest/ecu_est/est_soc/enc_hog/enemdu.
- [16] Instituto Nacional de Estadísticas y Censos del Ecuador (2007-2010) “Estadísticas de Empleo, Mercado Laboral”. Disponible en: www.ecuadorencifras.com/cifras-inec/bancInf.html.
- [17] Instituto Nacional de Estadísticas y Censos del Ecuador (2010) “Manual del Encuestador”, Estadística de Hogares, pp. 54–88.
- [18] B. Jovanovic (1979) “Job Matching and the Theory of Turnover”, *The Journal of Political Economy*, Vol. 87, No. 5, Part 1, pp. 972–990.
- [19] E. Loh (1994) “Employment probation as a sorting mechanism”, *Industrial and Labour Relations Review*, Vol. 47, No. 3, pp. 471–486.
- [20] Mandato Constituyente N° 8 (2008) “Eliminación y prohibición de la tercerización, intermediación laboral, contratación laboral por horas y cualquier forma de precarización de las relaciones de trabajo”, Suplemento del Registro Oficial 330, 6-V-2008.
- [21] E. Medina (2003) “Modelos de elección discreta”, *Estadística de Hogares*, www.eva.medinaam.es, pp. 9–21.
- [22] Ley N° 100 (2003) “Código de la niñez y adolescencia”, Registro Oficial 737, enero, pp. 24–26.
- [23] Ley de Maquilas y contratación laboral a tiempo parcial (1990) “Código del Trabajo”, Registro Oficial 493, agosto, Título II.
- [24] Ley para la Transformación Económica del Ecuador “Trole 1” (2000) “Código del Trabajo”, Registro Oficial 33, marzo, Capítulo XII.

A. Anexo: análisis descriptivo

A.1. Características personales

En las Tablas 32, 33, 34 y 35 se encuentra la información desglosada en lo que respecta a las variables que hacen referencia a las características personales que discriminan de mejor manera la proporción de contratos temporales *vs.* contratos indefinidos. El contrato temporal se concentra en mayor medida en los hombres en relación a las mujeres, que presentan una menor propensión en la concentración del contrato temporal con una concentración del 18,59%, recalcando de igual manera que en la mayor parte del universo analizado está conformado por hombres. También se puede evidenciar que el rango de edad en el que se concentra el contrato temporal es de 15 a 35 años de edad, con un porcentaje en conjunto del 34,40%, mientras que el contrato indefinido se encuentra más en personas de una edad comprendida entre 36 a 65 años de edad, que en conjunto representan el 25,64%.

Es claro que al tener un menor nivel de instrucción (ninguna, centro de alfabetización, educación básica, educación media) las personas se concentran en una relación laboral con contrato temporal,

mientras que cuando la persona tiene nivel de instrucción mayor (superior no universitario, superior, post-grado) se agrupan en contratos indefinidos; en general, del universo analizado, el porcentaje de individuos que presentan un nivel de educación superior no universitario y de post-grado tan solo corresponde al 3,53%. Por otro lado, el contrato temporal se concentra en los individuos solteros, mientras que los individuos casados tienen mayor tendencia a tener contrato indefinido.

Sexo	Contrato indefinido		Contrato temporal		Total	
Categoría	Número	%	Número	%	Número	%
Hombre	17447	26,18	23152	34,73	40599	60,91
Mujer	13663	20,50	12393	18,59	26056	39,09
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 32: Distribución por tipo de contrato de la variable *Sexo*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Edad	Contrato indefinido		Contrato temporal		Total	
Categoría	Número	%	Número	%	Número	%
15-25	5484	8,23	13102	19,66	18586	27,88
26-35	8535	12,80	9822	14,74	18357	27,54
36-45	7565	11,35	6648	9,97	14213	21,32
46-55	6338	9,51	4062	6,09	10400	15,60
56-65	3188	4,78	1911	2,87	5099	7,65
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 33: Distribución por tipo de contrato de la variable *Edad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Nivel de instrucción	Contrato indefinido		Contrato temporal		Total	
Categoría	Número	%	Número	%	Número	%
Ninguna	157	0,24	662	0,99	819	1,23
Centro de alfabetización	44	0,07	127	0,19	171	0,26
Educación básica	7295	10,94	16611	24,92	23906	35,87
Educación media	8671	13,01	10486	15,73	19157	28,74
Superior no universitaria	672	1,01	302	0,45	974	1,46
Superior universitaria	13142	19,72	7106	10,66	20248	30,38
Post-grado	1129	1,69	251	0,38	1380	2,07
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 34: Distribución por tipo de contrato de la variable *Nivel de Instrucción*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Estado civil	Contrato indefinido		Contrato temporal		Total	
Categoría	Número	%	Número	%	Número	%
Casado/Unión libre	18690	28,04	16785	25,18	35475	53,22
Soltero	8861	13,29	14744	22,12	23605	35,41
Separado/Divorciado	3035	4,55	3533	5,30	6568	9,85
Viudo	524	0,79	483	0,72	1007	1,51
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 35: Distribución por tipo de contrato de la variable *Estado Civil*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

A.2. Características laborales

En las Tablas 36, 37, 38, 39 y 40 se encuentra la información desglosada en lo que respecta a las variables que hacen referencia a las características laborales que discriminan de mejor manera la proporción de contratos temporales *vs.* contratos indefinidos.

Dentro del análisis se puede evidenciar que tanto el contrato temporal como el indefinido se concentran en los empleados privados; es decir, la distribución de las individuos en las categorías de ocupación es proporcional tanto para el contrato temporal como para el contrato indefinido.

En lo que respecta a la antigüedad, la capacitación que recibe el empleado en las empresas y el tamaño del establecimiento, se puede destacar que el contrato temporal se concentra cuando el individuo tiene una antigüedad inferior o igual a un año, se encuentra trabajando en una empresa de menos de 100 trabajadores y no recibe ningún tipo de capacitación.

Ocupación Categoría	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
Empleado/Obrero privado	18781	28,18	21578	32,37	40359	60,55
Empleado/Obrero de gobierno	10490	15,74	2997	4,50	13487	20,23
Jornalero o peón	145	0,22	7788	11,68	7933	11,90
Empleado/a doméstico/a	1581	2,37	2832	4,25	4413	6,62
Empleado tercerizado	113	0,17	350	0,53	463	0,69
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 36: Distribución por tipo de contrato de la variable *Ocupación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Antigüedad Categoría	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
Mayor que un año	26450	39,68	19719	29,58	46169	69,27
Menor igual que un año	4660	6,99	15826	23,74	20486	30,73
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 37: Distribución por tipo de contrato de la variable *Antigüedad*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Jornada laboral Categoría	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
Jornada completa	28269	42,41	27632	41,46	55901	83,87
Jornada parcial	2841	4,26	7913	11,87	10754	16,13
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 38: Distribución por tipo de contrato de la variable *Jornada Laboral*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Tamaño del establecimiento Categoría	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
Menos de 100	14583	21,88	28469	42,71	43052	64,59
100 y mas	16527	24,79	7076	10,62	23603	35,41
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 39: Distribución por tipo de contrato de la variable *Tamaño del Establecimiento*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

Capacitación Categoría	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
No	18700	28,05	32238	48,37	50938	76,42
Sí	12410	18,62	3307	4,96	15717	23,58
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 40: Distribución por tipo de contrato de la variable *Capacitación*. Fuente: elaboración propia a partir de la encuesta ENEMDU.

A.3. Características geográficas

En la Tabla 41 se encuentra la información desglosada en lo que respecta a la variable que hace referencia a residencia que discriminan de mejor manera la proporción de contratos temporales *vs.* contratos indefinidos.

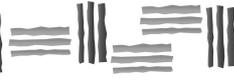
Se puede notar que la región Resto de la Costa presenta una mayor agrupación del contrato temporal, mientras que la región Resto de la Sierra presenta una distribución similar entre contratos temporales e indefinidos.

Ciudades autorepresentadas y regiones	Contrato indefinido		Contrato temporal		Total	
	Número	%	Número	%	Número	%
Resto de la Sierra	8449	12,68	8623	12,94	17072	25,61
Resto de la Costa	5307	7,96	10240	15,36	15547	23,32
Guayaquil	3583	5,38	5247	7,87	8830	13,25
Quito	4142	6,21	2715	4,07	6857	10,29
Cuenca	3526	5,29	1531	2,30	5057	7,59
Machala	1606	2,41	3266	4,90	4872	7,31
Ambato	2754	4,13	1711	2,57	4465	6,70
Amazonía	1671	2,51	1948	2,92	3619	5,43
Zonas no delimitadas	72	0,11	264	0,40	336	0,50
Total	31110	46,67	35545	53,33	66655	100,00

Tabla 41: Distribución por tipo de contrato de la variable *Ciudades Autorepresentadas y Regiones*. Fuente: elaboración propia a partir de la encuesta ENEMDU.



UNIVERSIDAD
PABLO DE OLAVIDE
SEVILLA



REVISTA DE MÉTODOS CUANTITATIVOS PARA
LA ECONOMÍA Y LA EMPRESA (13). Páginas 54-72.
Junio de 2012. ISSN: 1886-516X. D.L: SE-2927-06.
URL: <http://www.upo.es/RevMetCuant/art.php?id=58>

Relación entre los factores institucionales y el emprendimiento: análisis mediante técnicas cuantitativas

SALMERÓN GÓMEZ, ROMÁN

Departamento de Métodos Cuantitativos para la Economía y la Empresa
Universidad de Granada (España)

Correo electrónico: romansg@ugr.es

GÓMEZ HARO, SAMUEL

Departamento de Organización de Empresas
Universidad de Granada (España)

Correo electrónico: samugh@ugr.es

RESUMEN

La mayoría de los trabajos publicados sobre la relación entre los factores institucionales y emprendedores adolecen de un adecuado estudio empírico. Debido a la importancia que ha adquirido en los últimos años el entorno institucional, consideramos necesario un análisis riguroso mediante técnicas estadísticas cuantitativas. El objetivo del presente trabajo es el análisis empírico de la relación entre el entorno institucional y la orientación emprendedora de la empresa española, partiendo de la identificación teórica y empírica de las dimensiones reguladora, cognitiva y normativa para el entorno institucional y las dimensiones de innovación, asunción de riesgos, proactividad y autonomía para analizar el comportamiento emprendedor.

Palabras clave: actividad emprendedora; entorno institucional; análisis factorial; regresión.

Clasificación JEL: C13; C38.

MSC2010: 62H25; 62J05; 62J12.

Relationship between Institutional Factors and Entrepreneurship: A Quantitative Analysis

ABSTRACT

The majority of published studies about the relation between institutional and entrepreneurial factors lack of an adequate empirical study. Due to the importance that the institutional environment has acquired in the last few years, we consider necessary a rigorous analysis using quantitative statistical techniques. The aim of this paper is the empirical analysis of the relationship between institutional and entrepreneurial orientation Spanish company, based on theoretical and empirical identification of regulatory, cognitive and normative dimensions for the institutional environment and the dimensions of innovation, risk assumption, proactivity, and autonomy to analyze the entrepreneurial behavior.

Keywords: entrepreneurial activity; institutional environment; factorial analysis; regression.

JEL classification: C13; C38.

MSC2010: 62H25; 62J05; 62J12.



1. INTRODUCCIÓN

El emprendimiento es un aspecto vital para el crecimiento, la rentabilidad y la supervivencia de las organizaciones, siendo considerado el motor de la economía moderna y del desarrollo social a través del crecimiento económico, la generación de empleo y el fomento de la innovación (Bosma *et al.*, 2009; Bosma y Levie, 2010). Las organizaciones se orientan hacia el emprendimiento a partir de una serie de procesos, prácticas y actividades que posibilitan dicha actividad emprendedora, medido generalmente en términos de proactividad, innovación, asunción de riesgos, agresividad competitiva y autonomía (Dess y Lumpkin 2005; Lumpkin y Dess, 1996).

En el desarrollo de la orientación emprendedora en las organizaciones y sus miembros influyen diferentes factores (Covin y Slevin, 1991; Knight, 1997), donde el entorno institucional ha sido generalmente excluido de este análisis. La importancia que adquiere en tiempos recientes el entorno institucional muestra la oportunidad de analizar cómo éste presiona e influye en las organizaciones para que adopten posturas más emprendedoras (Bruton *et al.*, 2010; Busenitz *et al.*, 2000). Dicho entorno se define en base a la existencia de una serie de reglas y requisitos a los que deben someterse las organizaciones si desean recibir apoyo y legitimidad (Scott y Meyer, 1991) e influye en la gestión organizativa, ya que proporciona una estabilidad y significado al comportamiento de las organizaciones y las decisiones que adoptan (Scott, 1995).

Los trabajos de investigación que relacionan entorno institucional y emprendimiento son recientes (Lim *et al.*, 2010; Urbano-Pulido *et al.*, 2007; Vaillant y Lafuente, 2007) y necesitan un mayor contraste empírico que facilite la creación de modelos explicativos entre este entorno y las manifestaciones del emprendimiento (Bruton *et al.*, 2010; Spencer y Gómez, 2004). Con la intención de un empleo formalizado de herramientas estadísticas que aporten una mayor evidencia empírica que en los estudios encontrados hasta la fecha, este trabajo pretende analizar cómo el entorno institucional, a través de sus dimensiones reguladoras, normativas y cognitivas, influyen en la orientación emprendedora de las empresas.

El artículo está estructurado de la siguiente forma: en la sección 2 se realiza un resumen sobre el concepto de orientación emprendedora de la organización, mientras que en la sección 3 se hace sobre la influencia que ejerce el entorno institucional. En la sección 4 se identifica, mediante el análisis factorial, algunos de los conceptos introducidos en la revisión teórica anterior, a partir de los cuales se plantea un modelo lineal que los relacione. Finalmente, en la sección 5 se presentan las conclusiones obtenidas.

2. EL PAPEL CLAVE DEL EMPRENDIMIENTO

La orientación emprendedora se define como un conjunto de procesos, prácticas y actividades relacionadas con la toma de decisiones que posibilitan la actividad emprendedora (Dess y Lumpkin, 2005; Covin y Slevin, 1991; Lumpkin y Dess, 1996; Miller, 1983), siendo un instrumento fundamental para el logro de éxito empresarial en términos de crecimiento, rentabilidad y creación de empleo (Bosma y Levie, 2010).

Dichas investigaciones han analizado de forma amplia que la orientación emprendedora de las empresas está relacionada de forma positiva con la rentabilidad y el crecimiento de la empresa (Covin, Green y Slevin, 2006; Covin y Slevin, 1991; Ireland, Covin y Kuratko, 2009; Wiklund y Shepherd, 2005), de ahí el papel fundamental que juega en el contexto de investigación sobre estrategia y éxito empresarial. Las primeras investigaciones sobre orientación emprendedora apuntan que las empresas emprendedoras tienden a asumir mayores riesgos que el resto de empresas, especialmente cuando se encuentran en situaciones de incertidumbre (Khandwalla, 1977; Mintzberg, 1973). Posteriormente, los investigadores operativizan el comportamiento emprendedor como el conjunto de innovaciones de producto-mercado, proactividad en la toma de decisiones y la capacidad de asumir riesgos (Miller, 1983; Miller y Friesen, 1978). El desarrollo de nuevos modelos de orientación emprendedora añade nuevas variables como la autonomía o la agresividad competitiva (Lumpkin y Dess, 1996). Así, se conceptualiza la orientación emprendedora como un constructo con varias dimensiones (Dess y Lumpkin, 2005):

Autonomía: se refiere a la acción independiente de un individuo o grupo para tener una idea o visión y llevarla hasta su desarrollo. Puede variar el grado en función del tamaño y el estilo de gestión de la organización. Es la dimensión que menor atención ha recibido en la literatura, aunque empiezan a aparecer trabajos teóricos y empíricos que justifican su importancia (Dess y Lumpkin, 2005; Lumpkin, Cogliser y Schneider, 2009).

