



UNIVERSIDAD
**PABLO DE
OLAVIDE**
SEVILLA

Revista de Métodos Cuantitativos para la Economía y la Empresa



Journal of Quantitative Methods for
Economics and Business Administration

ISSN: 1886-516 X
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 23. Junio de 2017.

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

URL: <https://www.upo.es/revistas/index.php/RevMetCuant/index>

Editores:

Dr. Francisco Javier Blancas Peral
Universidad Pablo de Olavide
Ctra. de Utrera, km 1 - 41013
Sevilla (Spain).
Correo-e: fjblaper@upo.es

Dr. Ángel F. Tenorio Villalón
Universidad Pablo de Olavide
Ctra. de Utrera, km 1 - 41013
Sevilla (Spain).
Correo-e: aftenorio@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, CINEVESTAV-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)



Stochastic Frontier Models with Dependent Errors based on Normal and Exponential Margins

GÓMEZ-DÉNIZ, EMILIO

Department of Quantitative Methods in Economics and TiDES Institute
University of Las Palmas de Gran Canaria (Spain)
E-mail: emilio.gomez-deniz@ulpgc.es

PÉREZ-RODRÍGUEZ, JORGE V.

Department of Quantitative Methods in Economics
University of Las Palmas de Gran Canaria (Spain)
E-mail: jvperez@dmc.ulpgc.es

ABSTRACT

Following the recent work of Gómez-Déniz and Pérez-Rodríguez (2014), this paper extends the results obtained there to the normal-exponential distribution with dependence. Accordingly, the main aim of the present paper is to enhance stochastic production frontier and stochastic cost frontier modelling by proposing a bivariate distribution for dependent errors which allows us to nest the classical models. Closed-form expressions for the error term and technical efficiency are provided. An illustration using real data from the econometric literature is provided to show the applicability of the model proposed.

Keywords: Technical and cost efficiencies; stochastic frontier; marginal distribution; dependence; Sarmanov model.

JEL classification: C01; C13; C21; C51.

MSC2010: 91B70; 62P20; 91G70.

Modelos de frontera estocástica con errores dependientes basados en márgenes normal y exponencial

RESUMEN

Continuando el reciente trabajo de Gómez-Déniz y Pérez-Rodríguez (2014), el presente artículo extiende los resultados obtenidos a la distribución normal-exponencial con dependencia. En consecuencia, el principal propósito de este artículo es mejorar el modelado de la frontera estocástica tanto de producción como de coste proponiendo para ello una distribución bivalente para errores dependientes que nos permitan encajar los modelos clásicos. Se obtienen las expresiones en forma cerrada para el término de error y la eficiencia técnica. Se ilustra la aplicabilidad del modelo propuesto usando datos reales existentes en la literatura econométrica.

Palabras claves: eficiencias técnica y de coste; frontera estocástica; distribución marginal; dependencia; modelo de Sarmanov.

Clasificación JEL: C01; C13; C21; C51.

MSC2010: 91B70; 62P20; 91G70.



1 Introduction

In general, the methods used for estimating technical and cost efficiency can be considered either parametric or non-parametric. The former involves the estimation of a stochastic production frontier (SPF) or a stochastic cost frontier (SCF) by imposing an explicit functional form and distribution assumption on the data (Aigner *et al.*, 1977; Meeusen and van den Broeck, 1977; Battese and Corra, 1977; Stevenson, 1980; Greene, 1980a, 1980b; Jondron *et al.*, 1982; Lee, 1983; Greene, 1990, 2003; Smith, 2008), where the output of a firm is a function of a set of inputs, plus inefficiency and random error. The second approach is the linear programming technique of data envelopment analysis (DEA), a non-parametric approach which does not impose any assumptions regarding functional form and which does not take into account random error (see Lovell and Schmidt, 1988, for an early survey). Both techniques have advantages and disadvantages; for example, SPF and SCF require the analyst to assume an underlying distribution about the error term, and independence between the inefficiency term and random error. On the other hand, DEA cannot take into account such statistical noise, and efficiency estimates may be biased if the production process is largely characterised by stochastic elements.

Between these two alternatives of modelling, our main interest is based on the the stochastic frontier model in a cross-section framework. The model in this scenario can be written as $y_i = f(x_i; \beta) + \nu_i \pm u_i$, $i = 1, 2, \dots, n$, $u_i \geq 0$, where the sign of the last term depends on whether the frontier describes costs (positive) or production (negative). For example, if we assume that $f(x_i; \beta)$ takes the log-linear Cobb-Douglas form, then the stochastic production frontier (SPF) model can be written as: $\log y_i = \beta_0 + \sum_{j=1}^k \beta_j \log x_{ij} + \nu_i - u_i$, $i = 1, 2, \dots, n$, where $\log y_i$ is the natural logarithm of the production of the i -th firm; $\log x_i$ is a $k \times 1$ vector of (natural log transformations of the) input quantities of the i -th firm; β is a vector of unknown parameters, and the disturbance term $\varepsilon_i = \nu_i \pm u_i$ (which is asymmetric) is assumed to have two components: one with a strictly non-negative distribution, u_i (which is a non-negative component often referred to as the inefficiency term), and another with a symmetric distribution, ν_i (which is termed the idiosyncratic error). Although it is not an assumption of the model, independence of ν and u makes it easy to obtain the density of ε . The density of ε is then used to conduct maximum likelihood estimation of the model parameters. In addition, it is possible to obtain the conditional density of $u|\varepsilon$ and $E(u|\varepsilon)$. These serve as a basis to obtain estimates

for firm-specific inefficiency.

The maximum likelihood method can be used to estimate β and u_i , the variances of the errors and the technical efficiency of each firm. Therefore, distributional assumptions are required for v_i and u_i . In terms of v_i , and in general, these random variables are assumed to be independently and identically distributed (iid) $N(0, \sigma_v^2)$. On the other hand, in terms of u_i , various assumptions may be made; for example, Meeusen and van den Broeck (1977) assigned the exponential distribution to u_i , Battese and Corra (1977) assumed a half-normal distribution, while Aigner *et al.* (1977) considered both distributions. However, since the half-normal and exponential distributions are both single-parameter specifications with modes at zero, some scepticism has been expressed regarding their generality. Thus, Stevenson (1980) suggested the truncated normal and gamma distribution for u_i . Greene (1980a, 1980b) proposed the gamma distribution, Lee (1983) proposed a four-parameter Pearson family of distributions and Greene (1990, 2003) proposed the two-parameter gamma density as a more general alternative.

More recently, another way to model SPF and SCF are based on dependence of error terms such as Smith (2008) and Wiboonpongse *et al.* (2015) with copulas and El Mehdi and Hafner (2014) and Gómez-Déniz and Pérez-Rodríguez (2014) with closed-form solutions by using bivariate distributions. On the other hand, Tran and Tsionas (2015) and Amsler *et al.* (2016) study the correlation between the inputs and statistical noise or inefficiency. The former one proposes an approach which is based on copula function to directly model the correlation between the endogenous regressors and the composed errors assumed to be independent and identically distributed.

Accordingly, the main aim of the present paper is to enhance SPF and SCF modelling by proposing a closed form of a bivariate distribution for dependent errors which allows us to nest the classical models. In particular, we follow Gómez-Déniz and Pérez-Rodríguez (2014) and extend their results by using Sarmanov's family of distributions (Sarmanov, 1966; Lee, 1996; Gómez-Déniz and Pérez-Rodríguez, 2014; among others) to obtain closed-form expressions for the error term and technical efficiency. More specifically, we built a bivariate dependent SPF and SCF models by using normal and exponential distributions (NE), and thus we construct a general extension of the classical stochastic frontier model with these distributions.

The remainder of this paper is structured as follows. Section 2 introduces a brief note on the Sarmanov family of distributions which is used to estimate the technical (cost) efficiency in a

cross-section framework. We analysed one parametric form, deriving in closed-form expression the log likelihood functions and technical (cost) efficiencies, based on the classical pdf distributions, by including the dependence structure. An application of the new model is discussed in Section 3. Finally, the main conclusions drawn are presented in Section 4.

2 Modelling the dependence

In addition to the distributional assumptions on the error terms, ν_i and u_i , in stochastic parametric frontier models, another important characteristic of the above cited models is the independence between them to construct the density and marginal distributions.

The classical stochastic frontier model with normal and exponential assumptions is described by the following stochastic representation: (i) $\nu_i \sim \text{iid } N(0, \sigma_\nu^2)$; (ii) $u_i \sim \text{iid exponential}$ with parameter $\sigma_u > 0$; and (iii) u_i and ν_i are distributed independently of each other and of the regressors. The probability density functions of ν_i and u_i are as follows

$$f_{\sigma_\nu}(\nu) = \frac{1}{\sigma_\nu \sqrt{2\pi}} e^{-\frac{\nu^2}{2\sigma_\nu^2}}, \quad f_{\sigma_u}(u) = \frac{1}{\sigma_u} e^{-\frac{u}{\sigma_u}},$$

where $-\infty < \nu < \infty$, $\sigma_\nu > 0$, $u > 0$ and $\sigma_u > 0$.

In this case, we have

$$f_{\sigma_u, \sigma_\nu}(\varepsilon) = \frac{1}{\sigma_u} \Phi\left(-\frac{\varepsilon}{\sigma_\nu} - \frac{\sigma_\nu}{\sigma_u}\right) \exp\left\{\frac{\varepsilon}{\sigma_u} + \frac{\sigma_\nu^2}{2\sigma_u^2}\right\}, \quad (1)$$

$$f(u|\varepsilon) = \frac{1}{\sqrt{2\pi} \sigma_\nu \Phi(\tilde{\mu}/\sigma_\nu)} \exp\left\{-\frac{1}{2\sigma_\nu^2}(u - \tilde{\mu})^2\right\}, \quad (2)$$

where $\tilde{\mu} = -\varepsilon - \sigma_\nu^2/\sigma_u$.

The marginal $f(\varepsilon)$ is asymmetrically distributed with given by $E(\varepsilon) = -\sigma_u$ and the variance by $\text{var}(\varepsilon) = \sigma_u^2 + \sigma_\nu^2$. On the other hand, $u|\varepsilon$ follows a half-normal distribution, $N^+(\tilde{\mu}, \sigma_\nu^2)$, with mean

$$E(u|\varepsilon) = \tilde{\mu} + \sigma_\nu \frac{\phi(-\tilde{\mu}/\sigma_\nu)}{\Phi(-\tilde{\mu}/\sigma_\nu)} = \sigma_\nu \left(\frac{\phi(A)}{\Phi(-A)} - A \right),$$

and where $A = -\tilde{\mu}/\sigma_\nu$.

Following Gómez-Déniz and Pérez-Rodríguez (2014), we obtain closed-form expression for the likelihood function and technical efficiency for SPF and SCF likelihoods based on the classical

mixture of normal and exponential distributions. We propose a broader, more general and flexible range of dependence, which is also easy to handle, for testing the independence between the inefficiency term and random error (the idiosyncratic component).

This family of distributions is implemented by assuming that $f_1(x_1)$ and $f_2(x_2)$ are univariate probability density functions, with supports defined on $A_1 \subseteq \mathbb{R}$ and $A_2 \subseteq \mathbb{R}$, respectively. Let $\varphi_s(t)$, $s = 1, 2$, be bounded nonconstant functions (the mixing functions) such that

$$\int_{-\infty}^{\infty} \varphi_s(t) f_s(t) dt = 0,$$

then the function defined by

$$f(x_1, x_2) = f_1(x_1) f_2(x_2) [1 + \omega \varphi_1(x_1) \varphi_2(x_2)] \quad (3)$$

is a bivariate joint density with margins $f_1(x_1)$ and $f_2(x_2)$, provided ω is a real number which satisfies the condition $1 + \omega \varphi_1(x_1) \varphi_2(x_2) \geq 0$, for all x_1 and x_2 . Some methods to obtain the mixing function φ when $f_s(x_s)$, $s = 1, 2$, are members of the natural exponential family of distributions are described in Lee (1996). The Farlie–Gumbel–Morgernstern (FGM) family of copulas can be viewed as a special case of the above–mentioned construction, by setting $\varphi(x_s) = 1 - 2F(x_s)$, $s = 1, 2$ and therefore one of the models proposed in Smith (2008). Here $F(\cdot)$ represents the cumulative distribution function of the random variables with pdf $f(\cdot)$.

As we will see this family provide analytical expressions for the marginal pdf of the random variable ε , the conditional pdf of $u|\varepsilon$, $cov(u, \nu)$ and technical efficiency are provided in the SPF model, for the NE model. The classical independent models are derived as a particular case when $\omega = 0$ while $\omega \neq 0$ measures the dependence structure.

As in the classical SPF model, let $\nu = u + \varepsilon$. Using (3) we get

$$f_{\sigma_u, \sigma_\nu, \omega}(u, \varepsilon) = f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) [1 + \omega \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(u + \varepsilon)], \quad (4)$$

by taking $\varphi_{\sigma_u}(u) = e^{-u} - \delta_u(\sigma_u)$, $\varphi_{\sigma_\nu}(\nu) = e^{-\nu^2 - 2\nu} - \delta_\nu(\sigma_\nu)$, with

$$\delta_u(\sigma_u) = \frac{1}{1 + \sigma_u}, \quad (5)$$

$$\delta_\nu(\sigma_\nu) = \frac{1}{\sqrt{1 + 2\sigma_\nu^2}} \exp \left\{ \frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}, \quad (6)$$

defines a bivariate distribution of (u, ν) with marginal distributions $f_{\sigma_\nu}(\nu)$ and $f_{\sigma_u}(u)$ as in the classical model and where $\omega_1 \leq \omega \leq \omega_2$, being

$$\omega_1 = \max \left\{ \frac{-1}{\delta_u(\sigma_u)\delta_\nu(\sigma_\nu)}, \frac{-1}{(1-\delta_u(\sigma_u))(e-\delta_\nu(\sigma_\nu))} \right\}, \quad (7)$$

$$\omega_2 = \min \left\{ \frac{1}{\delta_u(\sigma_u)(e-\delta_\nu(\sigma_\nu))}, \frac{1}{(1-\delta_u(\sigma_u))\delta_\nu(\sigma_\nu)} \right\}. \quad (8)$$

To see this, observe that because $\frac{d}{du}\varphi_{\sigma_u}(u) < 0$ we have that $\varphi_{\sigma_u}(u)$ is a decreasing function on u and the range of variation of $\varphi_{\sigma_u}(u)$ is $(-\delta_u(\sigma_u), 1 - \delta_u(\sigma_u))$. In the same way, it is simple to see that $\varphi_{\sigma_\nu}(\nu)$ has a maximum in $\nu = -1$ and therefore the range of variation of $\varphi_{\sigma_\nu}(\nu)$ results $(-\delta_\nu(\sigma_\nu), e - \delta_\nu(\sigma_\nu))$. Now, we have that $f_{\sigma_u, \sigma_\nu, \omega}(u, \nu)$ represents a probability density function if $1 + \omega \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(\nu) \geq 0$, and this occurs if

$$\begin{aligned} \omega &\geq \frac{-1}{\varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(\nu)}, \quad \text{for } \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(\nu) > 0, \\ \omega &\leq \frac{-1}{\varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(\nu)}, \quad \text{for } \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(\nu) < 0, \end{aligned}$$

from which it is a simple exercise to see that range of ω is given by (ω_1, ω_2) .

Although any other mixing functions satisfying that $\int_0^\infty f_{\sigma_u}(u)\varphi_{\sigma_u}(u) du = 0$ and that $\int_{-\infty}^\infty f_{\sigma_\nu}(\nu)\varphi_{\sigma_\nu}(\nu) d\nu = 0$ can be considered we have chosen the mixing functions above since: (i) The presence of the exponential term which is also present in the probability density function of the normal and exponential distributions facilitates the computations in order to obtain closed-form expressions for the marginal of ε as we will see in the next section; (ii) The square term in the exponential part of $\varphi_{\sigma_\nu}(\nu)$ is important to ensure appropriate bounds for the ω parameter. Dependence assumption is now depending on ω and the Sarmanov family with normal and half normal marginals studied here can also be considered as an extension of the classical Sarmanov family of distributions dealt in Lee (1996) for the normal case.

Some algebra provides the correlation coefficient, which is given by

$$\rho = \frac{2\omega\sigma_u\sigma_\nu}{(1+\sigma_u)^2(1+2\sigma_\nu^2)^{3/2}} \exp \left\{ \frac{2\sigma_\nu^2}{1+2\sigma_\nu^2} \right\},$$

This correlation coefficient is bounded by (see Lee, 1996)

$$|\rho| \leq |\omega| \left[E \left(\varphi_{\sigma_u}^2(u) \right) E \left(\varphi_{\sigma_\nu}^2(\nu) \right) \right]^{1/2}.$$

Now, we have the following result which can be used to build the likelihood function for SPF (see Appendix 4 for the SCF model).

Theorem 1 *In the SPF model for NE distributions and assuming dependence, the marginal pdf of ε is given by*

$$\begin{aligned}
f_{\sigma_u, \sigma_\nu, \omega}(\varepsilon) &= \frac{\Upsilon_{\sigma_u, \sigma_\nu}^1(\varepsilon)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} \left\{ \Phi \left(\frac{\tilde{\mu}}{\sigma_\nu} \right) \left[\sqrt{1 + 2\sigma_\nu^2} + \frac{\omega e^{\frac{2\sigma_\nu^2}{1+2\sigma_\nu^2}}}{1 + \sigma_u} \right] \right. \\
&\quad - \omega \Upsilon_{\sigma_u, \sigma_\nu}^2(\varepsilon) \Phi \left(\frac{\tilde{\mu}}{\sigma_\nu} - \sigma_\nu \right) \\
&\quad - \frac{\omega}{1 + \sigma_u} \Upsilon_{\sigma_u, \sigma_\nu}^3 \Phi \left(-\frac{\sigma_\nu}{\sigma_u} \frac{1 + 2\sigma_u}{\sqrt{1 + 2\sigma_\nu^2}} - \frac{\varepsilon}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} \right) \\
&\quad \left. + \omega \Upsilon_{\sigma_u, \sigma_\nu}^4(\varepsilon) \Phi \left(-\frac{\sigma_\nu}{\sigma_u} \frac{1 + 3\sigma_u}{\sqrt{1 + 2\sigma_\nu^2}} - \frac{\varepsilon}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} \right) \right\}, \tag{9}
\end{aligned}$$

where

$$\begin{aligned}
\Upsilon_{\sigma_u, \sigma_\nu}^1(\varepsilon) &= \exp \left\{ \frac{\varepsilon}{\sigma_u} + \frac{\sigma_\nu^2}{2\sigma_u^2} \right\}, \\
\Upsilon_{\sigma_u, \sigma_\nu}^2(\varepsilon) &= \exp \left\{ \varepsilon + \frac{\sigma_\nu^2}{\sigma_u} + \frac{\sigma_\nu^2}{2} + \frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}, \\
\Upsilon_{\sigma_u, \sigma_\nu}^3 &= \exp \left\{ \frac{\sigma_\nu^2}{2\sigma_u^2} \frac{4\sigma_u - 2\sigma_\nu^2}{1 + 2\sigma_\nu^2} + \frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}, \\
\Upsilon_{\sigma_u, \sigma_\nu}^4(\varepsilon) &= \exp \left\{ \varepsilon + \frac{\sigma_\nu^2}{2\sigma_u^2} \frac{5\sigma_u^2 + 6\sigma_u - 2\sigma_\nu^2}{1 + 2\sigma_\nu^2} + \frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}.
\end{aligned}$$

Proof: See Appendix 1. ■

Observe that when $\omega = 0$, i.e. the independence case, pdf (9) reduces to (1). Simple computations provide that the mean for the marginal pdf given in (9) is equal to $E(\varepsilon) = -\sigma_u$ while the variance results

$$var(\varepsilon) = \sigma_u^2 + \sigma_\nu^2 \left[1 - \frac{4\omega\sigma_u^2}{(1 + \sigma_u^2)(1 + 2\sigma_\nu^2)^{3/2}} e^{\frac{2\sigma_\nu^2}{1+2\sigma_\nu^2}} \right].$$

Figure 1 shows examples of the marginal pdf (9) for different values of the model parameters.

Having obtained the main result of the likelihood function, we now show the conditional distribution of u given ε .

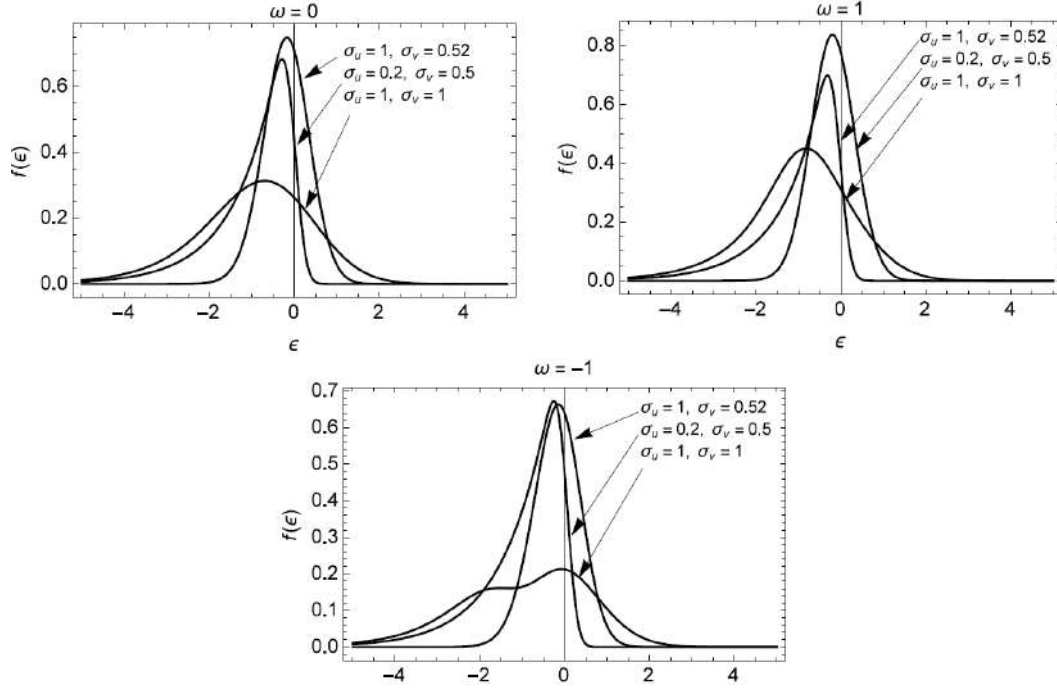


Figure 1: Marginal distribution in the SPF-NE model for selected parameter values.

Proposition 1 *In the SPF model under NE distributions, assuming dependence, the conditional pdf of u given ε is given by*

$$f_{\sigma_u, \sigma_v, \omega}(u|\varepsilon) = \frac{f_{\sigma_u, \sigma_v}(u|\varepsilon) + \omega \Psi_{\sigma_u, \sigma_v}^0(u|\varepsilon)}{1 + \omega \Psi_{\sigma_u, \sigma_v}^1(\varepsilon)}, \quad (10)$$

for $u > 0$, where $f_{\sigma_u, \sigma_v}(u|\varepsilon)$ is given in (2) and

$$\Psi_{\sigma_u, \sigma_v}^0(u|\varepsilon) = \varphi_{\sigma_u}(u) \varphi_{\sigma_v}(u + \varepsilon) f_{\sigma_u, \sigma_v}(u|\varepsilon), \quad (11)$$

$$\begin{aligned} \Psi_{\sigma_u, \sigma_v}^1(\varepsilon) = & \frac{1}{\sqrt{1 + 2\sigma_v^2}} \frac{1}{\Phi\left(\frac{\tilde{\mu}}{\sigma_v}\right)} \left[\Phi\left(\frac{\tilde{\mu}}{\sigma_v}\right) \frac{e^{\frac{2\sigma_v^2}{1+2\sigma_v^2}}}{1 + \sigma_u} - \Upsilon_{\sigma_u, \sigma_v}^2(\varepsilon) \Phi\left(\frac{\tilde{\mu}}{\sigma_v} - \sigma_v\right) \right. \\ & - \frac{1}{1 + \sigma_u} \Upsilon_{\sigma_u, \sigma_v}^3 \Phi\left(-\frac{\sigma_v}{\sigma_u} \frac{1 + 2\sigma_u}{\sqrt{1 + 2\sigma_v^2}} - \frac{\varepsilon}{\sigma_v} \sqrt{1 + 2\sigma_v^2}\right) \\ & \left. + \Upsilon_{\sigma_u, \sigma_v}^4(\varepsilon) \Phi\left(-\frac{\sigma_v}{\sigma_u} \frac{1 + 3\sigma_u}{\sqrt{1 + 2\sigma_v^2}} - \frac{\varepsilon}{\sigma_v} \sqrt{1 + 2\sigma_v^2}\right) \right]. \quad (12) \end{aligned}$$

Proof: See Appendix 2. ■

The point estimation for technical efficiency used in this study is the expression given by Battese and Coelli (1988), which is related to the conditional expectation of e^{-u_i} , that is, $E(e^{-u_i}|\varepsilon = \hat{\varepsilon})$ which is given in the next result.¹

Proposition 2 *The point estimation for the technical efficiency of the i -th producer in the SPF model under NE distributions, assuming dependence, is given by*

$$TE_i = \frac{\Psi_{\sigma_u, \sigma_\nu, \omega}^{01}(\varepsilon_i) + \omega \Psi_{\sigma_u, \sigma_\nu}^{02}(\varepsilon_i)}{1 + \omega \Psi_{\sigma_u, \sigma_\nu}^1(\varepsilon_i)}, \quad i = 1, 2, \dots, n,$$

where

$$\Psi_{\sigma_u, \sigma_\nu, \omega}^{01}(\varepsilon_i) = [1 + \omega \delta_u(\sigma_u) \delta_\nu(\sigma_\nu)] \frac{\Phi(-A - \sigma_\nu)}{\Phi(-A)} e^{\sigma_\nu(A + \sigma_\nu/2)},$$

and

$$\begin{aligned} \Psi_{\sigma_u, \sigma_\nu}^{02}(\varepsilon_i) = & \frac{1}{\Phi(-A)} \left[\frac{1}{\sqrt{1 + 2\sigma_\nu^2}} \Upsilon_{\sigma_u, \sigma_\nu}^{11}(\varepsilon_i) \Phi\left(-\frac{A + 2(2 + \varepsilon_i)\sigma_\nu}{\sqrt{1 + 2\sigma_\nu^2}}\right) \right. \\ & - \delta_\nu(\sigma_\nu) \Upsilon_{\sigma_u, \sigma_\nu}^{22}(\varepsilon) \Phi(-A - 2\sigma_\nu) \\ & \left. - \frac{1}{\sqrt{1 + 2\sigma_\nu^2}} \delta_u(\sigma_u) \Upsilon_{\sigma_u, \sigma_\nu}^{33}(\varepsilon_i) \Phi\left(-\frac{A + \sigma_\nu(3 + 2\varepsilon_i)}{\sqrt{1 + 2\sigma_\nu^2}}\right) \right], \end{aligned}$$

with

$$\begin{aligned} \Upsilon_{\sigma_u, \sigma_\nu}^{11}(\varepsilon_i) &= \exp \left\{ -\frac{\varepsilon_i^2 - A\sigma_\nu(4 - A\sigma_\nu) - 8\sigma_\nu^2 + 2\varepsilon_i(1 - \sigma_\nu(A + 2\sigma_\nu))}{1 + 2\sigma_\nu^2} \right\}, \\ \Upsilon_{\sigma_u, \sigma_\nu}^{22}(\varepsilon_i) &= \exp \{ 2\sigma_\nu(A + \sigma_\nu) \}, \\ \Upsilon_{\sigma_u, \sigma_\nu}^{33}(\varepsilon_i) &= \exp \left\{ -\frac{\varepsilon_i^2 - A\sigma_\nu(3 - A\sigma_\nu) - 9/2\sigma_\nu^2 + 2\varepsilon_i(1 - \sigma_\nu(A + \sigma_\nu))}{1 + 2\sigma_\nu^2} \right\}, \end{aligned}$$

while $\delta_u(\sigma_u)$, $\delta_\nu(\sigma_\nu)$ and $\Psi_{\sigma_u, \sigma_\nu}^1(\varepsilon_i)$ are given in (5), (6) and (12), respectively.

Proof: See Appendix 3. ■

¹This estimator is particularly useful when u_i is not close to zero. However, the estimate of technical efficiency is inconsistent because the variation associated with the distribution of $u_i|\varepsilon$ is independent of i (Kumbhakar and Lovell, 2000, p.78)

3 Numerical application

In this section, we use the theoretical results obtained in Section 3 to estimate and test independence between the inefficiency term and the idiosyncratic error, using one empirical framework.

In the example we estimate the proposed model normal–exponential by using data obtained from several bank branches of a large Spanish commercial bank during from January 2011 to December 2014 (monthly data). Specifically, data corresponds to gross operating annual cost as output and the inputs we use are labour and capital prices but also the annual revenues (total income) for each bank branch.

We estimate a log-linear Cobb-Douglas cost function for the 5009 pooled data and without imposing linear homogeneity in the input prices. The estimated model is written as follows

$$\log c_i = \beta_0 + \beta_1 \log l_i + \beta_2 \log k_i + \beta_3 \log y_i + \nu_i + u_i, \quad i = 1, 2, \dots, 5009,$$

where the variable $\log c_i$ is the natural log–transformed annual operating cost, $\log l_i$ is the natural log-transformed labour price, $\log k_i$ is the natural log-transformed price of capital and $\log y_i$ is the natural log-transformed annual revenue.

The maximum likelihood estimates for the cost frontier NE model and for a sample of n bank branches can be obtained by maximizing the log–likelihood function derived from (17) restricted by (7) and (8), respectively. After some algebra, it is given by

$$\begin{aligned} \log L = & \frac{n}{2} \left[\frac{\sigma_\nu^2}{2\sigma_u^2} - 2 \log \sigma_u - \log(1 + \sigma_\nu^2) \right] + \frac{1}{\sigma_u} \sum_{i=1}^n \varepsilon_i \\ & + \sum_{i=1}^n \log \left\{ \Phi \left(\frac{\tilde{\mu}}{\sigma_\nu} \right) \left[\sqrt{1 + 2\sigma_\nu^2} + \frac{\omega}{1 + \sigma_u} \exp \left(\frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right) \right] \right. \\ & - \omega \left[\sqrt{1 + 2\sigma_\nu^2} \Xi_{\sigma_u, \sigma_\nu}^2(\varepsilon_i) \delta_\nu(\sigma_\nu) \Phi \left(\frac{\tilde{\mu}}{\sigma_\nu} - \sigma_\nu \right) \right. \\ & + \Xi_{\sigma_u, \sigma_\nu}^3(\varepsilon_i) \delta_u(\sigma_u) \Phi \left(\frac{\sigma_\nu(2\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon_i \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu} \right) \\ & \left. \left. - \Xi_{\sigma_u, \sigma_\nu}^4(\varepsilon_i) \Phi \left(\frac{\sigma_\nu(\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon_i \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu} \right) \right] \right\}, \end{aligned} \quad (13)$$

where $\Xi_{\sigma_u, \sigma_\nu}^1(\varepsilon_i)$, $\Xi_{\sigma_u, \sigma_\nu}^2(\varepsilon_i)$, $\Xi_{\sigma_u, \sigma_\nu}^3(\varepsilon_i)$ and $\Xi_{\sigma_u, \sigma_\nu}^4(\varepsilon_i)$, $i = 1, \dots, n$, are given in (18), (19), (20) and (21), respectively.

Table 1 shows three estimation methods applied to the data: ordinary least squares (OLS), unrestricted maximum likelihood (UML) (by taking $\omega = 0$ in (13)) and restricted maximum likelihood (RML) (by using directly (13)).

Table 1: Stochastic cost frontier estimates

Variable	Coeff	t-Stat	Coeff	t-Stat	Coeff	t-Stat
	<i>OLS</i>		<i>UML-NE</i>		<i>RML-NED</i>	
β_0	0.1013	0.28	0.2744	21.43	0.2220	1.75
β_1	0.3238	9.16	0.2847	29.79	0.2835	21.33
β_2	0.2278	13.29	0.2549	14.14	0.2515	13.90
β_3	0.5072	36.07	0.4666	35.36	0.4709	33.76
σ_u			0.4336	31.95	0.4757	30.52
σ_v			0.2441	10.91	0.2573	11.42
ω					-1.3970	-119.36
ρ					-0.14648	(0.0141)
ρ_S					-0.01246	(0.0093)
L_{\max}		-5508.38		-3221.14		-3092.30

Note: Between parenthesis standard error is indicated.