Innovación: es la capacidad para innovar. Refleja la tendencia de una organización al apoyo de nuevas ideas, novedosas, experimentales y creativas que puedan resultar en nuevos productos, servicios o procesos tecnológicos. Ser innovador tiene muchos grados dentro de un continuo, desde la innovación ocasional hasta un compromiso por convertirlo en la piedra angular de la organización (Lumpkin y Dess, 1996).

Asunción de riesgos: es la voluntad de la empresa para emprender estrategias o negocios cuyo resultado está sometido a un nivel de incertidumbre (Miller, 1983). El riesgo implica comprometer recursos donde no existen resultados ciertos y los costes de un posible fracaso son altos (Wiklund y Shepherd, 2003).

Proactividad: refleja la postura de anticiparse al actuar sobre futuras carencias y necesidades del mercado, creando ventaja competitiva por ser el primero en actuar. Hace referencia a la libertad de experimentación, tanto a nivel individual como organizativo (Stopford y Baden-Fuller, 1994).

Agresividad competitiva: es la propensión de la empresa a retar a sus competidores para obtener una entrada o una mejora de posición, es decir, mejorar respecto a los rivales en el mercado (Dess y Lumpkin, 2005).

En el desarrollo de la orientación emprendedora en las organizaciones y sus miembros influyen diferentes factores, entre los que el entorno tiene un papel clave (Covin y Slevin, 1991; Kreiser y Justin, 2010; Zahra, 1993). Normalmente el análisis del entorno se ha realizado a través de las dimensiones más clásicas (hostilidad, dinamismo) obviando la existencia de un entorno institucional que influye igualmente en la gestión organizativa y en su nivel de emprendimiento (Baumol *et al.*, 2009; Bruton *et al.*, 2010). No obstante, el peso de las instituciones nacionales

y comunitarias en este aspecto es cada vez mayor, ya que existe una fuerte tendencia a diseñar por parte de las entidades públicas medidas para incentivar el emprendimiento, la mejora del clima emprendedor y la creación de una sociedad más innovadora y creativa que aproveche las oportunidades existentes en el mercado (European Commission, 2003, 2004). La importancia que adquiere en tiempos recientes el entorno institucional muestra la oportunidad de analizar cómo éste presiona e influye en las organizaciones para que adopten posturas más o menos emprendedoras. Por tanto, el análisis desde una perspectiva institucional es fundamental y necesario para la mejor comprensión del fenómeno emprendedor en sí mismo (Baumol *et al.*, 2009; Bruton *et al.*, 2010; Spencer y Gómez, 2004).

3. LA INFLUENCIA DE LOS FACTORES INSTITUCIONALES EN LA ACTIVIDAD ORGANIZATIVA

La teoría institucional se centra en los aspectos del contexto en el que están inmersas las organizaciones, haciendo hincapié en el conjunto de valores, normas o creencias que actúan como mitos racionales guiando el comportamiento de las organizaciones (Meyer y Rowan, 1977). Esta teoría es útil para el entendimiento de las relaciones sociales que establece la organización con otras organizaciones y con las instituciones con las que coexiste, señalando como factores de éxito la conformidad con las reglas y normas institucionales (DiMaggio y Powell, 1983; Meyer y Rowan, 1977). Asimismo, la teoría institucional ha demostrado ser un marco teórico idóneo para el análisis organizativo (DiMaggio y Powell, 1991).

Scott (1995) analiza el entorno institucional en base a tres dimensiones que proporcionan estabilidad y significado al comportamiento social. Kostova (1999) interpreta las dimensiones institucionales de Scott y plantea a través del concepto *Country Institutional Profile* (CIP) la relación de las diferentes dimensiones con la gestión organizativa. Estas tres dimensiones son:

Reguladora: las leyes, regulaciones y políticas gubernamentales que proveen apoyo a los nuevos negocios, reducen el riesgo para las personas que empiezan una nueva empresa y facilitan los esfuerzos de los emprendedores para adquirir recursos (Busenitz *et al.*, 2000).

Cognitiva: se refiere al conocimiento y a las habilidades que poseen los habitantes de un país determinado respecto a aspectos relacionados con la gestión empresarial. Determinados temas y conocimientos se pueden institucionalizar convirtiéndose en un conocimiento social compartido, como por ejemplo el nivel de conocimiento requerido para iniciar y gestionar una empresa (Busenitz *et al.*, 2000).

Normativa: se refiere al grado en el que los habitantes de un país determinado valoran tanto la actividad emprendedora como la mentalidad creativa e innovadora de las personas (Spencer y Gómez, 2004).

En base a lo expuesto, desde un punto de vista institucional, un mayor desarrollo del emprendimiento pasa por la mejora de determinados elementos estructurales, institucionalizando un entorno donde la capacidad para asumir riesgos, la innovación e iniciativa en la gestión

organizativa sean elementos estructurales comunes en todas las organizaciones del campo. En base a esta premisa, se considera que la existencia de un entorno institucional que fomente el emprendimiento influye positivamente en la capacidad emprendedora de sus organizaciones y ciudadanos.

Los trabajos de investigación que relacionan entorno institucional y emprendimiento han atraído mucha atención en los tiempos recientes (Amine y Staub, 2009; Lim *et al.*, 2010; Vaillant y Lafuente, 2007). Sin embargo, los investigadores reclaman mayor evidencia empírica que facilite la creación de modelos explicativos entre las distintas dimensiones institucionales y manifestaciones del emprendimiento (Bruton *et al.*, 2010; Spencer y Gómez, 2004). Esta necesidad de profundizar en el tema se debe a que la mayoría de los análisis sobre la influencia del entorno institucional en el emprendimiento han sido descriptivos y fragmentados (Stephen *et al.*, 2005; Urbano *et al.*, 2007).

La actual situación económica demanda empresas con iniciativas y creatividad; en el caso de España es más acentuada esta necesidad. Actualmente, España cuenta con una tasa de desempleo del 20,5 % frente al 8,5 % de la media de los países OECD (OECD Harmonised Unemployment Rates, 2010). Aunque muchos factores influyen en esta situación, diferentes expertos y organismos nacionales vienen subrayando los problemas estructurales que presenta el país relacionados con la debilidad de la capacidad de emprendimiento, situándola en las últimas posiciones en las clasificaciones elaboradas por la Comisión Europea sobre los índices de actividad y clima emprendedor en los países europeos (De la Vega *et al.*, 2007, 2008, 2009). Esta situación es problemática porque la orientación emprendedora es la forma de dotar a las economías desarrolladas de su entorno de una mayor capacidad competitiva (European Commission, 2003, 2004). Concretamente, en el presente trabajo se analiza cómo el entorno institucional está afectando al comportamiento emprendedor de las empresas españolas, en base a su orientación emprendedora.

4. ANÁLISIS EMPÍRICO

Con el objetivo de confirmar empíricamente la hipótesis de que el entorno institucional influye en la actividad emprendedora, a continuación se va a realizar un análisis estadístico de los datos obtenidos mediante un cuestionario. Los pasos a seguir son los siguientes: descripción de la muestra, análisis factorial en busca de factores reconocibles a partir de la literatura tradicional y análisis de regresión para analizar la influencia de unos factores a otros.

4.1. Cuestionario y muestra considerada

Como población objetivo se han considerado empresas dentro del ámbito geográfico español, lo cual asegura un marco político, económico, sociocultural, tecnológico y legal homogéneo entre las empresas analizadas en el trabajo, como sugiere la literatura sobre el tema (Busenitz *et al.*, 2000). En el trabajo, se han usado a los directores generales como fuente clave de datos, debido a que reciben información de un amplio rango de unidades de la empresa y son una fuente valiosa para evaluar cada una de las variables de la empresa (Baer y Frese, 2003). Los datos generales

de las empresas fueron obtenidos de la base de datos DUNS¹. Esta base de datos ofrece uno de los directorios más detallados de empresas del país, incluyendo algunos datos específicos como la muestra y el tamaño. Se realizó un muestreo aleatorio simple de 1455 empresas con las que se tomaría contacto de forma telefónica para realizar el cuestionario. El uso de la entrevista telefónica es especialmente útil para estar seguros de la identidad de la persona que responde el cuestionario. Finalmente, se obtuvieron 150 cuestionarios respondidos de forma completa, teniendo una tasa de respuesta del 10,30 % que, comparándola con los datos de la literatura de gestión empresarial, es una tasa de respuesta adecuada.

A pesar de que el uso de respuestas de una única persona es una efectiva aproximación en diversos contextos de investigación (Liao, 2007; Ling *et al.*, 2008), se han tenido en cuenta la existencia de posibles prejuicios que deben ser evitados. Para ello, se han seguido las recomendaciones de Podsakoff *et al.* (2003). En primer lugar, había que asegurarse de que los datos del entrevistado eran anónimos y que no existían respuestas erróneas en el cuestionario, por lo que se solicitaba la máxima honestidad posible; estos detalles sirven para reducir cualquier temor del entrevistado respecto a responder aquella respuesta que cree que es la esperada. Y en segundo lugar, hay que evitar ambigüedades en la redacción de las preguntas buscando una máxima claridad y precisión; si había términos que podían ser menos conocidos por los entrevistados se les daba todas las aclaraciones necesarias.

Respecto a los datos de la muestra, los sectores más representados entre las empresas que respondieron fueron: comerciales, metalúrgicas y maquinaria, y transporte, con 13 %, 10 % y 9 %, respectivamente. Respecto a su tamaño, un 5 % de las empresas que respondieron tenían menos de 50 trabajadores, un 16 % entre 50 y 250 y el resto (79 %) más de 250 trabajadores. Finalmente, respecto a las ventas, un 26 % tuvieron un volumen de ventas inferior al millón de euros en 2009, mientras que el resto tuvo un volumen superior.

Por otro lado, resulta factible considerar un posible sesgo de la muestra que no responde; es decir, los que deciden no responder pueden ser significativamente diferentes de los que sí lo hacen. Esto podría suponer un problema para extrapolar los resultados de la muestra de análisis a toda la población. Para determinar si esto ocurría se ha realizado un análisis de la varianza en las variables que lo permitían, no encontrándose diferencias significativas. De esta forma se puede afirmar que las empresas que no respondieron al cuestionario no introducen un sesgo significativo en los resultados finales del estudio, no existiendo un problema para extrapolar los resultados a la población total.

El cuestionario (disponible en Gómez, 2010) está formado por preguntas cerradas, ya que al tratarse de una encuesta telefónica, éstas garantizan una mayor tasa de respuesta. Además, todas se miden con escalas de tipo Likert de siete puntos, lo que permite el uso de procedimientos estadísticos que se pueden aplicar a escalas de intervalos. Dichas preguntas son adaptaciones de estudios científicos anteriores, recogidas en una secuencia lógica que facilita la contestación. Por dicho motivo, se divide el cuestionario en tres secciones. La primera sección busca obtener información sobre el contexto institucional de la organización, la segunda recoge información respecto

¹ *Duns 50.000 Principales Empresas Españolas*, es un directorio de empresas que recoge información comercial y de marketing sobre las 75.000 mejores compañías del país en función de su cifra de ventas. Más información sobre dicha base de datos está disponible en <http://www.duns50000.com/servlet/app>.

a la actividad emprendedora de la organización analizando la proactividad, innovación, renovación estratégica y creación de nuevos negocios. En la tercera sección se plantean las cuestiones relacionadas con el desempeño. Las preguntas consideradas² son las siguientes:

VAR1: las instituciones públicas de este país ayudan a las empresas a realizar nuevas actividades empresariales.

VAR2: los organismos gubernamentales ofrecen fondos públicos suficientes para nuevas actividades de negocio de las empresas.

VAR3: las instituciones nacional y local apoyan preferentemente a las empresas que quieren realizar nuevas actividades empresariales.

VAR4: incluso después de fracasar en la actividad empresarial, las instituciones estimulan a las empresas que intentan poner en marcha nuevas actividades empresariales.

VAR5: las empresas saben cómo hacer frente a situaciones de riesgo.

VAR6: las empresas saben cómo gestionar el riesgo.

VAR7: las empresas saben dónde encontrar información para el desarrollo de nuevas actividades empresariales.

VAR8: en este país, el pensamiento innovador y creativo es visto como una ruta hacia el éxito.

VAR9: los emprendedores son personas admiradas en nuestro país.

VAR10: la sociedad admira a aquellas organizaciones que son emprendedoras.

VAR11: en general, la dirección de la empresa es propensa a aceptar proyectos de elevado riesgo.

VAR12: cuando la empresa se enfrenta a situaciones de toma de decisiones que implican incertidumbre, normalmente se adopta una postura atrevida y agresiva que permita maximizar la probabilidad de explotar el potencial de oportunidades.

VAR13: en general, la dirección de su empresa cree que, dada la naturaleza del entorno, es necesario actuar con valentía y atrevimiento para lograr los objetivos de la organización.

VAR14: en general, nuestra empresa suele tomar la iniciativa en los procesos de negociación.

VAR15: comparado con los competidores, es frecuente que su empresa sea la primera en introducir nuevas herramientas en su gestión (productos/servicios, técnicas administrativas, tecnologías, etc.).

VAR16: comparado con nuestros competidores, la empresa adopta una postura de mercado muy competitiva.

VAR17: el énfasis de la organización en el desarrollo de nuevos productos o servicios.

VAR18: el ratio de introducción de nuevos productos o servicios en el mercado.

VAR19: el gasto de la organización en actividades de desarrollo de nuevos productos o servicios.

²Conviene destacar que el cuestionario completo consta de más variables que las enumeradas aquí. La eliminación de cuestiones responde bien a que no son relevantes en este estudio o bien a que su baja medida de suficiencia en el análisis factorial realizado aconsejaba eliminarlas.

- VAR20:** el número de nuevos productos introducidos por la organización ya existentes en el mercado
- VAR21:** el número de nuevos productos introducidos por primera vez en el mercado por la organización.
- VAR22:** el porcentaje de ingresos de la organización generados por productos que no existían hace tres años.
- VAR23:** el número de cambios drásticos en las líneas de productos o servicios.
- VAR24:** el énfasis de la dirección en I+D, liderazgo tecnológico e innovación.
- VAR25:** mi empresa apoya los esfuerzos de los individuos y/o equipos que trabajan autónomamente en comparación con aquellos individuos y/o equipos que confían en que los directivos guíen su trabajo.
- VAR26:** los directivos de la empresa creen que los mejores resultados ocurren cuando los individuos y/o equipos deciden por sí mismos qué oportunidades de negocio se deben realizar (mejores que cuando la dirección general y los altos directivos son quienes impulsan dicha búsqueda de oportunidades).
- VAR27:** en mi empresa, los individuos y/o equipos que buscan oportunidades de negocio toman decisiones propias sin tener constantemente que consultar a sus supervisores (en vez de tener que obtener aprobación para de sus supervisores antes de tomar decisiones).

Las preguntas anteriores están precedidas por las instrucciones para su cumplimiento, las cuales eran reproducidas telefónicamente, mientras que se finaliza con una serie de cuestiones de control como denominación social, sector, beneficio neto en el último año y gasto en I+D.

Finalmente, calculado el error muestral de cada variable a partir de la expresión: $1,96 \cdot \sqrt{\frac{S^2}{n}}$, donde S es la varianza, se obtiene que el error muestral mínimo es del 22,26 % y el máximo es de 27,4 %.

4.2. Análisis factorial

En el análisis factorial se supone que cuando entre varios fenómenos hay interrelaciones éstas se deben a factores de un mismo fenómeno subyacente no medible directamente, siendo el objetivo de esta técnica localizar dicho fenómeno, denominado factor. De esta forma, se obtiene una estructura más simple, con menos dimensiones y que proporciona la misma información. Por tanto, el primer paso a realizar para comprobar si es posible la realización de un análisis factorial es estudiar el grado de correlación existente entre las variables. Si existe un alto grado de correlación éste puede deberse a la existencia de variables no observables directamente por el modelo, es decir, por factores.

El grado de correlación viene marcado por la matriz de correlaciones de todas las variables incluidas en el análisis. Puesto que se dispone de 27 variables, dicha matriz contendrá 351 elementos distintos. Para evitar comprobar la significatividad de cada una de las correlaciones indicadas, la comprobación del grado de correlación entre las variables se realiza a través de la

Variable	Inicial	Final	Variable	Inicial	Final	Variable	Inicial	Final
VAR1	1,000	0,731	VAR11	1,000	0,817	VAR21	1,000	0,794
VAR2	1,000	0,730	VAR12	1,000	0,869	VAR22	1,000	0,648
VAR3	1,000	0,791	VAR13	1,000	0,854	VAR23	1,000	0,635
VAR4	1,000	0,768	VAR14	1,000	0,729	VAR24	1,000	0,657
VAR5	1,000	0,648	VAR15	1,000	0,755	VAR25	1,000	0,674
VAR6	1,000	0,817	VAR16	1,000	0,769	VAR26	1,000	0,777
VAR7	1,000	0,795	VAR17	1,000	0,698	VAR27	1,000	0,664
VAR8	1,000	0,803	VAR18	1,000	0,788			
VAR9	1,000	0,869	VAR19	1,000	0,830			
VAR10	1,000	0,808	VAR20	1,000	0,754			

Tabla 1: Comunalidades

Componente	Total	% de Varianza	% Acumulado
1	9,262	34,303	34,303
2	3,433	12,713	47,016
3	2,251	8,336	55,352
4	1,816	6,725	62,078
5	1,581	5,854	67,932
6	1,276	4,726	72,658
7	0,854	3,163	75,822

Tabla 2: Varianza total explicada

prueba de esfericidad de Bartlett y el índice KMO de medida de suficiencia en el muestreo. La primera prueba contrasta la hipótesis nula de si la matriz de correlaciones es la matriz identidad, ya que en caso afirmativo esto supondría ausencia de correlación entre las variables. Por otro lado, el índice KMO compara las magnitudes de los coeficientes de correlación observados con los coeficientes de correlación parcial. Un índice KMO bajo indicaría entonces que la intercorrelación entre las variables no es grande.

En el presente estudio se obtiene un p-valor menor que $0,0005 < 0,05$ para la prueba de Bartlett, por lo que se rechaza que la matriz de correlaciones sea la identidad, y un índice KMO de 0,869, muy próximo a 1. Estos dos indicadores informan de que todas las variables consideradas son relevantes y de la existencia de correlación suficiente entre las mismas para realizar el análisis factorial. En la Tabla 1 se muestran las comunalidades iniciales (siempre iguales a 1) y finales. Dichas comunalidades expresan el porcentaje de varianza que se ha explicado de cada una de las variables mediante el análisis factorial. Como se observa, al estar todas por encima de 0,6 se consideran satisfactorias.

La Tabla 2 proporciona información sobre los factores. Realizando una extracción de factores con autorvalores mayores que 1 se obtiene una estructura factorial a partir de 6 factores, si bien a partir del gráfico de sedimentación (Figura 1) parece más adecuado seleccionar 7 factores. Así, por ejemplo, el primero explica un 34,303 % del total de la información (varianza) o el segundo

Gráfico de sedimentación

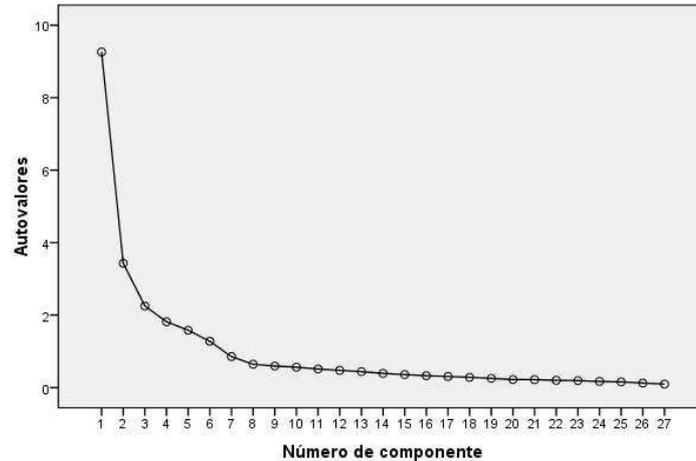


Figura 1: Gráfico de sedimentación

un 12,713 %, para un total del 75,822 % entre los 7 factores. Es decir, con los 7 factores se consigue explicar un 75,822 % de la información contenida en las variables.

Para la estructura factorial considerada, en la Tabla 3 se muestra la matriz de componentes rotados por el método Promax³, que es una rotación oblicua que permite que los factores estén correlacionados. Dicha rotación será clave para la comprensión de los factores sin más que tener en cuenta que una variable debe tener saturaciones elevadas en un solo factor (en la Tabla 3 se muestran solo cargas factoriales superiores a 0,5). Siguiendo estas premisas, el primer componente está claramente relacionado con las variables **VAR17** a **VAR24**, el segundo con **VAR1** a **VAR4**, el tercero con **VAR8** a **VAR10**, el cuarto con **VAR11** a **VAR13**, el quinto con **VAR5** a **VAR7**, el sexto con **VAR25** a **VAR27** y el séptimo con **VAR14** a **VAR16**.

Teniendo en cuenta las preguntas del cuestionario enumeradas en la sección anterior y lo expuesto en las secciones 2 y 3, es claro que:

- El factor 1 está relacionado con la innovación en la empresa.
- El factor 2 está relacionado con la dimensión reguladora del entorno institucional.
- El factor 3 está relacionado con la dimensión normativa del entorno institucional.
- El factor 4 está relacionado con el riesgo en la empresa.
- El factor 5 está relacionado con la dimensión cognitiva del entorno institucional.
- El factor 6 está relacionado con la autonomía de la empresa.
- El factor 7 está relacionado con la proactividad de la empresa.

³Adviértase que se han considerado otros tipos de rotación obteniendo la misma estructura, por lo que los factores obtenidos se pueden considerar estables.

Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
VAR1		0,851					
VAR2		0,844					
VAR3		0,849					
VAR4		0,840					
VAR5					0,777		
VAR6					0,874		
VAR7					0,846		
VAR8			0,845				
VAR9			0,923				
VAR10			0,918				
VAR11				0,877			
VAR12				0,914			
VAR13				0,852			
VAR14							0,735
VAR15							0,824
VAR16							0,727
VAR17	0,689						
VAR18	0,915						
VAR19	0,906						
VAR20	0,905						
VAR21	0,885						
VAR22	0,687						
VAR23	0,634						
VAR24	0,630						
VAR25						0,779	
VAR26						0,892	
VAR27						0,803	

Tabla 3: Matriz de componentes rotados

	Ecuación 1	Ecuación 2	Ecuación 3	Ecuación 4
p-valor	$1,45061 \cdot 10^{-5}$	$2,18409 \cdot 10^{-9}$	$7,82879 \cdot 10^{-9}$	$2,8826 \cdot 10^{-11}$

Tabla 4: Contraste de heteroscedasticidad de White

4.3. Modelo de regresión

A partir de las puntuaciones factoriales del análisis anterior, se consideran las siguientes variables: **FAC1** para el factor 1, **FAC2** para el factor 2, **FAC3**, para el factor 3, **FAC4** para el factor 4, **FAC5** para el factor 5, **FAC6** para el factor 6 y **FAC7** para el 7. De manera que para analizar la influencia del entorno institucional en el comportamiento emprendedor se plantea el siguiente análisis de regresión:

$$FAC1 = \beta_1 \cdot FAC2 + \beta_2 \cdot FAC3 + \beta_3 \cdot FAC5 + u_1, \quad (1)$$

$$FAC4 = \alpha_1 \cdot FAC2 + \alpha_2 \cdot FAC3 + \alpha_3 \cdot FAC5 + u_2, \quad (2)$$

$$FAC6 = \gamma_1 \cdot FAC2 + \gamma_2 \cdot FAC3 + \gamma_3 \cdot FAC5 + u_3, \quad (3)$$

$$FAC7 = \delta_1 \cdot FAC2 + \delta_2 \cdot FAC3 + \delta_3 \cdot FAC5 + u_4, \quad (4)$$

donde u_i , $i = 1, 2, 3, 4$, son perturbaciones aleatorias que cumplen los supuestos básicos de media cero, homocedasticidad e incorrelación.

En un primer acercamiento a los modelos anteriores, se comprueba la existencia de heteroscedasticidad en cada uno de ellos. En la Tabla 4 se muestran los p-valores de cada ecuación para el contraste de heterocedasticidad de White. Puesto que en todos los casos dicho p-valor es inferior a 0,05, se rechaza la hipótesis nula de que no hay heteroscedasticidad.

Por tanto, se procede a estimar dichas ecuaciones múltiples corrigiendo la heteroscedasticidad, obteniéndose los siguientes resultados:

$$FAC1 = 0,05065 \cdot FAC2 + 0,21705 \cdot FAC3 + 0,08465 \cdot FAC5, \quad R^2 = 0,13579$$

(0,3722) (0,000021) (0,0859)

$$FAC4 = 0,05879 \cdot FAC2 + 0,47321 \cdot FAC3 - 0,17398 \cdot FAC5, \quad R^2 = 0,24191$$

(0,3849) (1,69 · 10⁻¹⁴) (0,0138)

$$FAC6 = -0,06959 \cdot FAC2 + 0,13183 \cdot FAC3 + 0,39166 \cdot FAC5, \quad R^2 = 0,26164$$

(0,1894) (0,0885) (2,59 · 10⁻¹¹)

$$FAC7 = 0,23819 \cdot FAC2 + 0,27817 \cdot FAC3 + 0,12088 \cdot FAC5, \quad R^2 = 0,27057$$

(0,0008) (1,97 · 10⁻⁶) (0,0544)

Una vez realizada la estimación y debido a que la rotación oblicua realizada en el análisis factorial puede conducir a correlaciones entre las variables explicativas de los modelos, es conveniente estudiar la posible aparición de multicolinealidad entre dichas variables. Con tal motivo se

	FAC2	FAC3	FAC5
FAV	1,076	1,1	1,071

Tabla 5: Factor de agrandamiento de la varianza

	Ecuación 1	Ecuación 2	Ecuación 3	Ecuación 4
p-valor	$9,4 \cdot 10^{-5}$	$9,24 \cdot 10^{-9}$	$1,42 \cdot 10^{-9}$	$5,6 \cdot 10^{-10}$

Tabla 6: Contraste de significación conjunta

ha calculado el factor de agrandamiento de la varianza, FAV, de cada regresor (ver la Tabla 5). Valores mayores que 10 pueden indicar un problema de colinealidad, siendo el 1 el mínimo valor posible. Como se observa, en este caso se descarta la posibilidad de que exista multicolinealidad entre las variables explicativas de los modelos.

Por otro lado, aunque se obtiene un coeficiente de determinación no muy alto, la prueba ANOVA (ver la Tabla 6) indica que las variables independientes realmente influyen en las dependientes, ya que en todos los casos el p-valor es menor que 0,05 y entonces se rechaza la hipótesis nula de que todos los coeficientes sean cero de forma simultánea. En tal caso, teniendo en cuenta los contrastes de significación individual, se obtiene, a un nivel de significación del 5%, que:

- El factor regulador, **FAC2**, influye positivamente en el factor ligado a la proactividad, **FAC7**.
- El factor normativo, **FAC3**, influye positivamente en el factor ligado a la innovación, **FAC1**, en el factor riesgo, **FAC4**, y en el factor ligado a la proactividad, **FAC7**.
- El factor cognitivo, **FAC5**, influye negativamente en el factor riesgo, **FAC4**, mientras que lo hace de forma positiva en el factor relacionado con la autonomía, **FAC6**.

Es decir, actualmente, dentro del entorno geográfico español, la dimensión reguladora está favoreciendo los comportamientos proactivos. Esto es, está favoreciendo de forma positiva la iniciativa de las organizaciones a través de los distintos programas de apoyo, tanto estatales como autonómicos. Mientras que el entorno normativo, basado en cómo la sociedad valora y percibe la creatividad y la innovación, influye de forma positiva en el comportamiento emprendedor de las organizaciones en sus dimensiones de innovación, riesgo y proactividad. Por último, la dimensión cognitiva, basada en el grado de conocimientos comportados acerca de la empresa, tiene una influencia negativa en el nivel de riesgo y positiva en el nivel de autonomía de las personas y grupos de la organización. A continuación, en las conclusiones, se desarrollan cada uno de estos puntos.

Finalmente, cabe destacar que en todos los casos se cumple el supuesto de normalidad en los residuos (ver la Tabla 7). El p-valor es mayor que 0,05 en todos los casos, por lo que no se rechaza la hipótesis nula de que el error se distribuye normalmente.

	Ecuación 1	Ecuación 2	Ecuación 3	Ecuación 4
p-valor	0,1986	0,42041	0,41063	0,46375

Tabla 7: Contraste de normalidad de los residuos

5. CONCLUSIONES

El emprendimiento ha demostrado ser una de las opciones más viables para el desarrollo económico y social de los últimos tiempos, por lo que cada vez recibe una mayor atención por parte de los investigadores que analizan su funcionamiento, sus consecuencias y la forma de estimularlo para un desarrollo efectivo. La teoría institucional sirve de sustento para tratar de desarrollar cómo las empresas que se encuentran en un marco económico determinado pueden estimular su comportamiento emprendedor, ya que el entorno institucional provee las reglas a seguir por las organizaciones para poder obtener legitimidad en sus acciones.

La contribución principal del presente trabajo es la demostración empírica de cómo el entorno institucional influye actualmente en el nivel de emprendimiento de la empresa española, medido a través de su capacidad de innovación, proactividad y asunción de riesgos (orientación emprendedora). Hasta la fecha los informes y trabajos que analizaban el emprendimiento desde una perspectiva institucional eran fragmentados o descriptivos en sus análisis, por lo que era necesario un trabajo empírico más complejo que analizase dicha relación.

Este trabajo subraya la importancia de las diferentes dimensiones del entorno institucional en su influencia en el comportamiento emprendedor de las empresas. Además de la influencia del entorno legal y regulador, la existencia de una cultura que aliente la iniciativa y la innovación y una base de esquemas cognitivos compartidos relativos a la gestión empresarial, influye en la orientación emprendedora de las empresas y, por tanto, en la puesta en funcionamiento de nuevas iniciativas empresariales.