The OLS estimates are compared with those obtained from stochastic frontier models, using the exponential distribution (UML-NE) and normal-exponential with dependence (RML-NED). Also shown is the maximum value of the log-likelihood function (L_{\max}), together with some correlation coefficients. It is known that if X and Y are two random variables with cdf $F(x)$ and $G(y)$, respectively, Spearman's coefficient, denoted by ρ_S , is given by $\rho_S = \text{Corr}(F(X), G(Y))$, i.e. the ordinary (Pearson) correlation coefficient of the random variables $F(x)$ and $G(y)$ (see Fre-dricks and Nelsen, 2007). This coefficient was computed numerically for the analyzed data, and is also shown in the tables, together with ρ , the classical coefficient of correlation. The standard errors of both, Spearman's coefficient and correlation coefficient, appear between parenthesis. The first was calculated by using the expression $\sigma_{\rho_S} = \sqrt{0.437/(n-4)}$ (see Bonnet, 2000).

The second was computed by using the expression $\sigma_\rho = 1/\sqrt{n-3}$, provided also in Bonnet (2000) which produces a similar result than the one obtained by using $\sigma_\rho = (1-\rho^2)/\sqrt{n^2-1}$ (see Dingman and Perry, 1956 for details). As pointed out by Shubina and Lee (2004), the Spearman correlation coefficient for a Sarmanov family of distribution is situated in the interval $[-3/4, 3/4]$.

The model was estimated by using restricted maximum likelihood in two-stages. The first stage is based on the simplex method, a search procedure that requires only function evaluations, not derivatives. To apply simplex, OLS initial values are used for β_0 , β_1 and β_2 and then values for σ_u , σ_ν and ω are determined (these values are equal to 2.0, 2.0 and 2.0, respectively). The most important use of simplex is to refine initial estimates before applying one of the derivative-based methods, which are more sensitive to the choice of initial estimates. For all models, we used 5 iterations in this stage. In the second stage, the BFGS (Broyden, Fletcher, Goldfarb and Shanno) algorithm was applied to obtain the final estimates of the parameters and the asymptotic variance-covariance matrix estimated by the final iteration of the approximation of the inverse Hessian. Finally, we computed regression standard errors and the covariance matrix allowing for heteroscedasticity.²

It is noteworthy that the estimated parameters are very similar between the UML-NE and RML-NED estimates and, in general, they are statistically significant at any significant level and positive, indicating a positive relationship between the total cost and all input prices and total revenue. On the other hand, the estimated value of the dependence parameter, that is, ω , it is statistically significant at any significant level and negative (t -statistic is equal to -119.36 , p -value = 0.0). Moreover, the correlation coefficient (ρ) and Spearman's ρ_S measure (the probabilistic concordance among errors) indicate that correlation is low and negative among errors but it is not zero. Finally, the maximum value of likelihood function is higher for RML-NED than UML-NE and OLS estimates.

In terms of models fitting, we compare both the RML-NED and UML-NE models by using a likelihood ratio test. Therefore, we can also show evidence of the dependence assumption

²All the computations were performed using RATS software.

being tested. Specifically, the likelihood ratio test for the null hypothesis where $\omega = 0$ is 257.68 (p -value = 0.0) which also indicates that independence between errors is rejected by our data.

As a final point, to illustrate the behavior of estimated cost efficiencies for RML–NED and UML–NE models by using results in Table 1, we show in Figure 2 their kernel densities (Epanechnikov kernel). As we can observe, there are clear differences among the estimated cost efficiencies. In general, we can see as UML–NE overestimates cost efficiencies regarding RML–NED estimates.

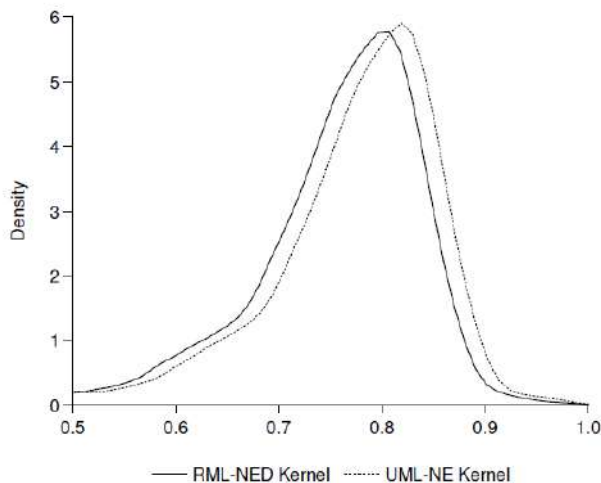


Figure 2: Kernel densities for estimated cost efficiencies.

Therefore, we can conclude that our data confirm a value added of our parametric specification and its practical relevance, that is, RML–NED is better fitted than UML–NE when the assumption of dependence is rejected.

4 Concluding remarks

In this paper, we have propose a new stochastic frontier model which introduces a flexible correlation structure between the probability of the normal error term, ν , and the inefficiency term, u . The formulae derived are closed form expressions for the marginal density of the estimated error, which allow us to introduce the classical assumptions applicable to the idiosyncratic error and the inefficiency term.

Our approach presents the following advantages: (i) Probabilistic interpretation, in the context of stochastic frontier models, is straightforward. (ii) It adds flexibility to the model by taking into account both the effects of the independent case and the correlation among variables. Furthermore, it allows a wider range of dependence. Any correlation sign is allowed, including the possibility of negative correlation among variables, thus reducing the possibility of misspecification. (iii) It is easily modelled in a maximum likelihood framework to test the dependence assumption, and it is computationally simpler than other models like, for example, Frank and Plackett copulas. (iv) It focuses on the classical case normal–exponential, but can also be extended to other distributions, including the normal–truncated normal and normal–gamma, among others.

Acknowledgements

The authors would like to express their gratitude to two anonymous referees for their relevant and useful comments.

The research was partially funded by Emilio Gómez–Déniz: ECO2013–47092 (Ministerio de Economía y Competitividad, Spain) and J.V. Pérez–Rodríguez: ECO2011–23189 (Ministerio de Economía y Competitividad, Spain). Also, J.V. Pérez–Rodríguez acknowledges the Department of Statistics and CRiSM, at the University of Warwick for their special support, since part of this paper was written while he was visiting the University of Warwick in 2013.

References

- Aigner, D., Knox Lovell, C.A. and Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Function Models. *Journal of Econometrics*, 6, 21–37.
- Amblard, C. and Girard, S. (2009). A new extension of bivariate FGM copulas. *Metrika*, 70, 1–17.
- Amsler, C., Prokhorov, A. and Schmidt P. (2016). Endogeneity in stochastic frontier models. *Journal of Econometrics*, 190:2, 280–288.
- Battese, G. and Corra, G. (1977). Estimation of a production frontier model: With application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics*, 21:3, 169–179.

- Battese, G. E. and T. J. Coelli. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 38, 387–99.
- Bonett, D.G. (2000). Sample size requirements for estimating Pearson, Kendall and Spearman correlations. *Psychometrika*, 65:1, 23–28.
- Dingman, H.F. and Perry, N.C. (1956). A comparison of the accuracy of the formula for the standard error of Pearson “r” with the accuracy of Fishers z-Transformation. *The Journal of Experimental Education*, 24:4, 319–321.
- El Mehdi, R. and Hafner, C.M. (2014). Inference in stochastic frontier analysis with dependent error terms. *Mathematics and Computers in Simulation*, 102, 104–116.
- Fredricks, G.A. and Nelsen, R.B. (2007). On the relationship between Spearman’s rho and Kendall’s tau for pairs of continuous random variables. *Journal of Statistical Planning and Inference*, 137, 2143–2150.
- Gómez-Déniz, E. and Pérez-Rodríguez, J.V. (2014). Closed-form solution for a bivariate distribution in stochastic frontier models with dependent errors. *Journal of Productivity Analysis*, 43:2, 215-223.
- Greene, W. (1980a). Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics*, 13:1, 27–56.
- Greene, W. (1980b). On the estimation of a flexible frontier production model. *Journal of Econometrics*, 13:1, 101–115.
- Greene, W. (1990). A Gamma distributed stochastic frontier model. *Journal of Econometrics*, 46:1, 141–164.
- Greene, W. (2003). Maximum simulated likelihood estimation of the Normal-Gamma stochastic frontier function. *Journal of Productivity Analysis*, 19:2–3, 179–190.
- Jondron, J.; Lovell, C.A.; Materov, I.S. and Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19, 233–238.

- Kumbhakar, S.C. and Lovell, C. A. (2000). *Stochastic Frontiers Analysis*. Cambridge University Press.
- Lee, L.-F. (1983). A test for distributional assumptions for the stochastic frontier functions. *Journal of Econometrics*, 22:3, 245–267.
- Lee, T.M.L. (1996). Properties and applications of the Sarmanov family of bivariate distributions. *Communications Statistics: Theory and Methods*, 25:6, 1207–1222.
- Meeusen, W. and Van Den Broeck, J. (1977). Efficiency estimation from Cobb–Douglas production function with composed error. *International Economic Review*, 18, 435–444.
- Park, Y.-H. and Fader, P.S. (2004). Modeling browsing behavior at multiple websites. *Marketing Science*, 23:3, 280–303.
- Sarabia, J.M. and Gómez–Déniz, E. (2011). Multivariate Poisson–Beta distributions with applications. *Communications in Statistics: Theory and Methods*, 40, 1093–1108.
- Sarmanov, O.V. (1966). Generalized normal correlation and two–dimensional Frechet classes. *Doklady (Soviet Mathematics)*, 168, 596–599.
- Shubina, M. and Lee, T.M.L. (2004). On maximum attainable correlation and other measures of dependence for the Sarmanov family of bivariate distributions. *Communications in Statistics: Theory and Methods*, 33:5, 1031–1052.
- Smith, M. (2008). Stochastic frontier models with dependent error components. *Econometrics Journal*, 11, 172–192.
- Stevenson, R. (1980). Likelihood functions for generalized stochastic frontier functions. *Journal of Econometrics*, 13, 57–66.
- Tran, K. and Tsionas, M. (2015). Endogeneity in stochastic frontier models: Copulas approach without external instruments. *Economics Letters*, 133, 85–88.
- Wiboonpongse, A.; Liu, J.; Sriboonchitta, S. and Denoeux, T. (2015). Modeling dependence between error components of the stochastic frontier model using copula: Application to intercrop coffee production in Northern Thailand. *International Journal of Approximate Reasoning*, 65, 34–44.

Appendix 1. Proof of Theorem 1

From (4) we have

$$\begin{aligned} f(\varepsilon) &= \int_0^\infty f_{\sigma_u, \sigma_\nu, \omega}(u, \varepsilon) du \\ &= \int_0^\infty f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) du + \omega \int_0^\infty f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(u + \varepsilon) du. \end{aligned} \quad (14)$$

The first integral in (14) coincides with (1). Then, we have

$$\begin{aligned} f(\varepsilon) &= [1 + \omega \delta_u(\sigma_u) \delta_\nu(\sigma_\nu)] \frac{1}{\sigma_u} \Phi \left(-\frac{\varepsilon}{\sigma_\nu} - \frac{\sigma_\nu}{\sigma_u} \right) \exp \left\{ \frac{\varepsilon}{\sigma_u} + \frac{\sigma_\nu^2}{2\sigma_u^2} \right\} \\ &\quad - \omega (\mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3), \end{aligned}$$

where $\delta_u(\sigma_u)$, $\delta_\nu(\sigma_\nu)$ are given in (5), (6), respectively, while

$$\begin{aligned} \mathcal{J}_1 &= \int_0^\infty e^{-u-(u+\varepsilon)^2-2(u+\varepsilon)} f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) du, \\ \mathcal{J}_2 &= \delta_\nu(\sigma_\nu) \int_0^\infty e^{-u} f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) du, \\ \mathcal{J}_3 &= \delta_u(\sigma_u) \int_0^\infty e^{-(u+\varepsilon)^2-2(u+\varepsilon)} f_{\sigma_u}(u) f_{\sigma_\nu}(u + \varepsilon) du. \end{aligned}$$

Again, simple but tedious computations lead to the following:

$$\begin{aligned} \mathcal{J}_1 &= \frac{1}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} \Upsilon_{\sigma_u, \sigma_\nu}^1(\varepsilon) \Upsilon_{\sigma_u, \sigma_\nu}^4(\varepsilon) \Phi \left(-\frac{\sigma_\nu}{\sigma_u} \frac{1 + 3\sigma_u}{\sqrt{1 + 2\sigma_\nu^2}} - \frac{\varepsilon}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} \right), \\ \mathcal{J}_2 &= \frac{1}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} \Upsilon_{\sigma_u, \sigma_\nu}^1(\varepsilon) \Upsilon_{\sigma_u, \sigma_\nu}^2(\varepsilon) \Phi \left(-\frac{\varepsilon}{\sigma_\nu} - \frac{\sigma_\nu}{\sigma_u} - \sigma_\nu \right), \\ \mathcal{J}_3 &= \frac{1}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} \Upsilon_{\sigma_u, \sigma_\nu}^1(\varepsilon) \Upsilon_{\sigma_u, \sigma_\nu}^{3p}(\varepsilon) \Phi \left(-\frac{\sigma_\nu}{\sigma_u} \frac{1 + 2\sigma_u}{\sqrt{1 + 2\sigma_\nu^2}} - \frac{\varepsilon}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} \right), \end{aligned}$$

from which (9) is obtained. ■

Appendix 2. Proof of Proposition 1

We start with the wellknown relation

$$f_{\sigma_u, \sigma_\nu, \omega}(u|\varepsilon) = \frac{f_{\sigma_u, \sigma_\nu, \omega}(u, \varepsilon)}{f_{\sigma_u, \sigma_\nu, \omega}(\varepsilon)}$$

and replace in this expression the numerator and the denominator by (4) and (9), respectively. Now by taking the common factor for ω in both, numerator and denominator, we get the result after some computations.

Appendix 3. Proof of Proposition 2

From (9) it is easy to see

$$E(e^{-u}|\varepsilon) \propto \int_0^\infty e^{-u} f_{\sigma_u, \sigma_\nu}(u|\varepsilon) du + \omega \int_0^\infty e^{-u} \Psi_{\sigma_u, \sigma_\nu}^0(u|\varepsilon) du, \quad (15)$$

where the proportionality factor is given by $[1 + \omega \Psi_{\sigma_u, \sigma_\nu}^1(\varepsilon)]^{-1}$.

The first integral in (15) is simple to solve by using the pdf of the half-normal provided in (2). For the second integral we use the expression given in (11) and with simple but tedious computations we get the result.

Appendix 4. The stochastic cost frontier model

The corresponding expressions for the stochastic cost frontier (SCF) are derived easily from the fact that now we assume $v = -u + \varepsilon$.

In this case and under the classical model we have that,

$$\begin{aligned} f_{\sigma_u, \sigma_\nu}(\varepsilon) &= \frac{1}{\sigma_u} \Phi\left(\frac{\varepsilon}{\sigma_\nu} - \frac{\sigma_\nu}{\sigma_u}\right) \exp\left\{-\frac{\varepsilon}{\sigma_u} + \frac{\sigma_\nu^2}{2\sigma_u^2}\right\}, \\ f_{\sigma_u, \sigma_\nu}(u|\varepsilon) &= \frac{1}{\sqrt{2\pi} \sigma_\nu \Phi(\tilde{\mu}/\sigma_\nu)} \exp\left\{-\frac{1}{2\sigma_\nu^2}(u - \tilde{\mu})^2\right\}, \end{aligned} \quad (16)$$

where $\tilde{\mu} = \varepsilon - \sigma_\nu^2/\sigma_u$.

Again, the marginal $f(\varepsilon)$ is asymmetrically distributed with mean $E(\varepsilon) = \sigma_u$ and variance $var(\varepsilon) = \sigma_u^2 + \sigma_\nu^2$.

The corresponding expressions under the dependence assumption are given bellow.

The marginal pdf of ε is given by

$$\begin{aligned} f_{\sigma_u, \sigma_\nu, \omega}(\varepsilon) &= \frac{\Xi_{\sigma_u, \sigma_\nu}^1(\varepsilon)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} \left\{ \Phi\left(\frac{\tilde{\mu}}{\sigma_\nu}\right) \left[\sqrt{1 + 2\sigma_\nu^2} + \frac{\omega}{1 + \sigma_u} \exp\left(\frac{2\sigma_\nu^2}{1 + 2\sigma_\nu^2}\right) \right] \right. \\ &\quad - \omega \left[\sqrt{1 + 2\sigma_\nu^2} \Xi_{\sigma_u, \sigma_\nu}^2(\varepsilon) \delta_\nu(\sigma_\nu) \Phi\left(\frac{\tilde{\mu}}{\sigma_\nu} - \sigma_\nu\right) \right. \\ &\quad \left. \left. + \Xi_{\sigma_u, \sigma_\nu}^3(\varepsilon) \delta_u(\sigma_u) \Phi\left(\frac{\sigma_\nu(2\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu}\right) \right] \right. \\ &\quad \left. - \Xi_{\sigma_u, \sigma_\nu}^4(\varepsilon) \Phi\left(\frac{\sigma_\nu(\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu}\right) \right\}, \end{aligned} \quad (17)$$

where

$$\Xi_{\sigma_u, \sigma_\nu}^1(\varepsilon) = \exp\left\{-\frac{\varepsilon}{\sigma_u} + \frac{\sigma_\nu^2}{2\sigma_u^2}\right\}, \quad (18)$$

$$\Xi_{\sigma_u, \sigma_\nu}^2(\varepsilon) = \exp \left\{ -\varepsilon + \frac{\sigma_\nu^2}{\sigma_u} + \frac{\sigma_\nu^2}{2} \right\}, \quad (19)$$

$$\Xi_{\sigma_u, \sigma_\nu}^3 = \exp \left\{ \frac{\sigma_\nu^2}{2\sigma_u^2} \frac{4\sigma_u(\sigma_u - 1) - 2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}, \quad (20)$$

$$\Xi_{\sigma_u, \sigma_\nu}^4(\varepsilon) = \exp \left\{ -\varepsilon + \frac{\sigma_\nu^2}{2\sigma_u^2} \frac{\sigma_u(\sigma_u - 2) - 2\sigma_\nu^2}{1 + 2\sigma_\nu^2} \right\}. \quad (21)$$

The conditional pdf of u given ε is given by

$$f_{\sigma_u, \sigma_\nu, \omega}(u|\varepsilon) = \frac{f_{\sigma_u, \sigma_\nu}(u|\varepsilon) + \omega \mathfrak{S}_{\sigma_u, \sigma_\nu}^0(u|\varepsilon)}{1 + \omega \mathfrak{S}_{\sigma_u, \sigma_\nu}^{1c}(\varepsilon)}, \quad (22)$$

for $u > 0$, where $f_{\sigma_u, \sigma_\nu}(u|\varepsilon)$ is the pdf given in (16) and

$$\begin{aligned} \mathfrak{S}_{\sigma_u, \sigma_\nu}^0(u|\varepsilon) &= \varphi_{\sigma_u}(u) \varphi_{\sigma_\nu}(u + \varepsilon) f_{\sigma_u, \sigma_\nu}^c(u|\varepsilon), \\ \mathfrak{S}_{\sigma_u, \sigma_\nu}^1(\varepsilon) &= \delta_u(\sigma_u) \delta_\nu(\sigma_\nu) - \frac{1}{\sqrt{1 + 2\sigma_\nu^2} \Phi\left(\frac{\tilde{\mu}}{\sigma_\nu}\right)} \times \left[\sqrt{1 + 2\sigma_\nu^2} \Xi_{\sigma_u, \sigma_\nu}^2(\varepsilon) \delta_\nu(\sigma_\nu) \Phi\left(\frac{\tilde{\mu}}{\sigma_\nu} - \sigma_\nu\right) \right. \\ &\quad + \Xi_{\sigma_u, \sigma_\nu}^3(\varepsilon) \delta_u(\sigma_u) \Phi\left(\frac{\sigma_\nu(2\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu}\right) \\ &\quad \left. - \Xi_{\sigma_u, \sigma_\nu}^4(\varepsilon) \Phi\left(\frac{\sigma_\nu(\sigma_u - 1)}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}} + \frac{\varepsilon \sqrt{1 + 2\sigma_\nu^2}}{\sigma_\nu}\right) \right]. \end{aligned}$$

The point estimation for the efficiency in the SCF model under NE distributions, assuming dependence, is given by

$$CE_i = \frac{\mathfrak{S}_{\sigma_u, \sigma_\nu, \omega}^{01}(\varepsilon_i) + \omega \mathfrak{S}_{\sigma_u, \sigma_\nu}^{02}(\varepsilon_i)}{1 + \omega \mathfrak{S}_{\sigma_u, \sigma_\nu}^1(\varepsilon_i)}, \quad i = 1, 2, \dots, n,$$

where

$$\mathfrak{S}_{\sigma_u, \sigma_\nu, \omega}^{01}(\varepsilon_i) = [1 + \omega \delta_u(\sigma_u) \delta_\nu(\sigma_\nu)] \frac{\Phi\left(\frac{\tilde{\mu}}{\sigma_\nu} - \sigma_\nu\right)}{\Phi\left(\frac{\tilde{\mu}}{\sigma_\nu}\right)} e^{-\varepsilon_i + \sigma_\nu^2/\sigma_u + \sigma_\nu^2/2}$$

and

$$\begin{aligned} \mathfrak{S}_{\sigma_u, \sigma_\nu}^{02}(\varepsilon_i) &= \frac{1}{\mathfrak{S}\left(\frac{\tilde{\mu}}{\sigma_\nu}\right)} \left[\frac{1}{\sqrt{1 + 2\sigma_\nu^2}} \Xi_{\sigma_u, \sigma_\nu}^{11}(\varepsilon_i) \Phi\left(\frac{\varepsilon_i}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} - \frac{\sigma_\nu}{\sigma_u \sqrt{1 + 2\sigma_\nu^2}}\right) \right. \\ &\quad - \delta_\nu(\sigma_\nu) \Xi_{\sigma_u, \sigma_\nu}^{12}(\varepsilon_i) \Phi\left(\frac{\tilde{\mu}}{\sigma_\nu} - 2\sigma_\nu\right) \\ &\quad \left. - \frac{\delta_u(\sigma_u)}{\sqrt{1 + 2\sigma_\nu^2}} \Xi_{\sigma_u, \sigma_\nu}^{13}(\varepsilon_i) \Phi\left(\frac{\varepsilon_i}{\sigma_\nu} \sqrt{1 + 2\sigma_\nu^2} + \frac{\sigma_u - 1}{\sigma_u} \frac{\sigma_\nu}{\sqrt{1 + 2\sigma_\nu^2}}\right) \right], \end{aligned}$$

with

$$\begin{aligned}\Xi_{\sigma_u, \sigma_\nu}^{11}(\varepsilon_i) &= \exp \left\{ -2\varepsilon_i - \frac{\sigma_\nu^4}{(1 + 2\sigma_\nu^2)\sigma_u^2} \right\}, \\ \Xi_{\sigma_u, \sigma_\nu}^{12}(\varepsilon_i) &= \exp \left\{ -2\varepsilon_i + \frac{2\sigma_\nu^2}{\sigma_u}(1 + \sigma_u) \right\}, \\ \Xi_{\sigma_u, \sigma_\nu}^{13}(\varepsilon_i) &= \exp \left\{ -\varepsilon_i + \frac{\sigma_\nu^2(\sigma_u(\sigma_u - 2) - 2\sigma_\nu^2)}{2\sigma_u^2(1 + 2\sigma_\nu^2)} \right\},\end{aligned}$$

while $\delta_u(\sigma_u)$ and $\delta_\nu(\sigma_\nu)$ are given in (5), (6), respectively.

Hacia una visión holística de la distancia en los negocios internacionales: el caso colombiano

CAICEDO MARULANDA, CAROLINA

Departamento de Ciencias Administrativas
Universidad Autónoma de Occidente, Cali (Colombia)
Correo electrónico: ccaicedo@uao.edu.co

PLA BARBER, JOSÉ

Departamento de Dirección de Empresas
Universidad de Valencia (España)
Correo electrónico: jose.pla@uv.es

LEÓN DARDER, FIDEL

Departamento de Dirección de Empresas
Universidad de Valencia (España)
Correo-e: fidel.leon@uv.es

MORA RODRÍGUEZ, JHON JAMES

Departamento de Economía
Universidad Icesi, Cali (Colombia)
Correo-e: jjmora@icesi.edu.co

RESUMEN

Este artículo presenta nuevas estimaciones de distancia en los negocios internacionales entre Colombia y 57 países de Europa, América, Asia, África y Oceanía. El concepto de distancia que utilizamos trasciende el plano de lo institucional e incorpora tanto aspectos institucionales como aquellos que aún no han sido considerados en la literatura como el mercado laboral, la práctica mercantil y la innovación. De esta forma, a partir de la construcción de un índice sintético de distancia nosotros calculamos las distancias entre Colombia y los países de la muestra usando el concepto de distancia de Mahalanobis en las dimensiones aquí analizadas. Nuestros resultados muestran que si bien hay avances en la globalización sobre los negocios internacionales entre Colombia y los países de la muestra, todavía existen restricciones a las que se enfrentan las empresas colombianas que están interesadas en incursionar estos mercados buscando beneficiarse de los acuerdos de libre comercio.

Palabras claves: distancia holística; análisis factorial; densidad de *Kernel*; negocios internacionales.

Clasificación JEL: C23; D22; J25.

MSC2010: 62E17; 62Q05; 62P20.

Artículo recibido el 3 de marzo de 2015 y aceptado el 30 de noviembre de 2016.

Towards Holistic Version of Distance in International Business: The Colombian Case

ABSTRACT

This article presents new estimations of the distance in international business between Colombia and 57 countries in Europe, America, Asia, Africa and Oceania. The concept of distance that we use transcends the institutional level and incorporates both institutional aspects and those that have not yet been considered in the literature such as labor market, business practice and innovation. In this way, we build a holistic distance index and discuss the results for the specific case of Colombia. Thus, from the construction of a synthetic index of distance, we calculated the distances between Colombia and the countries of the sample using the Mahalanobis distance in the dimensions analyzed here. Our results show that although there is progress in the globalization of international business between Colombia and the countries of the sample there are still restrictions faced by Colombian companies that are interested in entering these markets seeking to benefit from the free trade agreements.

Keywords: Holistic distance; factor analysis; kernel density; international business.

JEL classification: C23; D22; J25.

MSC2010: 62E17; 62Q05; 62P20.



1.- Introducción

Con el fin de analizar la distancia en los negocios internacionales hay que entender el conjunto de situaciones que enfrentan las empresas cuando incursionan en mercados extranjeros y enfrentan situaciones y barreras que limitan su posición competitiva en los mercados. La investigación y el estudio de las actividades de las empresas multinacionales en el extranjero, en los últimos años, han mostrado que tanto la distancia espacial (debido a los costos derivados de las barreras arancelarias y fletes internacionales) como las condiciones del mercado bajo los cuales compite y los comportamientos discriminatorios que las empresas domésticas exhiben (a través de competidores directos, proveedores, distribuidores e instituciones gubernamentales, que colocan en franca desventaja a las empresas extranjeras frente a las domésticas) afectan la actividad internacional de la firma. Esta condición se conoce como *la desventaja de ser extranjero* que es el efecto agregado de la interacción de la firma extranjera con todos los elementos del ambiente de los negocios internacionales (Sethi y Guisinger, 2002). Este efecto surge debido no solo a la falta de conocimiento o mala lectura del ambiente de los negocios internacionales sobre aspectos tales como el mercado local, las prácticas mercantiles domésticas y las regulaciones propias del nuevo mercado en el que incursiona la firma (Hymer, 1960), sino también por la formulación e implementación de estrategias que no son compatibles con ese ambiente.

Para mitigar la desventaja de ser extranjero, las empresas deben adquirir conocimiento institucional, el cual se define como el conocimiento que una empresa posee de la cultura de un país, de su lengua, tradiciones, normas de comportamiento, sistemas educativos, leyes, regulaciones y costumbres comerciales (Kostova y Zaheer, 1999; Eriksson *et al.*, 1997). Las estrategias para adquirir este conocimiento pueden ser varias: por ejemplo, la experiencia propia de la empresa y su operación en el mercado huésped, lo que le permite a la empresa acumular directamente conocimiento sobre sus clientes, los recursos necesarios para su operación, las regulaciones, leyes y aspectos legales y las prácticas de los negocios en ese mercado (Johanson y Vahlne, 1977). El problema que surge con esta estrategia es que requiere de una gran cantidad de recursos, un alto compromiso y un tiempo considerable para la adquisición de conocimiento y desarrollo de capacidades. Otra estrategia que implementan las firmas consiste en las formas (modos) de entrada en los países. De esta forma, las alianzas, por ejemplo, ofrecen una alternativa para las empresas extranjeras que tienen poco conocimiento sobre el mercado huésped y que no pueden comprometer grandes recursos. Las alianzas con actores locales que tienen poder de mercado eliminan, de manera paulatina, pero más rápida, la desventaja que enfrentan estas empresas (Kostova *et al.*, 2008). Hacer fusiones y adquisiciones ofrece los mismos beneficios, sin embargo requieren de un compromiso mayor de recursos y, si la fusión o adquisición implica la desaparición de la empresa local, esto aumentará la desventaja de ser extranjero. La empresa también puede optar por aquellos mercados que tengan ambientes similares al de su país de origen, comerciar los mismos productos y desarrollar economías de escala. En la medida en que las empresas ajusten su conocimiento y capacidades a los requerimientos del ambiente local, logran mitigar los costos, disminuir el riesgo y bajar los niveles de incertidumbre de las operaciones en el extranjero.

Cabe observar que, aun cuando la distancia en los negocios internacionales se ha investigado considerablemente tanto desde la economía internacional como desde los negocios

internacionales, no existe un consenso en torno a una definición única y una medida uniforme para medir las diferencias en los ambientes de los negocios internacionales en los cuales operan las empresas multinacionales, lo cual constituye un reto en la investigación sobre estas empresas.

De esta forma, este artículo contribuye a la literatura sobre los negocios internacionales al considerar que la distancia en los negocios trasciende el plano de lo institucional y desarrolla una forma de medición que incorpora, además de aspectos institucionales, características del mercado laboral, la práctica mercantil y la innovación.

Este artículo se estructura de la siguiente manera: la sección 1 es introductoria; en la sección 2 se revisa la literatura previa sobre distancia; en la sección 3 se plantean las dimensiones de la distancia; en la sección 4, la metodología; en la sección 5 se analizan los datos y los principales resultados; y, finalmente, la sección 6 presenta las conclusiones sobre el caso colombiano y líneas futuras de investigación.

2.- La investigación sobre la distancia en los negocios internacionales

El concepto de distancia asociado a las actividades internacionales ha sido tratado desde la economía internacional y los negocios internacionales. Desde la economía internacional la investigación se ha centrado en el papel de la distancia geográfica sobre el comercio internacional. Disdier y Head (2008) consideran que en el modelo de proporciones factoriales existen fuertes presiones para la igualación de precios salvo en el caso de separación espacial, rendimientos crecientes a escala, aislamiento geográfico, diferencias espaciales en salarios y cambios tecnológicos (Berthelon y Freund, 2008).

La inclusión de la distancia en la literatura económica sobre comercio internacional es reciente (De Groot *et al.*, 2004; Bénassy-Quéré *et al.*, 2007). Se ha incluido en estudios sobre las barreras no observadas en el comercio, elementos a los que North (1990, 1995) atribuye el hecho de que las personas formen instituciones, y reglas que reducen la incertidumbre y los costos de transacción (De Groot *et al.*, 2004). Es así como Hall y Jones (1999), Olson (1996) y Knack y Keefer (1995) discuten el efecto que tienen las instituciones sobre el crecimiento y el desarrollo económico y concluyen que una pobre gobernanza generará externalidades negativas sobre las transacciones privadas y, por ende, sobre los costos de transacción. Anderson y Marcouiller (2002) analizan el efecto de las instituciones sobre el comercio y sus resultados muestran que instituciones inadecuadas limitan al comercio tal como lo haría el establecimiento de aranceles. Koukhartchouk y Maurel (2003) analizan los efectos de instituciones internacionales conjuntas como la Organización Mundial de Comercio y la Unión Europea sobre los patrones de comercio. Sus resultados muestran que existe una relación de auto-reforzamiento entre mejores instituciones y comercio y que la apertura al comercio tendrá un impacto sobre el crecimiento solo si las instituciones hacen al comercio atractivo y rentable para todas las partes creando el ambiente que conduzca hacia intercambios seguros y fiables. Para De Groot *et al.* (2004), la calidad institucional tiene un efecto significativo y positivo sobre los flujos de comercio bilaterales, al igual que la similaridad en la calidad de gobernanza. Adicionalmente, la literatura sobre instituciones e inversión extranjera directa (IED en adelante) se ha orientado principalmente al estudio del impacto de las “buenas” instituciones sobre los flujos de entrada de la IED y, en menor medida, sobre el impacto de la distancia institucional entre el país emisor y el país receptor. Wheeler y Mody (1992) y Wei (2000) desarrollan un conjunto de variables para medir la calidad de las

instituciones tales como la corrupción, la inestabilidad política y la calidad del sistema legal pero no encuentran un impacto significativo sobre las decisiones de localización (Wheeler y Mody, 1992), salvo la corrupción, la cual resulta un impedimento significativo para la entrada de la IED (Wei, 2000). Globerman y Shapiro (2002), empleando los indicadores de gobernanza de Kaufman *et al.* (1999), encontraron un impacto positivo del buen gobierno sobre los flujos de entrada y salida de la IED pero solo en los países relativamente grandes y desarrollados.

Por su parte, la inclusión de la distancia en la literatura sobre los negocios internacionales parte de considerar las diferencias entre países con respecto a las actividades de emprendimiento, lo cual muestra la relevancia de entender la distinción entre las dimensiones del perfil institucional (Busenitz *et al.*, 2000). Kostova y Roth (2002) muestran que la implementación e internalización de la adopción de una práctica organizacional varía entre las filiales en el extranjero debido a dos factores: el ambiente institucional en el país huésped y el contexto relacional dentro de la multinacional. Xu *et al.* (2004) muestran que a mayor distancia normativa y regulatoria, niveles más bajos de propiedad accionaria y baja presencia de expatriados. Por su parte, Chao y Kumar (2010) encuentran que la distancia regulatoria es el factor que más obstaculiza las operaciones de las empresas en los mercados extranjeros; adicionalmente encuentran que los efectos culturales sobre el desempeño de la firma no tienen una relación significativa con la elección del modo de entrada, la diversidad internacional y el desempeño de la multinacional.

3.- Las dimensiones de la distancia

El concepto de distancia se ha centrado en aspectos institucionales como el normativo, el regulatorio, el cognitivo y la distancia geográfica. Bae y Salomon (2010) encuentran que las dimensiones de distancia institucional más estudiadas son la política, regulatoria, económica, cultural y cognitiva. Berry *et al.* (2010) considera tres dimensiones: demográfica, conectividad global y conocimiento. Berry *et al.* (2010) separan los aspectos administrativos de la dimensión política mientras que los aspectos financieros son considerados independientemente de la dimensión económica.

Las dimensiones identificadas en los estudios varían dependiendo del concepto de instituciones empleado y de la disciplina desde la que se define. Los estudios sociológicos emplean más las dimensiones normativas y cognitivas, mientras que los estudios económicos tienden a emplear las dimensiones económicas, políticas y geográficas (Ghemawat, 2001). A continuación, se analizarán estas dimensiones.

La *dimensión política* refleja una parte de las instituciones formales que se refiere a los gobiernos e instituciones políticas. Aspectos tales como políticas de gobierno, la inestabilidad política, la calidad de las instituciones de gobierno, la debilidad institucional, la hostilidad política, la transparencia, los vínculos coloniales y los acuerdos monetarios o pertenencia a asociaciones políticas, son variables a considerar cuando se mide la distancia política (Martin *et al.*, 2010). Cuando las empresas extranjeras entran a mercados de países políticamente distantes es más difícil desarrollar sus actividades comerciales. En este sentido, los países pueden establecer medidas unilaterales que aumenten la distancia política convirtiéndose en las barreras más comunes a la competencia a través de las fronteras (Ghemawat, 2001).

La *dimensión regulatoria* refleja la otra parte de las instituciones formales y mide las diferencias a través de los países en términos de la promulgación y ejecución de las

regulaciones sancionadas por el Estado (North, 1990). Las regulaciones normalizan las actividades de las organizaciones (Scott, 1995) y por tanto, una comparación de las regulaciones a través de los países es importante ya que las empresas podrán tratar de desarrollar ventajas de la estandarización de prácticas cuando operan en ambientes institucionales con pocas diferencias regulatorias o si por el contrario, las características regulatorias son muy distintas, entonces desarrollar prácticas muy específicas para ese contexto. Los aspectos institucionales considerados con mayor frecuencia en la dimensión regulatoria son los regímenes de derechos de propiedad intelectual, el sistema judicial, el sistema de gobierno y las regulaciones antimonopolio. En términos del nivel de análisis, es más común que se realice a nivel de industrias debido a la alta especificidad en el ambiente regulatorio a la que están expuestas las actividades industriales y de servicios; ejemplo de esto son las regulaciones existentes para el sector bancario (Miller y Parkhe, 2002), telecomunicaciones (Perkins, 2014) y “*utilities*” (Zelner *et al.*, 2009). Sin embargo, a nivel país también se han realizado estudios que incluyen la dimensión regulatoria considerando empresas multisectoriales pero sobre aspectos organizacionales diversos de las filiales (Gaur *et al.*, 2007).

La *dimensión cultural* compara países basados en sus perfiles culturales (Kogut y Singh, 1988; Hofstede, 2001). La cultura puede definirse como “el conjunto de normas y valores implícitos a un colectivo humano que lo dotan de identidad y dan sentido a las conductas de sus integrantes” (Pla-Barber y León-Darder, 2004:231). Las diferencias fundamentales entre las normas y valores del país de origen y sus operaciones en el país huésped a menudo crean dificultades operacionales para las empresas e incrementan los esfuerzos requeridos para entrar a dicho país, aspectos tales como diferencias idiomáticas, religiosas, manejo del tiempo, actitud hacia el trabajo, afectan las condiciones de negociación de las firmas entrantes y pueden deteriorar su posición competitiva. Esta dimensión es, quizás, la más empleada en los estudios sobre los negocios internacionales, para evaluar su impacto o establecer su relación con múltiples aspectos, tales como modos de entrada (Drogendijk y Slangen, 2006; Chang *et al.*, 2011), aprendizaje (Barkema *et al.*, 1996), diversidad internacional (Chao y Kumar, 2010), modernización y cambio cultural (Inglehart y Baker, 2000), y desempeño (Morosini *et al.*, 1998). Esta dimensión se ha operacionalizado a través del constructo de distancia cultural que mide el grado en el cual las culturas son similares o diferentes (Shenkar, 2001). La *dimensión cognitiva* es la dimensión que refleja la diferencia en estructuras cognitivas arraigadas en la sociedad (Scott, 1995). Dichas estructuras comprenden rutinas, marcos y guiones empleados por los individuos en una sociedad para juzgar y asignar significado a un fenómeno y resolver problemas (Markus y Zajonc, 1985). Esta dimensión suele ser empleada en los estudios sobre transferencias de prácticas entre unidades organizacionales y entre filiales. Bae y Salomon (2010) muestran que es común tratar los aspectos de la dimensión cognitiva conjuntamente con los aspectos de la dimensión cultural, tal como se hará en este artículo.

La *dimensión económica* considera las diferencias a través de los países en patrones de intercambio, estructura económica, orientación del mercado y estabilidad del mercado (Ghemawat, 2001). Desde la teoría sobre el comercio internacional se han usado variables macroeconómicas para mostrar las diferencias económicas entre los países del mundo como el producto interno bruto (PIB en adelante), el PIB per cápita, tasas de crecimiento del PIB, las exportaciones y las importaciones como porcentaje del PIB, la intensidad en inversión, características del mercado cambiario, pertenencia a bloques comerciales, entre otras. Bae y Salomon (2010) argumentan que la dimensión económica no ha tenido mucha atención debido

a que la forma en que los mercados trabajan afecta directamente a la dimensión política y regulatoria e indirectamente a la cultural.

La *dimensión financiera* considera los diferentes sistemas financieros que evolucionan en el tiempo y que impactan la forma de hacer negocios de las compañías (La Porta *et al.*, 1998) en aspectos tales como la cobertura del riesgo, los medios de pago, el costo de los productos financieros, la sofisticación de los servicios prestados y su intermediación en el mercado de divisas. Su instrumentación la hacen a través de variables como el crédito doméstico al sector privado, la capitalización del mercado y el número de compañías.

La *dimensión demográfica*, introducida inicialmente por Whitley (1992), se define como las diferencias demográficas entre los países en términos del tamaño, crecimiento, estructura de edad y cualidades de sus poblaciones. Estos aspectos demográficos son vitales en los negocios internacionales ya que dan cuenta del tamaño del mercado objetivo, de la segmentación del mercado y de su potencial de explotación en el futuro.

La *dimensión geográfica* plantea que existen diferencias en los costos de la distribución física internacional, la información y comunicación. Berry *et al* (2010) calculan la distancia geográfica usando el método del gran círculo propuesto por Shaxby (1945); el método usa la latitud y longitud de la ciudad principal en cada región o país y se basa sobre el supuesto de que la tierra es una esfera real (Fitzpatrick y Modlin, 1986). Sin embargo algunos autores han argumentado que las rutas del gran círculo a menudo difieren sustancialmente de las rutas reales de carga (Disdier y Head, 2008).

La dimensión *innovación* plantea que los países difieren en términos de su capacidad para crear conocimiento e innovar y cómo esto tiene fuertes implicaciones sobre la economía global (Furman *et al.*, 2002). La innovación se constituye como un determinante del crecimiento económico, la competitividad nacional y como un elemento clave para la atracción de IED de empresas innovadoras (Hasan y Tucci, 2010).

La *dimensión conectividad global* se define como la habilidad de los residentes de un país y de sus empresas para interactuar con otras partes del mundo, obtener información y difundir sus propias actividades (Oxley y Yeung, 2001). Ghemawat y Altman (2014) consideran como variables fundamentales los flujos transfronterizos de personas, información, bienes, servicios y capital.

En este artículo consideraremos tres dimensiones nuevas: la dimensión práctica mercantil, la dimensión comercial y la dimensión mercado laboral.

La *dimensión práctica mercantil* considera una serie de aspectos relacionados intrínsecamente con “el hacer” de los negocios internacionales, tales como los costos burocráticos, la facilidad de hacer negocios, el tiempo requerido para iniciar un negocio, etc. Deardorff (1987) plantea que este tipo de barreras al comercio surgen debido a limitaciones de tipo institucional tales como las reglas de la Organización Mundial de Comercio y las constituciones nacionales que limitan el uso de tarifas o medidas arancelarias o en consideración de las posibles reacciones o retaliaciones en términos de las políticas comerciales que podrían desarrollar los socios comerciales. Deardorff y Stern (1985) consideran que las barreras no arancelarias más comunes son los procedimientos administrativos, las regulaciones del gobierno, la estructura del mercado y los factores institucionales.

La *dimensión comercial* generalmente ha estado incluida en la dimensión económica trayendo consigo un problema de agregación, en la medida en que mezcla las características de la política económica interna de un país como la inflación y el crecimiento económico con su política comercial, la cual regula el mantenimiento de sus relaciones económicas con el

resto del mundo. Los aranceles, las cuotas de importación, las restricciones a la inversión extranjera, las políticas de compras gubernamentales y los subsidios, entre otros, son barreras comerciales que tienen su origen en la política comercial arancelaria aplicada por los diferentes gobiernos. La dimensión comercial considera tanto las importaciones como las exportaciones, la tasa de protección arancelaria y la divisa del país.

La *dimensión mercado laboral* es la más novedosa de las dimensiones propuestas. La entrada de empresas multinacionales que poseen ventajas tecnológicas sobre las firmas domésticas genera diferencias en productividad entre las firmas nacionales y extranjeras que influyen directamente los salarios. Otro aspecto importante consiste en que la entrada de estas firmas genera desbordamiento (*spillover*) sobre el mercado laboral. Esto se relaciona con el grado en el cual la nueva tecnología que acompaña a la IED es asimilada por el sector doméstico; si esto ocurre, se incrementará la productividad de los trabajadores cualificados aún más en el sector doméstico lo cual incrementará la demanda de trabajadores cualificados a expensas de los trabajadores no cualificados, lo cual genera un aumento de la brecha salarial entre ambos grupos de trabajadores (Driffield y Taylor, 2000).

De esta forma, la visión holística de distancia en los negocios internacionales que aquí se propone va mucho más allá que las consideraciones que hasta ahora han sido hechas, no solo por la inclusión de nuevas dimensiones sino también por el instrumento cuantitativo propuesto y la información disponible para 57 países del mundo, lo cual aportará a la investigación empírica en negocios internacionales al disponer de doce matrices de distancia en los principales aspectos que afectan a los negocios internacionales.

4.- Metodología

Si bien no hay consenso en la literatura sobre cuál es la mejor forma de medir la distancia entre dos puntos u objetos (Sokal, 1979; Seber, 1984), es claro que cualquier medida de distancia deberá tener las propiedades de: *simetría*, *no-negatividad* y, en el caso de las distancias métricas, *identificación*, *definitud* y *desigualdad triangular* (Mardia *et al.*, 1979; Krzanowski, 1988).

La metodología que proponemos con el fin de medir la distancia entre Colombia y el resto de mundo es la siguiente:

1. Cálculo de índices sintéticos para cada una de las dimensiones propuestas.
2. A partir de las dimensiones, se calculará la distancia utilizando el método de distancia de Mahalanobis.

Con respecto al primer punto, para la asignación de los pesos o de las puntuaciones factoriales en cada dimensión (índice sintético) se utilizará el procedimiento de regresión. Este método supone el cálculo de ecuaciones de regresión múltiple, una por cada dimensión latente, en cada ecuación donde los factores son la variable dependiente y los indicadores las variables independientes. A partir de la regresión anterior se obtienen los coeficientes de regresión parciales y éstos se multiplican por el valor estandarizado de la variable (Cea D'Ancona, 2002) con el fin de calcular el índice sintético para cada dimensión.

Con respecto al segundo punto, en los estudios sobre negocios internacionales uno de los métodos más empleados para medir la distancia es el de Mahalanobis (1930, 1936). Considérese la distancia entre dos poblaciones distintas i y j , y un número de características

relevantes de los individuos en esas poblaciones, x_i y x_j , donde x denota un vector (aleatorio) que contiene las mediciones realizadas bajo estudio. Dado esto, la distancia Mahalanobis vendrá dada por:

$$d_{ij}^2 = (x_i - x_j)^T S^{-1} (x_i - x_j)$$

donde el superíndice T denota la matriz traspuesta, S es la matriz de covarianza común (no singular) en cada grupo i y j , y se asume S^{-1} existe. Aparte de tener en cuenta las correlaciones entre las variables y las diferencias en las varianzas, la distancia de Mahalanobis deberá cumplir con las propiedades de simetría, no negatividad, identificación, definitud, desigualdad triangular, estar relacionada al logaritmo de la verosimilitud de una distribución normal multivariada y tener una escala multidimensional. Mitchell y Krzanowski (1988) sugieren que la distancia de Mahalanobis es una medida apropiada de distancia entre dos distribuciones elípticas que tienen diferentes locaciones pero una forma común. Perkins (2014) empleó esta medida para capturar la distancia en el ambiente regulatorio, planteando que cada país representa un espacio elíptico de dimensiones que tienen localizaciones varias con relación a un punto de referencia.

5.- Datos y resultados

Los datos provienen de la base de datos QoG Standard de *The Quality of Government Institute*, QoG. La base de datos contiene un corte transversal con cobertura global para el año 2002 (o el año más cercano disponible). Sin embargo, algunas variables se obtuvieron de otras fuentes de datos tales como el Banco Mundial o de *Economic Freedom of the World* del *Fraser Institute*. También se emplearon los datos de Diego Puga y Nathan Nunn disponibles en la página web diegopuga.org. Las variables seleccionadas y las dimensiones que componen se presentan en la Tabla 1.

Los datos se encuentran disponibles para 57 países en total y los países incluidos en el estudio por continente presentan la siguiente distribución: 28 países europeos, 12 de América, 10 de Asia, 5 de África y 2 de Oceanía.

Para la asignación de los pesos o de las puntuaciones factoriales para cada dimensión, se utilizó el procedimiento de regresión, que como ya se mencionó en la sección anterior, implica el cálculo de ecuaciones de regresión múltiple, uno por cada dimensión latente. Como las variables están estandarizadas, la puntuación media para cada dimensión latente es cero. Por esta razón, las puntuaciones negativas se interpretan como puntuaciones bajas. Las fórmulas de los índices calculados se presentan en la Tabla 2.

Tabla 1: Variables seleccionadas y dimensiones asignadas

Variable	Nombre de etiqueta	Dimensión
h_polcon5 van_index wbg_i_pse hf_Fiscal hf_Govt	Political Constraints Index V Index of Democratization Political Stability - Estimate Fiscal Freedom Freedom from government	Política (PolDist)
wbg_i_rle wbg_i_rqe wbg_i_cce wbg_i_gee wbg_i_vae fi_reg	Rule of Law – Estimate Regulatory Quality - Estimate Control of Corruption - Estimate Government Effectiveness - Estimate Voice and Accountability - Estimate Regulation of credit labor and business	Regulatoria (RegDist)
wdi_tou_rec wdi_tou_exp wdi_inet wdi_pl	International Tourism, receipts International Tourism, expenditures Internet Users (per 100 People) Phone Lines (per 100 People)	Conectividad (ConDist)
unna_pop_u14 unna_pop_b15_65 wdi_fr wdi_nnm2 unna_cdr wdi_lifexp	Population under 14 years Population between 15-65 years Fertility rate (births per woman) Net number of migrants Crude death rate Life expectancy	Demográfica (DemDist)
wdi_gdpc wdi_infl pwt_isg hf_efiscore wdi_ittr wdi_csg wdi_gsg	GDP, PPP (Constant International USD) Inflation (%) Investment Share of GDP (%) Economic Freedom Index International trade tax revenues (% of trade sector) Consumption share of GDP (%) Government share of GDP (%)	Económica (EconDist)
al_religion wvs_auth wvs_lifsat wvs_proud wvs_rs wvs_trust	Religious fractionalization Respect for authority Life satisfaction National Pride Religiosity Scale Interpersonal Trust	Cultural (CulDist)
Soil desert land_area dist_coast	% Fertile soil % Desert Land area (1000 Ha) Average distance to nearest ice-free coast (1000 km)	Geográfica (GeoDist)
efw_fofcfa efw_owbn wdi_icmc wdi_fo_ir wdi_cc	Freedom to own foreign currency bank accounts Ownership of banks International capital market controls Foreign ownership/investment restrictions Capital controls	Financiera (FinDist)
wdi_hte wdi_stja wdi_tma	High-technology exports (% of manufactured exports) Scientific and technical journal articles Trademark applications	Innovación (InnDist)
wdi_icpinf wdi_cb wdi_eodb wdi_rtb wdi_nttb wdi_lr wdi_trsb	Quality of port infrastructure, WEF Bureaucracy costs Ease of Doing Business Regulatory trade barriers Non-tariff trade barriers Licensing restrictions Time Required to Start Business (Days)	Práctica mercantil (PracMerDist)
wdi_mtr wdi_exp wdi_imp unna_cu	Mean tariff rate Exports (% of GDP) Imports (% of GDP) Currency	Comercial (CommDist)
wdi_ue wdi_lmr wdi_mw wdi_hfr	Unemployment (%) Labor Market Regulations Minimum wage Hiring and firing regulations	Mercado laboral (LabMarDist)

Fuente: Elaboración propia.

Con el fin de determinar si las variables incluidas en el cálculo de los índices tienen relación con las dimensiones de distancia, se calculó la alpha de Cronbach, α . Los resultados se muestran en la Tabla 3.¹²

¹ Los valores aceptables deberán ser superiores a 0.50 (Rockefeller College, 2007).

² Las variables fueron normalizadas debido a las diferencias en las escalas (OECD, 2008).

Tabla 2: Índice sintético

Dimensión	Cálculo del índice
Política	$\text{PolIndex} = (z1h_polcon5 * 0.16979) + (wbgi_pse * 0.21201) + (z1van_index * 0.19904) + (z1hf_govt * -0.40779) + (z1hf_fiscal * -0.16327)$
Regulatoria	$\text{RegIndex} = (wbgi_rle * 0.20593) + (z1fi_reg * 0.03126) + (wbgi_rqe * 0.22893) - (wbgi_cce * 0.20754) - (wbgi_gee * 0.29007) + (wbgi_vae * 0.06135)$
Práctica mercantil	$\text{PraIntBusIndex} = (z1wdi_cb * 0.02301) + (z1wdi_lr * 0.8063) + (z1wdi_rtb * 0.54223) + (z1wdi_nttb * 0.05559) + (z1wdi_iqpinfr * 0.10420) + (z1wdi_eodb * 0.-0.23332) + (z1wdi_trsb * -0.08698)$
Conectividad	$\text{ConnIndex} = (z1Lwdi_tou_re * 0.24446) + (z1Lwdi_tou_exp * 0.33470) + (z1wdi_inet * 0.23327) + (z1wdi_pl * 0.29901)$
Demográfica	$\text{DemIndex} = (z1wdi_fr * 0.15591) + (z1unna_nnm2 * -0.05397) + (z1Lunna_pop_u14 * 1.08833) + (z1Lunna_pop_u15_65 * -0.24758) + (z1unna_cdr * -0.01413) + (z1wdi_lifexp * -0.04638)$
Comercial	$\text{CommIndex} = (z1wdi_mtr * 0.08076) + (z1Lwdi_exp * 0.47758) + (z1Lwdi_imp * 0.44653) + (z1unna_cu * 0.07710)$
Cultural	$\text{CulIndex} = (z1wvs_lifsat * 0.03498) + (z1wvs_auth * 0.19555) + (z1wvs_proud * 0.21259) + (z1wvs_rs * -0.44172) + (z1wvs_trust * 0.15558) + (al_religion * 0.03512) + (al_language * -0.06717) + (al_ethnic * -0.16386)$
Económica	$\text{EcoIndex} = (z1wdi_ittr * 0.13025) + (z1pwt_csg * -0.10445) + (z1pwt_gsg * -0.01218) + (z1hf_efiscore * 0.16078) + (z1pwt_isg * 0.20722) + (z1Lwdi_gdpc * 0.51204)$
Geográfica	$\text{GeoIndex} = (z1Isoil * 0.13025) + (z1InvLand_area * -0.10445) + (z1InvDist_coast * -0.01218)$
Financiera	$\text{FinIndex} = (z1Iefw_fofba * 0.05454) + (z1Iefw_owbn * 0.07844) + (z1Swdi_icmc * -0.69120) + (z1Cwdi_fo_ir * 0.06390) + (z1Swdi_cc * 0.17629)$
Mercado Laboral	$\text{LabIndex} = (z1wdi_ue * (-0.08722)) + (z1wdi_lmr * 0.38234) + (z1wdi_mw * 0.42989) + (z1wdi_hfr * 0.15515)$
Innovación	$\text{InnIndex} = (z1Swdi_hte * 0.12930) + (z1Lwdi_stja * 0.44847) + (z1Lwdi_tma * 0.42201)$

Fuente: Cálculos propios.

Tabla 3: Estadísticas por dimensión

Dimensión	Contraste de esfericidad de Barlett*	Kaiser-Meyer-Olkin (KMO)	Alpha de Cronbach
Política	113.272	0.777	0.8422
Regulatoria	574.645	0.932	0.9702
Práctica Mercantil	241.309	0.814	0.8866
Conectividad	182.057	0.646	0.8838
Demográfica	226.555	0.54	0.6960
Comercial	112.102	0.567	0.7058
Cultural	147.846	0.592	0.7307
Económica	134.675	0.734	0.7923
Geográfica	37.536	0.521	0.6293
Financiera	236.382	0.624	0.8700
Mercado laboral	42.047	0.63	0.6394
Innovación	54.781	0.617	0.7583

Fuente: Cálculos propios. Nota: * χ^2 ; significativo: 0.05

Como se puede observar en la Tabla 3, todas las dimensiones tienen puntajes para los estadísticos que permiten aceptar y validar las dimensiones propuestas. Los resultados muestran que las dimensiones de distancia consideradas son el resultado de procesos

estadísticos multivariados y no una selección sesgada por parte del autor (Cea D’Ancona, 2002).

En la Tabla 4 se presenta la matriz de distancia de Mahalanobis de Colombia frente al resto del mundo. Con estos datos se prueba el cumplimiento de las propiedades de la distancia de Mahalanobis y que la distribución es unimodal para acercarse a la normalidad.

Tabla 4. Distancia de Mahalanobis por dimensión desde Colombia hacia el resto del mundo

País	PolDist	RegDist	IntBusDist	ConnDist	DemDist	CommDist	CulDist	EcoDist	GeoDist	FinDist	LabDist	InnDist
Albania	0,55	1,03	1,11	0,44	2,46	0,93	0,52	0,02	7,28	0,17	0,04	2,67
Algeria	0,92	0,43	0,05	1,05	0,01	0,61	0,11	0,45	0,60	0,70	0,79	0,52
Argentina	0,54	1,73	0,31	0,10	0,05	0,01	0,11	0,04	0,00	0,05	0,75	0,97
Australia	5,37	0,19	1,25	3,65	1,17	0,13	2,26	3,83	0,56	1,37	1,80	1,98
Austria	9,66	1,04	1,74	2,47	2,69	3,94	1,67	2,88	4,29	4,22	0,44	0,58
Bangladesh	0,02	0,14	1,00	1,40	0,55	0,06	0,09	3,52	0,53	0,00	0,02	0,74
Belgium	11,2	2,07	1,75	2,80	2,06	8,59	2,30	3,04	1,25	4,92	0,30	1,02
Brazil	1,17	0,12	0,42	0,13	0,43	0,31	0,04	0,00	0,43	1,17	2,02	2,48
Bulgaria	3,03	1,88	0,25	0,07	2,86	3,91	2,37	0,06	8,27	1,32	0,14	0,00
Canada	4,62	1,33	3,21	5,05	0,64	2,14	0,24	3,36	0,59	4,89	0,18	2,08
Chile	1,53	0,03	1,31	0,01	0,62	1,27	0,11	1,16	0,69	4,31	0,00	0,27
China	0,02	3,63	1,05	1,08	1,78	0,25	8,67	0,01	0,14	0,09	0,23	6,52
Colombia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Croatia	4,49	0,00	2,40	0,32	3,57	2,85	0,70	1,04	3,84	0,05	0,58	0,02
Czech Republic	4,65	1,99	0,55	0,49	2,40	5,09	4,31	2,09	4,52	3,36	0,06	0,41
Denmark	13,1	1,81	4,87	3,80	2,81	2,85	5,49	3,89	1,90	4,89	2,96	0,38
Dominican Republic	0,10	0,54	0,00	0,05	0,57	1,49	0,06	0,00	3,16	1,37	0,24	1,80
Egypt	1,31	0,01	0,07	0,00	0,23	0,01	0,13	0,57	0,81	1,44	0,26	0,33
Estonia	2,73	3,47	3,41	0,04	6,13	6,82	4,44	2,40	2,17	4,10	0,03	0,10
Finland	8,71	2,13	5,40	1,84	2,99	2,26	3,46	3,41	0,46	4,18	0,33	0,34
France	9,56	1,19	1,52	4,57	0,32	0,96	2,30	2,09	3,91	4,63	1,15	3,44
Germany	7,96	1,27	3,14	6,51	0,40	1,96	4,75	2,40	1,25	3,64	1,56	3,53
Greece	1,96	0,87	1,06	1,37	2,44	1,05	1,25	1,49	3,69	3,77	0,93	0,24
Guatemala	0,01	0,12	0,56	0,39	0,01	1,20	0,32	0,00	4,71	3,42	0,45	0,92
Hungary	7,16	1,86	0,99	0,47	2,25	5,45	1,72	1,32	1,92	4,75	1,06	0,37
India	0,95	0,68	0,02	0,03	3,18	0,31	0,26	0,87	2,05	0,10	0,98	2,50
Indonesia	0,01	0,01	2,14	0,10	0,50	0,84	1,03	0,29	0,17	0,05	0,88	0,05
Ireland	3,71	0,26	1,79	1,30	4,58	9,21	0,03	4,83	1,88	5,52	1,15	0,05
Israel	6,44	0,18	0,71	0,72	1,25	1,72	0,21	2,16	2,23	4,18	0,06	0,31
Italy	7,67	0,10	0,34	3,34	0,45	0,83	0,80	2,84	3,49	3,22	0,14	2,36
Japan	3,52	0,04	1,60	3,60	0,10	0,65	10,3	2,92	2,56	3,29	3,08	5,09
Jordan	0,62	0,00	0,59	0,08	0,67	3,74	0,30	0,13	0,76	4,11	1,39	0,49
Korea, South	1,88	0,05	3,01	3,65	0,33	1,10	6,07	3,47	3,19	0,14	0,11	3,66
Latvia	3,02	2,31	0,69	0,11	4,90	3,01	1,28	0,75	3,40	3,54	0,59	0,29
Malaysia	0,61	0,97	0,24	0,70	0,11	9,18	0,11	0,65	3,16	0,04	2,33	0,29
Mexico	0,61	0,07	0,79	0,55	0,28	0,49	0,02	0,77	0,90	0,39	0,01	1,56
Morocco	1,47	0,05	0,11	0,06	0,08	0,52	0,28	0,31	2,32	0,00	5,07	0,34
New Zealand	4,01	0,45	5,11	1,30	2,88	1,16	1,92	2,74	2,84	5,33	2,47	0,41
Norway	8,43	3,20	0,36	2,78	2,89	1,48	5,77	4,04	0,09	3,29	0,19	0,31
Pakistan	0,01	0,14	0,02	0,39	1,45	0,13	0,25	1,15	0,07	0,07	0,14	0,22
Peru	0,16	0,63	0,00	0,11	0,05	0,01	0,18	0,00	0,45	4,33	0,00	0,20
Philippines	0,10	0,28	0,34	0,00	0,24	3,55	0,23	0,36	7,75	0,28	0,58	0,04
Poland	3,76	1,04	0,00	0,66	0,60	1,05	0,00	0,59	0,54	0,20	0,18	0,62
Portugal	5,85	0,17	0,14	0,88	2,08	1,73	0,00	2,51	2,21	2,98	0,07	0,21
Romania	3,42	0,64	0,01	0,18	0,99	1,62	0,01	0,08	5,53	1,65	0,83	0,01
Russia	0,37	0,16	0,85	0,66	0,00	0,74	2,36	0,01	0,70	0,12	0,15	2,21
Slovenia	6,23	0,14	0,73	0,15	5,77	5,11	1,95	2,32	7,26	1,97	0,52	0,01
Spain	4,81	1,21	0,50	2,38	0,76	1,24	1,18	3,24	4,54	4,77	1,64	2,58
Sweden	12,4	1,98	3,61	4,14	2,20	2,43	6,79	2,54	0,08	3,36	0,01	0,77
Switzerland	5,30	1,45	0,90	4,63	2,80	2,38	3,07	4,17	2,84	4,62	3,31	1,17
Thailand	0,51	1,68	0,56	0,26	0,01	4,37	1,13	0,27	0,60	0,02	0,03	0,54
Trinidad and Tobago	1,35	3,61	0,02	0,23	4,32	2,90	0,25	1,33	2,36	3,31	1,81	2,59
Turkey	2,74	0,13	0,38	0,44	0,01	0,41	0,00	0,19	4,41	0,21	1,42	0,81
Ukraine	2,26	0,07	1,74	0,02	0,67	3,68	2,05	0,15	1,59	1,77	0,02	0,22
United Kingdom	4,11	0,49	2,12	6,32	0,13	0,72	2,02	2,62	0,76	5,67	2,82	2,94
United States	3,63	0,42	4,56	8,94	0,27	0,47	0,03	3,37	0,93	4,22	5,66	8,09
Uruguay	1,49	0,03	0,05	0,13	2,60	0,03	0,56	0,03	2,17	3,41	0,12	0,13

Fuente: Cálculos propios.