Los resultados afirman que la dimensión reguladora del entorno institucional, compuesto por las reglas y los programas de apoyo a las actividades empresariales, no influye actualmente en la orientación emprendedora de las empresas, a excepción de la dimensión “proactividad”, lo que quiere decir que esas normativas y programas de apoyo facilitan la iniciativa a través de los distintos reglamentos y programas de apoyo a la iniciativa empresarial.

El marco cognitivo influye negativamente en la capacidad de asumir riesgos en la empresa y las personas que la componen. Esto se explica por el déficit de educación sobre aspectos relacionados con la empresa y el emprendimiento en España. Este déficit no crea un ámbito cognitivo común donde se compartan planteamientos y conocimientos de temas de empresa, limitando la orientación emprendedora, y donde se tolere el fracaso empresarial, lo que limita dicha asunción de riesgos. Sin embargo, sí influye de forma positiva en el nivel de autonomía del personal y grupos de la organización. Aunque pueda parecer contradictorio con lo anterior, no tiene por qué serlo. La existencia de esquemas cognitivos puede favorecer la independencia en la toma de decisiones de la empresa, sin ser contradictorio con que no favorezca la asunción de riesgos dentro de la organización.

Por último, pese a que históricamente la percepción social del emprendedor no ha sido la

más adecuada, los resultados muestran que actualmente el marco normativo, compuesto por la cultura y la valoración que hace la sociedad del empresario y de las empresas con iniciativa y creatividad, influye de forma positiva en la orientación emprendedora de las organizaciones en términos de innovación, riesgo y proactividad. Esto confirma la tendencia de cambio de una sociedad española que históricamente estaba en las últimas posiciones sobre la percepción positiva del empresario y la actividad económica (Valdaliso-Gago, 2005).

El reto actual de la empresa española, de las instituciones y de la sociedad española, en general, es el cambio de paradigma en su postura ante el emprendimiento. Desde la perspectiva institucional, este cambio debe originarse mediante la existencia de leyes, reglamentos y programas de apoyo empresarial que faciliten la puesta en práctica de nuevas actividades corporativas y que la sociedad esté compuesta por personas con iniciativa, capaces de asumir riesgos, creativas e innovadoras, y dotados de autonomía en su toma de decisiones ya que es una de las vías de éxito empresarial. Por ello, una mayor formación empresarial en todos los niveles educativos sería imprescindible para lograr el cambio social referido. Esta orientación del entorno institucional, el cambio de actitud respecto a esta materia de las organizaciones y la mayor agresividad y competitividad en las posturas empresariales de los ciudadanos en general logrará potenciar la capacidad emprendedora de la sociedad en su conjunto, algo que redundará en dicha capacidad a nivel organizacional.

El trabajo presenta una serie de limitaciones: un solo informante, la naturaleza transversal del trabajo y los resultados parten de una muestra de empresas españolas. No obstante, la literatura institucional avisa de la dificultad de utilizar muestras agregadas de diferentes países, debido a los distintos entornos institucionales existentes y la dificultad de establecer un único entorno institucional común que generalice los resultados (Busenitz *et al.*, 2000).

Los futuros trabajos empíricos deben superar las limitaciones expuestas con anterioridad, profundizando y ampliando los resultados aquí obtenidos. Sería de gran interés analizar las mismas relaciones en organizaciones de diferentes países para establecer comparaciones y, si es posible, establecer un modelo general para la literatura emprendedora. La posibilidad de ampliar el estudio empírico a ámbitos europeos, daría la oportunidad de verificar los resultados obtenidos a una escala superior, despuntando los análisis meramente descriptivos actuales y obtener modelos explicativos del fenómeno emprendedor corporativo que incluyan la influencia institucional, aspecto que ha sido obviado por los investigadores.

AGRADECIMIENTOS

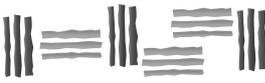
Los autores desean agradecer a los evaluadores anónimos sus comentarios, ya que han ayudado a mejorar el trabajo realizado.

REFERENCIAS

- Amine, L.S.; Staub, K.M. (2009) *Women entrepreneurs in Sub-Saharan Africa: An institutional theory análisis from a social marketing point of view*. Entrepreneurship & Regional Development, Vol. 21, n.º. 2, pp. 183–211.
- Baer, M.; Frese, M. (2003) *Innovation is not enough: climate for initiative and psychological safety, process innovations, and firm performance*. Journal of Organizational Behavior, 24, pp. 45–68.
- Baumol, W.J.; Litan, R.E.; Schramm, C.J. (2009) *Good capitalism, bad capitalism, and the economics of growth and prosperity*. New Haven, CT: Yale University Press.
- Bosma, N.; Acs, Z.J.; Autio, E.; Coduras, A.; Levie, J. (2009) *Global Entrepreneurship Monitor. 2008 Executive Report, GEM*.
- Bosma, N.; Levie, J. (2010) *Global Entrepreneurship Monitor. 2009 Executive Report, GEM*.
- Bruton, G.D.; Ahlstrom, D.; Li, H-L. (2010) *Institutional theory and entrepreneurship: Where are we now and where do we need to move in the future?* Entrepreneurship Theory and Practice, Vol. 34, n.º.3, pp. 421–440.
- Busenitz, L.W.; Gómez, C.; Spencer, J.W. (2000) *Country institutional profiles: Interlocking entrepreneurial phenomena*. Academy of Management Journal, Vol. 43, n.º. 5, pp. 994–1003.
- Covin, J.G.; Green, K.M.; Slevin, D.P. (2006) *Strategic process effects on the entrepreneurial orientation-sales growth rate relationship*. Entrepreneurship Theory and Practice, Vol. 30, n.º. 1, pp. 57–82.
- Covin, J.G.; Slevin, D.P. (1991) *A conceptual model of entrepreneurship as firm behaviour*. Entrepreneurship Theory and Practice, Vol. 16, n.º. 1, pp. 7–25.
- Dess, G.G.; Lumpkin, G.T. (2005) *The role of entrepreneurial orientation in stimulating effective corporate entrepreneurship*. Academy of Management Executive, 19(1), pp. 147–156.
- De la Vega García-Pastor, I.; Coduras, A.; Cruz, C.; Justo, R.; González, I. (2007) *Global Entrepreneurship Monitor. 2007 Spain Executive Report. GEM*.
- De la Vega García-Pastor, I.; Coduras, A.; Cruz, C.; Justo, R.; González, I. (2008) *Global Entrepreneurship Monitor. 2008 Spain Executive Report. GEM*.
- De la Vega García-Pastor, I.; Coduras, A.; Cruz, C.; Justo, R.; González, I. (2009) *Global Entrepreneurship Monitor. 2009 Spain Executive Report. GEM*.
- DiMaggio, P.J.; Powell, W.W. (1983) *The iron cage revisited: Institutional isomorphism and collective rationality*. American Sociological Review, Vol. 48, pp. 147–160.
- DiMaggio, P.J.; Powell, W.W. (1991) *The New Institutionalism in Organizational Analysis*, The University of Chicago Press.
- European Commission (2003) *Green Paper: Entrepreneurship in Europe*. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions, COM (98) 222.

- European Commission (2004) Action Plan: The European Agenda for Entrepreneurship. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions, COM (03) 70.
- Gómez Haro, S. (2010) Influencia de los factores institucionales sobre la actividad emprendedora corporativa. Un análisis causal. Tesis Doctoral, Universidad de Granada.
- Ireland, R.D.; Covin, J.G.; Kuratko, D.F. (2009) *Conceptualizing corporate entrepreneurship strategy*. Entrepreneurship Theory and Practice, Vol. 33, n^o. 1, pp. 19–46.
- Khandwalla, P.N. (1977) The Design of Organizations. New York: Harcourt Brace Jovanovich.
- Knight, G.A. (1997) *Cross cultural reliability and validity of a scale to measure firm entrepreneurial orientation*. Journal of Business Venturing, Vol. 12, pp. 213–225.
- Kostova, T. (1999) *Transnational transfer of strategic organizational practices: A contextual perspective*. Academy of Management Review, Vol. 24, n^o. 2, pp. 308–324.
- Kreiser, P.M.; Davis, J. (2010) *Entrepreneurial orientation and firm performance: The unique impact of innovativeness, proactiveness, and risk-taking*, Journal of Small Business and Entrepreneurship, Vol. 23, n^o. 1, pp. 39–51.
- Liao, Y-S. (2007) *The effects of knowledge management strategy and organization structure on innovation*. International Journal of Management, 24, n^o. 1, pp. 53–60.
- Lim, D.S.K.; Morse, E.A.; Mitchell, R.K.; Seawright, K.K. (2010) *Institutional environment and entrepreneurial cognitions: A comparative business perspective*. Entrepreneurship Theory and Practice, Vol. 34, n^o. 3, pp. 491–516.
- Ling, Y.; Simsek, Z.; Lubatkin, M.H.; Veiga, J.F. (2008) *Transformational Leadership's Role in Promoting Corporate Entrepreneurship: Examining the CEO-TMT Interface*. Academy of Management Journal, 51, vol. 3, pp. 557–576.
- Lumpkin, G.T.; Cogliser, C.C.; Schneider, D.R. (2009) *Understanding and measuring autonomy: An entrepreneurial orientation perspective*. Entrepreneurship Theory and Practice, 33(1), pp. 47–68.
- Lumpkin, G.T.; Dess, G.G. (1996) *Clarifying the entrepreneurial orientation construct and linking it to performance*, Academy of Management Review, Vol. 21, n^o. 1, pp. 135–172.
- Meyer, J.W.; Rowan, B. (1977) *Institutionalized organizations: Formal structure as myth and ceremony*. American Journal of Sociology, Vol. 83, pp. 340–363.
- Miller, D. (1983) *The correlates of entrepreneurship in three types of firms*. Management Science, Vol. 29, n^o. 7, pp. 770–791.
- Miller D.; Friesen, P.H. (1978) *Archetypes of strategy formulation*. Management Science, Vol. 24, pp. 921–933.
- Mintzberg, H. (1973) The Structuring of Organizations. Englewood Cliffs, NJ: Prentice-Hall.
- OCDE (2010) OECD Harmonised Unemployment Rates, August 2010, disponible en: <http://www.oecd.org/dataoecd/45/21/46175570.pdf>

- Podsakoff P.M.; MacKenzie, S.B.; Lee, J.Y.; Podsakoff, N.P. (2003) *Common method biases in behavioral research: a critical review of the literature and recommended remedies*. Journal of Applied Psychology, 88, no. 5, pp. 879—903.
- Scott, W.R. (1995) *Institutions and Organizations*, Sage Publications.
- Scott, W.R.; Meyer, J.W. (1991) *The organization of societal sectors: Propositions and early evidence*. En W.W. Powell y P. DiMaggio (Eds.) *The New Institutionalism in Organizational Analysis*. The University of Chicago Press.
- Spencer, J.W.; Gómez, C. (2004) *The relationship among national institutional structures, economics factors and domestic entrepreneurial activity: a multicountry study*. Journal of Business Research, Vol. 57, pp. 1098–1107.
- Stephen, F.H.; Urbano, D.; van Hemmen, S. (2005) *The Impact of institutions on entrepreneurial activity*. Managerial and Decisions Economics, Vol. 26, pp. 413–419.
- Stopford, J.M.; Baden-Fuller, C.W.F. (1994) *Creating corporate entrepreneurship*. Strategic Management Journal, 15(7), pp. 521–536.
- Urbano-Pulido, D.; Díaz-Casero, J.C.; Hernández-Mogollón, R. (2007) *Evolución y principios de la teoría económica institucional. Una propuesta de aplicación para el análisis de los factores condicionantes de la creación de empresas*. Investigaciones Europeas de Dirección y Economía de la Empresa, Vol. 13, n.º. 3, pp. 183–198.
- Vaillant, Y.; Lafuente, E. (2007) *Do different institutional frameworks condition the influence of local fear or failure and entrepreneurial examples over entrepreneurial activity?* Entrepreneurship & Regional Development, Vol. 19, n.º. 4, pp. 313–337.
- Valdaliso-Gago, J.M. (2005) *El espíritu emprendedor en España: un análisis histórico*. En Cuervo-García, J.A. y Sastre-Castillo (Eds.), *La empresa y el espíritu emprendedor de los jóvenes*, Ministerio de Educación.
- Wiklund, J.; Shepherd, D. (2005) *Entrepreneurial orientation and small business performance: A configurational approach*. Journal of Business Venturing, Vol. 20, n.º. 1, pp. 71–91.
- Zahra, S.A. (1993) *A conceptual model of entrepreneurship as firm behavior: A critique and extension*. Entrepreneurship Theory and Practice, Vol. 17, n.º. 4, pp. 5–21.
- Zahra, S.A. (1995) *Corporate entrepreneurship and financial performance: The case of management leveraged buyouts*. Journal of Business Venturing, Vol. 10, pp. 225–247.



Análisis de la evolución temporal de las desigualdades con datos irregulares

CALLEALTA BARROSO, FRANCISCO JAVIER
Departamento de Estadística Económica, Estructura Económica y O.E.I.
Universidad de Alcalá de Henares
Correo electrónico: franciscoj.callealta@uah.es

FEDRIANI MARTEL, EUGENIO M.
Departamento de Economía, Métodos Cuantitativos e Historia Económica
Universidad Pablo de Olavide, de Sevilla (España)
Correo electrónico: efedmar@upo.es

MARTÍN CARABALLO, ANA M.
Departamento de Economía, Métodos Cuantitativos e Historia Económica
Universidad Pablo de Olavide, de Sevilla (España)
Correo electrónico: ammarcar@upo.es

SÁNCHEZ SÁNCHEZ, ANA MARÍA
Departamento de Economía, Métodos Cuantitativos e Historia Económica
Universidad Pablo de Olavide, de Sevilla (España)
Correo electrónico: amsansan@upo.es

RESUMEN

En este artículo se proporciona una herramienta que permite estudiar las distribuciones de renta en una población a lo largo del tiempo, sin importar que los datos se distribuyan de forma irregular. Incluso, se posibilita la aproximación de curvas de Lorenz en instantes en los que no se cuenta con los datos necesarios. Finalmente, se presenta un ejemplo de aplicación de la técnica en un país europeo en que las desigualdades son especialmente relevantes: España.

Palabras clave: curva de Lorenz; índice de Gini; series temporales irregulares; distribución de la renta.

Clasificación JEL: D31; D63; C81.

MSC2010: 91B82; 60E15; 52A38.

Analyzing the Income Inequalities with Irregular Time Series

ABSTRACT

This article provides a tool to study over time the distributions of income in a population, regardless the irregular distribution of data. We even approach Lorenz curves at the times when we lack the required data. Finally we apply the present technique to a European country where inequalities are particularly relevant: Spain.

Keywords: Lorenz curve; Gini index; irregular time series; income distribution.

JEL classification: D31; D63; C81.

MSC2010: 91B82; 60E15; 52A38.



1. Introducción

Las curvas de Lorenz, complementadas con el índice de Gini, constituyen probablemente la herramienta más utilizada para estudiar el grado de desigualdad económica que se manifiesta en las distribuciones de renta en una población. Su solidez está bien documentada por los numerosos autores que se refieren a ellas. A este respecto pueden consultarse, por ejemplo, [13] y [15]. No obstante, aún no se ha resuelto satisfactoriamente el problema de la medición del grado de desigualdad registrado globalmente a lo largo de un intervalo de tiempo. Si bien, cuando se dispone de datos regularmente espaciados y fiables en todas las unidades de análisis, se pueden utilizar a este fin numerosas técnicas de estadística descriptiva básica o de análisis funcional; sin embargo, cuando los datos están irregularmente espaciados la aplicación de tales técnicas no resulta tan inmediata, debiéndose adoptar otras alternativas como, por ejemplo, considerar variantes de las ideas propuestas en [14], que fueron posteriormente utilizadas en [4] para analizar las desigualdades en cuatro países europeos. En lo que sigue, además de introducir algunas mejoras a la técnica original y de analizar su aplicación a las series que presentan específicamente problemas de regularidad, se llega a proponer otra posible utilidad de la misma: la obtención de curvas de Lorenz en casos en los que no se dispone de suficiente información para construirla por los procedimientos habituales.

En esencia, el primer objetivo de este trabajo es introducir un instrumento que permita expresar el grado de desigualdad de una población medido globalmente a lo largo de un período de tiempo; y, en consecuencia, también se pretende comparar los grados de desigualdad de varias poblaciones (por ejemplo, regiones) medidos globalmente a lo largo de un mismo período. Para ello, se propondrá un indicador que generaliza el índice de Gini, incorporando la dimensión temporal. El segundo objetivo, como ya se ha insinuado, es proporcionar un procedimiento que permita obtener curvas de Lorenz en instantes no observados, a partir de la información observada en el resto del período de tiempo del estudio.

Este artículo consta de cuatro secciones, además de esta breve introducción. La siguiente se dedica a la presentación de los conceptos esenciales para desarrollar la técnica. Sigue un resumen de las propiedades que pueden resultar más interesantes y útiles al investigador que desee aplicarla. Después, la parte teórica se ilustra con datos de España en el período 2003-2009 y se propone su aplicación en diferentes situaciones. Finalmente, se resumen algunas conclusiones. Para facilitar la lectura del artículo, se ha decidido incorporar las cuestiones más técnicas en un Anexo, tras las referencias bibliográficas.

2. Definiciones básicas

El interés de la comunidad científica en la problemática de la desigualdad económica experimentó un notable incremento tras la publicación de los emblemáticos trabajos de Atkinson en 1970 [1] y de Sen en 1973 [19]. Desde entonces los artículos y libros sobre este campo han proliferado enormemente, derivando hacia otros cercanos de gran interés social, como los estudios sobre pobreza, movilidad, polarización y privación.

En concreto, Rawls definió la desigualdad en [16] a partir de las diferencias entre los beneficios y cargas vinculados directa o indirectamente a ellos, tales como prestigio o riqueza, o sujeción a imposición fiscal y a servicios obligatorios. Por tanto, para medir la desigualdad, beneficios y cargas han de ser cuantificados, lo que resulta problemático; habitualmente, y quizás de una forma excesivamente simplista, esta cuantificación se suele hacer a través de las distintas componentes y variantes de la renta. Pese a la existencia de voces críticas al respecto, es evidente que la distribución personal de la renta está relacionada, lógicamente, con la desigualdad y el reparto de la riqueza. Así, al estudiar la distribución de la renta, podemos ver qué parte de la población concentra más o menos recursos (acercándonos al estudio de la desigualdad); o qué otra parte de la población no tiene acceso a niveles de vida que se consideren aceptables en una región (lo que nos acerca al estudio de la pobreza). Este trabajo se centrará estrictamente en el campo de la desigualdad y se basa en la concepción de desigualdad subyacente en los trabajos de Gini y Lorenz, por lo que se parte de los instrumentos que específicamente aportaron ellos y que se exponen a continuación.

La curva de Lorenz de una distribución de rentas se define como una función que, para cualquier proporción p ($0 \leq p \leq 1$), indica la proporción de renta, respecto del total, poseída por la subpoblación cuyas rentas no superan el valor del cuantil de orden p (obviamente, dicha subpoblación está formada por la proporción p de la población con menores rentas). Dicha curva suele denotarse por $L_x(p)$ y, desde un punto de vista empírico-descriptivo, para un vector de rentas observadas, $x = (x_1, x_2, \dots, x_n)$, la curva de Lorenz se aproxima mediante una poligonal cuyos vértices son los puntos $(p, L_x(p))$, siendo:

$$L_x(p) = L_x\left(\frac{k}{n}\right) = \frac{\sum_{i=1}^k x_i}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^k x_i}{n\mu_x},$$

donde μ_x es la media de la distribución x , $k = 0, 1, 2, \dots, n$ y n representa el tamaño de la población (ver [3], [17] y [18]).

Algunas de las propiedades más relevantes que verifica la curva de Lorenz son que $L_x(0) = 0$, que $L_x(1) = 1$ y que, además, es siempre no-decreciente y convexa en el intervalo $[0, 1]$. Por otro lado, prácticamente cualquier curva que verifique estas condiciones

puede considerarse una curva de Lorenz de una población hipotética.

Las curvas de Lorenz se pueden emplear para comparar poblaciones, desde el punto de vista de la desigualdad que presentan, en el conocido como “sentido de Lorenz”. Así, al comparar dos distribuciones, x e y , se dice que la distribución x domina en el sentido de Lorenz a la distribución y cuando la curva de la primera distribución queda siempre por encima de la segunda, en todo el intervalo abierto $(0, 1)$. Y si la distribución x domina en el sentido de Lorenz a la y , eso significa que la primera población presenta menos desigualdad que la segunda. Pero llegar a una conclusión de este tipo exige que las respectivas curvas de Lorenz no se corten, lo que suele ser bastante infrecuente en la gran mayoría de situaciones de interés.

Tratando de dar una respuesta parcial al problema de comparar la desigualdad entre distribuciones que no se dominan estrictamente en el sentido de Lorenz y empleando la información que se recoge en sus respectivas curvas de Lorenz, se puede medir numéricamente el grado de desigualdad dentro de cada distribución mediante el índice de Gini, que mide el grado de concentración de la renta. Usualmente se calcula multiplicando por 2 la llamada “área de desigualdad” (área de la región comprendida entre la curva de Lorenz y la línea de distribución igualitaria $L_x(x) = x$). Si G es el índice de Gini asociado a la curva $L_x(p)$:

$$G = 2 \int_0^1 (p - L_x(p)) dp = 1 - 2 \int_0^1 L_x(p) dp.$$

En el caso extremo de desigualdad absoluta, el índice de Gini sería igual a 1; en el caso contrario, en el que se da una distribución totalmente igualitaria de la renta, el índice de Gini sería igual a 0. Cuanto más cercano a 0 se encuentre el índice de Gini, más igualitario será el reparto de los ingresos, mientras que cuanto más cercano esté a 1, menos igualitario resultará ese reparto y la renta estará más concentrada.

Debe tenerse en cuenta que, además de la anterior, hay muchas formas alternativas de expresar el índice de Gini, como por ejemplo las utilizadas en [5], [2], [6] o [12]. A los efectos de este artículo, no hay problema en considerar cualquiera de dichas definiciones.

Hay que reconocer que el Índice de Gini no es el único indicador de desigualdad posible y que hay investigadores que prefieren utilizar en su lugar otros índices que verifican ciertas propiedades deseables, que el índice de Gini no cumple, como ocurre por ejemplo con: el índice de Theil [20], que verifica la propiedad de descomposición aditiva (es decir, que el índice de desigualdad de una población se puede obtener a partir de los índices de desigualdad de cada una de las subpoblaciones en que se subdivide); o el de Atkinson [1], que recoge de forma flexible la llamada “aversión al riesgo”, que puede variar de una sociedad a otra; entre otros. Sin embargo, por su simple interpretación y aceptación más

generalizada, este artículo se ocupa especialmente del índice de Gini, como una primera experiencia, con la esperanza de que los resultados aquí obtenidos puedan servir para alumbrar otras extensiones a otros índices en un futuro. Así pues, a continuación se presenta el índice de desigualdad en el tiempo (*IDT*), que se propone como instrumento para medir el grado de desigualdad de las distribuciones de renta, que se aprecia globalmente a lo largo de un intervalo de tiempo fijado.

2.1. Indicador de las desigualdades a lo largo del tiempo

Inicialmente supondremos que se dispone de sendos conjuntos de datos conocidos (de renta) para cada una de las unidades de análisis del estudio (por ejemplo, regiones) y para diferentes instantes de tiempo del período que se desea analizar. Para facilitar la siguiente exposición, se fijará una de dichas unidades de análisis y se considerarán sus distribuciones de renta a lo largo del tiempo (posteriormente se generalizaría el mismo proceso para las otras unidades de análisis).

Se puede representar conjuntamente las correspondientes curvas de Lorenz de dichas distribuciones de renta en un espacio de 3 dimensiones, determinadas por tres variables (x, y, z) . Así, las variables x e y (sobre sus correspondientes ejes OX y OY) representarán, respectivamente, a modo de variables independientes, las proporciones p y el instante de tiempo t que determinarán la ordenada de la curva de Lorenz para la proporción p (de población con menor renta) sobre la distribución de renta observada en el instante t ; reservando la variable z (sobre su correspondiente eje OZ) para representar, a modo de variable dependiente, la ordenada de dicha curva de Lorenz.

Así pues, fijada una de las unidades de análisis, se dispone de datos en diferentes instantes de tiempo que, desde un punto de vista empírico-descriptivo, determinan nodos sobre el plano (X, Y) , cuyas coordenadas (x, y) se componen con el instante de observación (y) y el orden (x) del cuantil de renta con el que se corresponde cada dato observado para la distribución de renta en dicho instante. Dichos nodos, junto con los valores de la variable dependiente en ellos (ordenadas de las curvas de Lorenz), se utilizarán para construir una superficie poliédrica a partir de técnicas sencillas de interpolación (en concreto, se trata de aproximaciones funcionales lineales). Una descripción más detallada de este proceso puede encontrarse en [14]. Si procedemos paso a paso, se tiene la siguiente descripción algorítmica:

Supóngase que, para cada unidad de análisis, se observa la distribución de la renta en $k + 1$ instantes de tiempo t (t_0, t_1, \dots, t_k), obteniéndose $k + 1$ conjuntos de datos con tamaños n (n_0, n_1, \dots, n_k) no necesariamente iguales. A cada una de las k curvas de

Lorenz, en cada instante t , las representaríamos por $L_t(x)$ para $t \in \{t_0, t_1, \dots, t_k\}$, con $x \in [0, 1]$ (por ser $L_t(x)$ una curva de Lorenz). Sin embargo, desde una perspectiva empírico-descriptiva, basta con conocer sus valores en los nodos de cada corte temporal (sobre cada recta $y = t$). Estos nodos vienen dados por $\Pi_X \equiv \{0 = x_0 < x_1 < \dots < x_n = 1\}$, una partición del intervalo $[0, 1]$. Obviamente, Π_X viene recomendado por los datos de los que se disponga o por los que fueron utilizados para construir las curvas de Lorenz. Realmente no existe ninguna limitación para elegir los datos, salvo que es necesario contar con algún x_i distinto del cero y el uno. Más adelante, en la Figura 1, se presentará un ejemplo de conjunto de nodos irregular. Fijado el valor (disponible) en $x_j \in \Pi_X$ para cada instante de tiempo t_i , se tienen los puntos $(x_j, t_i, L_{t_i}(x_j))$ para $i = 1, \dots, k$ y $j = 1, \dots, n$, donde $\Pi_T \equiv \{t_0 = 0 < t_1 < \dots < t_{k-1} < t_k = T\}$, es una partición del intervalo $[0, T]$, que normalmente vendrá determinada por la disponibilidad de los datos.

A partir de aquí, se puede calcular una función de interpolación para los puntos de \mathbb{R}^3 que surgen de los nodos y sus respectivas ordenadas. Para ello, sería necesario utilizar una interpolación en dos variables y en la malla rectangular $[0, 1] \times [0, T]$. Supongamos que decidimos resolver este problema de interpolación mediante técnicas sofisticadas, como el método de interpolación de Lagrange (luego se recomendará un procedimiento mucho más sencillo). A la función de interpolación obtenida utilizando la interpolación de Lagrange para los puntos $(x_j, t_i, L_{t_i}(x_j))$ se representa ahora por $L(x, t)$ (que es un polinomio en las variables x y t de grado n para la variable x y de grado k para la variable t , luego es un polinomio de grado $n + k$), donde $x \in [0, 1]$ y $t \in [0, T]$. Por supuesto, este procedimiento debería repetirse para cuantas unidades de análisis compongan el estudio.

A partir de las funciones interpoladas $L(x, t)$ (que luego propondremos reemplazar por otras más simples, que producirán una o varias superficies poliédricas), se va a definir el indicador propuesto IDT , que permitirá comparar la desigualdad en distintas zonas durante un período temporal de amplitud T .

Construida dicha superficie, $z = L(x, t)$, se calcula el volumen que queda por debajo, aunque nuestra propuesta es aproximarlo por la suma de los volúmenes de cada uno de los prismas triangulares truncados en que se puede descomponer el volumen que queda por debajo de la superficie poliédrica.

Definición 2.1 *En las condiciones anteriores y con la notación establecida,*

$$IDT = T - 2 \int_0^1 \int_0^T L(x, t) dt dx.$$

En cualquier caso,¹ el indicador IDT se puede interpretar como una generalización del índice de Gini, cuyo valor se corresponde con el doble del área comprendida entre

¹Como se verá enseguida, también tiene sentido definir el IDT sin multiplicar por dos la integral.

la diagonal del cuadrado $[0, 1] \times [0, 1]$ y la curva de Lorenz. Por su parte, el indicador IDT , tal y como queda determinado por la Definición 2.1, representa el doble del volumen comprendido entre el plano $z = x$ (la “diagonal” del cubo $[0, 1] \times [0, T] \times [0, 1]$) y la superficie interpolada por las curvas de Lorenz en cada instante. Lógicamente, cuanto menor sea tal volumen, menor desigualdad existirá en el período de tiempo estudiado.

Recordemos que el índice de Gini es un ejemplo de indicador de desigualdad construido a partir de la curva de Lorenz, el cual calcula la desigualdad de cierta distribución en un determinado instante de tiempo. En este caso, al definir el indicador IDT , se persigue un objetivo distinto, ya que una vez calculadas (o aproximadas) las curvas de Lorenz para una misma distribución en instantes de tiempo distintos, proponemos caracterizar el grado de desigualdad globalmente observado, o “típico”, en el período de observación, a partir de cómo han evolucionado dichas curvas; es decir, de cómo ha evolucionado la desigualdad a lo largo de un período de tiempo. Este es el primer objetivo del IDT .

3. Algunas propiedades

Cuando el número de unidades de análisis (por ejemplo, zonas geográficas) objetos del estudio que se realiza es reducido y el número de nodos es suficiente, sucede que el indicador IDT se comporta de forma similar a la media de los índices de Gini en cada instante de tiempo para cada una de las zonas. Esto puede hacer pensar que el IDT no es suficientemente útil como para justificar su estudio, pero la ventaja más importante del uso de este indicador frente a la media aritmética de los índices de Gini es que permite la realización de un análisis temporal comparativo incluso cuando no se tengan datos con estructuras análogas en todas las unidades de análisis; es decir, cada unidad de análisis puede tener distinto número de datos, estos pueden estar desigualmente distribuidos (diferente espaciamiento o un espaciamiento no uniforme), pueden incluirse nuevos datos conforme se vayan obteniendo, etc. y, además, el proceso propuesto es independiente de la irregularidad temporal de los datos con los que se cuenta, algo que no se verifica para otras técnicas estadísticas.