Cada uno de los índices de distancia cumple con las propiedades de simetría, no negatividad, definitud y desigualdad triangular. Es así como, por ejemplo, la distancia política desde Colombia hasta España es la misma que la distancia política desde España hasta Colombia. Por tanto,

$$PolDist_{Colombia-España}(4.8084) = PolDist_{España-Colombia}(4.8084)$$

lo cual prueba la *simetría*. La *no negatividad* se prueba al observar la Tabla 4 y comprobar que todas las distancias para todas las dimensiones son medidas como cantidades no negativas. Por tanto,

$$Dist_{Colombia-Resto\ del\ mundo} \geq 0.$$

También se puede evidenciar que las propiedades de *identificación* y *definitud* se cumplen al observar que para cada dimensión de distancia entre *Colombia* y ella misma el resultado es cero. Por tanto,

$$Dist_{Colombia-Colombia} = 0$$

Finalmente, se analizó si las distribuciones eran *unimodales*, para ello se presentan los resultados para cada una de las dimensiones de distancia propuestas. Siguiendo a Silverman (1986) no existe un valor específico a partir del cual se pueda determinar el número de modas. Sin embargo, se ha sugerido como una regla para determinar el número de modas un *p*-valor mayor o igual a 0.4 (Izenman y Sommer, 1988) o incluso es suficiente considerar que el *p*-valor sea mayor a 0.1 (Salgado- Ugarte *et al.*, 1997). Se tomó como valor de referencia para el *p*-valor 0.1 y los cálculos fueron realizados con la modificación del programa *silvtest.ado* realizada por Mora y Nuñez (2009). Los resultados se condensan en la Tabla 5.

Tabla 5: Número de modas para los índices de distancia (Colombia-Resto del mundo)

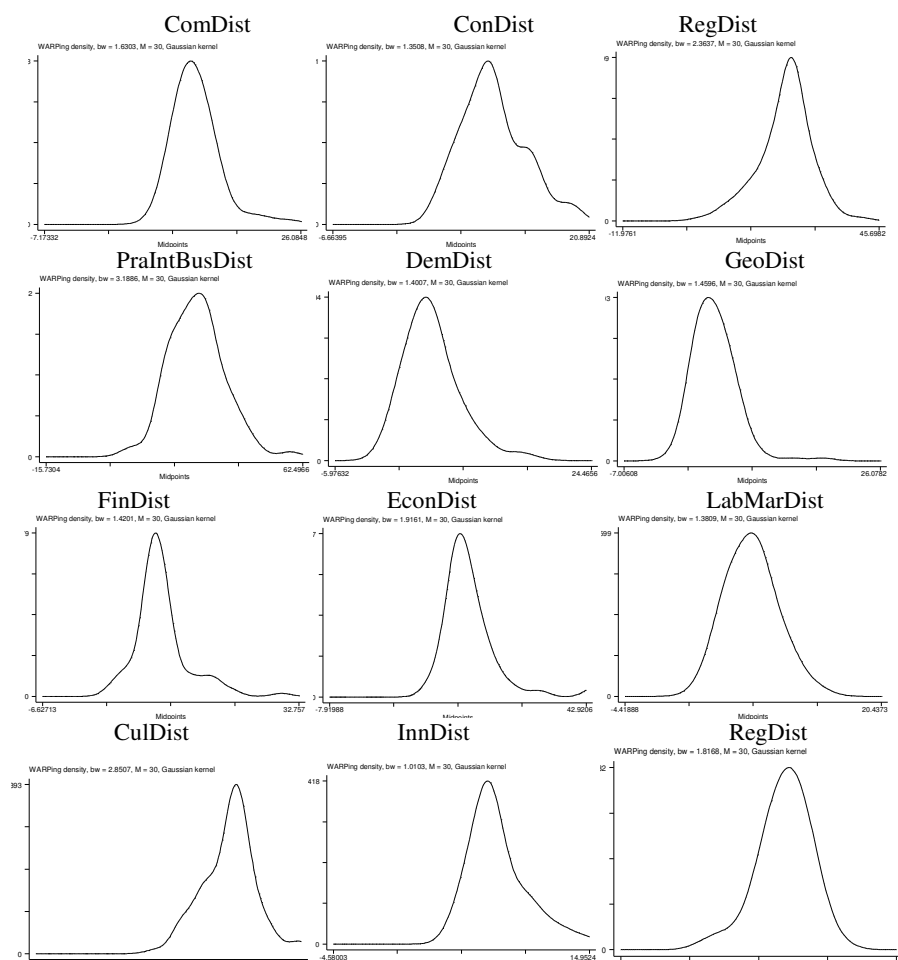
Variable	Nº de modas	P-Valor
Distancia Política (PolDist)	1	0.39
Distancia Regulatoria (RegDist)	1	0.11
Distancia Práctica mercantil (PracMerDist)	1	0.94
Distancia Comercial (CommDist)	1	0.97
Distancia en Conectividad (ConnDist)	1	0.84
Distancia Demográfica (DemDist)	1	0.52
Distancia Geográfica (GeoDist)	1	0.23
Distancia Financiera (FinDist)	1	0.87
Distancia Económica (EcoDist)	1	0.71
Distancia en Mercado laboral (LabDist)	1	0.96
Distancia Cultural (CulDist)	1	0.92
Distancia en Innovación (InnDist)	1	0.67

Fuente: Cálculos propios. Programa *silvtest.ado* (Mora y Nuñez, 2009).

De acuerdo con los resultados, se puede concluir que la distribución es unimodal para todos los índices de distancia calculados entre Colombia y el resto del mundo. La Gráfica 1 muestra el resultado de la densidad de Kernel para cada uno de los índices de distancia propuestos.

Burdenski (2000) considera que la distancia de Mahalanobis es una medida de distancia del puntaje de cada persona en cada variable, (observación o país) desde el grupo de medias de cada variable, ajustado por cómo se correlacionan dichas variables. En términos de los resultados de la distancia de Mahalanobis es de esperarse en este caso que un individuo con desviaciones redundantes respecto a la media (ind_1) tenga un puntaje bajo comparado con un individuo que no tenga desviaciones redundantes respecto a la media (ind_2). La estadística descriptiva de los resultados de la distancia de Mahalanobis considerando todas las dimensiones propuestas para el caso Colombia-Resto del mundo se presentan en la Tabla 6.

Gráfica 1: Densidad de Kernel



Fuente: Cálculos propios usando silvtest.ado (Mora y Nuñez, 2009)

Tabla 6: Estadísticas descriptivas de la distancia de Mahalanobis

Variable	Obs	Mean	Std. Dev.	Min	Max
PolDist	57	3.492058	3.426884	0	13.13379
RegDist	57	.9876711	1.179796	0	5.764992
PracMerDist	57	1.319607	1.483906	0	5.395597
ConnDist	57	1.53908	1.986798	0	8.937593
DemDist	57	1.556045	1.58872	0	6.128626
CommDist	57	2.390374	3.051001	0	17.70387
CulDist	57	1.688566	2.303378	0	10.3105
EcoDist	57	1.675382	1.561999	0	6.412694
GeoDist	57	2.258232	2.086456	0	8.27081
FinDist	57	2.470882	1.941173	0	5.665264
LabDist	57	.9983458	1.28004	0	5.658009
InnDist	57	1.268264	1.643123	0	8.092216

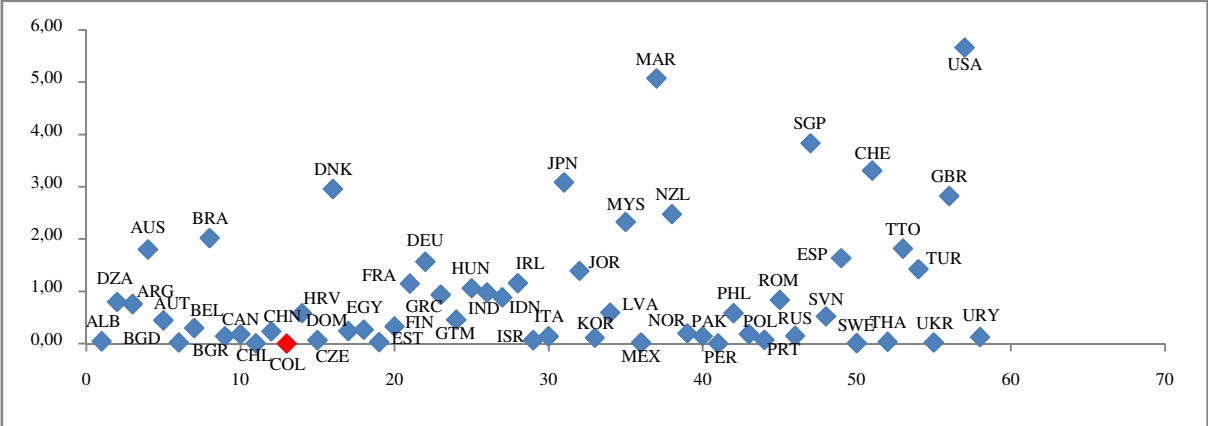
Fuente: Cálculos propios.

Como se puede observar, los índices muestran medidas de distancia relativamente pequeñas, lo cual es un indicio del alto grado de integración de las economías a nivel global. El rango de distancias más amplio se encuentra en la dimensión comercial, seguida por la

dimensión política y la dimensión cultural; mientras que el rango más estrecho se encuentra en la dimensión de práctica mercantil.

Al considerar cada una de las dimensiones de distancia en un diagrama de dispersión, donde el eje Y representa el índice de distancia y el eje X muestra simplemente el ordenamiento de los países, podemos apreciar mejor cuáles países son más cercanos y cuáles más lejanos en términos de la distancia a la que hace referencia cada dimensión, lo cual resulta una herramienta fundamental cuando se desarrollan los negocios internacionales en el día a día y se trata de tomar decisiones acerca de los procesos de expansión internacional y procesos de deslocalización de la firma e incluso decisiones relacionadas a la gestión de los recursos de una compañía multinacional.

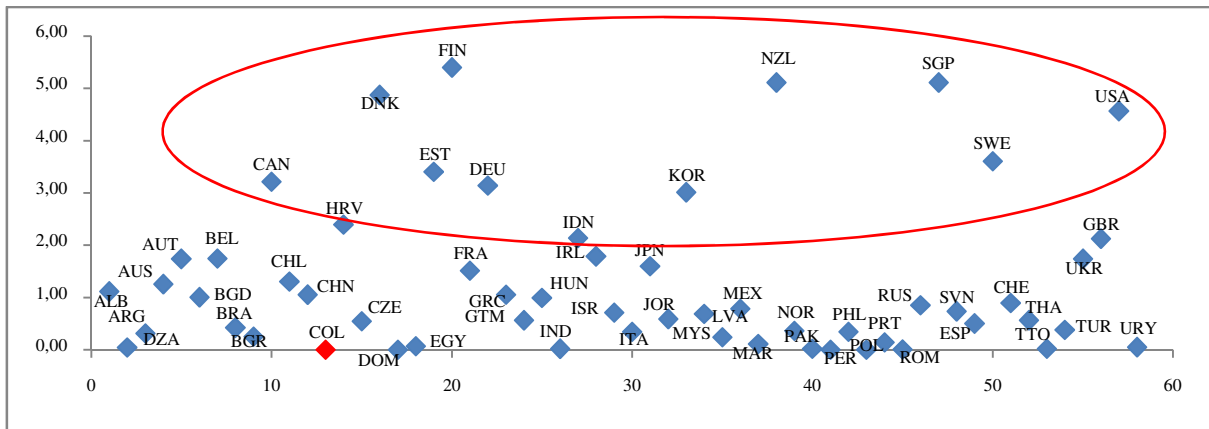
Gráfica 2: Distancias en el mercado laboral



Fuente: Cálculos propios. Datos de distancia del resto del mundo respecto a Colombia. Códigos ISO disponibles en el Anexo 2.

La distancia en el mercado laboral no había sido considerada antes en la literatura sobre negocios internacionales pese a sus fuertes implicaciones sobre procesos tan relevantes para la firma como lo son la contratación de mano de obra local y despidos en los países hacia los cuales se ha deslocalizado la firma y que pueden generar fuertes efectos negativos en términos de la desventaja de ser extranjero. Un claro ejemplo de ello en Colombia fueron los fuertes efectos negativos sobre la imagen corporativa de Hyundai y Pacific Rubiales en el 2015 ante el eventual despido de 4000 y 3200 trabajadores, respectivamente. Respecto al comportamiento del índice de distancia, vemos que la dispersión no es muy alta y esto podría atribuirse a la existencia de la Organización Internacional del Trabajo, una institución supranacional que desde hace casi un siglo ha abogado por la estandarización de prácticas en los mercados laborales del mundo, aun cuando esto depende de la buena voluntad de los países miembro de esta organización.

Gráfica 3: Distancias en la práctica mercantil

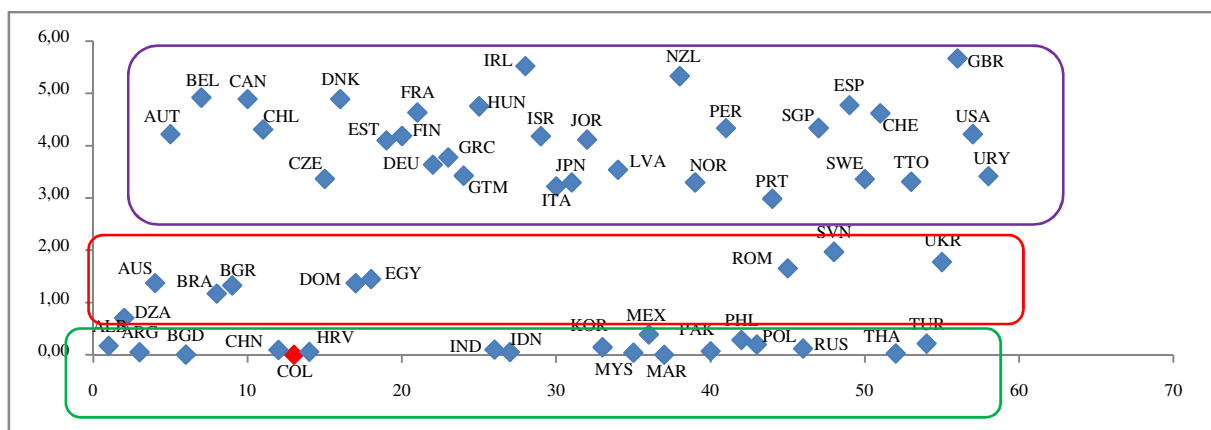


Fuente: Cálculos propios. Datos de distancia del resto del mundo respecto a Colombia. Códigos ISO disponibles en el Anexo 2.

Respecto a la dimensión Práctica Mercantil, esta representa una de las dimensiones más relevantes ya que agrupa una serie de factores que hoy día se conocen como barreras paraarancelarias o no cuantitativas al comercio internacional. La Gráfica 3 muestra evidencia de que cada día las prácticas de los negocios están más estandarizadas y que Colombia no ha sido ajena a esta tendencia, lo cual evidencia avances en la globalización sobre los negocios internacionales. Sin embargo, aún es posible identificar a algunos países que se salen de esta tendencia como lo son Dinamarca, Finlandia, Alemania, Estonia y Suecia en Europa, Singapur, Nueva Zelanda y Corea del Sur en el Asia Pacifico y Estados Unidos y Canadá en América del norte. Cabe aclarar que el índice de distancia nos entrega información acerca de la similitud o diferencia entre los países en el aspecto de la dimensión pero no nos permite decir si las características de la dimensión se deben porque unos países están mejor posicionados en cada una de las variables índice que componen la dimensión. En este caso particular, podría afirmarse que estos países europeos tienen unas prácticas mercantiles que se diferencian mucho de las que se practican en Colombia, lo cual podría representar ciertas restricciones adicionales que enfrentarían las empresas colombianas interesadas en incursionar estos mercados buscando beneficiarse de los acuerdos alcanzados en los Tratados de Libre Comercio con la Unión Europea y que las empresas europeas deberán adaptarse a las prácticas locales, lo cual podría suponer un aumento en los costos de transacción que se derivarían de su participación en el mercado local.

La dimensión financiera considera aspectos como la libertad del sistema financiero, el control sobre los mercados de capitales y las restricciones a la inversión entre otros aspectos. En el caso particular de Colombia, su comportamiento es considerablemente diferente al que se ha presentado en otras dimensiones de distancia donde Colombia generalmente se encontraba cercana a la mayoría de países; en este índice, el comportamiento muestra una concentración de los países a valores altos en términos de distancia respecto a Colombia.

Gráfica 4: Distancias en la dimensión financiera



Fuente: Cálculos propios. Datos de distancia del resto del mundo respecto a Colombia. Códigos ISO disponibles en el Anexo 2.

El primer gran agrupamiento lo encontramos concentrado en los valores más altos de la distancia financiera y su composición incluye a los grandes países desarrollados. Gran Bretaña se encuentra en el *top* de las economías con fuertes y eficientes mercados financieros de allí que este país y su comportamiento en esta dimensión respecto a Colombia nos permita inferir el rezago sustancial que Colombia tiene en este mercado. Un notable caso a resaltar es el de Perú o el de Chile, únicos países latinoamericanos en ese grupo. Tenemos algunos países en una posición media como algunas economías de la Europa oriental y Brasil. Y al último grupo, con pequeñas distancias respecto a Colombia y compuesto por economías de países en vía de desarrollo (véase Gráfica 4).

El rezago en el mercado financiero o su bajo nivel de sofisticación se constituye en una clara desventaja para las empresas que inician o tratan de mantener su expansión internacional ya que el acceso al crédito para pequeñas y medianas empresas es fundamental para apalancar sus operaciones en el extranjero; también lo es el tener una amplia disponibilidad de medios de pagos y sistemas de cobertura del riesgo cambiario tales como *forward* y opciones de compra y venta de divisas, necesarios para desarrollar la actividad exportadora por parte de las empresas colombianas y en general de empresas latinoamericanas que se enfrentan a mercados de divisas muy volátiles (por ejemplo, Venezuela y Colombia).

6.- Conclusiones

Este artículo presenta estimaciones de la distancia en los negocios internacionales entre Colombia y los continentes de Europa, Asia, África y Oceanía. El concepto de distancia que utilizamos incorpora aspectos institucionales y aspectos no considerados hasta ahora como son el mercado laboral, la práctica mercantil y la innovación.

La distancia en el mercado laboral no ha sido considerada antes en la literatura sobre negocios internacionales pese a las implicaciones que esta dimensión tiene sobre los procesos de la firma, como lo son la contratación de mano de obra local y los despidos en los países hacia los cuales se ha deslocalizado la firma y que pueden generar fuertes efectos negativos en términos de la desventaja de ser extranjero. La dispersión en esta dimensión no es muy alta y esto podría atribuirse a la existencia de la Organización Internacional del Trabajo, quien

durante casi un siglo ha abogado por la estandarización de prácticas en los mercados laborales del mundo, aun cuando esto depende de la buena voluntad de los países miembros de esta organización.

Con respecto a la distancia en las prácticas mercantiles, la misma agrupa una serie de factores que hoy día se conocen como barreras paraarancelarias o no cuantitativas al comercio internacional. Nuestros resultados con respecto a Colombia muestran evidencia de que cada día las prácticas de los negocios están más estandarizadas y que Colombia no ha sido ajena a esta tendencia lo cual evidencia avances en la globalización sobre los negocios internacionales. Sin embargo, países como Dinamarca, Finlandia, Alemania, Estonia y Suecia en Europa, Singapur, Nueva Zelanda y Corea del Sur en Asia Pacífico y Estados Unidos y Canadá en América del norte, muestran resultados que sugieren que sus prácticas mercantiles se diferencian mucho de las que se practican en Colombia, lo cual podría representar restricciones adicionales que enfrentarían las empresas colombianas interesadas en incursionar estos mercados buscando beneficiarse de los acuerdos alcanzados en los Tratados de Libre Comercio con la Unión Europea y que las empresas europeas deberán adaptarse a las prácticas locales, lo cual podría suponer un aumento en los costos de transacción que se derivarían de su participación en el mercado local.

La dimensión financiera considera aspectos como la libertad del sistema financiero, el control sobre los mercados de capitales y las restricciones a la inversión entre otros aspectos. En el caso particular de Colombia, su comportamiento es considerablemente diferente al que se ha presentado en otras dimensiones de distancia donde Colombia generalmente se encontraba cercana a la mayoría de países; en este índice, el comportamiento muestra una concentración de los países a valores altos en términos de distancia respecto a Colombia. El primer gran agrupamiento lo encontramos concentrado en los valores más altos de la distancia financiera y su composición incluye a los grandes países desarrollados. Gran Bretaña se encuentra en el *top* de las economías con fuertes y eficientes mercados financieros de allí que este país y su comportamiento en esta dimensión respecto a Colombia nos permita inferir el rezago sustancial que Colombia tiene en este mercado. Un notable caso a resaltar es el de Perú o el de Chile, únicos países latinoamericanos en ese grupo. Tenemos algunos países en una posición media como algunas economías de la Europa oriental y Brasil. Y al último grupo, con pequeñas distancias respecto a Colombia y compuesto por economías de países en vía de desarrollo (véase Gráfica 4). El rezago en el mercado financiero o su bajo nivel de sofisticación se constituye en una clara desventaja para las empresas que inician o tratan de mantener su expansión internacional ya que el acceso al crédito para pequeñas y medianas empresas es fundamental para apalancar sus operaciones en el extranjero, también lo es el tener una amplia disponibilidad de medios de pagos y sistemas de cobertura del riesgo cambiario tales como *forward* y opciones de compra y venta de divisas, necesarias para desarrollar la actividad exportadora por parte de las empresas colombianas y en general de empresas latinoamericanas que se enfrentan a mercados de divisas muy volátiles (por ejemplo, Venezuela y Colombia).

De esta forma, aunque nuestros resultados muestran que hay avances en la globalización sobre los negocios internacionales entre Colombia y los países de la muestra todavía existen restricciones a las que se enfrentan las empresas colombianas que están interesadas en incursionar estos mercados buscando beneficiarse de los acuerdos de libre comercio.

REFERENCIAS BIBLIOGRÁFICAS

1. Anderson, J. & Marcouiller, D. (2002): "Insecurity and the Pattern of Trade: An Empirical Investigation", *The Review of Economics and Statistics*, 84(2): 342-352.
2. Bae, J. & Salomon, R. (2010): "Institutional distance in international business research", en Devinney, T., Pedersen, T. & Tihanyi, L. (eds.): *The Past, Present and Future of International Business & Management (Advances in International Management, 23)*, Emerald Group Publishing Limited, pp. 327-349.
3. Barkema, H., Bell, J. & Pennings, J. (1996): "Foreign entry, cultural barriers, and learning", *Strategic Management Journal*, 17(2): 151-166.
4. Bénassy-Quéré, A., Coupet, M. & Mayer, T. (2007): "Institutional Determinants of Foreign Direct Investment", *World Economy*, 30(5): 764-782.
5. Berry, H., Guillen, M. & Zhou, N. (2010): "An institutional approach to cross-national distance", *Journal of International Business Studies*, 41(9), 1460-1480.
6. Berthelon, M. & Freund, C. (2008): "On the Conservation of Distance in International Trade", *Journal of International Economics*, 75(2): 310-320.
7. Burdinski, T. (2000): "Evaluating univariate, bivariate, and multivariate normality using graphical and statistical procedures", *Multiple Linear Regression Viewpoints*, 26(2): 15-28.
8. Busenitz, L., Gómez, C. & Spencer, J. (2000): "Country Institutional Profiles: Unlocking Entrepreneurial Phenomena", *Academy of Management Journal*, 43(5): 994-1003.
9. Chang, Y-C., Kao, M-S. & Chiu, C-F. (2011): "How cultural distance influences entry mode choice: The contingent role of host country's governance quality", *Journal of Business Research*, 65(8): 1160-1170.
10. Cea D'Ancona, M. (2002): *Análisis Multivariable Teoría y Práctica en la Investigación Social*, Madrid: Editorial Síntesis.
11. Chao, M. & Kumar, V. (2010): "The impact of institutional distance on the international diversity-performance relationship", *Journal of World Business*, 45(1): 93-103.
12. Deardorff, A. (1987): "Why Do Governments Prefer Nontariff Barriers?", *Carnegie-Rochester Conference Series on Public Policy*, 26: 191-216.
13. Deardorff, A. & Stern, R. (1985): *Methods of Measurement of Non-Tariff Barriers*, Geneva: UNCTAD.
14. De Groot, H.L.F., Linders, G-J., Rielveld, P. & Subramanian, U. (2004): "The Institutional Determinants of Bilateral Trade Patterns", *Kyklos*, 57(1): 103-123.
15. Disdier, A-C. & Head, K. (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade", *The Review of Economics and Statistics*, 90(1): 37-48.
16. Driffield, N. & Taylor, K. (2000): "FDI and the Labor Market: A review of the evidence and policy implications", *Oxford Review of Economic Policy*, 16(3): 90-103.
17. Drogendijka, R. & Slangen, A. (2006): "Hofstede, Schwartz, or managerial perceptions? The effects of different cultural distance measures on establishment mode choices by multinational enterprises", *International Business Review*, 15(4): 361-380.
18. Eriksson, K., Johanson, J., Majkgard, A. & Sharma, D. (1997): "Experiential Knowledge and Cost in the Internationalization Process", *Journal of International Business Studies*, 28(2): 337-360.

19. Fitzpatrick, G. & Modlin, M. (1986): *Direct-line Distances: International Edition*, Metuchen: Scarecrow Press.
20. Furman, J., Porter, M. & Stern, S. (2002): "The determinants of national innovative capacity", *Research Policy*, 31(6): 899-933.
21. Gaur, A., Delios, A. & Singh, K. (2007): "Institutional Environments, Staffing Strategies, and Subsidiary Performance", *Journal of Management*, 33(4): 611-636.
22. Ghemawat, P. & Altman, S. (2014): *DHL Global Connectedness Index 2014*, Deutsche Post DHL, Headquarters.
23. Ghemawat, P. (2001): "Distance Still Matters. The Hard Reality of Global Expansion", *Harvard Business Review*, 79(8): 137-147.
24. Globerman, S. & Shapiro, D. (2002): "Global Foreign Direct Investment Flows: The Role of Governance Infrastructure", *World Development*, 30(11): 1899-1919.
25. Hall, R. & Jones, Ch. (1999): "Why Do Some Countries Produce So Much More Output per Worker than Others?", *Quarterly Journal of Economics*, 114(1), 83-116.
26. Hasan, I. & Tucci, C. (2010): "The innovation-economic growth nexus: Global evidence", *Research Policy*, 39(10), 1264-1276.
27. Hymer, S. (1960): *The International Operations of National Firms: A Study of Direct Foreign Investment*. Cambridge: MIT Press.
28. Inglehart, R. & Baker, W. (2000): "Modernization, Cultural Change and the Persistence of Traditional", *American Sociological Review*, 65(1): 19-51.
29. Izenman, A. & Sommer, C. (1988): "Philatelic mixtures and multimodal densities", *Journal of the American Statistical Association*, 83(404): 941-953.
30. Johanson, J. & Vahlne, J. (1977): "The Internationalization Process of the Firm-A Model of Knowledge Development and Increasing Foreign Market Commitments", *Journal of International Business Studies*, 8(1): 23-32.
31. Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations across Nations*, Thousand Oaks: Sage.
32. Kaufmann, D., Kraay, A. & Zoido-Lobaton, P. (1999): "Aggregating governance indicators", *Policy Research Working Paper*, 2195: 39 pp.
33. Knack, S. & Keefer, P. (1995): "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures", *Economics & Politics*, 7(3): 207-227.
34. Kogut, B. & Singh, H. (1988): "The Effect of National Culture on the Choice of Entry Mode", *Journal of International Business Studies*, 19(3): 411-432.
35. Kostova, T. & Roth, K. (2002): "Adoption of an Organizational Practice by Subsidiaries of Multinational Corporations: Institutional and Relational Effects", *Academy of Management Journal*, 45(1): 215-233.
36. Kostova, T., Roth K. & Dacin, T. (2008): "Institutional Theory in the Study of Multinational Corporations: A Critique and New Directions", *Academy Management Review*, 33(4): 994-1006.
37. Kostova, T. & Zaheer, S. (1999): "Organizational legitimacy under conditions of complexity: The case of the multinational enterprise", *Academy of Management Review*, 24(1): 64-81.
38. Koukhartchouk, O. & Maurel, M. (2003): "Accession to the WTO and EU enlargement: what potential for trade increase?", *CEPR Discussion Paper*, 3944, London: Centre for Economic Policy Research.

39. Krzanowski, W. J. (1988): *Principles of Multivariate Analysis: A User's Perspective*, New York: Oxford University Press.
40. La Porta, R., López-de-Silanes, F., Shleifer, A. & Vishny, R. (1998): "Law and finance", *Journal of Political Economy*, 106(6): 1113-1155.
41. Mahalanobis, P. (1930): "On tests and measures of group divergence", *Journal of the Asiatic Society of Bengal*, 26, 541-588.
42. Mahalanobis, P. (1936): "On the generalized distance in statistics", *Proceedings of the National Institute of Sciences of India*, 2, 49-55.
43. Mardia, K., Kent, J. & Bibby, J. (1979): *Multivariate Analysis*, London: Academic Press.
44. Markus, H. & Zajonc, R. (1985): "The cognitive perspective in social psychology", en Lindzey, G. & Aronson, E. (eds.): *The Handbook of Social Psychology*, New York: Random House, pp. 137-229.
45. Martin, X., Salomon, R. & Wu, Z. (2010): "The institutional determinants of agglomeration: a study in the global semiconductor industry", *Industrial and Corporate Change*, 19(6): 1769-1800.
46. Miller, S. & Parkhe, A. (2002): "Is there a Liability of Foreignness in Global Banking? An empirical test of Banks' X-Efficiency", *Strategic Management Journal*, 23(1): 55-75.
47. Mitchell, A. & Krzanowski, W. (1985): "The Mahalanobis distance and elliptic distributions", *Biometrika*, 72(2): 464-467.
48. Mora, J.J. & Nuñez, J. (2009): "Economic convergence of income distribution worldwide from 1986 to 2000", *Journal of Economic Studies*, 36(6): 675-691.
49. Morosini, P., Shane, S. & Singh, H. (1998): "National cultural distance and cross-border acquisition performance", *Journal of International Business Studies*, 29(1): 137-158.
50. North, D. (1990): "Institutions", *The Journal of Economic Perspectives*, 5(1), 97-112.
51. North, D. (1995): "Five Propositions about Institutional Change", en Knight, J. & Sened, I. (eds.): *Explaining Social Institutions*, Ann Arbor: University of Michigan Press, pp. 15-26.
52. OECD (2008): *Handbook on Constructing Composite Indicators: Methodology and user guide*, Organization for Economic Cooperation and Development (OECD).
53. Olson, M. (1996): "Big Bill a Left on the Sidewalk: Why Some Nations are Rich and Others Poor", *Journal of Economic Perspectives*, 10(2): 3-24.
54. Oxley, J. & Yeung, B. (2001): "E-commerce readiness: Institutional environment and international competitiveness", *Journal of International Business Studies*, 32(4): 705-723.
55. Pla-Barber, J. & León-Darder, F. (2004): *Dirección de Empresas Internacionales*. Madrid: Prentice Hall.
56. Perkins, S. (2014): "When does prior experience pay? Institutional experience and the case of multinational corporation", *Administrative Science Quarterly*, 59(1): 145-181.
57. Rockefeller College (2007): "PAD 705 Handout: Factor Analysis", disponible en: <http://www.albany.edu/faculty/kretheme/PAD705/SupportMat/FactorAnalysisTheory.pdf>.
58. Salgado-Ugarte, I., Shimizu, M. & Taniuchi, T. (1997): "Nonparametric assessment of multimodality for univariate data", *Stata Technical Bulletin*, 38: 27-35.
59. Seber, G. (1984): *Multivariate observations*. New York: Wiley.