Pero, ciertamente, una desventaja del indicador IDT (en su primera versión, tal y como se presentó en la Definición 2.1) es su coste computacional, ya que para su cálculo sería necesario determinar las curvas de Lorenz para varios instantes de tiempo, además de calcular una función de interpolación polinómica en dos variables.

Por otro lado, es posible normalizar este indicador dividiendo por T . Pudiera, incluso, considerarse la posibilidad de dividir por T antes de realizar la triangulación, para que el rectángulo de la malla base sea un cuadrado.

Hemos detectado, además, que el cálculo del *IDT* (en su primera versión) puede presentar algunos problemas adicionales, como pueden ser: la sensibilidad a que algunos datos sean incorrectos (por ejemplo, si los datos son agregados, pueden haberse producido errores significativos por falta de información al agregarlos); errores significativos de pérdida de información; se incrementa el riesgo de errar el valor del indicador por la complejidad de su cálculo; y, finalmente, puede ocurrir que la “superficie de Lorenz”, $z = L(x, t)$, obtenida mediante interpolación, y en la que se basa el cálculo del *IDT*, no se ajuste convenientemente conforme a lo deseado (esto es bastante probable, pues tal superficie, en general, no es convexa). Para resolver estos problemas, se propondrán pautas alternativas de cálculo del *IDT* que evitan las complejas interpolaciones y que permiten calcular el *IDT* mediante operaciones muy sencillas y fácilmente implementables con un programa de computación simbólica o, incluso, en una hoja de cálculo, lo que mejora y dota de mayor potencia a este indicador de desigualdad.

Una primera pauta sería, en vez de trabajar con los datos agregados, utilizar directamente todos los datos (o microdatos, procedentes de las encuestas), de forma que se estarían reduciendo los posibles errores por pérdida de información en los datos. No obstante, es obvio que un gran número de datos, además de proporcionar una mejor aproximación final, también genera un mayor coste computacional. A continuación se verá cómo solucionar este problema gracias a las superficies poliédricas de las que ya se había comentado algo con anterioridad.

Consideremos que se tiene una malla de puntos en el rectángulo $[0, 1] \times [0, T]$ (los nodos en los que se cuenta con datos para la correspondiente curva de Lorenz); si los denotamos por (x_j, t_i) , los datos son $L_{t_i}(x_j)$ y la función $L(x, t)$ debe pasar por todos ellos. Recordemos que (considerando las curvas de Lorenz en tantos por uno y no en tantos por ciento) siempre se verifica que: $L(0, t) = 0$ para todo $t \in [0, T]$; y $L(1, t) = 1$ con $t \in [0, T]$. Por tanto, siempre podrían considerarse nodos informados, tanto los $(1, t_i)$ como los $(0, t_i)$, cualquiera que sea el $t_i \in [0, T]$.

A continuación se desea establecer una triangulación en el rectángulo $[0, 1] \times [0, T]$ que utilice todos los nodos anteriores y que pueda ser útil para dividir el volumen bajo la superficie $z = L(x, t)$ en prismas triangulares truncados. De todas las triangulaciones posibles, las más convenientes para nuestros intereses son aquellas que minimizan las áreas de los triángulos (más precisamente, con el mayor número posible de triángulos y, a la vez, minimizando el área del mayor triángulo) y, de entre todas ellas, las que minimizan sus perímetros, porque todo ello supondrá un mejor ajuste a la “superficie de Lorenz” por parte de la superficie poliédrica que se va a definir. Además, merece la pena tener en cuenta que dicha superficie poliédrica no es necesariamente convexa (a pesar de serlo todas las curvas de Lorenz en cada instante), por lo que tampoco es una buena idea definir

triángulos con vértices en dos instantes temporales no consecutivos. Según todo lo anterior, para encontrar estas triangulaciones óptimas, se puede recurrir al siguiente algoritmo, que presenta una posible variante en algunos de los pasos, por lo que puede llegar a producir varias triangulaciones distintas en la misma malla:

Algoritmo para triangularizar la malla de puntos:

1. Dos nodos de la forma (x_j, t_i) y (x_{j+1}, t_i) están siempre conectados para todos los instantes.
2. Por las restricciones explicadas anteriormente, se reduce el problema a obtener una triangulación entre los nodos correspondientes a dos instantes consecutivos, esto es, t_h y t_i , donde $h = i - 1$ o $h = i + 1$.
3. Se ordenan los valores de la primera variable de los nodos correspondientes en t_h y t_i ; es decir, todos los nodos en esos dos instantes de tiempo determinarían un conjunto de valores distintos en la primera variable: $\Pi_X = \{x_0 = 0, x_1, \dots, x_n = 1\}$. Téngase en cuenta que algunos x_j pueden corresponder a nodos en ambos instantes y otros pueden corresponder a uno solo.
4. Se conectan (x_0, t_i) y (x_0, t_h) .
5. Supongamos conectados los puntos (x_j, t_i) y (x_k, t_h) , donde j y k pueden ser iguales o distintos, pero consecutivos; esto significa que en el instante t_i se tienen conectados todos los nodos hasta el x_j (pongamos que el siguiente nodo en el instante t_i , aún no conectado, se llama x_r) y que en el instante t_h tenemos conectados todos los nodos hasta el x_k (pongamos que el siguiente, aún no conectado, se llama x_s).
6. Se calcula la distancia entre (x_j, t_i) y (x_s, t_h) y la distancia entre (x_k, t_h) y (x_r, t_i) . Se elige la menor de esas distancias y se conectan los nodos correspondientes. En el caso en que ambas distancias sean iguales, hay que elegir una de las parejas de nodos y la elección de la otra pareja produciría otra triangulación distinta.
7. Se vuelve al paso 5 mientras queden nodos por conectar (es decir, hasta que j y k sean los dos 1).

En la Figura 1 se representa gráficamente la triangulación en un ejemplo en el que surgen únicamente dos posibilidades si se aplica el algoritmo anterior. El que el algoritmo de triangulación anterior pueda proporcionar varias triangulaciones diferentes no parece producir graves problemas desde el punto de vista aplicado, pese a que después de la Proposición 3.1 será evidente que las triangulaciones obtenidas no son equivalentes para nuestros propósitos. Como en general no será posible determinar cuál de las triangulaciones proporciona un mejor ajuste de la superficie poliédrica a la superficie teórica que se desea aproximar, se podría elegir una de ellas, pero lo más razonable parece ser realizar los cálculos siguientes para cada una de las triangulaciones (lo que permitirá alcanzar varias

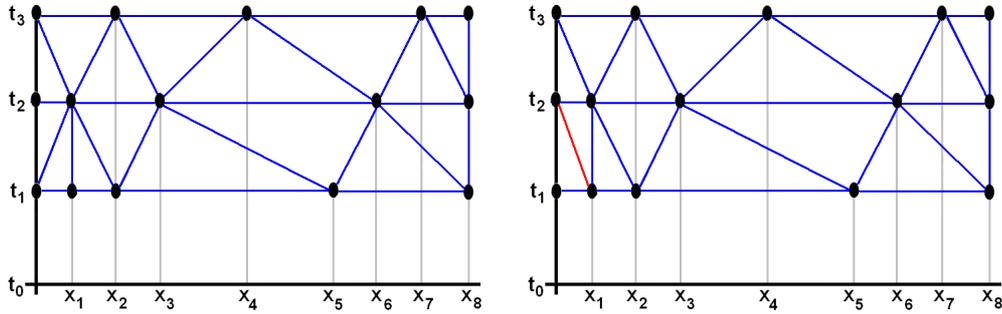


Figura 1: Dos posibles triangulaciones para el mismo conjunto de nodos.

aproximaciones al volumen real) y al final calcular una media aritmética entre todos los volúmenes obtenidos.

La función bidimensional quedaría definida en cada triángulo como aquella función lineal cuya representación gráfica es el plano que pasa por los datos de los tres vértices. Una vez conocida dicha función (para cada una de las triangulaciones, si así se ha decidido), se debería calcular el volumen bajo la función lineal bidimensional en cada uno de los triángulos, para después calcular las sumas de todos los volúmenes de los prismas truncados y, finalmente, la media de las aproximaciones al volumen (utilizando las distintas triangulaciones).

Por supuesto, el cálculo de los volúmenes puede realizarse mediante integrales dobles en regiones triangulares, pero este procedimiento es tedioso y puede provocar errores de computación (aparte de que, a pesar de que las funciones son lineales en cada triángulo, puede tener un coste elevado si el número de nodos es considerable). Por eso, se trata de evitar errores en los cálculos de los volúmenes aplicando la Proposición 3.1, que permitirá calcular el volumen de cada prisma truncado multiplicando el área del triángulo (base) por una determinada altura, también fácil de calcular.

Proposición 3.1 *En la situación anterior, consideremos el prisma triangular truncado cuya base viene determinada por los nodos $N_1(x_1, t_1)$, $N_2(x_2, t_2)$ y $N_3(x_3, t_3)$ y cuya altura en cada nodo (x_j, t_i) es el valor de $L_{t_i}(x_j)$. Entonces, se tiene que el volumen de dicho prisma se puede calcular como el producto del área de la base por la media de las tres alturas en los nodos; o, lo que es lo mismo, el producto del área de la base por la altura que tiene el prisma en el baricentro del triángulo.*

La demostración de la Proposición 3.1 puede consultarse en el Anexo.

Corolario 3.2 *Dos triangulaciones distintas (de las generadas por el algoritmo anterior) pueden producir prismas cuyas sumas de volúmenes son distintas.*

La demostración del Corolario 3.2 también puede consultarse en el Anexo.

Por otra parte, el cálculo de las áreas de los triángulos es bastante sencillo en nuestro caso, por las características de la malla en la que trabajamos. No obstante, por si en algún caso fuera conveniente el cálculo del área del triángulo de la base de cada prisma, en el Anexo se recuerda una fórmula clásica: el Lema 5.1.

Pese a lo anterior, la Proposición 3.1 proporciona una forma alternativa y directa de calcular los volúmenes de los prismas triangulares, que son fácilmente identificables y sus volúmenes fácilmente acumulables, a partir de la representación gráfica de la malla de nodos. En lo que resta, nos centraremos en las propiedades más útiles del *IDT* en cuanto a los objetivos de este artículo: su significado (en el siguiente párrafo) y su relación con la estimación de curvas de Lorenz (en el apartado 3.1).

Evidentemente, el *IDT* que se ha propuesto presentará un mayor valor, cuanto mayor sea la desigualdad en la distribución de la renta en los distintos instantes del período de tiempo observado. El valor 0 se corresponde con el caso de total equidad y el valor T (o 1, si se normaliza convenientemente el *IDT*) corresponde con el caso teórico de total desigualdad. Finalmente, conviene tener en cuenta que el *IDT* determina mejor la desigualdad cuanto mayor sea el número de nodos (datos) con los que se cuente, aunque también es relevante que los nodos estén suficientemente distribuidos en el rectángulo $[0, 1] \times [0, T]$.

3.1. Estimación de una curva de Lorenz

Otra posible aplicación de esta técnica puede ser la de estimar la curva de Lorenz en un instante de tiempo t dentro del período en estudio, aunque no se disponga de observaciones en dicho instante; para ello, una vez calculada la superficie poliédrica, se realizará el corte de dicha superficie poliédrica en el instante de tiempo t con un plano vertical de la forma $y = t$. Esta idea proporciona una forma relativamente asequible de obtener estimaciones de curvas de Lorenz sin necesidad de aplicar exigentes técnicas de interpolación funcional, lo que hace aún más interesante la utilización de la metodología empleada para el cálculo del *IDT* y, en concreto, de la superficie poliédrica subyacente. A continuación se verán las propiedades más relevantes que verifica la curva que se obtiene mediante el corte.

Consideremos la superficie poliédrica $z = L(x, y)$ obtenida por el procedimiento descrito en esta sección. Para que el corte de dicha superficie con $y = \tilde{t}$ (curva de Lorenz estimada) sea realmente una curva de Lorenz, deben verificarse varias condiciones ya comentadas:

La primera propiedad se cumple trivialmente y es que la curva de Lorenz estimada

pasa por $(0, 0)$ y por $(1, 1)$, ya que la superficie $z = L(x, y)$ contiene, por construcción al eje OY y la recta que pasa por $(1, 0, 1)$ y $(1, 1, 1)$.

La segunda cuestión que hay que comprobar es que la curva estimada es no decreciente. La verificación de este aspecto puede consultarse en el Anexo.

Finalmente, en lo que respecta a la tercera propiedad, sería muy deseable que la curva estimada fuese convexa. Sin embargo, es necesario exigir condiciones adicionales a los datos (aunque habitualmente se verifican cuando no hay grandes modificaciones al variar el instante temporal considerado) para poder garantizar algebraicamente la convexidad. A modo de boceto del razonamiento necesario para establecer las condiciones adicionales, se puede consultar la última parte del Anexo.

4. Aplicación al análisis de la desigualdad en España

A continuación se presenta, a modo de ilustración de la técnica anteriormente descrita, una valoración de la desigualdad en España en el período 2003-2009. Para ello, hemos empleado la información sobre la distribución personal de la renta que se encuentra, concretamente, en la Encuesta de Presupuestos Familiares y en la Encuesta Continua de Presupuestos Familiares (la Encuesta de Presupuestos Familiares, iniciada en enero de 2006, sustituye a la Encuesta Continua de Presupuestos Familiares base 1997, que se realizó con periodicidad trimestral desde 1997 hasta 2005), utilizando como unidad básica de análisis el hogar. Sin embargo, tienen características muy diferentes respecto al número de individuos que los forman, sus edades, el número de componentes que realizan un trabajo remunerado, los niveles educativos, etc. No obstante, sin perjuicio de las múltiples formas en que se pueden aplicar las escalas de equivalencia, por motivos didácticos y de simplicidad, en este ejemplo se utilizará la renta per cápita, al dividir la renta del hogar por el número de individuos del mismo.

Las Encuestas de Presupuestos Familiares (en adelante EPF) a lo largo del tiempo han ido sufriendo diversos cambios, tanto en lo que se refiere a su periodicidad como a cambios de tipo metodológicos, pero todas ellas informan sobre diversas características (o variables) relativas a las condiciones de vida de los hogares así como a los ingresos y gastos de ellos. Tales cambios se han realizado por recomendaciones metodológicas de distintos organismos internacionales en general y de EUROSTAT en particular.

Desde hace unos años el INE pone a disposición de los usuarios ficheros de microdatos para que se puedan analizar los datos contenidos en esta encuesta, siendo estos ficheros los que se han utilizado en este trabajo, explotando los datos que en ellos se proporcionaban.

En particular, interesa la información recogida referente a los ingresos de los hogares.

La variable ingreso del hogar se ha construido de la siguiente forma: en la encuesta se pregunta el ingreso mensual neto regular de todos los miembros del hogar y, en caso de que no sepan dar el valor puntual, el intervalo al cual pertenece; así, en algunos casos se dispone del punto y en otros del intervalo. Por ello, se ha imputado a todos los hogares de los cuales se dispone de intervalo pero no de un valor puntual, unos ingresos netos regulares, condicionado a que dicho valor esté situado dentro del intervalo que se ha dado. Esto se ha llevado a cabo según el programa de imputación IVE (Imputation and Variance Estimator), del Institute of Social Research de la Universidad de Michigan, corregidos después de falta de respuesta e inflación.

Una vez obtenidos tales datos se ha tenido en cuenta el tamaño del hogar y se ha calculado el ingreso mensual neto regular per cápita del hogar, que es la variable que se utiliza en el presente estudio, a modo de ejemplo. De forma análoga se podría replicar el caso usando otras escalas de equivalencia si el investigador lo creyera más conveniente para su propósito de análisis.