60. Scott, W. (1995): *Institutions and Organizations*. Thousand Oaks: Sage.
61. Sethi, D. & Guisinger, S. (2002): "Liability of foreignness to competitive advantage: How multinational enterprises cope with the international business environment", *Journal of International Management*, 8: 223-240.
62. Silverman, B. (1986): *Density Estimation for Statistics and Data Analysis*, London: Chapman & Hall.
63. Shaxby, J.H. (1945): "Great circles routes", *Geography*, 30(4): 112-115.
64. Shenkar, O. (2001): "Cultural distance revisited: Towards a more rigorous conceptualization and measurement of cultural differences", *Journal of International Business Studies*, 32(3): 519-535.
65. Sokal, R. (1979): "Testing Statistical Significance of Geographic Variation Patterns", *Systematic Zoology*, 28(2): 227-232.
66. Wei, S-J. (2000): "Natural openness and good government", *NBER Working Paper* no. 7765: 32 pp.
67. Wheeler, D. & Mody, A. (1992): "International investment location decisions: The case of U.S. firms", *Journal of International Economics*, 33(1-2): 57-76.
68. Whitley, R. (1992): *Business systems in East Asia: Firms, markets, and societies*, London: Sage.
69. Xu, D., Pan, Y. & Beamish, P. (2004): "The Effect of Regulative and Normative Distances on MNE Ownership and Expatriate Strategies", *Management International Review*, 44(3): 285-307.
70. Zelner B., Henisz W. & Holburn, G. (2009): "Contentious Implementation and Retrenchment in Neoliberal Policy Reform: The Global Electric Power Industry, 1989-2001", *Administrative Science Quarterly*, 54(3): 379-412.

Anexo 1. Índices sintéticos de distancia

País	PolIndex	RegIndex	PracMerIndex	ConnIndex	DemIndex	CommIndex	CulIndex	EcoIndex	GeoIndex	FinIndex	LabIndex	InnIndex
Albania	23,5	79,4	12,7	12,4	21,3	35,3	34,0	46,0	95,3	24,8	44,2	0,0
Algeria	29,3	40,6	45,5	3,8	55,6	31,7	6,1	57,7	3,4	13,3	29,4	20,4
Argentina	23,3	25,2	25,6	35,8	52,8	17,6	23,9	38,0	24,9	30,0	29,9	58,5
Australia	65,4	45,8	68,9	74,1	32,7	23,2	54,8	87,0	4,0	67,4	77,6	67,9
Austria	86,3	32,1	74,0	66,0	19,6	55,6	49,2	81,0	78,7	91,2	63,0	53,5
Bangladesh	7,9	64,4	14,1	0,0	75,6	11,4	7,1	0,0	43,1	37,0	51,6	17,3
Belgium	93,0	22,4	74,1	68,4	24,4	74,5	55,1	82,1	53,4	95,6	60,4	59,0
Brazil	32,5	64,0	23,2	37,0	73,5	5,0	20,3	42,3	6,4	65,0	18,0	71,7
Bulgaria	50,0	87,7	53,0	34,6	18,4	55,5	55,7	48,2	100,0	66,9	56,7	35,5
Canada	60,9	29,0	86,3	82,2	39,4	45,2	27,9	84,1	3,5	95,4	57,7	68,6
Chile	36,7	59,7	69,5	31,1	39,6	38,5	23,8	66,9	45,8	91,8	47,4	48,1
China	0,0	11,5	13,5	53,2	89,5	26,0	92,9	44,5	33,8	27,7	59,0	93,5
Colombia	3,9	55,8	40,0	28,3	58,1	16,1	14,9	42,6	23,9	35,9	48,6	36,5
Croatia	60,1	55,0	0,0	41,8	13,8	49,8	37,0	65,7	75,7	42,1	32,2	33,7
Czech Republic	61,1	88,6	59,1	45,1	21,8	61,1	69,9	75,4	80,2	85,3	54,0	50,8
Denmark	100,0	24,5	97,0	75,1	18,8	49,8	77,0	87,3	60,4	95,4	85,8	50,3
Dominican Republic	12,3	72,9	38,7	22,7	40,5	40,4	21,1	42,8	70,9	67,4	38,1	6,5
Egypt	34,2	53,3	33,2	29,8	69,3	18,1	5,5	25,5	0,0	68,3	59,7	23,6
Estonia	47,7	99,2	87,7	33,3	0,0	68,2	70,8	77,7	62,9	90,4	52,1	29,3
Finland	82,1	21,9	100,0	60,9	17,5	46,1	64,2	84,4	6,0	91,0	36,3	49,5
France	85,9	30,4	71,8	79,6	44,8	35,6	55,1	75,4	76,2	93,9	25,5	77,9
Germany	78,7	29,6	85,8	89,5	43,4	44,0	72,7	77,7	53,5	87,2	21,6	78,4
Greece	41,0	77,5	66,5	56,4	21,4	36,5	44,5	70,3	74,7	88,2	27,8	47,4
Guatemala	0,7	63,9	20,6	13,5	55,6	37,9	0,0	41,1	81,3	85,7	34,1	15,0
Hungary	74,8	87,6	65,7	44,8	22,9	62,6	49,6	68,6	60,6	94,6	70,8	50,1
India	29,8	75,1	36,6	32,8	100,0	5,0	28,4	21,4	61,8	27,5	70,0	71,8
Indonesia	5,9	53,8	2,2	35,9	74,8	34,3	41,8	30,3	34,9	41,9	28,3	41,6
Ireland	55,0	67,6	74,5	55,7	7,9	76,6	19,4	92,4	60,2	99,2	71,8	41,4
Israel	71,2	46,0	61,8	48,7	31,9	42,3	26,9	75,9	63,4	91,0	54,1	48,9
Italy	77,3	63,3	55,2	72,1	42,4	34,3	38,6	80,8	73,4	84,2	40,7	70,8
Japan	53,7	51,4	72,6	73,8	50,9	0,0	100,0	81,3	66,2	84,8	86,5	86,8
Korea, South	40,2	61,1	84,8	74,1	44,6	37,0	80,2	84,8	71,1	46,1	55,7	79,2
Latvia	50,0	91,2	61,4	20,3	6,2	50,7	44,9	62,2	72,7	86,5	32,0	24,5
Malaysia	24,6	32,9	52,7	48,4	50,4	76,5	6,3	60,9	70,9	41,1	81,6	48,4
Mexico	24,5	62,2	62,9	46,1	70,5	30,1	10,9	62,4	49,0	52,7	46,0	64,4
Morocco	36,1	61,0	48,6	22,5	51,3	30,4	0,9	29,9	64,2	35,6	0,0	23,5
New Zealand	57,0	40,2	98,4	55,7	18,3	37,6	51,6	80,1	68,5	98,1	82,6	50,8
Norway	80,9	14,2	55,5	68,3	18,2	40,3	78,6	88,2	31,8	84,8	39,1	49,0
Pakistan	0,7	47,2	36,5	13,3	86,4	8,9	1,5	18,2	30,7	42,9	40,6	26,0
Peru	14,4	74,4	40,1	20,5	52,7	13,7	3,6	43,8	41,6	91,9	48,0	26,5
Philippines	12,3	68,1	55,1	26,9	69,7	53,7	2,2	28,9	97,5	21,6	32,2	41,1
Poland	55,3	79,6	39,7	47,8	40,0	36,5	14,6	60,0	43,4	47,8	57,8	54,1
Portugal	68,0	65,3	49,8	50,8	24,3	42,3	14,9	78,5	63,2	82,4	42,9	46,7
Romania	52,9	74,5	42,3	18,3	34,8	41,5	17,4	48,9	86,1	70,5	28,9	38,7
Russia	20,0	46,4	16,1	47,8	57,0	33,2	55,6	40,6	1,8	45,1	56,9	69,7
Slovenia	70,1	64,5	62,1	37,6	1,7	61,2	51,9	77,1	95,2	73,7	33,1	34,5
Spain	62,0	30,3	58,3	65,3	37,7	38,3	43,7	83,4	80,3	94,7	21,0	72,4
Sweden	97,4	23,1	89,1	77,1	23,3	47,2	83,9	78,7	16,5	85,3	46,9	56,1
Switzerland	64,9	27,9	64,4	79,9	18,8	46,8	61,3	88,9	68,5	93,8	87,9	60,7
Thailand	22,9	86,0	59,4	40,7	55,4	57,7	43,1	54,2	44,3	40,0	52,4	52,9
Turkey	47,8	64,1	56,0	44,3	60,4	28,9	15,8	52,3	79,4	48,2	22,9	56,5
Ukraine	43,8	62,2	6,0	25,1	38,9	54,4	52,9	33,7	57,3	0,0	51,6	46,9
United Kingdom	57,6	39,5	77,6	88,6	49,8	33,0	52,5	79,3	46,9	100,0	84,9	74,8
United States	54,4	40,8	95,2	100,0	70,4	2,5	19,8	84,2	49,4	91,2	100,0	100,0
Uruguay	36,2	51,8	45,7	19,5	20,3	19,7	34,7	46,5	62,9	85,7	56,2	28,4

FUENTE: Cálculos propios.

Anexo 2. Códigos ISO

Nombre ISO del país	Codigo alfa 2	Codigo alfa 3
Albania	AL	ALB
Algeria	DZ	DZA
Argentina	AR	ARG
Australia	AU	AUS
Austria	AT	AUT
Bangladesh	BD	BGD
Belgium	BE	BEL
Brazil	BR	BRA
Bulgaria	BG	BGR
Canada	CA	CAN
Chile	CL	CHL
China	CN	CHN
Colombia	CO	COL
Croatia	HR	HRV
Czech Republic	CZ	CZE
Denmark	DK	DNK
Dominican Republic	DO	DOM
Egypt	EG	EGY
Estonia	EE	EST
Finland	FI	FIN
France	FR	FRA
Germany	DE	DEU
Greece	GR	GRC
Guatemala	GT	GTM
Hungary	HU	HUN
India	IN	IND
Indonesia	ID	IDN
Ireland	IE	IRL

Nombre ISO del país	Codigo alfa 2	Codigo alfa 3
Norway	NO	NOR
Pakistan	PK	PAK
Peru	PE	PER
Philippines	PH	PHL
Poland	PL	POL
Portugal	PT	PRT
Romania	RO	ROM
Russia	RU	RUS
Slovenia	SI	SVN
Spain	ES	ESP
Sweden	SE	SWE
Switzerland	CH	CHE
Thailand	TH	THA
Trinidad and Tobago	TT	TTO
Turkey	TR	TUR
Ukraine	UA	UKR
United Kingdom	GB	GBR
United States	US	USA
Uruguay	UY	URY
Israel	IL	ISR
Italy	IT	ITA
Japan	JP	JPN
Jordan	JO	JOR
Korea, South	KR	KOR
Latvia	LV	LVA
Malaysia	MY	MYS
Mexico	MX	MEX
Morocco	MA	MAR
New Zealand	NZ	NZL



Extreme Value Theory: An Application to the Peruvian Stock Market Returns

RODRÍGUEZ, GABRIEL

Pontificia Universidad Católica del Perú (Perú)

E-mail: gabriel.rodriguez@pucp.edu.pe

ABSTRACT

Using daily observations of the index and stock market returns for the Peruvian case from January 3, 1990 to May 31, 2013, this paper models the distribution of daily loss probability, estimates maximum quantiles and tail probabilities of this distribution, and models the extremes through a maximum threshold. This is used to obtain the better measurements of the Value at Risk (VaR) and the Expected Short-Fall (ES) at 95% and 99%. One of the results on calculating the maximum annual block of the negative stock market returns is the observation that the largest negative stock market return (daily) is 12.44% in 2011. The shape parameter is equal to -0.020 and 0.268 for the annual and quarterly block, respectively. Then, in the first case we have that the non-degenerate distribution function is Gumbel-type. In the other case, we have a thick-tailed distribution (Fréchet). Estimated values of the VaR and the ES are higher using the Generalized Pareto Distribution (GPD) in comparison with the Normal distribution and the differences at 99.0% are notable. Finally, the non-parametric estimation of the Hill tail-index and the quantile for negative stock market returns shows quite instability.

Keywords: Extreme Value Theory; Value-at-Risk (VaR); Expected Short-Fall (ES); Generalized Pareto Distribution (GPD); Gumbel Distribution; Exponential Distribution; Fréchet Distribution; Extreme Loss; Peruvian Stock Market.

JEL classification: C22; C58; G32.

MSC2010: 62M10; 62P20; 91G70; 91G80.

Teoría de valores extremos: una aplicación a los retornos bursátiles peruanos

RESUMEN

Usando observaciones diarias del índice y los retornos bursátiles para el caso Peruano desde el 3 de enero de 1990 hasta el 31 de mayo de 2013, este documento modela la distribución de las probabilidades de pérdidas diarias, estima los cuantiles máximos y las colas de la distribución y finalmente, modela los valores extremos usando un umbral máximo. Todo esto es usado para obtener una mejor medida del valor en riesgo (VaR) y de la pérdida esperada (ES) al 95% y 99% de confianza. Uno de los resultados de calcular el bloque máximo anual de los retornos bursátiles negativos es la observación que el retorno bursátil más negativo (diario) es 12.44% en el 2011. El parámetro de forma es igual a -0.020 y 0.268 para los bloques anuales y trimestrales, respectivamente. En consecuencia en el primer caso tenemos una distribución de tipo Gumbel. En el otro caso se tiene una distribución de cola pesada (Fréchet). Los valores estimados para el VaR y el ES son más elevados utilizando una distribución de tipo Pareto Generalizada (GPD) en comparación con la distribución normal y las diferencias al 99% son remarcables. Finalmente, la estimación no paramétrica del índice de cola de Hill y del cuantil para retornos negativos muestra ser bastante inestable.

Palabras claves: Teoría de valores extremos; valor en riesgo (VaR); pérdida esperada (ES); distribución de Pareto Generalizada (GPD); distribución de Gumbel; distribución exponencial; distribución de Fréchet; pérdida extrema; mercado bursátil peruano.

Clasificación JEL: C22; C58; G32.

MSC2010: 62M10; 62P20; 91G70; 91G80.



1 Introduction

As part of the Peruvian economy's good performance in recent years, the financial sector has played a significant role in terms of the objective of economic growth and capital accumulation. Nonetheless, the global financial crisis that began in the fourth quarter of 2007 affected the Peruvian capitals market and brought about a sharp fall in the General Index of the Lima Stock Exchange (IGBVL) of 59.78%, and in the Selective Index (ISBVL) of 59.73%. This event illustrates that big losses occur as a result of extreme movements in the markets, and hence that financial risk is related to the possible losses that investors can suffer in these markets; see Jorion (2001).

In general, the series of stock market returns have heavy-tailed distribution, due to which, unlike traditional distributions, the distribution of stock market returns possess greater probabilistic density on the tails. The above has, as a consequence, greater probability of extreme losses and it is necessary to analyze the tails of the distribution through the use of methodologies in the context of the Extreme Values Theory (EVT). I seek to capture in the best way possible the sudden movements of the performances of financial assets associated with the tails of the distribution, and thus allow better measurement of the behavior of financial asset performance¹. The recent financial crisis put in evidence the existence of multiple faults in the form of risk modeling, and this in turn prompted notable criticism of the different mathematical models and traditional statistics employed by companies in attempts to predict the risk. In 1993, the members of the Bank for International Settlements (BIS) gathered in Basel and amended the Basel Accords to require that banks and other financial institutions keep sufficient capital in reserve to cover ten days of potential losses based on the 10-day Value at Risk (VaR)².

The estimation of VaR by way of traditional models is not entirely adequate, because many of the techniques employed are based on the assumption that the financial returns follow a normal distribution. In this context, the measurement of risk through traditional measures occasions large losses to market participants because of the unexpected falls in financial market returns. Another measure of risk is that proposed by Artzner *et al.* (1999), called expected shortfall or expected loss (Expected Shortfall - ES) which is an expectation of loss conditioned to exceeding the indicated VaR level. One of the objectives of financial risk management is the exact calculation of the magnitudes and probabilities of big financial losses that are produced at times of financial crisis. It is thus of relevance to model the probability of loss distribution and estimate the maximum quantiles and tail probabilities associated with this distribution; see Zivot and Wang (2006).

The modern EVT started with von Bortkiewicz (1922). Thereafter, Fisher and Tippett (1928) laid the foundations of the asymptotic theory of the distributions of extreme values. Hill (1975) introduces a general approach for inference around the behavior of the tail of a distribution, while Danielsson and De Vries (1997) believe that a specific estimation of the form of the tail of foreign currency returns is of vital importance for adequate risk assessment.

¹Important texts include Embrechts *et al.* (1997) and Coles (2001). Other references applied to finance and financial risk management are Diebold *et al.* (1998), Danielsson and De Vries (1997), McNeil (1998a, 1998b) and Longin (2000).

²Danielsson *et al.* (2001) hold that the Committee on Banking Supervision was wrong to consider the risk to be endogenous and affirm that VaR can destabilize an economy and generate breaks which would not otherwise occur. In this way, the authors leave open the possibility that traditional financial models employed to measure and diagnose the risk have a certain degree of inconsistency, primarily because certain assumptions of these models are incapable of capturing the behavior of the indices that are used to measure risk. In particular, it is found that traditional models have a poor performance against sudden movements of these indices in a context of crisis.

On the other hand, Embrechts *et al.* (1997) present the probabilistic models and techniques with the aim of mathematically describing extreme events in the unidimensional case. McNeil (1998a) reduces data from the S&P500 index to 28 annual maximums corresponding to the period 1960-1987, and adjusts them to a Fréchet distribution. In this way, they calculate the estimations of various levels of returns, as well as the confidence interval at 95% for a 50-year level of return—which on average must be exceeded in just one year— every fifty years. The most probable calculated value is 7.4, but there is a great deal of uncertainty in the analysis as the confidence interval is approximately [4.9, 24].

Moreover, McNeil (1998b) considers the estimation of quantiles in the marginal distribution tail in the series of financial returns, utilizing statistical methods of extreme values based on the distribution limit of maximum blocks of stationary time series. The author proposes a simple methodology for the quantification of the worst possible scenarios, with losses of ten or twenty years.

Diebold *et al.* (1998) hold that the literature on the EVT is more accurate for the exact estimation of the extreme quantiles and tail probabilities of the financial assets³. McNeil (1999) shows a general vision of the EVT in the management of risks as a method for modeling and measuring extreme risks, concentrating on the peaks through a threshold. McNeil and Frey (2000) propose a method for estimating VaR and relate it to the risk measurements that describe the conditional distribution tail of a series of heteroskedastic financial yields.

Moreover, Longin (2000) present an application of the EVT to calculate VaR of a position in the market. For Embrechts *et al.* (2002), the modern risk management requires an understanding of stochastic dependence. The authors conduct a discussion on joint distributions and the use of copulas as descriptions of dependency among random variables.

Tsay (2002) applies the EVT to the logarithm of profitability of IBM shares for the period from July 3, 1962 to December 31, 1998 and finds that the range of fluctuation of the daily yields, excluding the crisis of 1987, fluctuates between 0.5% and 13%. He also estimates the Hill estimator and finds stable results for a minimum and a maximum value of the biggest n -th observation of this estimator. Tsay (2002) performs the estimation for different sample sizes (monthly, quarterly, weekly, and yearly) and concludes that the estimation of the scale and location parameters increase in modulus when the sample size increases. The shape parameter is stable for extreme negatives values when the sample size is greater than 62 and is approximately equal to a -0.33. The estimator of the shape parameter is small, significantly different to zero, and less stable for positive extremes. The result for the annual sample size has high variability when the number of subperiods is relatively small.

According to Delfiner and Gutiérrez Girault (2002), the returns in developing markets are characterized by being more leptokurtic compared to the returns of more developed economies; see also Humala and Rodríguez (2013) for stylized facts in the Peruvian stock market. The authors estimate an autoregressive AR-GARCH model of stochastic volatility, and then apply the EVT to the distribution tail of standardized residuals of the model by estimating a generalized Pareto distribution with a view to obtaining a better estimation of the probability when extreme losses are presented.

Finally, McNeil *et al.* (2005) provide two main types of models of extreme values. The most traditional models are maximum block, which are models for the biggest ordered observations of

³Diebold *et al.* (1998) demonstrate the existence of a trade off between the bias error and the variance when the largest n -th observation increases in Hill's tail index estimator (Hill, 1975).

big samples of identically distributed observations. The other group of models are for threshold exceedances and apply to all big observations that exceed a high level. They are generally considered very useful for practical applications, given their more efficient use (often limited) of the data on the extreme results.

Using daily observations of the index and stock market returns for the Peruvian case from January 3, 1990 to May 31, 2013, this paper models the distribution of daily loss probability, estimates maximum quantiles and tail probabilities of this distribution, and models the extremes through a maximum threshold. This is used to obtain the best measurements of the Value at Risk (VaR) and the Expected Short-Fall (ES) at 95% and 99%. One of the results on calculating the maximum annual block of the negative stock market returns is the observation that the largest negative stock market return (daily) is 12.44% in 2011. The shape parameter is equal to -0.020 and 0.268 for the annual and quarterly block, respectively. Then, in the first case, we have that the non-degenerate distribution function is Gumbel-type. In the other case, we have a thick-tailed distribution (Fréchet). Estimations of VaR and ES are higher using the Generalized Pareto Distribution (GPD) in comparison with the normal distribution and the differences at 99.0% are notable. Finally, the non-parametric estimation of the Hill tail-index and the quantile for negative stock market returns shows quite instability.

This paper is structured as follows: Section 2 describes the main definitions associated with EVT, as well as the method for estimating the main measurements of risk, VaR and ES. Section 3 presents the results, utilizing a sample of daily returns of the Peruvian stock market. Section 4 presents the main conclusions.

2 Methodology

In this Section, we closely follow and employ the notation in Zivot and Wang (2006). The EVT provides the statistical tools to model the unknown cumulative distribution function of the random variables that represent the risk or losses, especially in those situations where large losses are produced. Let $\{X_1, X_2, \dots, X_n\}$ independent and identically distributed (*i.i.d.*) random variables that symbolize the risk or expected losses, which have an unknown cumulative distribution function $F(x) = \Pr[X_i \leq x]$. $M_n = \max[X_1, X_2, \dots, X_n]$ is specified as the worst loss in a n -size sample of losses. In virtue of the assumption of *i.i.d.*, the cumulative distribution function of M_n is

$$\Pr[M_n \leq x] = \Pr[X_1 \leq x, X_2 \leq x, \dots, X_n \leq x] = \prod_{i=1}^n F(x) = F^n(x).$$

It is assumed that the function F^n is unknown, and, moreover, it is known that the function of empirical distribution is not a good approximation of $F^n(x)$. According to Fisher-Tippett Theorem (Fisher and Tippett, 1928)⁴, an asymptotic approximation is obtained for F^n based on the standardization of the maximum value; that is, $Z_n = \frac{M_n - \mu_n}{\sigma_n}$ where $\sigma_n > 0$ and μ_n are measurements of scale and position, respectively. In this way, for Fisher and Tippett (1928), the maximum standardized value converges to a distribution function of generalized extreme value (GEV), defined as:

$$H_\xi(z) = \begin{cases} \exp[-(1 + \xi z)^{-1/\xi}], & \text{for } \xi \neq 0 \text{ and } 1 + \xi z > 0 \\ \exp[-\exp(-z)], & \text{for } \xi = 0 \text{ and } -\infty \leq z \leq \infty \end{cases}$$

⁴This theorem is analogous to the Central Limit Theorem for extreme values.

where ξ is denominated the shape parameter and determines the behavior of the tail of $H_\xi(\cdot)$ ⁵. This distribution is not degenerated and is generalized in the sense that the parametric shape summarizes three types of known distributions. Moreover, if $\xi = 0$, H is a Gumbel distribution; if $\xi > 0$, H is a Fréchet distribution; and finally if $\xi < 0$, H is a Weibull distribution.

The parameter shape ξ is associated with the behavior of the tail of the distribution F and decays exponentially for a function of power $1 - F(x) = x^{-1/\xi}L(x)$ where $L(x)$ changes slowly. The GEV distribution is not changed for the transformations of location and scale: $H_\xi(z) = H_\xi(\frac{x-\mu}{\sigma}) = H_{\xi,\mu,\sigma}(x)$. For a large size n , Fisher-Tippett Theorem (Fisher and Tippet, 1928) can be interpreted as follows: $\Pr[Z_n < z] = \Pr[\frac{M_n - \mu_n}{\sigma_n} < z] \approx H_\xi(z)$. Assuming $x = \sigma_n z + \mu_n$, then: $\Pr[M_n < x] \approx H_{\xi,\mu,\sigma}(\frac{x-\mu_n}{\sigma_n}) = H_{\xi,\mu_n,\sigma_n}(x)$. This expression is useful for performing inference related to the maximum loss M_n . The expression depends on the parameter of ξ in form and the standardized constants σ_n and μ_n , which are estimated for maximum likelihood.

To perform the estimation of maximum likelihood, it is supposed to be a set of identically distributed losses from a sample of size T represented for $\{X_1, X_2, \dots, X_T\}$ that have an cumulative density function F . A sub-sample method is utilized to form the likelihood function for the parameters ξ, σ_n and μ_n from the GEV distribution for M_n . In this way, the sample is divided into m non-overlapping blocks of equal size $n = T/m$, with which we have $[X_1, \dots, X_n | X_{n+1}, \dots, X_{2n} | \dots | X_{(m-1)n+1}, \dots, X_{mn}]$ and where $M_n^{(j)}$ is defined as the maximum value of X_i in the block $j = 1, \dots, m$. The likelihood function for the parameters ξ, σ_n and μ_n of the GEV distribution is constructed from the maximum block sample of $\{M_n^{(1)}, \dots, M_n^{(m)}\}$. The likelihood log function assuming *i.i.d.* observations of the GEV distribution when $\xi \neq 0$ is

$$\log(\mu, \sigma, \xi) = -m \log(\sigma) - (1 + \frac{1}{\xi}) \sum_{i=1}^m \log[1 + \xi(\frac{M_n^{(i)} - \mu}{\sigma})] - \sum_{i=1}^m [1 + \xi(\frac{M_n^{(i)} - \mu}{\sigma})]^{-1/\xi}$$

with the restriction $1 + \xi(\frac{M_n^{(i)} - \mu}{\sigma}) > 0$. When $\xi = 0$, we obtain a Weibull distribution⁶.

It is important to discuss the limit distribution of extremes on high thresholds and the generalized Pareto distribution (GPD). When there is a succession of *i.i.d.* random variables $\{X_1, X_2, \dots, X_n\}$ associated with an unknown function of distribution $F(x) = \Pr[X_i \leq x]$, the extreme values are defined as the X_i values that exceed the high threshold κ . So, the variable $X - \kappa$ represents the excesses on this threshold. The distribution of excesses on the threshold κ is defined as a conditional probability: $F_\kappa(y) = \Pr[X - \kappa \leq y | X > \kappa] = \frac{F(y+\kappa) - F(\kappa)}{1 - F(\kappa)}$ for $y > 0$. This is interpreted as the probability that a loss exceeds the threshold κ for a value that is equal to or less than y , given that the threshold of κ has been exceeded. For $M_n = \max\{X_1, X_2, \dots, X_n\}$, defined as the worst loss in a n -sized sample of losses, the distribution function F satisfies Fisher-Tippett Theorem (Fisher and Tippet, 1928) and, for a sufficiently large κ , there is a positive function $\beta(\kappa)$. Thus, the surplus distribution is approximated through the GPD

$$G_{\xi,\beta(\kappa)}(y) = \begin{cases} 1 - [1 + \xi y / \beta(\kappa)]^{-1/\xi}, & \text{for } \xi \neq 0 \text{ and } \beta(\kappa) > 0 \\ 1 - \exp[-y / \beta(\kappa)], & \text{for } \xi = 0 \text{ and } \beta(\kappa) > 0 \end{cases}$$

⁵The expression $H_\xi(\cdot)$ is continuous in the shape parameter ξ .

⁶Distribution in the domain of attraction of the Gumbel-type distribution are thin-tailed distributions where practically all moments exist. If they are Fréchet-type, they include fat-tailed distributions such as Pareto, Cauchy or t-Student, among others. Some moments do not exist for these distributions.

defined for $y \geq 0$ when $\xi \geq 0$ and $0 \leq y \leq -\beta(\kappa)$ when $\xi < 0$. For a sufficiently high threshold κ , it is found that $F_\kappa(y) \approx G_{\xi, \beta(\kappa)}(y)$ for a wide range of loss functions $F(\cdot)$. To apply this result, the value of the threshold must be specified and the estimates of ξ and $\beta(\kappa)$ can be obtained.

There is a close connection between the GEV limit distribution for maximum blocks and the GPD for excesses with respect to the threshold. For a given value of κ , the parameters ξ , μ and σ of the GEV distribution determine the parameters ξ and $\beta(\kappa)$. It is clear that the shape parameter ξ of the GEV distribution is the same parameter ξ in the GPD and is independent of the threshold value κ . In consequence, if $\xi > 0$, the function F is Fréchet-type and the expression $G_{\xi, \beta(\kappa)}(y)$ is denominated classic Pareto distribution; when $\xi = 0$, the function F is Gumbel-type and $G_{\xi, \beta(\kappa)}(y)$ follows an exponential distribution. Finally, it is found that $0 \leq y \leq -\beta(\kappa)/\xi$ when $\xi < 0$, so the function F is Weibull-type and $G_{\xi, \beta(\kappa)}(y)$ is a type-II Pareto distribution. The parameter ξ is the shape or tail-index parameter and is associated with the rate of decay of the tail of the distribution, and the decreasing parameter β is the shape parameter and is associated with the position of the threshold κ .⁷

Now, assuming that the parameter of form is $\xi < 1$, the mean excess function above the threshold κ_0 will be $E[X - \kappa_0 | X > \kappa_0] = \frac{\beta(\kappa_0)}{1-\xi}$ for any $\kappa > \kappa_0$, and it is found that the excess function of the mean $e(\kappa) = E[X - \kappa | X > \kappa] = \frac{\beta(\kappa_0) + \xi(\kappa - \kappa_0)}{1-\xi}$. Analogously, for any value of $y > 0$, the following conditions hold: $e(\kappa_0 + y) = E[X - (\kappa_0 + y) | X > \kappa_0 + y] = \frac{\beta(\kappa_0) + \xi y}{1-\xi}$. Therefore, to graphically deduce the threshold value for the GPD, we get the excess function of the empirical mean: $e_n(\kappa) = \frac{1}{n_\kappa} \sum_{i=1}^{n_\kappa} [x_{(i)} - \kappa]$, where $x_{(i)}$ ($i = 1, 2, \dots, n_\kappa$) is the value of x_i such that $x_i > \kappa$. With the previous expression, a graph representation of $e_n(\kappa)$ is constructed with the mean excess on the vertical axis. This graph can be interpreted as follows: if the slope is rising, it indicates thick tail behavior; but if there is a downward trend, this shows thin tail behavior in the distribution; and finally, if the slope of the line is equal to zero, the behavior of the tail is exponential. If the line is straight and has a positive slope located above the threshold, then it is an example of Pareto-type tail behavior.

For the values of the maximum losses that exceed the threshold (that is, when $x_i > \kappa$), the threshold excess is defined as $y_i = x_{(i)} - \kappa$ for $i = 1, \dots, k$, in which the values of x_1, \dots, x_n have been denoted as $x_{(1)}, \dots, x_{(k)}$. When the threshold value is sufficiently large, then the sample $\{y_1, \dots, y_k\}$ can be expressed within a likelihood that is based on the unknown parameters ξ and $\beta(\kappa)$; that is, a random sample of a GPD.

When $\xi \neq 0$, the log likelihood function of $G_{\xi, \beta(\kappa)}(y)$ has the following form

$$\log[\beta(\kappa)] = -k \log[\beta(\kappa)] - \left[1 + \frac{1}{\xi}\right] \sum_{i=1}^k \log\left[1 + \frac{\xi y_i}{\beta(\kappa)}\right]$$

where $y_i \geq 0$ when $\xi > 0$ and $0 \leq y_i \leq -\beta(\kappa)/\xi$ when $\xi < 0$. If the parameter of form is $\xi = 0$, then the log likelihood function is

$$\log[\beta(\kappa)] = -k \log[\beta(\kappa)] - \beta(\kappa)^{-1} \sum_{i=1}^k y_i.$$

⁷For $\xi > 0$ (the most relevant case for risk administration purposes), it can be shown that $E[X^k] = \infty$ for $k \geq \alpha = 1/\xi$. If $\xi = 0.5$, $E[X^2] = \infty$ and the distribution of losses X does not have finite variance. Analogously, if $\xi = 1$, then $E[X] = \infty$.

To estimate the tails of the loss distribution for $F(x)$, and where $x > \kappa$, we use $F(x) = [1 - F(\kappa)]G_{\xi, \beta(\kappa)}(y) + F(\kappa)$. The previous expression is fulfilled for a sufficiently large threshold and in which $F_\kappa(y) \approx G_{\xi, \beta(\kappa)}(y)$.⁸

There are two common risk measurements: Value at Risk (VaR) and Expected Shortfall (ES). The VaR is the largest quantile of the distribution of loss; that is, $VaR_q = F^{-1}(q)$.⁹ For a given probability $q > F(\kappa)$, it is found that $\widehat{VaR}_q = \kappa + \frac{\widehat{\beta}(\kappa)}{\widehat{\xi}}[\frac{n}{k}(1-q)^{-\widehat{\xi}} - 1]$. The ES is the expected size loss, given that VaR_q is exceeded: $ES_q = E[X|X > VaR_q]$. This equation is related to VaR_q in accordance with $ES_q = VaR_q + E[X - VaR_q|X > VaR_q]$, where the second term is the mean of excess of the distribution $F_{VaR_q}(y)$ on a threshold VaR_q . The approximation of the GPD (due to the translation property) to $F_{VaR_q}(y)$ has the shape parameter ξ and the scale parameter $\beta(\kappa) + \xi[VaR_q - \kappa] : E[X - VaR_q|X > VaR_q] = \frac{\beta(\kappa) + \xi(VaR_q - \kappa)}{1 - \xi}$ provided that $\xi > 1$. Moreover, it is found that the GPD approximates $\widehat{ES}_q = \frac{\widehat{VaR}_q}{1 - \widehat{\xi}} + \frac{\widehat{\beta}(\kappa) - \widehat{\xi}\kappa}{1 - \widehat{\xi}}$.¹⁰ It is also possible to perform the non-parametric estimation of the shape parameter ξ or the tail-index parameter $a = 1/\xi$ of the distributions $H_\xi(z)$ and $G_{\xi, \beta(\kappa)}(y)$ utilizing Hill method (Hill, 1975), in which $\xi > 0$ ($\alpha > 0$), is generated by the same thick-tailed distributions in the domain of attractions of a Fréchet GEV. Considering a sample of losses $\{X_1, X_2, \dots, X_T\}$, the statistical order is defined as $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(T)}$ for a positive whole k , and the Hill estimator of ξ is defined as $\widehat{\xi}^{Hill}(k) = \frac{1}{k} \sum_{j=1}^k [\log X_{(j)} - \log X_{(k)}]$. The Hill estimator of α is $\widehat{\alpha}^{Hill}(k) = \frac{1}{\widehat{\xi}^{Hill}(k)}$.¹¹

3 Empirical Evidence

Figure 1 shows the series for the closing prices of the General Index of the Lima Stock Exchange (IGBVL)¹². The series is of daily frequency and covers the period from January 3, 1990 to May 30, 2013. The returns are defined as $r_t = \log[\frac{P_t}{P_{t-1}}]$, which are shown in Figure 2. Empirically, the returns display certain properties as marginal thick-tailed distributions, nonexistence of correlation, and dependency across these; though they are highly correlated if it concerns the squared results or their absolute value; see Humala and Rodríguez (2013) for a more detailed description about the stylized facts.

By way of motivation, Figure 3 shows the GEV cumulative distribution function for the distribution function $H_\xi(\cdot)$, which adopts Fréchet, Weibull or Gumbel form of distribution when the shape parameter is $\xi = 0.5$, $\xi = -0.5$ or $\xi = 0$, respectively, and for general values of z , the parameter of position μ and the parameter of scale σ . In this particular case, the Fréchet distribution is defined for $z > -2$ and the Weibull distribution is only defined for $z < 2$. Figure 4 shows the GEV probability density function $H_\xi(\cdot)$ for the non-degenerate Fréchet, Weibull and Gumbell dis-

⁸It is assumed that $x = \kappa + y$.

⁹If $X \sim N(\mu, \sigma^2)$, then $VaR_{0.99} = \mu + \sigma q_{0.99}$.

¹⁰If $X \sim N(\mu, \sigma^2)$, it is found that $ES_{0.99} = \mu + \sigma \frac{\phi(z)}{1 - \Phi(z)}$.

¹¹It can be seen that if F is located in the domain of attraction of a GEV distribution, then $\widehat{\xi}^{Hill}(k)$ converges in probability to ξ when $k \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$, and $\widehat{\xi}^{Hill}(k)$ is normally asymptotically distributed with asymptotic variance: $avar[\widehat{\xi}^{Hill}(k)] = \frac{\xi^2}{k}$. Via the delta method, $\widehat{\alpha}^{Hill}(k)$ is normally asymptotically distributed with asymptotic variance $avar[\widehat{\alpha}^{Hill}(k)] = \frac{a^2}{k}$.

¹²The closing prices of the General Index of the Lima Stock Exchange are taken into account from Monday until the closure price on Friday. Moreover, it should be recalled that non-working days are not considered, and more generally the days on which the market was closed.

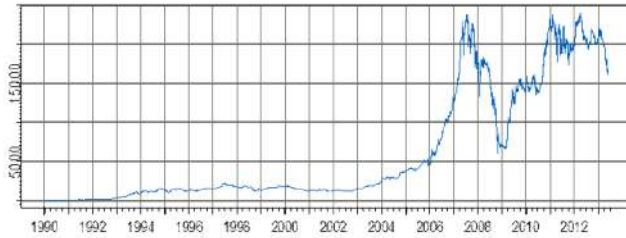


Figure 1: Closing prices of the General Index of the Lima Stock Exchange.

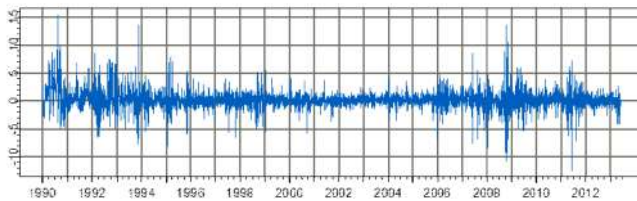


Figure 2: Daily Percentage Returns of the Stock Market of Peru.

tribution functions depending on the values $\xi = 0.5$, $\xi = -0.5$ and $\xi = 0$ for the shape parameter, respectively. The Fréchet and Weibull distributions are defined for $z > -2$ and $z < 2$, respectively.

Figure 5 shows the GEV density function for negative stock market returns. The horizontal axis represents the standardized value Z_n of the maximum value of the block M_n with respect to the measurements of scale and position. The vertical axis shows the probability associated with the GEV density function. It is observed that this distribution does not have the form of a known distribution and the maximum probability shows positive asymmetry.

The p -quantile of a distribution function G is defined by the value X_p such that $G(p) = X_p$; that is, the value of X_p that leaves the p -percentile of probability to its left. If a distribution function G is continuous and thus strictly growing, the quantile function is the inverse of the distribution function G and is usually denoted as G^{-1} . Figure 6 shows the q-q plot, taking the normal distribution as theoretical distribution to be contrasted with the distribution of stock market returns. Let us note that a straight diagonal line is not observed (approximately), and so it is concluded that the distribution of the variable is not the same as the comparison distribution, showing evidence that the distribution of negative stock market returns is unknown.

Subsequently, the annual maximum block of the negative stock market returns is calculated. Figure 7 shows four representations for this annual maximum block. In the upper-left graph, it is shown the largest negative return of the period analyzed, which reaches 12.44% in 2011. The upper-right graph shows the histogram where the horizontal axis represents the annual maximum blocks. In the lower-left representation, the q-q plot is shown, contrasting again the distribution of stock market returns for the period of analysis. In the vertical axis, the quantiles of the referential theoretical distribution are represented (Gumbel distribution, H_0), which satisfies $H_0^{-1}(p) = -\log[-\log(p)]$ and the horizontal axis represents the empirical quantiles for the annual

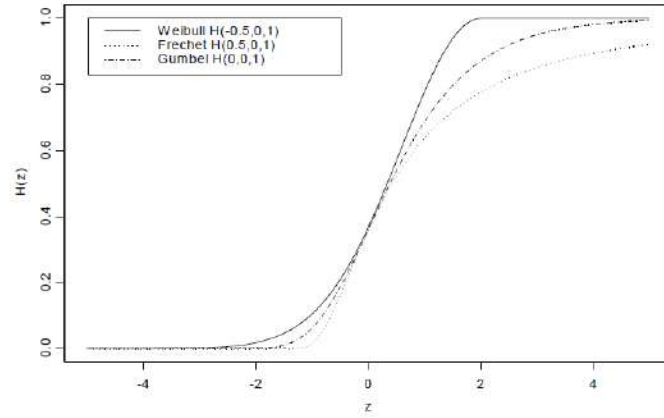


Figure 3: Generalized Extreme Value CDFs for Fréchet, Weibull and Gumbel.

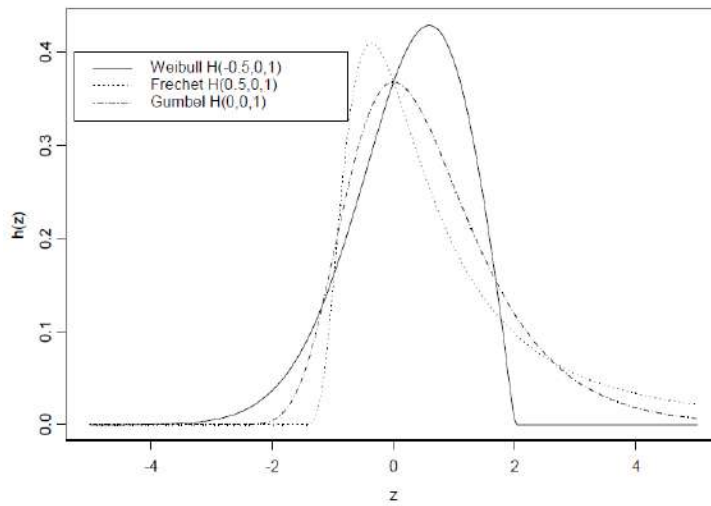


Figure 4: Generalized Extreme Value pdfs for Fréchet, Weibull and Gumbel.

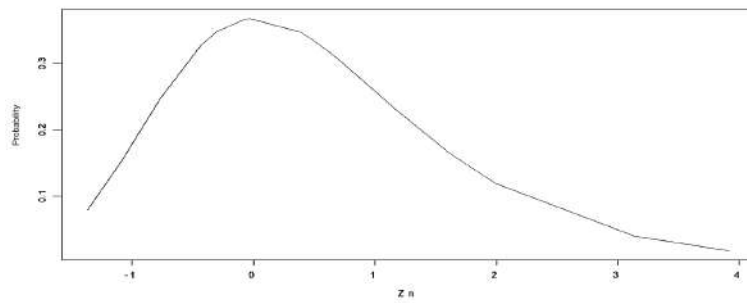


Figure 5: GEV pdf for Daily Returns in Peru.

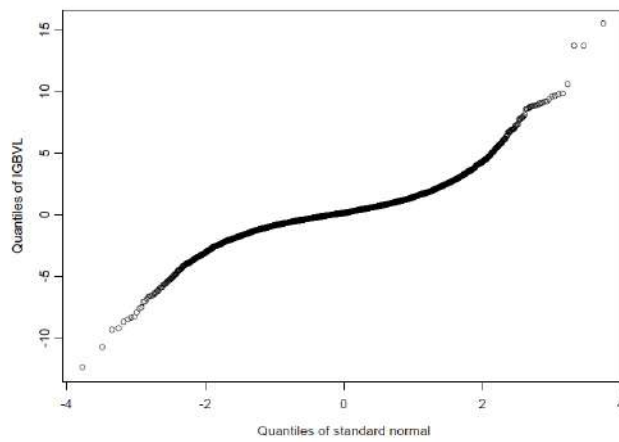


Figure 6: Normal q-q plot for the Daily Percentage Returns in Peru.

maximum blocks of the distribution of stock market returns. It is observed that the point cloud fits approximately the bisector for the axes, which suggests that the distribution of the variable of real data (empirical distribution) is the same as the distribution of comparison (Gumbel distribution). Finally, the lower-right graph shows the development of the records (new maximum) for the negative stock market returns, together with the expected number of returns for the *i.i.d.* data. In this graph, it is observed that the data was not within the confidence interval (delimited by dotted lines), due to which it can be concluded that the data is not consistent with the *i.i.d.* behavior.

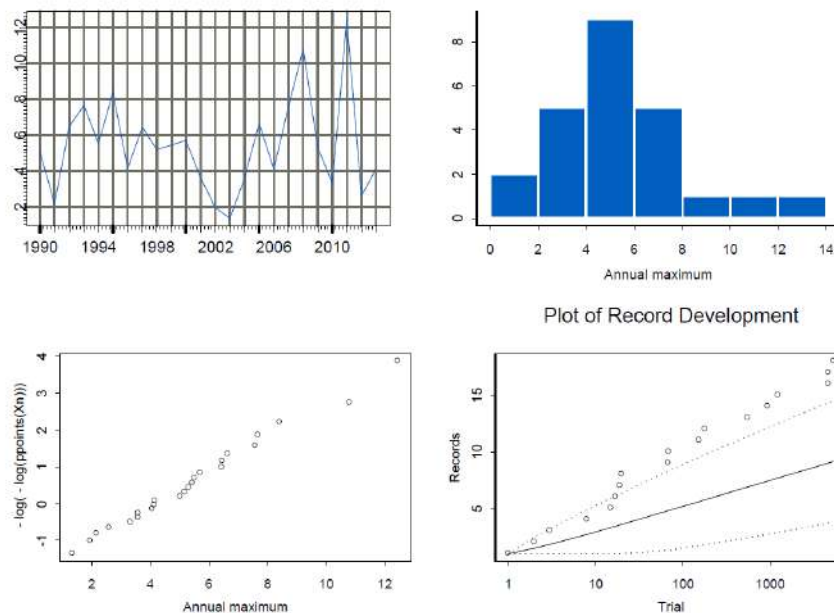


Figure 7: Annual block maxima, histogram, Gumbel q - q plot and records summary for the Daily Stock Returns in Peru.

Analogously to the lower-left graph in Figure 7, Figure 8 shows the q - q plot, using as referential distribution the Gumbel distribution H_0 . Unlike Figure 7, the horizontal axis represents the standardization of maximum value Z_n . As shown previously for the Gumbel distribution, the quantiles satisfy $H_0^{-1}(p) = -\log[-\log(p)]$ and the points of the quantiles correspond to the standardization of the maximum value Z_n and indicate a GEV distribution with $\xi = 0$.

Then, the entire annual value of the number of observations in each maximum block is determined by $M_n^{(i)}$ $i = 1, \dots, m$ for the stock market returns, with $m = 24$. The shape parameter ξ is statistically insignificant ($\hat{\xi} = -0.020$, $t_{\hat{\xi}} = -0.126$) and so the value of this parameter is equal to zero ($\xi = 0$). Moreover, the asymptotic interval at 95% of confidence for ξ is $[-0.337, 0.2968]$ and indicates the considerable uncertainty related to the value of ξ . This result determines the tail behavior of the GEV distribution function of stock market returns, and it is concluded that the non-degenerate distribution function is Gumbel-type. The position and scale parameters (standardized constants) are statistically significant: $\hat{\mu}_n = 4.232$, $t_{\hat{\mu}_n} = 8.713$ and $\hat{\sigma}_n = 2.098$, $t_{\hat{\sigma}_n} = 5.954$, respectively.

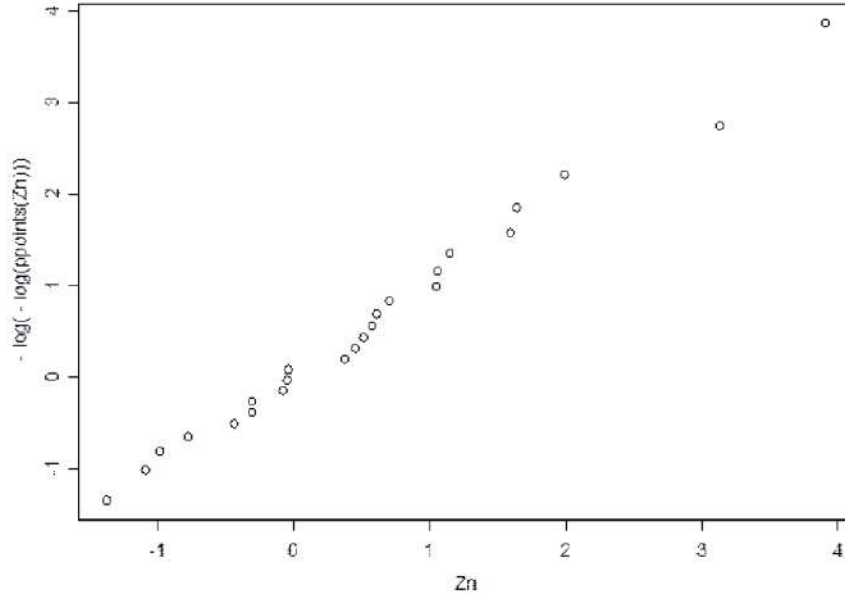


Figure 8: Gumbel q-q plot of Stock Daily Returns in Peru.

Utilizing the estimation by maximum likelihood of the adjusted GEV distribution for the maximum annual block of negative stock market returns, the following question can be answered: How probable is it that the maximum annual negative profitability for the following year exceeds the above negative returns? This probability is calculated utilizing the expression $H_{\xi, \mu_n, \sigma_n}(x)$ where the maximum block is equal to 1.68%, and so there is a 1.68% possibility that a new maximum record of negative performance will be established during the following year.

A similar analysis is possible by considering the GEV distribution adjusted for the quarterly maximum block for the data from the series of stock market returns. The maximum block for the return of this series is $m = 94$. It is observed that estimated standard asymptotic errors are much lower when quarterly blocks are employed. The shape estimator is $\hat{\xi} = 0.268$ ($t_{\hat{\xi}} = 2.031$) and in this case, the asymptotic interval at 95% of confidence for ξ is $[-0.004, 0.532]$ and contains only positive values for the shape parameter, indicating a thick-tailed distribution, with the estimated probability equal to 0.0172. Finally, the estimations of the position and scale parameters are significant: $\hat{\mu}_n = 2.419$, $t_{\hat{\mu}_n} = 9.186$ and $\hat{\sigma}_n = 2.419$, $t_{\hat{\sigma}_n} = 13.705$, respectively.

In Figure 9, the asymmetric form of the asymptotic confidence interval can be observed. Figure 9 allows us to give response to the following question: What is the level of stock market return for the last forty years? The estimated point of the level of return (11.67%) is at the point where the vertical line cuts at the maximum point of the asymmetric curve. The upper extreme point of the confidence interval of 95% is approximately 22%; this point is located where the asymmetrical curve cuts at the straight horizontal line. In addition, Figure 10 shows the estimation of the expected yield level of the negative stock market returns for forty years with a confidence level band of 95% based on the model of GEV for an annual maximum block. In Figure 10, a horizontal line is drawn dividing the graph into two halves and corresponding to the expected level of return (11.67%). In addition, the dot horizontal line below the line for the expected level of return corresponds to the

lowest level of return (9.33%); whereas the dot horizontal line above the expected level of return corresponds to the highest level of return (22.21%). In this Figure, the 24 annual maximum blocks ($m = 24$) obtained from the real data of the stock market returns (point cloud) can also be seen, in which only two points exceed the lower extreme.

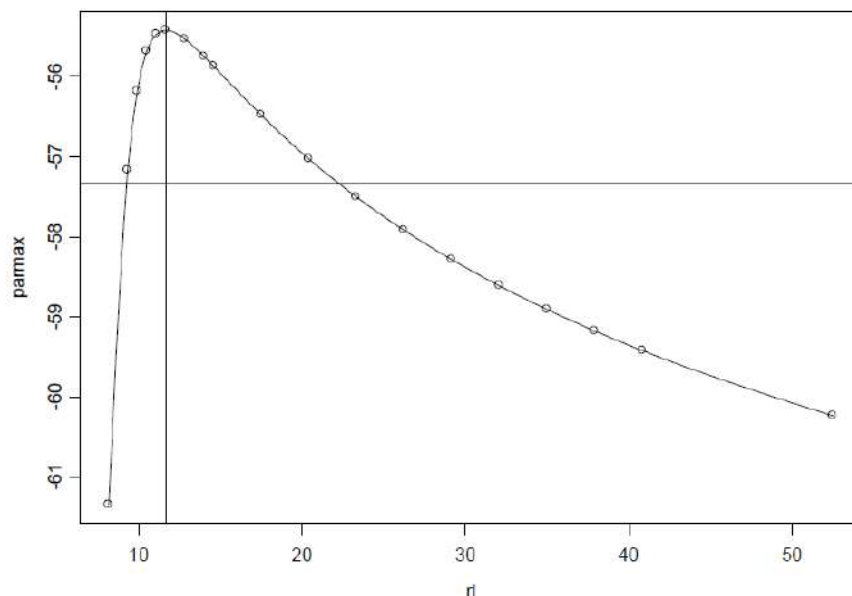


Figure 9: Asymptotic 95% Confidence Interval for the 40 Year Return Level.

Following Zivot and Wang (2006), the 40-year level of return can also be estimated based on the GEV fitted to quarterly periods as a maximum, where forty years correspond to 160 quarters, obtaining the lowest and highest level of return; see Figures 11 and 12. In Figure 12, the horizontal line located on the mean corresponds to the expected level of return (17.18%); the dot horizontal line below the level of expected return corresponds to the lowest level of return (10.88%) and the dotted line above the expected level of return corresponds to the highest level of return, being equal to 40.68%. Figure 12 also shows the 94 ($m = 94$) quarterly maximum blocks obtained from the data on stock market returns (point cloud) below the lower confidence band of the confidence interval, except for two points, which means that the return for these 160 quarters must be above these values.

According to Zivot and Wang (2006), modeling only the maximum block of data is inefficient if there are other data being available on the extreme values. A more efficient, alternative approach that utilizes more observations is to model the behavior of extreme values above a given high threshold. This method is called peaks over threshold (POT). Another advantage of the POT method is that the common risk measurements, such as VaR and ES, can be calculated easily¹³.

To motivate the importance of the foregoing in Figure 13, the calculation of the cumulated

¹³For risk administration purposes, insurance companies may be interested in the frequency of occurrence of a large demand above a certain threshold, as well as the average value of the demand that exceeds the threshold. In addition, they may be interested in the daily VaR and ES. The statistical models for extreme values above a threshold can be used to tackle these questions.

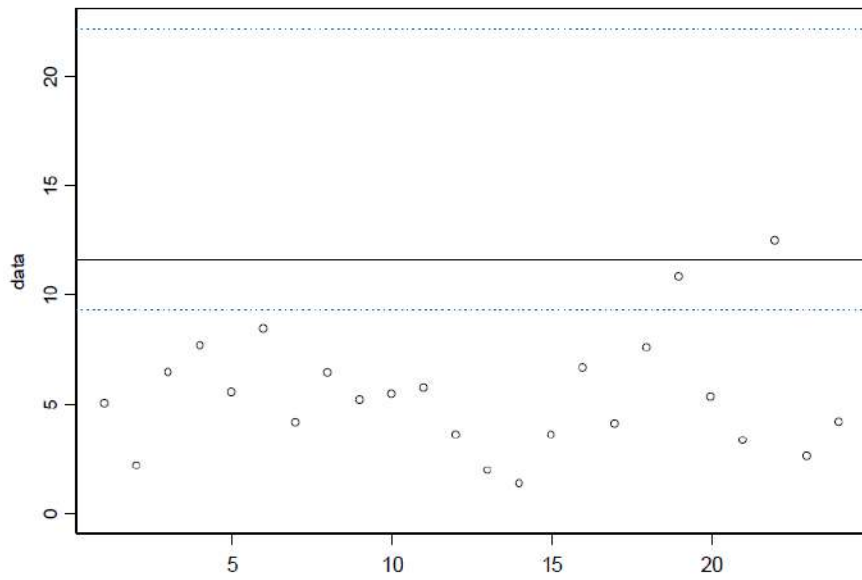


Figure 10: Estimated 40-Year Return Level with 95% Confidence Band for the Stock Daily Returns in Peru.

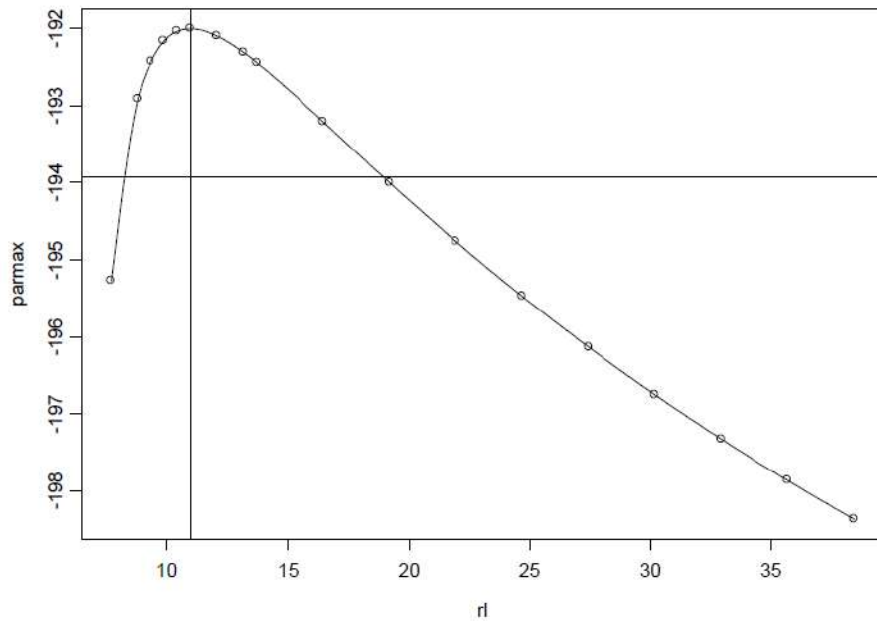


Figure 11: Asymptotic 95% Confidence Interval for the 160-Quarterly Return Level.

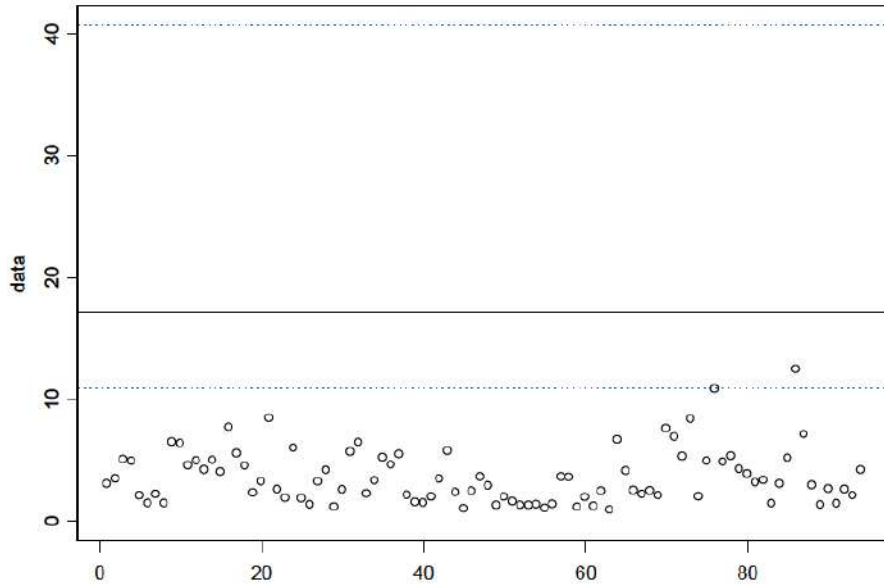


Figure 12: Estimated 160-Quarters Return Level with 95% Confidence Band for the Stock Daily Returns in Peru.

distribution and probability functions are shown with $\beta(\kappa) = 1$ for Pareto ($\xi = 0.5$), exponential ($\xi = 0$), and Pareto type II ($\xi = -0.5$) distributions. The Pareto type II distribution is defined only for $y < 2$. According to Zivot and Wang (2006), to infer the tail behavior of the observed losses, a q - q plot is created using the exponential distribution as reference distribution. If the excess on the threshold is a thin-tailed distribution, then the generalized Pareto distribution is exponential with $\xi = 0$ and the q - q plot should be linear. Deviations from the linearity in the q - q plot indicate thick-tailed behavior ($\xi > 0$) or bounded tails ($\xi < 0$).

In Figure 14, the q - q plot is observed for the distribution of negative stock market returns through the threshold when this is equal to one ($\kappa = 1$). The selection of the threshold under this methodology is complicated. Hence, to identify the threshold, there are a number of methodologies, such as parametric and graphic methods¹⁴. Figure 14 shows a slight deviation from the linearity for negative stock market returns, which leads us to conclude that the distribution of negative stock market returns is a thick-tailed distribution.

The main distributional model for excess through the threshold is the GPD. So, when defining the excess function of the empirical sample mean, a graph can be prepared in which the expectation of the values above the threshold κ is represented, once the threshold has been exceeded on the vertical axis associated with each of the thresholds. This is useful for discerning tails of a distribution against the different possible levels of threshold κ on the horizontal axis. Figure 14 must be approximately linear at the level of the selected threshold, and it is possible to determine intervals that allow for selecting the threshold. In general, the thick-tailed distributions give way to a mean excess function that tends toward the infinite for high values of κ and displays a linear form with a

¹⁴One of these methods is the mean excess plot.

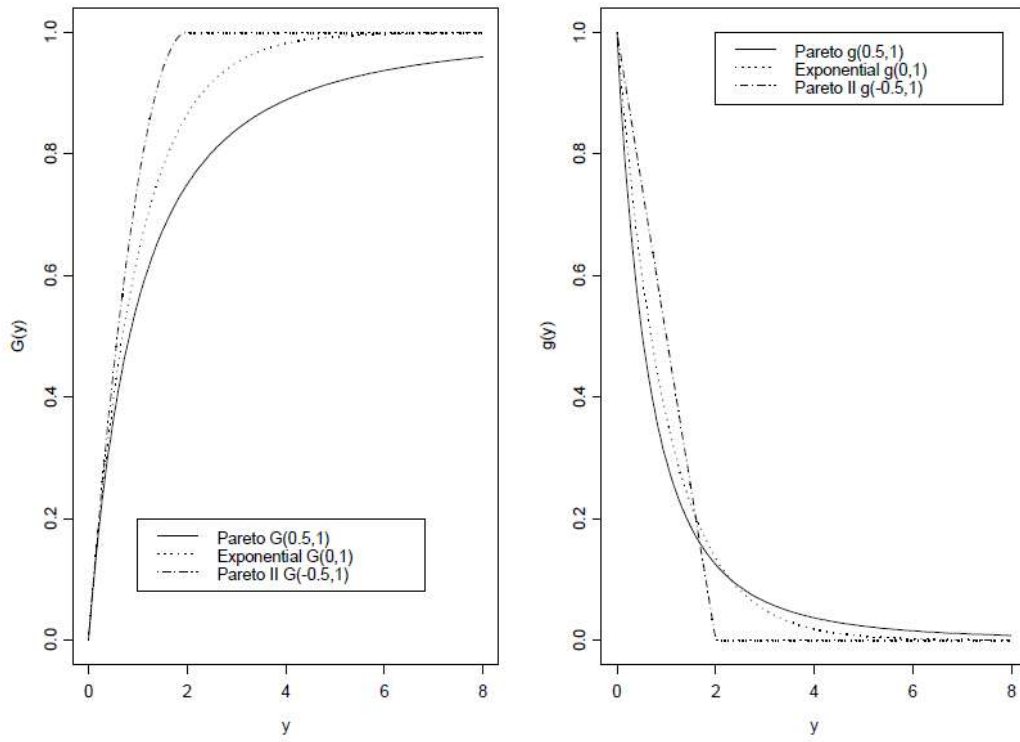


Figure 13: Generalized Pareto CDFs and pdfs for Pareto.