En los años estudiados (2003 y de 2006 a 2009) la encuesta cambia sustancialmente en el año 2006. Las principales diferencias entre la encuesta del año 2003 [7] y las de los años de 2006 a 2009 [8, 9, 10, 11] son las siguientes:

- Cambia la periodicidad, que pasa de ser trimestral a anual.
- Aumenta el tamaño muestral; en la encuesta base 1997 (la utilizada en el año 2003) se entrevistaban unos 8 000 hogares cada trimestre, lo que suponía (al tener en cuenta el diseño de panel rotante de la encuesta) una muestra anual de aproximadamente 11 000 hogares. En la EPF base 2006 (la utilizada a partir del año 2006) la muestra de hogares está en torno a los 24 000.
- El tiempo de colaboración de los hogares seleccionados para formar parte de la muestra cambia sutilmente; pasan de colaborar en esta durante ocho trimestres a colaborar durante dos años consecutivos.

En las encuestas utilizadas, la representatividad de la muestra es buena tanto a nivel nacional como de comunidades autónomas, por lo que, aunque en esta ilustración se utilizan datos a nivel nacional, análogamente se podría replicar para cualquiera de las comunidades autónomas del país.

Sin embargo, los datos también presentan irregularidades. La primera que salta a la vista es la temporal, puesto que la serie comienza en el año 2003 y después faltan dos años. Aunque en 2004 sí se realizó la encuesta, no incluía la variable ingresos, mientras que en 2005 no se realizó dicha encuesta. Además, los datos son irregulares en cuanto al número,

Año	Hogares	Índice de Gini
2003	9 132	0,2998
2006	19 391	0,3177
2007	21 514	0,3149
2008	22 027	0,3137
2009	22 272	0,3211

Tabla 1: Número de hogares de la muestra e índices de Gini calculados para España en el período 2003-2009.

ya que cada año se realizó la encuesta a un número distinto de individuos (los tamaños muestrales con los que se trabaja son exactamente los que se recogen en la Tabla 1).

Antes de calcular el *IDT* de España para el período de análisis, decidimos realizar el ejercicio de determinar las curvas de Lorenz y los índices de Gini anuales. El procedimiento seguido para determinar las curvas de Lorenz para cada año, así como su correspondiente índice de Gini, fue el siguiente:

1. Se calculan los ingresos regulares netos mensuales per cápita para cada hogar, en cada año de la encuesta.
2. Se ordenan, para cada año, los hogares según la variable anterior y se le asigna un “orden” (entre 1 y el tamaño muestral).
3. Se añade un hogar con orden 0 e ingreso 0.
4. Se calcula una columna de ingresos acumulados.
5. En el eje X se colocan los valores orden/tamaño muestral; la altura viene dada por el ingreso acumulado hasta el hogar/suma total del ingreso acumulado. Con estos datos, se dibuja la curva (poligonal) de Lorenz.
6. También se puede aproximar el índice de Gini: primero se calcula la distancia entre dos puntos consecutivos del eje X (por construcción, son equiespaciados para cada año) y se multiplica por la diferencia entre el valor de X y el de la altura; después se suman todos esos productos y se multiplica por dos.

En la Figura 2 se muestran las curvas de Lorenz obtenidas para cada período (nótese que son muy similares al nivel de resolución de la representación, pero los índices de Gini demuestran que son distintas). El valor correspondiente de los índices de Gini calculados se puede consultar en la Tabla 1.

Después de este ejercicio, calculamos el *IDT* para el período. Puesto que las operaciones eran relativamente sencillas, decidimos realizarlas con la hoja de cálculo Excel 2003,

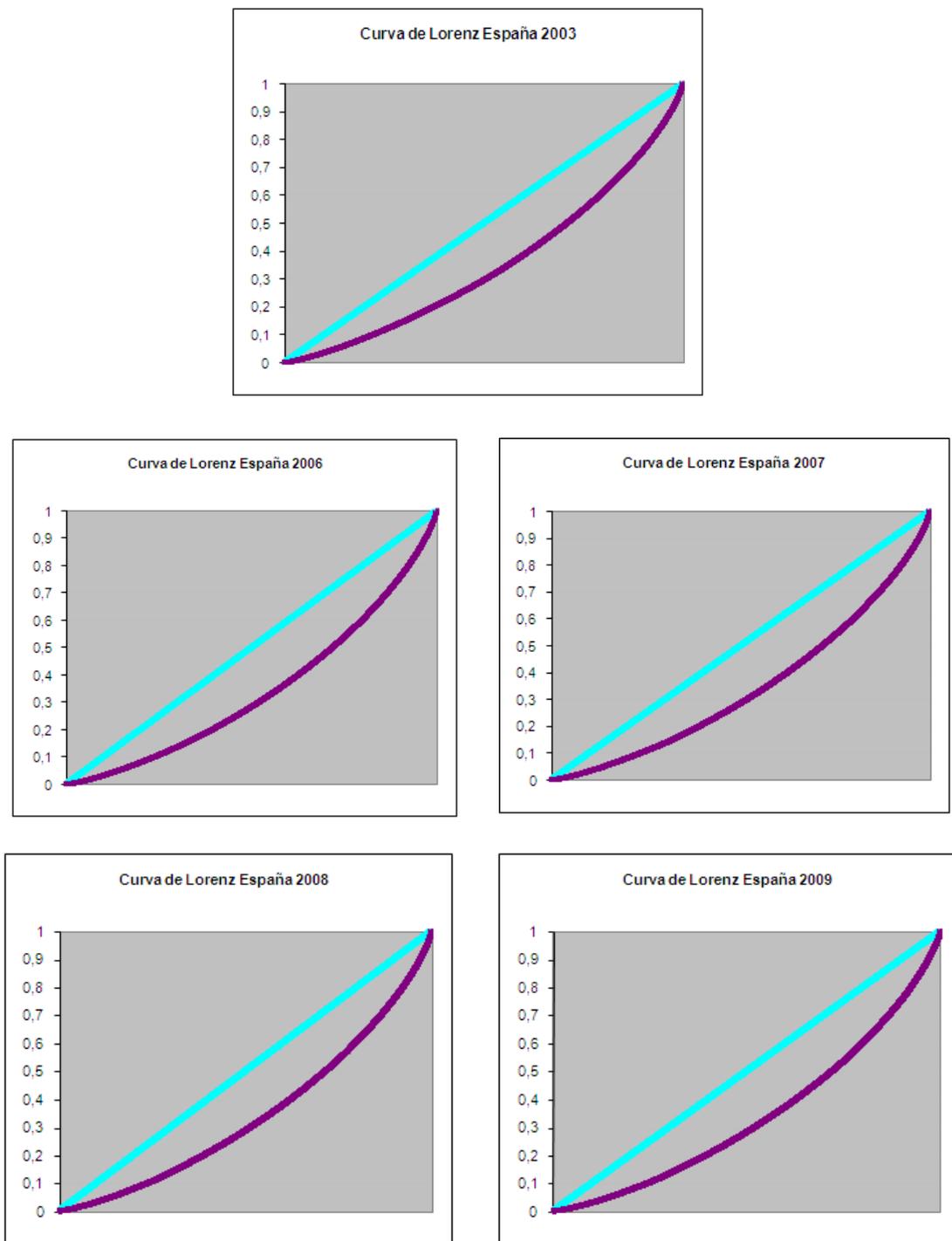


Figura 2: Curvas de Lorenz calculadas para España durante los años 2003, 2006, 2007, 2008 y 2009.

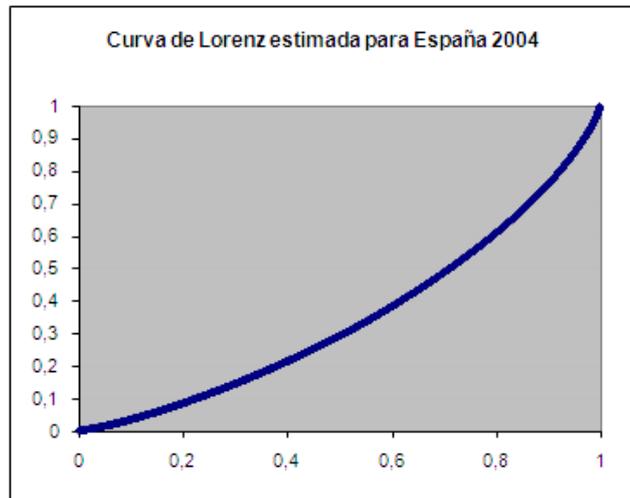


Figura 3: Estimación de la curva de Lorenz para España en el año 2004.

calculando el volumen para cada par de instantes temporales consecutivos. El resultado final fue de 1,0368 para el período 2003-2006, de 0,3418 para 2006-2007, de 0,3428 para 2007-2008 y de 0,3413 para 2008-2009, con lo que el *IDT* global resultó 2,0628. Este resultado se mueve en el rango de $[0, 6]$, siendo 0 el valor de máxima igualdad y 6 el de mayor desigualdad.

A continuación, utilizamos la misma hoja de cálculo para estimar la curva de Lorenz en un año del que no se tenían datos. En concreto, se eligió 2004. Para realizar dicha estimación, se calculó el corte del plano vertical que representaba a 2004 (a un tercio de distancia de 2003 respecto de la distancia total entre 2003 y 2006, instantes de los que sí disponíamos de datos) con la “superficie de Lorenz” calculada anteriormente y eso produjo la curva de la Figura 3. A partir de la curva de Lorenz estimada, repitiendo el ejercicio anteriormente realizado para los años de los que se tenían datos, se calculó el índice de Gini para 2004, que resultó ser de 0,3058, muy similar a los estimados oficialmente para dicha fecha.

5. Conclusiones

En lo que respecta a la cuestión metodológica, en este artículo se ha presentado un conjunto de conceptos y técnicas que consideramos que pueden ser útiles para las futuras investigaciones sobre desigualdad, en muy diversos ámbitos y zonas geográficas. Como se ha visto, además de proporcionar una forma de medir las desigualdades en un período de tiempo, se posibilita una estimación de la curva de Lorenz en un instante del que no se tengan datos suficientes. A pesar de que el comportamiento del indicador y del método de estimación parecen adecuados, hay que advertir de que su funcionamiento correcto depende de la coherencia de los datos a los que se aplica, pues en ocasiones el *IDT* puede compensar o enmascarar situaciones anómalas puntuales y, por otro lado, la estimación de una curva de Lorenz, aunque aproxime a la función efectiva, puede no ser totalmente convexa en casos extremos. Queda para el futuro la determinación explícita de las condiciones que deben verificarse para garantizar el funcionamiento correcto de los procedimientos presentados.

Asimismo, queda pendiente la utilización de las ideas anteriores para generalizar (en cuanto a su aspecto temporal) otros índices de desigualdad distintos del de Gini.

Agradecimientos

Este trabajo ha sido parcialmente subvencionado por el proyecto 2012DEC002 de AACID.

Referencias

- [1] Atkinson, A.B. “On the Measurement of Inequality”, *Journal of Economic Theory*, **2** (1970), 244–263.
- [2] Cowell, F.A. “Measuring inequality”, *LSE Handbooks in Economics. Prentice Hall/Harvester Wheatsheaf*, 2^a ed., Baltimore, 1995.
- [3] Cowell, F.A. y Victoria-Feser, M.P. “Modeling Lorenz Curves: Robust and Semi-Parametric Issues”, *Distributional Analysis Research Programme*, Londres, 2007.
- [4] Domínguez, J., Fedriani, E.M., Martín, A.M. y Núñez, J.J. “Una herramienta para el estudio de la evolución temporal de la desigualdad con datos incompletos. Una

- aplicación a los países del sur de Europa”, *Anales de Economía Aplicada*, **XXV** (2011), 737–749.
- [5] Dorfman, R. “A Formula for the Gini Coefficient”, *Review of Economics and Statistics*, **61** (1979), 146–149.
- [6] Duclos, J.Y. y Araar, A. “Poverty and Equity: Measurement, Policy and Estimation with DAD”, *Springer*, Nueva York, 2006.
- [7] INE. “Encuesta de Presupuestos Familiares. Base 1997”, *Ficheros longitudinales del usuario*, 2003.
- [8] INE. “Encuesta de Presupuestos Familiares. Base 2006”, *Ficheros del usuario*, 2006.
- [9] INE. “Encuesta de Presupuestos Familiares. Base 2006”, *Ficheros del usuario*, 2007.
- [10] INE. “Encuesta de Presupuestos Familiares. Base 2006”, *Ficheros del usuario*, 2008.
- [11] INE. “Encuesta de Presupuestos Familiares. Base 2006”, *Ficheros del usuario*, 2009.
- [12] Liao, T.F. “Measuring and Analyzing Class Inequality with the Gini Index Informed by Model-Based Clustering”, *Sociological Methodology*, **36** (2006), 201–224.
- [13] Lorenz, N.C. “Methods of measuring the concentration of wealth”, *Publications of the American Statistical Association*, **9** (1905), 209–219.
- [14] Martín, A. “Valoración de la pobreza mediante técnicas de agregación de datos de diferente naturaleza”, *Tesis doctoral*, Universidad Pablo de Olavide, Sevilla, 2005.
- [15] Núñez, J.J. “La desigualdad económica medida a través de las curvas de Lorenz”, *Revista de Métodos Cuantitativos para la Economía y la Empresa*, **2** (2006), 67–108.
- [16] Rawls, J. “Justicia como equidad”, *Tecnos*, Madrid, 1986; traducido de Rawls, J. “Justice and Fairness”, *Philosophical Review*, Vol. **LXVII** (1958).
- [17] Schluter, C. y Trede, M. “Statistical Inference for Inequality and Poverty Measurement with Dependent Data”, *International Review of Economics*, **43** (2002), 493–508.
- [18] Schluter, C. y Trede, M. “Tails of Lorenz Curves”, *Journal of Econometrics*, **109** (2002), 151–166.
- [19] Sen, A. “On Economic Inequality”, *Clarendon Press, Paperbacks*, Oxford, 1973.
- [20] Theil, H. “Economics and Information Theory”, *North-Holland*, Ámsterdam, 1967.

Anexo: algunos resultados geométricos y demostraciones

Demostración de la Proposición 3.1. Sin pérdida de generalidad, y a fin de simplificar la demostración analítica, podemos descomponer el prisma triangular truncado en dos volúmenes: un prisma triangular regular que se obtiene desde su base al truncarlo regularmente a la altura del menor de sus aristas verticales (que suponemos la del nodo N_1) y la pirámide irregular que quedaría en el extremo desgajado al truncar, que en general puede tener base trapezoidal. Al realizar esta operación pueden presentarse dos casos, A y B, como ilustra la Figura 4.

En primer lugar calcularemos el volumen de la pirámide trapezoidal, para lo que, sin pérdida de generalidad, podemos suponer que el vértice sobre el nodo N_1 está situado en el origen del sistema de referencia, por lo que sus coordenadas serían $(0, 0, 0)$. Análogamente, aplicando un movimiento conveniente, podemos suponer que las coordenadas de los otros dos vértices situados sobre los nodos N_2 y N_3 son (e, c, a) y $(0, c + d, b)$, respectivamente.