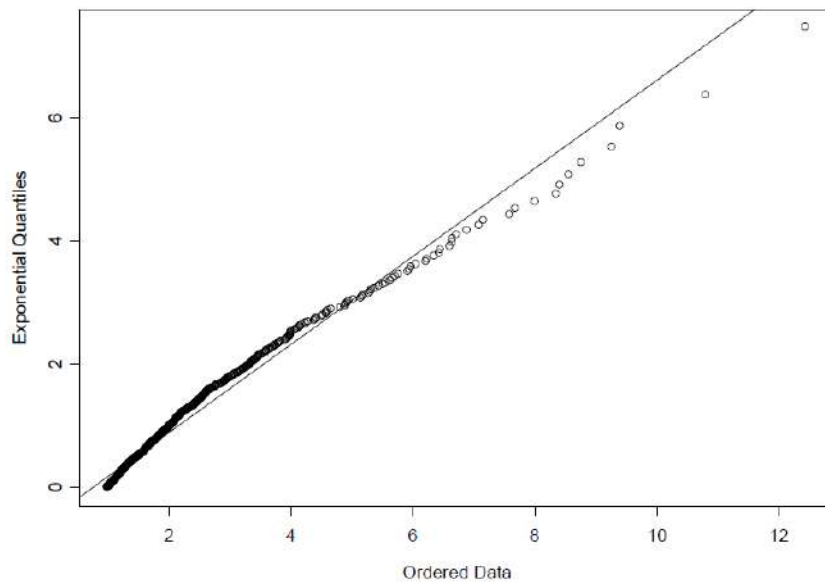


Figure 14: Q-q plot with Exponential Reference Distribution for the Stock Daily Negative Returns over the Threshold $\kappa = 1$.

positive slope. All the above mentioned is shown in Figure 15. On the vertical axis, the empirical mean excess is represented for the series of stock market returns, and on the horizontal axis, the threshold κ is represented. If the points that are represented have an upward trend (upward slope), this indicates thick-tailed behavior in the sample represented, as well as a GPD with positive shape parameter $\xi > 0$. If there is a downward trend (negative slope), this involves the thin-tailed behavior of the GPD with negative parameter $\xi < 0$. Finally, if an approximately linear graph is obtained (tending toward the horizontal axis), this indicates a GPD and the tail behavior is exponential (an exponential excess distribution), with the shape parameter approximately equal to zero ($\xi = 0$). From the observation of Figure 15 on mean excess, a declining trend for the data up to the value of the threshold $\kappa = -1$ is detected, which indicates a thin-tailed distribution therein; but from this value for the threshold, there is an upward trend for the data, indicating the thick-tailed behavior in the sample represented¹⁵.

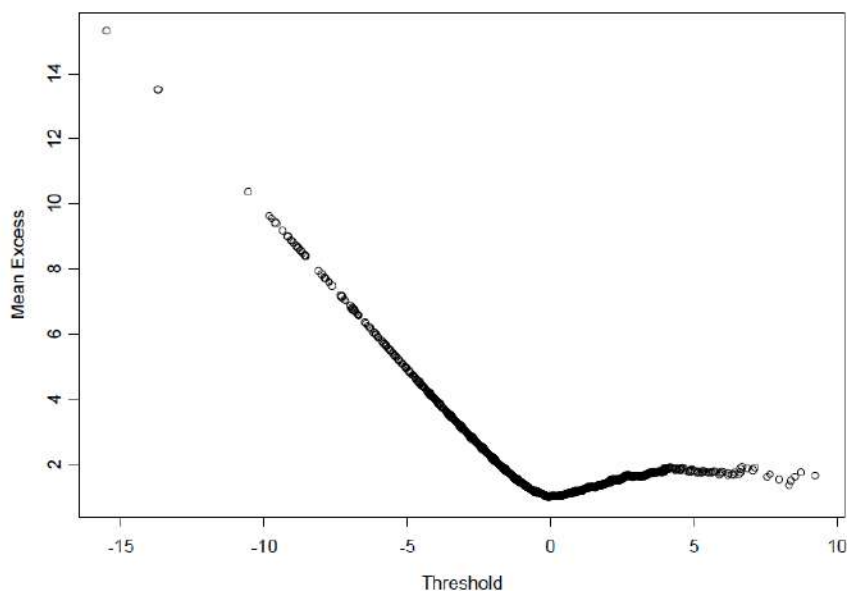


Figure 15: Mean Excess Plot for the Stock Daily Negative Returns.

Once the mean excess function is determined, the tails of the distribution of negative stock market losses are estimated for the period of analysis by way of the maximum likelihood estimation of the parameters $\beta(\kappa)$ and ξ of the GPD. To determine this estimation, a threshold κ must be specified, which must be big enough for the approximation of the GPD to be valid, but must also be small enough so that a sufficient number of observations is available for an exact fit; see Carmona (2004). In Figure 16 on the excess of the mean for stock market returns, it is observed that the threshold has a value of one (that is $\kappa = 1$) and may be appropriate for the GPD to be valid. The estimation of the parameters indicates $\hat{\xi} = 0.185$ ($t_{\hat{\xi}} = 4.463$) and $\hat{\beta}(1) = 0.941$ ($t_{\hat{\beta}(1)} = 18.801$). If the estimated shape parameter for the GPD ($\hat{\xi} = 0.185$) is compared with the GEV estimations of

¹⁵Empirical evidence on different behavior in the tails of the Peruvian stock market returns is also found in Rodríguez (2017) and Lengua Lafosse *et al.* (2014).

the yearly and quarterly maximum blocks, it is seen that this is higher in the case where the analysis is based on quarterly data ($\hat{\xi} = 0.268$), but less if annual data are used ($\hat{\xi} = -0.020$), being close to zero in the latter case. According to Carmona (2004), Figure 16 shows the underlying distribution. In the left-side graph, the survival function $1 - F(x)$ is represented on the vertical axis instead of the cumulative distribution function $F(x)$ and it is seen that the curve moves very close to the horizontal axis, so it is extremely difficult to correctly quantify the quality of fitting. Since this graph is not very useful, the right-side graph represents the survival function in logarithmic scale on the vertical axis, which helps ensure that the fit of the distribution is adequate by taking into account the available data. Observing both graphs in Figure 16, it is concluded that the fitting is good.

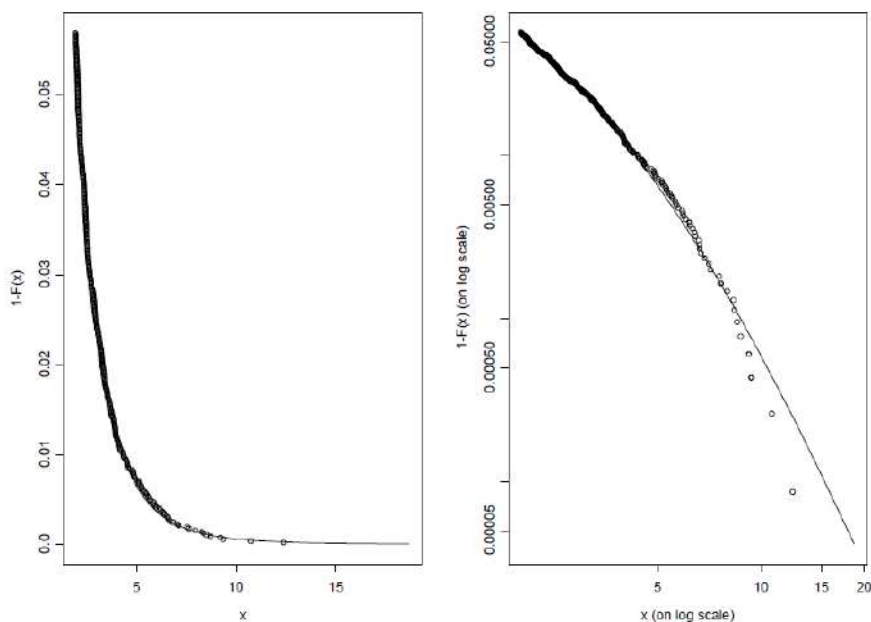


Figure 16: Diagnostic Plots for GPD Fit to Daily Negative Returns on Stock Index.

Changing the value of the threshold brings about changes in the estimation of ξ , so the stability of the shape parameter must be considered. It is optimal not to depend on a procedure that is too sensitive to small changes in the threshold selection. In effect, since there is no clear procedure for the selection of the threshold with a high level of accuracy, the estimation of the shape parameter must remain robust in the face of variations in the errors in the selection of this threshold. The best way to verify the stability of the parameter is through visual inspection. Now, to show how the estimation by maximum likelihood in the shape parameter ξ varies with the threshold which has been selected, we can observe Figure 17 where the lower horizontal axis represents the maximum number of threshold excesses, and it is assumed to be equal to six hundred. On the upper horizontal axis, the threshold is represented, whereas the estimation of the shape parameter with a confidence of 95% is represented on the vertical axis. Figure 17 shows that ξ has a very stable behavior close to 0.185 for threshold values lower than 1.91.

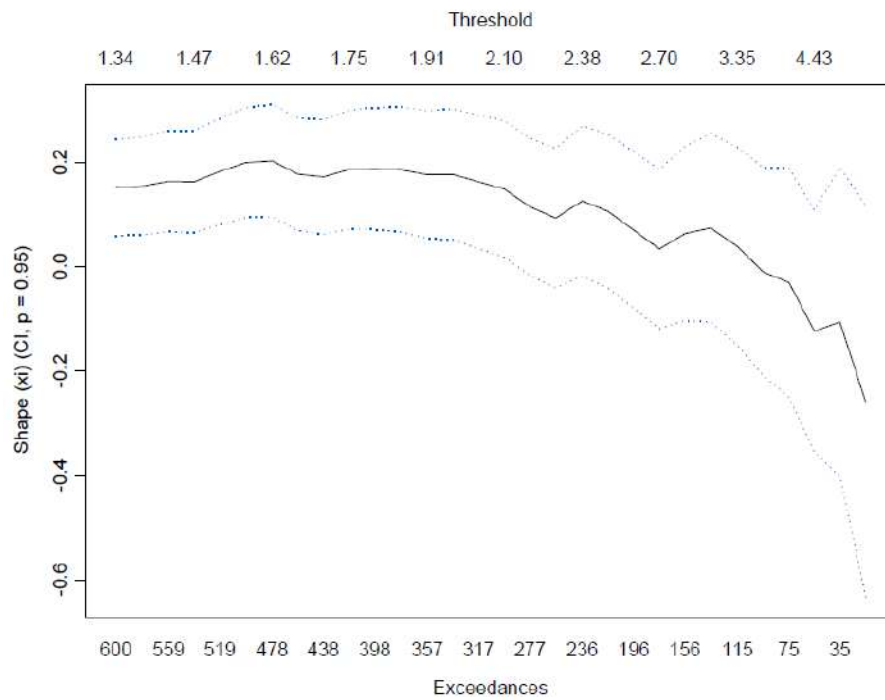


Figure 17: Estimates of the Shape Parameter for the Daily Negative Returns as a Function of the Threshold.

In accordance with the above, Figure 18 shows how the estimator of the GPD shape parameter varies with the threshold, where starting value of the quantile has been specified based on the data equal to 0.9, to be used as a threshold that fits the model. In the upper horizontal axis in Figure 18, the proportion of the points included in the estimation is represented. This information is useful to decide whether it is necessary or not to take seriously some of the estimations of ξ that appear on the left and right extremes of Figure 18. The central part of the graph should essentially be horizontal, though this does not always result in a straight line, when the empirical distribution of the data can be reasonably explained by a GPD. Finally, on the lower horizontal axis in Figure 18, the threshold is represented. The leftmost part of the graph should be ignored because of the following reason: If the threshold is too small, much of the data (that must be included in the center of the distribution) contributes to the estimation of the tail, skewing the result. Analogously, the rightmost part of the graph should also be ignored as few points will contribute to the estimation. This is the case in the current situation, and a value of $\xi = 0.185$ appears to be a reasonable estimation for the intersection of a horizontal line fitted to the central part of the graph.

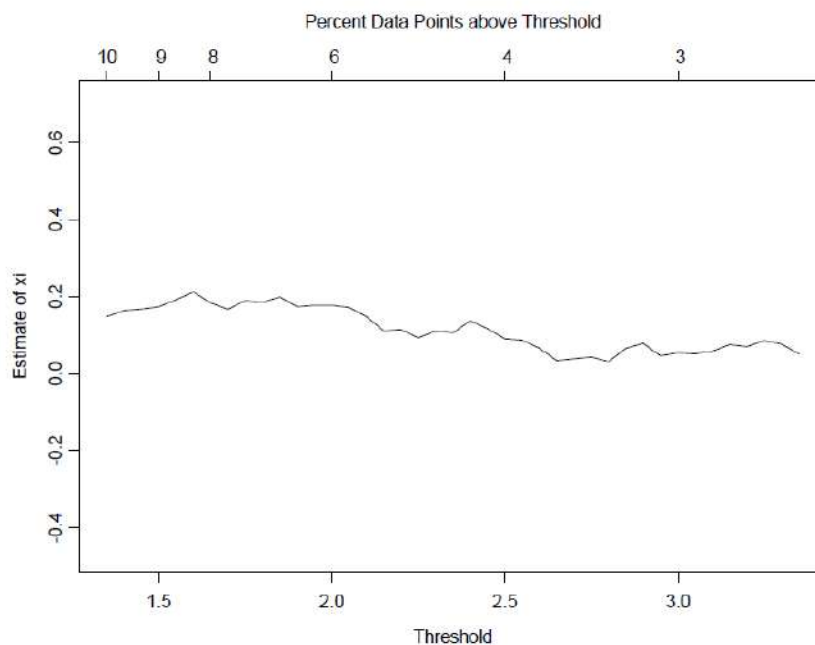


Figure 18: Estimates of the Shape Parameter with Time-Varying Threshold.

On the other hand, according to Zivot and Wang (2006), it is often desirable to estimate the parameters ξ and $\beta(\kappa)$ through the maximum likelihood estimation of the GPD separately for the upper and lower tails of the negative returns (POT analysis). In the analysis of the mean excess through the threshold (see Figure 15), the lower threshold is determined, which is equal to -1 . Analogously, with the help of Figure 15, the upper threshold is selected, which is equal to 1 . The estimations for the lower threshold are $\hat{\xi} = 0.185$ and $\hat{\beta}(\kappa) = 0.912$, while for the upper threshold they are $\hat{\xi} = 0.217$ and $\hat{\beta}(\kappa) = 1.087$. Note that the estimated values of the parameters ξ and $\beta(\kappa)$ are the same as the estimates in the analysis of excess on the previously realized threshold when

the threshold equal to minus one was estimated (see Figure 15).

The next analysis is very similar to the previous one: the difference lies in the presence of two tails instead of one. Figure 19 shows the q - q plots of the excess on the specified threshold versus that of the quantiles of the GPD by employing the estimated shape parameters of the upper and lower tails. In this case, the lower tail (above graph) could start at minus one, and the upper tail (below graph) at one¹⁶. In both representations in Figure 19, it is seen that the point sets form a straight line up to a certain stage, so it is reasonable to assume that a GPD fits the data. Moreover, the two estimations for the shape parameter ξ are not the same based on the particular selections of the upper (0.217) and lower (0.185) thresholds. If the distribution is not symmetrical, there is no special reason for the two values of ξ to be the same; that is, there is no particular reason why, in general, the polynomial decay of the right and left tails must be identical.

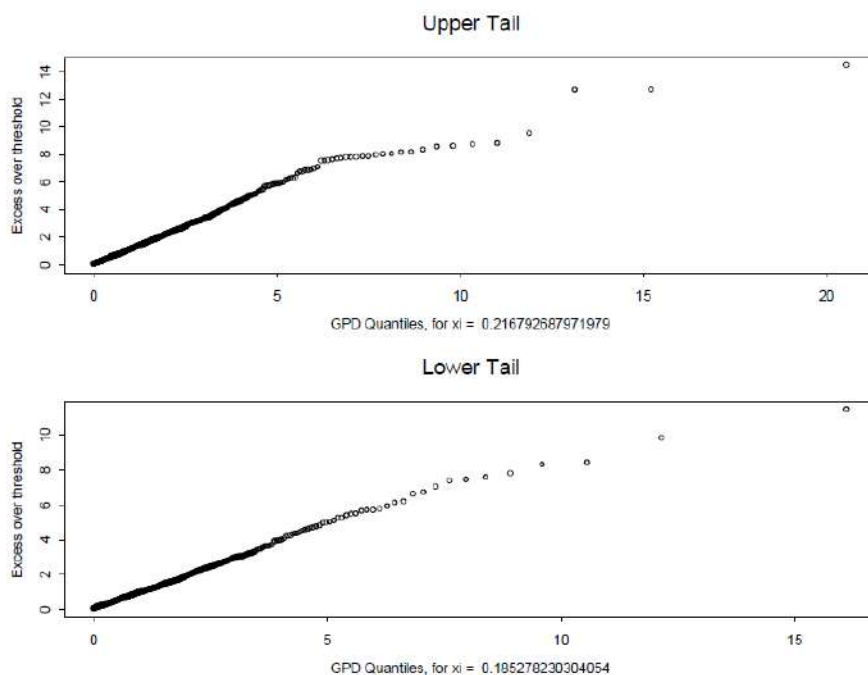


Figure 19: Estimated Tails when Distributions does not have Lower or Upper Limit.

At the beginning of this paper it is held that, for a better understanding of the risk, VaR and ES should be borne in mind to quantify the financial risks. The estimation of these risk measurements is performed for negative stock market returns for the quantiles $q = 0.95$ and $q = 0.99$, which are based on the GPD¹⁷. For the case of the GPD, it is inferred that with probability of 5%, $\widehat{VaR}_{0.95} = -2.146\%$ and, given that the return is less than -2.146% , then $\widehat{ES}_{0.95}$ is -3.562% . Analogously, with probability of 1%, $\widehat{VaR}_{0.99} = -4.309\%$ and $\widehat{ES}_{0.99} = -6.217\%$ given that the

¹⁶It should be recalled that the upper and lower thresholds do not necessarily have to be equal in absolute value, as they are in this case.

¹⁷Under the assumption of normally distributed returns, it is found that $VaR_{0.99} = \mu + \sigma \times q_{0.99}$ and $ES_{0.99} = \mu + \sigma \times \frac{\phi(z)}{1-\Phi(z)}$ for the case of the quantile 0.99.

return is less than -4.309%. Compared with the results obtained utilizing a normal distribution, $\widehat{VaR}_{0.95}$ is less than the estimation of the GPD. Nonetheless, $\widehat{VaR}_{0.99}$ is higher in the case of the GPD in comparison with the normal distribution. In the case of $\widehat{ES}_{0.95}$ and $\widehat{ES}_{0.99}$, both are higher using the GPD approximation. The difference at 99.0% is remarkable and significant (6.217% in the GPD compared with 4.375% for the normal distribution).

Once adjusted to a model of GPD for the excess of stock market returns above a threshold, we proceed to the estimation of valid asymptotic confidence intervals for VaR_q and ES_q .¹⁸ These intervals can be visualized in Figure 20 with the tail estimate $\widehat{F}(x) = 1 - \frac{k}{n} [1 + \widehat{\xi} \times \frac{x-\kappa}{\beta(\kappa)}]$. The confidence intervals for VaR are [2.062, 2.240] and [4.048, 4.643] for 95% and 99%, respectively. With respect to ES, the intervals are [3.358, 3.839] and [5.595, 7.156] for 95% and 99%, respectively.

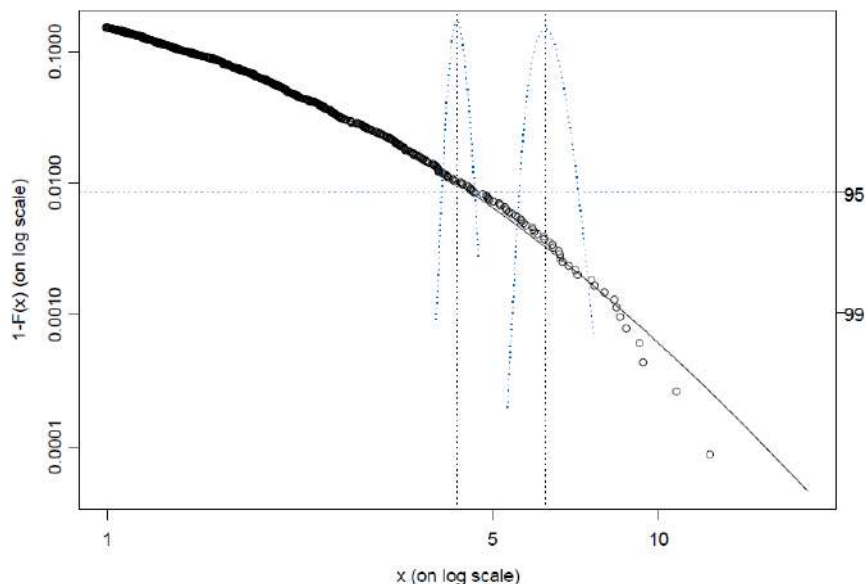


Figure 20: Asymptotic Confidence Intervals for $VaR_{0.99}$ and $ES_{0.99}$ based on the GPD Fit.

Figure 21 allows for an analysis of the sensitivity of VaR_q estimated in response to changes in the threshold κ . It is observed how the estimation by maximum likelihood of the parameter of form ξ varies with the threshold. In Figure 21, it is estimated that the behavior of the shape parameter is very stable and close to the estimated value of the Value at Risk (4.309) for threshold values less than four.

According to McNeil *et al.* (2005), the GPD method is not the only way to estimate the tails of a distribution as has been performed above. The other methodology for the selection of the threshold is based on the Hill estimator, estimating, in a non-parametric way, the Hill tail index $\alpha = 1/\xi$ and the quantile $x_{q,k}$ for the negative stock market returns. This estimator is often a good estimator of α , or its reciprocal ξ . In practice, the general strategy is to graph the Hill estimator for all possible values of k (numbers of excesses through the threshold). Practical experience suggests that the best options for k are relatively small —for example, between 10 and 50 of statistical

¹⁸ VaR_q and ES_q are based on the delta method of the likelihood log function profile.

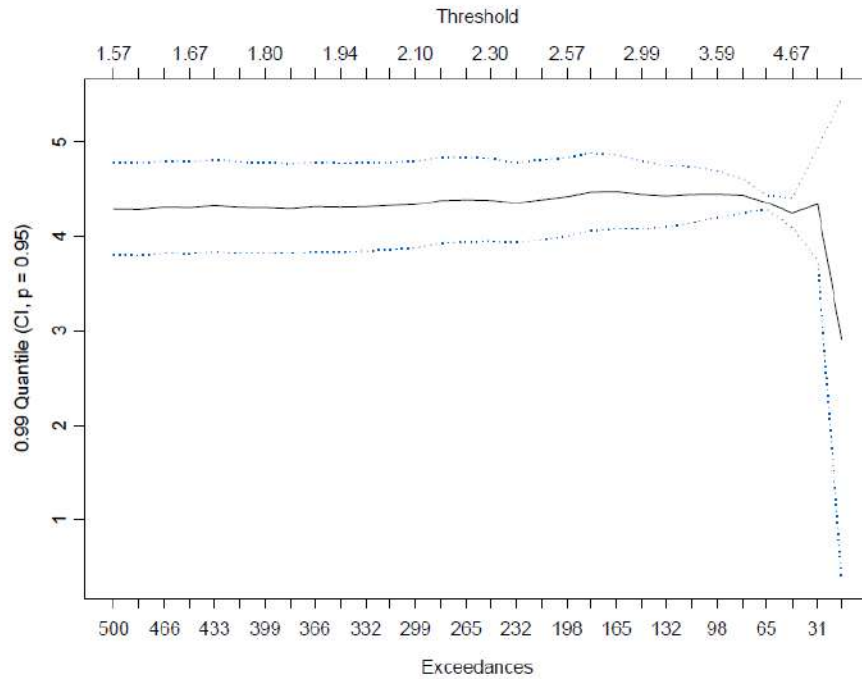


Figure 21: Estimation of $VaR_{0.99}$ as a function of the Threshold.

orders in a sample of size 1000. In Figure 22, the Hill estimator $\{k, \hat{\alpha}_{k,n}^{(H)} : k = 2, \dots, n\}$ is estimated for negative stock market returns of the shape parameter ξ . We expect to find a stable region for the Hill estimator where estimations are constructed based on the different numbers of statistical order. In Figure 22, the upper horizontal axis represents the threshold associated with the possible values of k ; in the lower horizontal axis, the number of observations included in the estimation is represented, and finally the confidence interval is observed at 95% (dotted lines). According to the results, it is observed that the estimation of the shape parameter does not stabilize as the statistical order increases hence, $\hat{\xi}^{Hill}(k)$ is quite unstable. It should be borne in mind that in practice, the ideal situation does not usually occur if the data does not come from a distribution with a tail that changes with regularity. If this occurs, the Hill method is not appropriate. The serial dependence on the data can also impair the performance of the estimator, although this can also be said of the estimator of the GPD.

4 Conclusions

Using daily observations of the index and stock market returns for the Peruvian case from January 3, 1990 to May 31, 2013, this paper models the distribution of daily loss probability, estimates maximum quantiles and tail probabilities of this distribution, and models the extremes through a maximum threshold. This is used to obtain the best measurements of VaR and ES at 95% and 99%.

One of the results on calculating the maximum annual block of the negative stock market returns

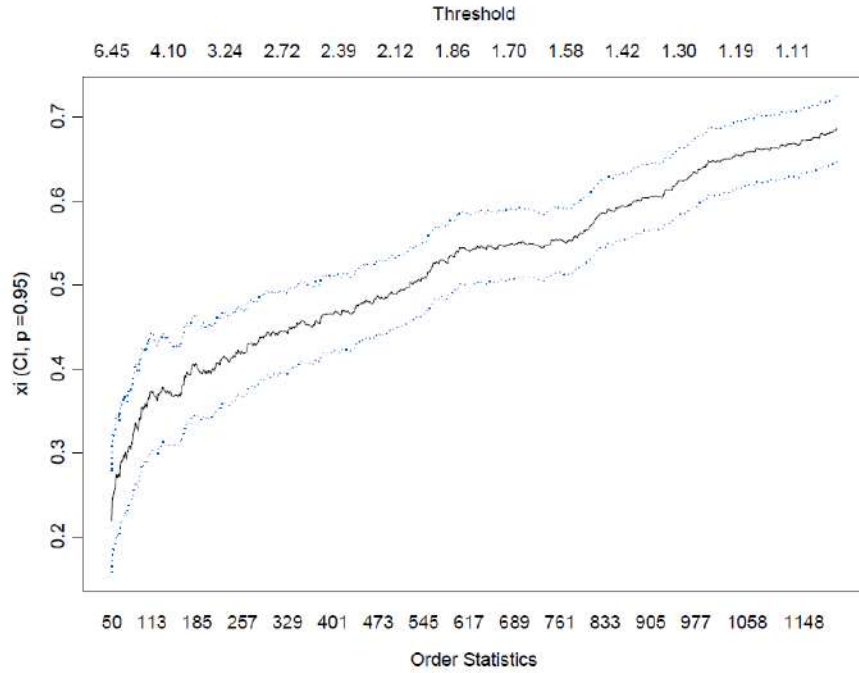


Figure 22: Estimates of the Hill ξ for the Daily Negative Returns.

is the observation of the fact that the largest negative stock market return (daily) is 12.44% in 2011. Moreover, if it is estimated that the probability of the maximum negative annual profitability for the following year exceeds all previous negative returns, turning out equal to 1.68, which means a probability of 1.68% of a negative maximum record of the negative yield being stabilized during the following year.

Then, by way of the estimator of maximum likelihood, the parameter of form and the asymptotic interval are estimated at 95% confidence level thereof for the annual and quarterly maximum block. The results indicate that the shape parameter is equal to -0.020 and 0.268, as well as the asymptotic interval $[-0.337, 0.2968]$ and $[-0.004, 0.532]$ for the maximum annual and quarterly block, respectively. The shape parameter estimation (-0.020) of the calculation of the maximum annual block of negative stock market returns is insignificant, due to which the value of this parameter is equal to zero and determines the tail behavior of the GEV distribution, and it is concluded that the non-degenerate distribution function is Gumbel-type. In the case of the estimation by maximum likelihood for the maximum quarterly block, a positive value was obtained for the shape parameter (0.268), with this being significant, indicating a thick-tailed distribution (Fréchet).

For the case of the GPD, it is inferred that with probability of 5%, the daily return would be as low as -2.146% and, given that the return is less than -2.146%, the average of the value of the return is -3.562%. Analogously, with probability of 1%, the daily returns could be as low as -4.309% with an average return of -6.217%, given that the return is less than -4.309%. Compared with the results obtained utilizing a normal distribution, the $\widehat{VaR}_{0.95}$ is smaller with the estimation of the GPD. Nonetheless, the $\widehat{VaR}_{0.99}$ is higher in the case of the GPD, in comparison with the normal distribution. In the case of $\widehat{ES}_{0.95}$ and $\widehat{ES}_{0.99}$, both are higher using the GPD approximation. The

difference in 99.0% is remarkable and significant (6.217% in the GPD, compared with 4.375% for the normal distribution).

Finally, the non-parametric estimation is performed for the Hill tail-index and the quantile for negative stock market returns, expecting to find a stable region for the Hill estimator. The results related to the estimation of the parameter do not stabilize as the statistical order increases, due to which the estimator of the Hill tail-index is quite unstable. This allows us to infer that the data do not come from a distribution with a tail that regularly changes, where the estimated values of the Hill parameter of form suggest a threshold close to one, according to their respective statistical order.

5 Acknowledgments

I appreciate the collaboration of Alfredo Calderón during the first part of the realization of this document when he was studying the Masters of Economics of the Graduate School at the Pontificia Universidad Católica del Perú (PUCP). I also thank useful comments by Paul Castillo (Central Reserve Bank of Peru and PUCP) and very constructive comments from the Editor of the Journal, Professor Ángel F. Tenorio and two anonymous referees. Any remaining errors are my responsibility.

References

- [1] Bortkiewicz, L. V. (1922), “Variationsbreite und mittlerer Fehler. Sitzungsber”, *Berli. Math. Gess* **3(21)**, 3-11.
- [2] Carmona, R. (2004), *Statistical Analysis of Financial Data in S-PLUS*, Springer-Verlang, London.
- [3] Coles, S. G. (2001), *An Introduction to Statistical Modeling of Extreme Values*, Springer-Verlang, London.
- [4] Danielsson, J. and De Vries, C. (1997), “Tail index and quantile estimation with very high frequency data”, *Journal of Empirical Finance* **4**, 241-257.
- [5] Danielsson, J., Embrechts, P., Goodhart, Ch., Keating, C., Muennich, F., Renault, O. and Shin, S. H. (2001), “An Academic Response to Basel II”, *LSE Financial Markets Group and ESCR Research Centre, Special Paper Series*, Special Paper No. 130.
- [6] Delfiner, M. T. and Gutiérrez Girault, M. A. (2002), “Aplicación de la teoría de valores extremos al gerenciamiento del riesgo”, *Serie Documentos de Trabajo, Universidad del CEMA: Área: finanzas*, No. 217.
- [7] Diebold, F. X., Schuermann, T. and Stroughair, J. D. (1998), “Pitfalls and opportunities in the use of extreme value theory in risk management”. In A.-P. N. Refenes, J.D. Moody and A.N. Burgess (eds.). *Advances in Computational Finance*, pp. 3-12. Amsterdam: Kluwer. Reprinted in *Journal of Risk Finance* **1**, 30-36.
- [8] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997), *Modeling Extremal Events for Insurance and Finance*, Springer-Verlang, London.