En cualquiera de los dos casos, A y B, la base trapezoidal de la pirámide irregular desgajada tendría por área $\frac{a+b}{2} \cdot \sqrt{e^2 + d^2}$, mientras que su altura h se podría obtener aplicando dos veces el Teorema de Pitágoras (en el caso A aparece “+ x ” en algunas de las expresiones siguientes, mientras que en el caso B aparece “- x ”; por lo que, para presentar los dos casos a la vez, escribimos el símbolo “ \pm ” delante de la x):

$$\begin{cases} h^2 + x^2 = c^2 + e^2 \\ h^2 + (\pm x + \sqrt{e^2 + d^2})^2 = (c + d)^2 \end{cases} \quad \begin{cases} h^2 + x^2 = c^2 + e^2 \\ h^2 + (\pm x + \sqrt{e^2 + d^2})^2 = c^2 + d^2 + 2cd \end{cases}$$

Luego $\pm 2x\sqrt{e^2 + d^2} + e^2 = 2cd - e^2$ y, por tanto, $x = \left| \frac{cd - e^2}{\sqrt{e^2 + d^2}} \right|$, lo que nos lleva a que $h^2 = c^2 + e^2 - \left(\frac{cd - e^2}{\sqrt{e^2 + d^2}} \right)^2 = \frac{e^2(c^2 + d^2 + 2cd)}{e^2 + d^2} = \frac{e^2(c + d)^2}{e^2 + d^2}$.

Con esto, la altura h de la pirámide trapezoidal resulta ser $h = \frac{e(c+d)}{\sqrt{e^2 + d^2}}$. De ahí que el volumen de la pirámide se puede calcular, multiplicando el área de la base por la altura y dividiendo por tres, como sigue:

$$V = \frac{\frac{a+b}{2} \sqrt{e^2 + d^2} \cdot \frac{e(c+d)}{\sqrt{e^2 + d^2}}}{3} = \frac{e(a+b)(c+d)}{6}.$$

O, lo que es lo mismo:

$$V = \frac{e(c+d)}{2} \cdot \left(\frac{a+b}{3} \right).$$

Para finalizar la demostración, el volumen total del prisma truncado es la suma del volumen de la pirámide anterior más el volumen de un prisma (no truncado) de base

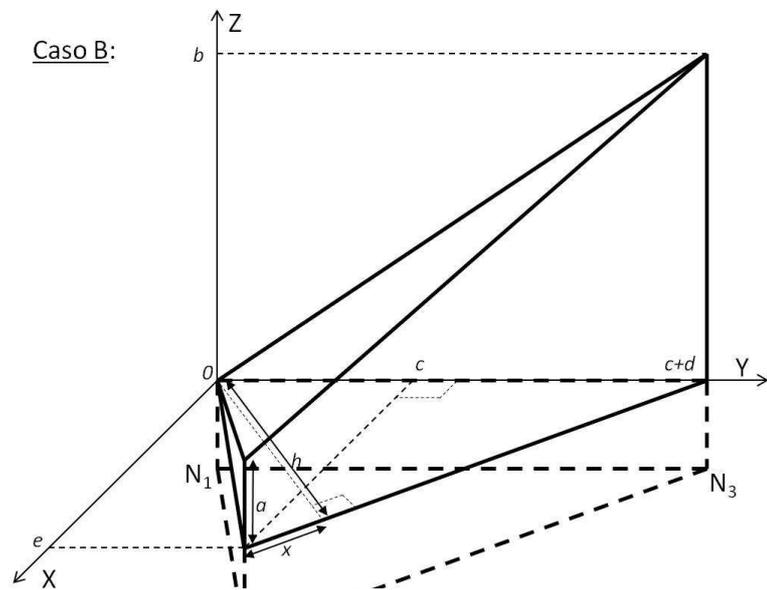
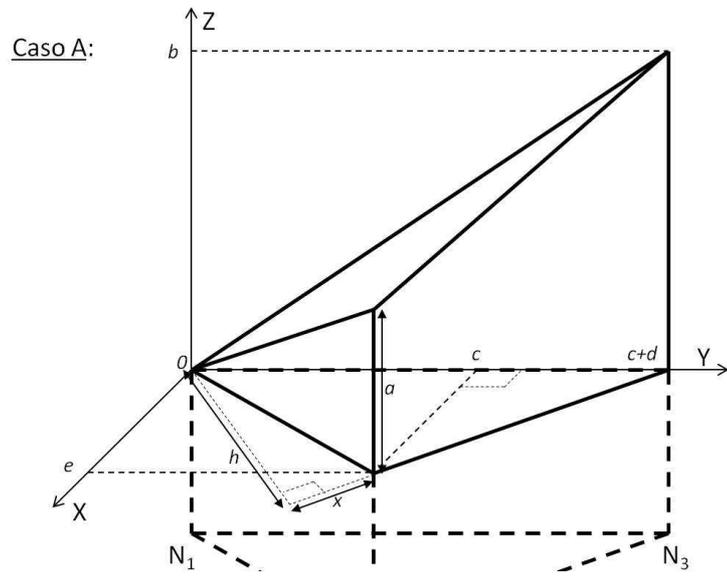


Figura 4: Dos casos posibles de pirámide irregular en la Proposición 3.1.

triangular y altura la de la menor de las aristas sobre los tres nodos:

$$\begin{aligned}
\text{Vol. prisma truncado total} &= \frac{e(c+d)}{2} \cdot \left(\frac{a+b}{3}\right) + \text{Vol. prisma no truncado} \\
&= \frac{(x_2 - x_1) \cdot (t_3 - t_1)}{2} \cdot \frac{(L_{t_2}(x_2) - L_{t_1}(x_1)) + (L_{t_3}(x_3) - L_{t_1}(x_1))}{3} \\
&\quad + \frac{(x_2 - x_1) \cdot (t_3 - t_1)}{2} \cdot L_{t_1}(x_1) \\
&= \frac{(x_2 - x_1) \cdot (t_3 - t_1)}{2} \cdot \frac{L_{t_1}(x_1) + L_{t_2}(x_2) + L_{t_3}(x_3)}{3}. \quad \square
\end{aligned}$$

La demostración anterior se podría haber afrontado desde un punto de vista geométrico (y no tan analítico), con solo comprobar que el prisma triangular tiene el mismo volumen que otro de igual base y que tenga por altura un determinado “punto intermedio” del triángulo inclinado de la base superior. Pero dicho punto ha de dejar el mismo volumen por encima y por debajo de él, lo que solo puede conseguirse en la intersección de las tres medianas del triángulo inclinado, es decir, en su baricentro, cuya proyección en el plano horizontal es también el baricentro del triángulo de la base. Para ver esto, conviene recordar que las medianas unen cada vértice con el punto medio del lado opuesto, por lo que cada mediana divide al triángulo en dos triángulos de igual área.

Demostración del Corolario 3.2. Basta con darse cuenta de que en los cálculos de los volúmenes interviene el valor en un nodo concreto tantas veces como el número de triángulos a los que pertenece dicho nodo y ese número de triángulos no se mantiene invariante en las diferentes triangulaciones (el área de los triángulos sí puede ser la misma, pero la altura variará en general).

□

Lema 5.1 *Considérese el triángulo cuyos vértices (A , B y C) tienen por coordenadas respectivas (x_A, y_A) , (x_B, y_B) y (x_C, y_C) . Entonces, el área de dicho triángulo viene dada por:*

$$\begin{aligned}
\frac{1}{2} \cdot \left| \text{Det} \begin{pmatrix} x_A & x_B & x_C \\ y_A & y_B & y_C \\ 1 & 1 & 1 \end{pmatrix} \right| &= \frac{1}{2} |x_A y_B + x_B y_C + x_C y_A - x_C y_B - x_B y_A - x_A y_C| \\
&= \frac{1}{2} |(x_A - x_C) \cdot (y_B - y_A) - (x_A - x_B) \cdot (y_C - y_A)|,
\end{aligned}$$

siendo Det el determinante e indicando el valor absoluto por líneas verticales. □

Comprobación de que la curva estimada en el apartado 3.1 es no decreciente

Para confirmar esta cuestión, conviene comentar algunos aspectos geométricos e introducir una notación sencilla y coherente con lo anterior. Como antes, supondremos los valores x_j (en los que se tienen datos para algún instante) ordenados de menor a mayor. Llamemos t_i y t_{i+1} a dos instantes en los que se tienen datos o las correspondientes curvas de Lorenz; esto es, en dichos instantes se conocen los valores $L_t(x)$ o, lo que es lo mismo, se conocen las coordenadas $(x_j, t_i, L_{t_i}(x_j))$. Supongamos que se desea calcular las coordenadas de un punto de corte de la superficie poliédrica con el plano $y = \tilde{t}$, siendo $\tilde{t} \in (t_i, t_{i+1})$. En esta situación, consideremos tres nodos conectados por una triangulación (por lo que dos de ellos corresponden a los dos instantes consecutivos). Sin pérdida de generalidad, se puede suponer que dos de dichos nodos son de la forma (x_j, t_i) , (x_{j+1}, t_{i+1}) (el que $x_j = x_{j+1}$ no es más que el caso límite y tampoco presenta mayor dificultad que el analizado) y el tercer nodo, para comprender la situación más fácilmente, podría ser (x_{j+2}, t_i) o (x_{j+2}, t_{i+1}) ; en el primer caso, el triángulo lo formarían los nodos (x_j, t_i) , (x_{j+1}, t_{i+1}) y (x_{j+2}, t_i) , mientras que en el segundo caso los tres vértices serían (x_j, t_i) , (x_{j+1}, t_{i+1}) y (x_{j+2}, t_{i+1}) (como antes, hay otros casos posibles, pero su comportamiento es similar). En ambas situaciones, el que las curvas de Lorenz sean no decrecientes en t_i y en t_{i+1} nos asegura que $L_{t_i}(x_j) \leq L_{t_i}(x_{j+2})$ y $L_{t_{i+1}}(x_{j+1}) \leq L_{t_{i+1}}(x_{j+2})$, respectivamente. Por consiguiente, La altura $L_{\tilde{t}}(x)$ vendría dada por la pendiente de un plano inclinado (el que pasa por $(x_j, t_i, L_{t_i}(x_j))$, $(x_{j+1}, t_{i+1}, L_{t_{i+1}}(x_{j+1}))$ y $(x_{j+2}, t_i, L_{t_i}(x_{j+2}))$ o, en el segundo caso, por $(x_j, t_i, L_{t_i}(x_j))$, $(x_{j+1}, t_{i+1}, L_{t_{i+1}}(x_{j+1}))$ y $(x_{j+2}, t_{i+1}, L_{t_{i+1}}(x_{j+2}))$) en el que $L_{t_i}(x_j) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L_{t_{i+1}}(x_{j+1}) - L_{t_i}(x_j)) \leq L_{t_i}(x_{j+2}) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L_{t_{i+1}}(x_{j+1}) - L_{t_i}(x_{j+2}))$ (equivalentemente, en el segundo caso sería $L_{t_i}(x_j) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L_{t_{i+1}}(x_{j+1}) - L_{t_i}(x_j)) \leq L_{t_i}(x_j) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L_{t_{i+1}}(x_{j+2}) - L_{t_i}(x_j))$) y, por tanto, la curva estimada siempre será no decreciente.

Condiciones para que la curva estimada en el apartado 3.1 sea convexa

Para comprobar que las condiciones que hay que exigir a los datos son naturales, comentaremos lo que sigue. Utilizando la notación anterior, podemos suponer, sin pérdida de generalidad (renumerando los índices y obviando casos similares), que todas las configuraciones posibles de nodos se pueden encuadrar en una de las dos representadas en la Figura 5: (x_1, t_i) , (x_2, t_{i+1}) , (x_3, t_i) y (x_4, t_{i+1}) (en la representación de la izquierda) o, en el segundo caso (representado a la derecha), (x_1, t_i) , (x_2, t_{i+1}) , (x_3, t_i) y (x_4, t_i) . En ambos casos, debido a la linealidad de la superficie poliédrica, probar la convexidad de la curva

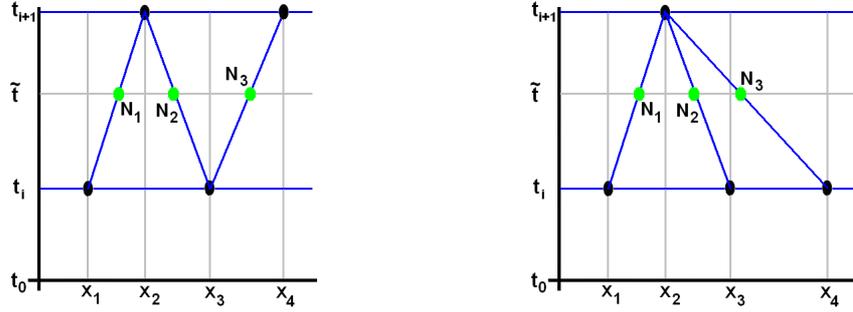


Figura 5: Dos triangulaciones de interés para comprobar la convexidad de las curvas de Lorenz estimadas.

de Lorenz estimada es equivalente a demostrar que se verifica lo siguiente:

$$\frac{L(N_2) - L(N_1)}{x(N_2) - x(N_1)} \leq \frac{L(N_3) - L(N_2)}{x(N_3) - x(N_2)},$$

donde, para $i = 1, 2, 3, 4$, $x(N_i)$ denota la primera coordenada en el nodo N_i , mientras que $L(N_i)$ es la altura propuesta para dicho nodo. En particular, $x(N_i)$ se puede calcular fácilmente a partir de la Figura 5; además, $L(N_1) = L(x_1, t_i) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L(x_2, t_{i+1}) - L(x_1, t_i))$ y $L(N_2) = L(x_3, t_i) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L(x_2, t_{i+1}) - L(x_3, t_i))$; en cuanto a $L(N_3)$, en el primer caso es $L(N_3) = L(x_3, t_i) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L(x_4, t_{i+1}) - L(x_3, t_i))$, mientras que en el segundo vale $L(N_3) = L(x_4, t_i) + \frac{\tilde{t}-t_i}{t_{i+1}-t_i} \cdot (L(x_2, t_{i+1}) - L(x_4, t_i))$. En la configuración de la izquierda de la Figura 5 sí es posible concluir la convexidad de la superficie poliédrica; sin embargo, dicha desigualdad no tiene que verificarse siempre en la configuración de la derecha, aunque es una condición muy razonable si las curvas de Lorenz tienen un comportamiento similar en i e $i + 1$.

Volumen 13 (junio de 2012)

Extreme Learning Machine to Analyze the Level of Default in Spanish Deposit Institutions

Análisis de la morosidad de las entidades financieras españolas mediante Extreme Learning Machine

Montero-Romero, Teresa; López-Martín, María del Carmen; Becerra-Alonso, David; Martínez-Estudillo, Francisco José

Páginas 3–23

Determinantes de la temporalidad en el mercado en el mercado laboral ecuatoriano

Determinants of Temporality in Ecuadorian Labor Market

Chávez, Yannira; Medina, Paúl

Páginas 24–53

Relación entre los factores institucionales y el emprendimiento: análisis mediante técnicas cuantitativas

Relationship between Institutional Factors and Entrepreneurship: A Quantitative Analysis

Salmerón Gómez, Román; Gómez Haro, Samuel

Páginas 54–72

Análisis de la evolución temporal de las desigualdades con datos irregulares

Analyzing the Income Inequalities with Irregular Time Series

Callealta Barroso, Francisco Javier; Fedriani Martel, Eugenio M.; Martín Caraballo, Ana M.; Sánchez Sánchez, Ana María

Páginas 73–96