- [9] Embrechts, P., Klüppelberg, C. and Straumann, D. (2002), “Correlation and Dependency in Risk Management: Properties and Pitfalls”. In M. Dempster (ed.). *Risk Management: Value at Risk and Beyond*, pp. 176-223, Cambridge University Press.
- [10] Fisher, R. and Tippett L. H. C. (1928), “Limiting Form of the Frequency Distribution of the largest or smallest member of a sample”, *Mathematical Proceedings of the Cambridge Philosophical Society* **24**, 180-190.
- [11] Hill, B. M. (1975), “A Simple General Approach to Inference about the Tail of a Distribution”, *Annals of Statistics* **3**, 1163-1174.
- [12] Humala, A. and Rodríguez, G. (2013), “Some Stylized Facts of Returns in the Stock and Foreign Exchange Markets in Peru”, *Studies in Economics and Finance* **30(2)**, 139-158.
- [13] Jorion, P. (2001), *Value at Risk*, Second Edition, McGraw-Hill, New York.
- [14] Lengua Lafosse, P., Bayes, C. and Rodríguez, G. (2014), “An Empirical Application of Stochastic Volatility Models to Latin-American Stock Returns using *GH* Skew Student’s *t*-Distribution”, Working Paper 405, Department of Economics, Pontificia Universidad Católica del Perú.
- [15] Longin, F. (2000), “From value at risk to stress testing: The extreme value approach”, *Journal of Banking & Finance* **24**, 1097-1130.
- [16] McNeil, A. (1998a), “On Extremes and Crashes”, *RISK*, **11**, 99-104.
- [17] McNeil, A. (1998b), “Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory”, unpublished manuscript, Departement Mathematik, ETH Zürich.
- [18] McNeil, A. (1999), “Extreme value theory for risk managers”. In *Internal Modelling and CAD II*, pp. 93-113. London: Risk Waters Books.
- [19] McNeil, A. and Frey, R. (2000), “Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach”, *Journal of Empirical Finance* **7**, 271-300.
- [20] McNeil, A., Frey, R. and Embrechts, P. (2005), *Quantitative Risk Management: Concepts, Techniques, Tools*, Princeton Series in Finance. Princeton University Press.
- [21] Rodríguez, G. (2017), “Selecting between Autoregressive Conditional Heteroscedasticity Models: An Empirical Application to the Volatility of Stock Markets Returns in Peru”, *Revista de Análisis Económico* **32(1)**, 69-74.
- [22] Tsay, R.S. (2002), *Analysis of Financial Time Series*, Wiley Series in Probability and Statistics. Wiley-Interscience.
- [23] Zivot, E. and Wang, J. (2006), *Modelling Financial Times Series with S-PLUS*, Second edition. Springer-Verlag, Carey.



El contagio en el fracaso empresarial como consecuencia de la proximidad geográfica: un análisis con los estadísticos *join-count* aplicado al sector servicios

RODRÍGUEZ FUENTES, CHRISTIAN CAMILO

Facultad de Ciencias de la Empresa

Universidad Politécnica de Cartagena (España)

Correo electrónico: chris.rodriquez.fuentes@gmail.com

MATÉ SÁNCHEZ-VAL, MARÍA LUZ

Departamento de Economía Financiera y Contabilidad

Universidad Politécnica de Cartagena (España)

Correo electrónico: mluz.mate@upct.es

LÓPEZ HERNÁNDEZ, FERNANDO A.

Departamento de Métodos Cuantitativos e Informáticos

Universidad Politécnica de Cartagena (España)

Correo-e: fernando.lopez@upct.es

RESUMEN

El objetivo de este trabajo es contrastar el efecto de contagio en el fracaso empresarial como consecuencia de la proximidad geográfica entre empresas del mismo sector. Para ello, desarrollamos una aplicación empírica basada en el análisis de los estadísticos espaciales *join-count* sobre una muestra de empresas del sector servicios en el municipio de Murcia (España). Los resultados muestran un patrón de autocorrelación espacial significativo, de tal forma que la probabilidad de fracaso de una empresa no solo depende de sus características internas sino también de su proximidad a otras empresas fracasadas. Estos resultados pueden ser un interesante punto de partida para el desarrollo de análisis que incluyan los efectos de la proximidad geográfica entre empresas en la literatura de fracaso empresarial.

Palabras claves: autocorrelación espacial; fracaso empresarial; tests *join-count*.

Clasificación JEL: C49; L80; R39.

MSC2010: 62P20; 91G70; 91B72; 91B82; 91B32.

The Contagion on Business Failure by the Geographical Proximity: An Analysis with the Join-Count Tests in the Service Sector

ABSTRACT

This paper aims to contrast the spillover effects in business failure derived from the geographic proximity among firms of the same sector. To get this purpose, we develop an empirical application based on the analysis of the join-count statistics on a sample of firms of service sector located in the municipality of Murcia (Spain). Our results show significant spatial autocorrelation pattern, therefore, the probability of failure in a firm not only depends on its internal characteristics, but also on the situation of failure of its vicinity peer companies. These results can be an interesting starting point in the development of papers which consider the interdependence effects among closer peer companies in business failure literature.

Keywords: Spatial autocorrelation; business failure; join-count test.

JEL classification: C49; L80; R39.

MSC2010: 62P20; 91G70; 91B72; 91B82; 91B32.



1. Introducción

En los últimos años, la economía ha sufrido un periodo de recesión económica que derivó en una destrucción del tejido productivo empresarial, dando lugar a una situación de fracaso en un gran número de empresas. Las importantes repercusiones económicas negativas del fracaso empresarial han producido que éste haya sido un tópico ampliamente estudiado (Lang y Stulz, 1992). Los estudios desarrollados desde esta línea de investigación han tenido como objetivo, no solo plantear modelos econométricos que permitan predecir el fracaso empresarial sino también identificar aquellos factores que influyen en la probabilidad del fracaso de una empresa. Hasta la fecha, las evidencias empíricas apuntan al efecto de varios factores - la edad de la empresa (Cinca *et al.*, 2005), su tamaño (Turetsky y McEwen, 2001) o el sector económico (Quintana *et al.*, 2012) - como claros determinantes en la probabilidad de fracaso empresarial. En este contexto una reciente línea de trabajo está aportando nuevas evidencias sobre la importancia del efecto de contagio entre agentes económicos (Kolay *et al.*, 2015; Addoum *et al.*, 2014; Le, 2012; Hertznel y Officer, 2012; Hertznel *et al.*, 2008).

Las repercusiones derivadas del análisis de estos efectos de contagio en el fracaso empresarial han motivado su estudio considerando también sus distintas vías de propagación: la red comercial, la pertenencia a un mismo sector económico y la proximidad geográfica. Así, encontramos estudios de contagio en el fracaso empresarial centrados en el concepto de red comercial. Desde esta perspectiva, se considera que el contagio en el fracaso empresarial se extiende a través de la cadena de suministros de las distintas empresas. Dicho de otro modo, esta línea de investigación sostiene que una empresa en situación de fracaso ejerce efectos económicos (en algún sentido) sobre las empresas con las que mantiene relaciones comerciales. Desde esta perspectiva, Giesecke y Weber (2004) estudian el contagio de crédito entre empresas interrelacionadas comercialmente y concluyen que la relación comercial facilita la transmisión del riesgo de crédito. Por tanto, cuando una empresa no cumple con sus obligaciones comerciales, la liquidez de sus proveedores se ve afectada, lo que a su vez origina el incumplimiento en el pago a otros proveedores, resultando una cadena de transmisión del riesgo de crédito. Battiston *et al.* (2007) investigan la propagación de los efectos del fracaso empresarial entre empresas interrelacionadas verticalmente por sus sistemas de producción y los créditos comerciales. Estos autores argumentan que cuando una empresa entra en situación de fracaso, la probabilidad de fracaso de sus empresas vinculadas aumenta. Hertznel *et al.* (2008) también analizan los efectos del fracaso empresarial que se transmiten a través de la cadena de suministros. En este caso, los autores diferencian entre clientes y proveedores. De su estudio se concluye que el fracaso afecta principalmente a los

proveedores y al valor de la cartera de sus acciones, mientras que los clientes apenas perciben los efectos del contagio en el fracaso empresarial. Kolay *et al.* (2015) estudian si los problemas económicos y financieros de una empresa afectan a sus socios comerciales. Sus resultados indican que la situación económica negativa de una empresa genera un coste de sustitución a las empresas vinculadas a ella, de forma que el coste que le supone a las empresas proveedoras tiene una relación positiva sobre la probabilidad de fracaso de su cliente.

Otra vía de propagación de fracaso empresarial es la que considera los efectos derivados de la situación de bancarrota de una empresa sobre las empresas del mismo sector que forman parte de su competencia. Desde esta perspectiva, Lang y Stultz (1992) analizan el efecto del fracaso empresarial sobre la competencia. Estos autores concluyen que la situación de bancarrota en una empresa puede originar dos tipos de efectos sobre las empresas de la competencia: por un lado, estas pueden ver reducido el valor de su cartera de acciones ya que la bancarrota genera una información desfavorable de la situación del sector y, por otro lado, las empresas competidoras pueden beneficiarse de la redistribución de riqueza que deja la empresa en bancarrota. Iqbal (2002) investiga si el fracaso de una empresa explica los beneficios en sus competidores y encuentra que éstos experimentan un incremento en su rentabilidad financiera, el cual es mayor cuando las empresas forman parte de un sector concentrado. En lo referente al ámbito financiero, Hertzell y Officer (2012) determinan que si la bancarrota tiene lugar entre empresas de la competencia, los proveedores de capital reaccionarán al contagio imponiendo mayores costes crediticios y términos contractuales más estrictos.

En esta línea de contagio en el fracaso empresarial entre empresas del mismo sector, encontramos algunos estudios que señalan que la proximidad geográfica entre empresas podría ser un instrumento que potencie los canales de transmisión del contagio entre empresas del mismo sector. Así, en el trabajo de Addoum *et al.* (2014), los autores analizan la influencia que tiene la bancarrota empresarial en las políticas de inversión y financiación de empresas localizadas geográficamente próximas entre sí. La conclusión de su análisis es que la proximidad a una empresa en bancarrota puede causar una disminución de los niveles de inversión y endeudamiento en las empresas sanas. Staber (2001) investiga los efectos que provoca la proximidad en la supervivencia empresarial en el sector textil y encontró que la proximidad geográfica entre empresas de un mismo sector juega un papel fundamental como transmisor de los efectos de contagio ya que si las empresas están localizadas próximas entre sí, las tasas de fracaso empresarial tenderán a aumentar. Encontramos resultados similares en

el trabajo de De Silva y McComb (2012). Estos autores identifican mayores índices de contagio en el fracaso empresarial cuando la proximidad geográfica entre empresas en un mismo sector comprende un área inferior a una milla. A pesar de la repercusión de esta línea de análisis en la estrategia de la empresa ante situaciones de fracaso de sus vecinas, el número de estudios desarrollados es bastante limitado.

Con el objetivo de realizar una aportación a esta literatura, este trabajo analiza los patrones de colocación entre empresas sanas y fracasadas del mismo sector. Para ello, se desarrolla una aplicación empírica sobre una muestra de pequeñas y medianas empresas (pyme) del sector servicios localizadas en el municipio Murcia (España). A partir de esta información, se realiza un análisis basado en los estadísticos *join-count* (Iriondo *et al.*, 2008). Los resultados obtenidos en este estudio confirman la existencia de efectos de contagio en el fracaso empresarial significativos derivados de la proximidad geográfica entre pymes. Además, a diferencia de estudios anteriores, se examina el contagio en el fracaso empresarial diferenciando entre subsectores de actividad. De esta diferenciación concluimos que el efecto derivado del contagio entre empresas presenta distinto signo en función del subsector considerado. Así, en los subsectores donde hay un efecto de sustitución más alto (comercio) se identifica un efecto negativo. Es decir, empresas sanas tienden a estar rodeadas de empresas fracasadas. En este caso, las empresas sanas absorben la clientela de los anteriores que han fracasado por un efecto de sustitución respecto a su oferta en el mercado, disminuyendo así su probabilidad de fracaso. Por otro lado, las empresas de los subsectores con mayor interdependencia con respecto a la situación de sus respectivos mercados (hostelería, servicios profesionales e inmobiliarios) presentan indicios de contagio en el fracaso empresarial de signo positivo. En este caso, existe alta probabilidad de encontrar patrones de colocación espacial constituidos por una empresa fracasada rodeada de empresas fracasadas del mismo subsector.

Este trabajo se estructura como sigue. La sección 2 presenta el concepto formal de fracaso empresarial, la sección 3 la muestra y la metodología, en la sección 4 se presentan los resultados y discusión y la última sección concluye.

2. Definición de fracaso empresarial

No existe una definición precisa de qué es lo que se entiende por empresa fracasada y tampoco hay consenso unánime en los investigadores sobre cómo identificar a estas empresas. La literatura previa muestra distintas definiciones de fracaso. Así, la definición de fracaso empresarial desde el punto de vista legal considera como empresas fracasadas aquellas que se

encuentran inmersas en procedimientos de insolvencia (Smaranda, 2014). Otros estudios consideran como empresas fracasadas aquellas que hayan presentado un expediente concursal según “*ley concursal 22/2003*” (Campillo *et al.*, 2013). Otros trabajos suponen, desde la perspectiva legal, en situación fracaso empresarial a aquellas empresas que tengan una declaración formal de quiebra conforme a la legislación de cada país (García y Mures, 2012; Pervan *et al.*, 2011; Carter y Auken, 2006), que hayan sido declaradas en bancarrota por el índice “*The Wall Street Journal*” (Gepp *et al.*, 2010) o que dejen de existir (Pušnik y Tajnikar, 2008). Otros autores condicionaron el fracaso empresarial desde la perspectiva legal, a una situación de moratoria de pagos a petición legal (Ooghe y Balcaen, 2007) y otros por el hecho de tener la situación legal “*disuelta*” (Puagwatana y Guanawardana, 2005).

Por otro lado, desde un enfoque económico, y mediante la información de los estados contables de una empresa, se define como fracaso empresarial el estado en el que se encuentran aquellas empresas que obtienen resultados del ejercicio negativos consecutivos durante los dos últimos años (Xu *et al.*, 2014; Li y Xiong, 2012; Mateos y López, 2011; Li y Sun, 2011) o en los últimos tres años (Youn y Gu, 2010), e incluso una empresa con situación de quiebra técnica, lo que significa que el valor de las deudas supere al de sus inversiones y, que en términos contables, equivale a un patrimonio neto negativo (Misas, 2008).

En base a esta literatura, en este trabajo aplicamos la definición derivada del enfoque económico. En particular, identificamos como empresas en situación de fracaso empresarial aquellas que presenten resultados del ejercicio negativos durante los tres últimos periodos consecutivos, y a empresas sanas en el caso contrario. A pesar de sus limitaciones, hemos seleccionado esta definición económica debido a su objetividad a la hora de determinar la situación de fracaso empresarial de una empresa. De este modo, evitamos los problemas asociados a la definición legal de fracaso, como los que se pueden producir al no existir una diferencia clara entre empresa fracasada y empresa no declarada aún en quiebra, pero en proceso de serlo cuando se aplica la definición legal. Además, de esta definición legal también surgen conflictos por los distintos criterios adoptados en las normativas según el tiempo¹ o país para una misma situación de empresa (Fernández y Gutiérrez, 2012). Por otro lado, Campillo *et al.* (2013) argumentan que la definición legal genera una reducción considerable del tamaño de la muestra y Balcaen, y Ooghe (2006) consideran que una empresa puede estar declarada legalmente en quiebra, y no mostrar signos reales de fracaso,

¹ Fernández y Gutiérrez (2012) exponen que, “de acuerdo con la normativa actual en España, una empresa puede solicitar voluntariamente la declaración de concurso de acreedores para beneficiarse de las ventajas que la normativa le ofrece para salir de una situación transitoria de dificultades, o bien, para evitar las consecuencias negativas de no haberlo hecho, si posteriormente llega a quebrar y se la declara en quiebra culpable”.

lo que contaminaría la muestra para la estimación del modelo de predicción, por ejemplo, cuando una empresa se declara en quiebra para deshacerse de sus deudas o sufre un desastre natural.

3. Muestra y metodología

Muestra

Con el objetivo de desarrollar nuestro estudio seleccionamos una muestra de pymes² españolas del sector servicios, localizadas en el municipio de Murcia (Figura 1) para las que se dispone de información de los resultados de sus ejercicios para el periodo 2011-2013 así como de su exacta localización geográfica (latitud y longitud). Esta información fue obtenida de la base de datos del Sistema de Análisis de Balances Ibéricos (SABI). La muestra está constituida por 2.796 pymes.

Figura 1. Comunidades Autónomas de España



Fuente: Elaboración propia.

El estudio se centra en el sector servicios debido a la importancia que tiene este sector en el tejido empresarial del municipio de Murcia concentrando, en el año 2012, más del 80%

² Consideramos pymes aquellas empresas que cumplen al menos dos de los tres criterios propuestos por la Comisión Europea (2007) en función del número de trabajadores, volumen de activos y cifra de ventas.

de la actividad económica durante el periodo de estudio, según datos del directorio de actividades económicas del Centro Regional de Estadística de Murcia (CREM). Además, realizar el análisis para un solo municipio elimina los problemas derivados de la heterogeneidad regional que pueden aparecer en este tipo de modelos. De este modo, se aísla el propio efecto de contagio entre empresas del efecto derivado de que empresas de distintas regiones que estén afectadas por sus propias características territoriales. El municipio de Murcia también es una región propicia para el desarrollo de este estudio debido a varios factores: (1) el tejido productivo está constituido por un porcentaje muy elevado de empresa de reducido tamaño; (2) la crisis económica ha tenido una repercusión muy importante en la situación de las empresas de la Región de Murcia. De este modo, la Región de Murcia pertenece al grupo de regiones con valores económicos por debajo de la media (Maté *et al.*, 2009). Por tanto, es de esperar que los efectos producidos por el fracaso empresarial sean más intensos que en otras regiones. Además, el periodo de análisis considerado se corresponde con el periodo de crisis financiera por lo que habrá una mayor intensidad en el contagio empresarial. Finalmente, nos centramos en la pyme debido a la alta dependencia de este tipo de empresas respecto a las características de su entorno (Maté *et al.*, 2013).

A partir de esta información, se define la variable dicotómica fracaso empresarial F para cada empresa que toma el valor 1 si es una empresa fracasada y cero en caso contrario.

Metodología

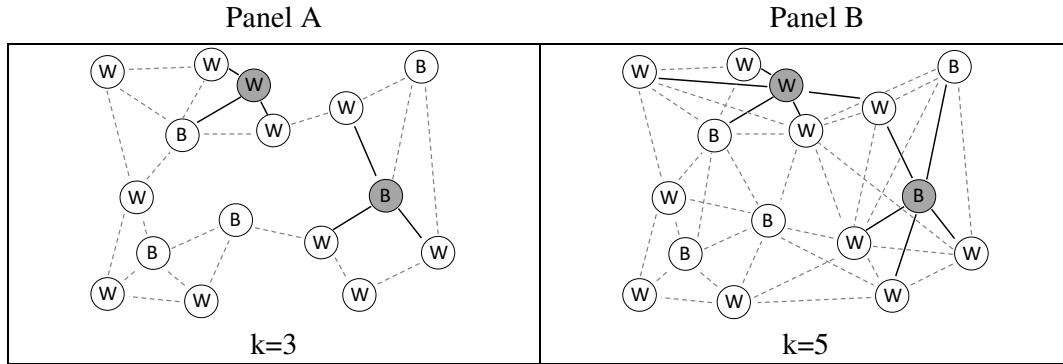
Se analiza la significatividad del contagio en el fracaso empresarial utilizando los estadísticos *join-count* que permiten contrastar la existencia de patrones de colocación espacial en la variable fracaso F (Iriondo *et al.*, 2008). En otras palabras, buscamos modelos de distribución empresarial con una estructura determinada respecto a su situación de fracaso. En particular, para nuestro análisis diferenciamos dos categorías: empresas fracasadas (B) y empresas sanas (W)³. A partir de estas categorías, resultan tres tipos posibles de conexiones entre cualesquiera dos unidades: BB, BW y WW. Los *joins* BB y WW representa el número de pares de unidades conectadas de la misma categoría, mientras que por BW se representa los pares de unidades muestrales conectadas de distinta categoría.

Para conectar las distintas unidades, se ha seleccionado el criterio de conectividad basado en la proximidad geográfica, de tal forma que cada empresa se encuentra conectada a las k empresas más próximas. Por ejemplo, si consideramos las tres empresas más próximas a

³ Tradicionalmente se utilizan las letras B (Black) y W (White) para denotar las dos categorías posibles aunque en este trabajo se utilizarán también las letras F (Fracaso) y S (Sana) dependiendo del contexto.

cada empresa analizada ($k = 3$) estaremos suponiendo un esquema de conexión como muestra la Figura 2 (Panel A), mientras que si consideramos un criterio de conectividad basado en los cinco vecinos más próximos ($k = 5$) tendremos un esquema de conexión como el mostrado en la Figura 2 (Panel B).

Figura 2. Ejemplo k-vecinos más cercanos



Otros criterios alternativos al de los k vecinos más próximos podrían usarse para establecer criterios de vecindad, sin que exista consenso sobre cuál es el más adecuado. Con el objeto de dar robustez a nuestros resultados, hemos considerado distintos valores de k (valores impares entre 3 y 19) en nuestro análisis. Por otra parte, al trabajar con la localización exacta de cada empresa (latitud y longitud), los criterios de vecindad basados en la distancia euclídea asignaban un número de vecinos extremadamente alto a empresas localizadas en zonas muy céntricas, mientras que en otras muchas situaciones no se asignaba ningún vecino. Esta elevada heterogeneidad en el número de vecinos podía distorsionar los resultados.

Los estadísticos *join-count* (J_{BB} , J_{BW} , J_{WW}) cuentan todos los pares de conexiones que unen empresas de la misma y diferente categoría de entre aquellos que se consideran vecinos. Estos estadísticos vienen dados por las siguientes expresiones:

$$J_{BB} = \frac{1}{2} \left[\sum_{\substack{i=1 \\ i \neq j}}^n \sum_{\substack{j=1 \\ j \neq i}}^n W_{ij} (BB)_{ij} \right] \quad (1)$$

$$J_{BW} = \frac{1}{2} \left[\sum_{\substack{i=1 \\ i \neq j}}^n \sum_{\substack{j=1 \\ j \neq i}}^n W_{ij} (BW)_{ij} \right] \quad (2)$$

donde, W_{ij} toma el valor 1 si las unidades i, j están conectadas por el criterio de los k vecinos más próximos y 0 en otro caso; $(BB)_{ij}=1$ si las unidades i y j son ambas de la categoría “B”, y

$(BB)_{ij}=0$ en caso contrario; $(BW)_{ij}=1$ si la unidad i y son de distinta categoría y $(BW)_{ij}=0$ en caso contrario. Por n se denota el número total de unidades.

El estadístico J_{WW} puede calcularse a partir de los dos estadísticos anteriores como

$$J_{WW} = \frac{1}{2} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n W_{ij} - (J_{BB} + J_{BW}) \quad (3)$$

Bajo la hipótesis nula de aleatoriedad en la distribución espacial de las observaciones muestrales se obtiene que los valores esperados de estos estadísticos y sus varianzas vienen dados por las expresiones:

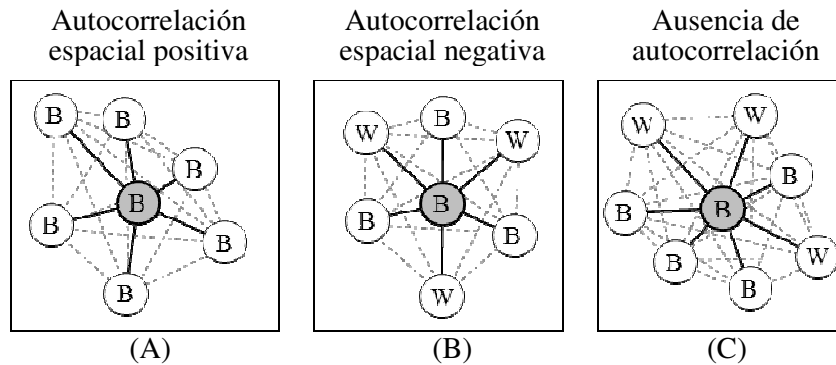
$$\begin{aligned} J_{BB} = E[J_{BB}] &= Tp_B^2 & E[J_{BB}] &= \sqrt{Tp_B^2 + 2mp_B^2 + (T - 2m)p_B^4} \\ J_{WW} = E[J_{WW}] &= Tp_W^2 & E[J_{WW}] &= \sqrt{Tp_W^2 + 2mp_W^2 + (T - 2m)p_W^4} \\ J_{BW} = E[J_{BW}] &= TP_B P_W & E[J_{BW}] &= \sqrt{2(Y + m)p_B p_W - 4(T + 2m)p_B^2 p_W^2} \end{aligned} \quad (4)$$

donde $p_B = n_B/n$ y $p_W = n_W/n$ ($n = n_B + n_W$) son las proporciones esperadas de número de observaciones de cada tipo, $T = \frac{1}{2} \sum_{i \neq j} W_{ij}$ es el total del número de conexiones y $m = \frac{1}{2} \sum_{i=1}^n k_i(k_i - 1)$. La significatividad de los estadísticos *join-count* se puede evaluar a través del estadístico dado por $Z = (\text{Observados-Esperados})/(\text{desviación estándar esperada})$.

Basándonos en la literatura previa (Iriondo *et al.*, 2008), los patrones de colocación contrastados a través de los estadísticos *join-count* pueden resultar positivos o negativos. Un patrón de colocación positivo vendría indicado por unos valores elevados de los estadísticos J_{BB} y/o J_{WW} mostrando la presencia de una estructura espacial donde existe una alta probabilidad de que unidades de categoría B y/o W estén rodeadas de unidades de la misma categoría (Figura 3.A), mientras que un valor significativamente elevado del estadístico J_{BW} sería síntoma de un patrón de asociación espacial negativo⁴ (Figura 3.B) indicando la alta probabilidad de encontrar conexiones de tipo WB (tomando como referencia una situación de aleatoriedad). En los casos que la distribución espacial presente una forma aleatoria, no habrá patrón de colocación (Figura 3.C).

⁴ Nótese que un valor del estadístico J_{WB} muy inferior a su valor esperado también podría indicar una estructura de autocorrelación positiva.

Figura 3. Tipos de autocorrelación espacial



4. Resultados

Análisis descriptivo

La Tabla 1 muestra los porcentajes de fracaso en función de las tres tipologías consideradas (edad, tamaño y sector) sobre el total de casos de fracaso empresarial. El principal resultado es una tasa global de fracaso del 17,2%. Respecto a la edad, los resultados obtenidos para esta variable muestran que la tasa de fracaso es mayor en empresas muy jóvenes y jóvenes con el 21,5% y 50,2%, respectivamente. La escasa información disponible de las empresas de menos de 5 años explica que los porcentajes de fracaso en la muestra puedan estar sesgados hacia las empresas de edad entre 5 y 15 años, aun así, estas cifras describen el comportamiento descrito en la literatura del fracaso (Honjo, 2000), que a menor edad la probabilidad de fracaso empresarial es mayor.

En cuanto al tamaño, al cruzar esta información con la situación de fracaso empresarial de las empresas de la muestra, nos encontramos con que el mayor porcentaje de fracaso, el 86,9%, se presenta en microempresas, mientras que para las empresas pequeñas este porcentaje supone un 11,5%, y en las medianas sólo el 1,7%. Por tanto, obtenemos una relación negativa con la tasa de fracaso. Es decir, a mayor tamaño nos encontramos con una disminución en la probabilidad de fracaso empresarial. Este resultado coincide con estudios previos (Gil y Giner, 2013; Vivel *et al.*, 2015).

Finalmente se analizan las tasas de fracaso considerando la variable sectorial. Para ello, el sector servicios se ha desagregado por subsectores según la Clasificación Nacional de Actividades Económicas 2009 (CNAE-2009). Así, podemos observar que más de la mitad de los casos de fracaso se encuentran concentrados en cuatro subsectores: empresas dedicadas a las actividades del comercio (32,1%), hostelería (10,6%), inmobiliarias (13,5%) y actividades profesionales científicas y técnicas (27,9%). El resto de subsectores presentaron una tasa de fracaso mucho más baja.

Tabla 1. Distribución de la muestra por edad, tamaño y subsectores. Valores medios para el periodo 2011-2013

<i>Factores</i>	<i>Total de Empresas</i>	<i>Tasa de Actividad empresarial</i>	<i>Casos de fracaso</i>	<i>Tasa de fracaso</i>
<i>Edad¹</i>				
Muy joven	565	20,2%	103	21,5%
Joven	1340	47,9%	241	50,2%
Madura	891	31,9%	136	28,3%
Total	2796	100%	480	100%
<i>Tamaño²</i>				
Micro	2336	83,5%	417	86,9%
Pequeña	387	13,9%	55	11,5%
Mediana	73	2,6%	8	1,6%
Total	2796	100%	480	100%
<i>Subsectores</i>				
(G) Comercio	1072	38,3%	154	32,1%
(I) Hostelería	175	6,3%	51	10,6%
(L) Inmobiliarias	317	11,3%	65	13,5%
(M) Científicas y técnicas	474	17,0%	76	15,8%
Resto subsectores ³	758	27,1%	134	27,9%
(H) Transporte	151	5,4%	24	5,0%
(J) Información	87	3,1%	15	3,1%
(K) Financieras	77	2,8%	10	2,1%
(N) Administración	151	5,4%	34	7,1%
(O) Defensa	5	0,2%	1	0,2%
(P) Educación	51	1,8%	5	1,0%
(Q) Sanitarias	100	3,6%	11	2,3%
(R) Artísticas	71	2,5%	17	3,5%
(S) Otros servicios	65	2,3%	17	3,5%
Total	2796	100%	480	100%

¹ Consideramos la edad de la empresa definiendo como empresa muy joven aquella que tiene menos de cinco años, joven la que tiene entre 5 y 15 años y madura aquella empresa con más de 15 años (Sánchez y García, 2003)

² El tamaño empresarial, hemos seguido como referencia los criterios de la recomendación del 6 de mayo de 2003 propuesta por la Comisión Europea para definir las empresas según sean micro, pequeñas medianas empresas.

³ El resto de subsectores agrupa actividades de las secciones H, J, K, N, O, P, Q, R y S.

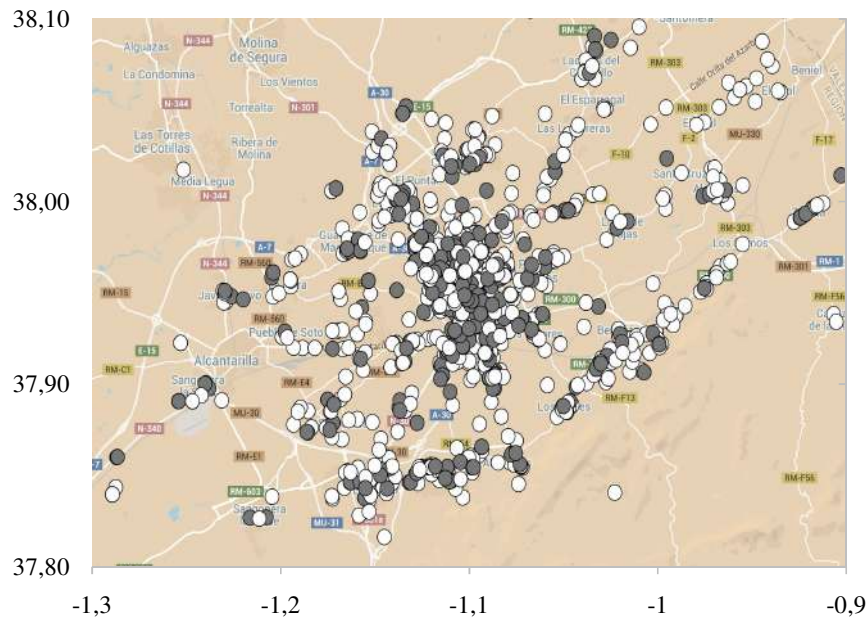
Resultados de los estadísticos join-count

En este apartado se determina la presencia de patrones de colocación espacial en la situación de fracaso empresarial. Para ello, en primer lugar, se analiza la distribución espacial de empresas de la muestra y, posteriormente, se contrastan estos resultados mediante los estadísticos *join-count*.

La Figura 4 muestra la localización de empresas de la muestra en el área del municipio de Murcia. Los puntos blancos identifican las empresas sanas y los puntos grises las

fracasadas. Se puede apreciar que las observaciones presentan una cierta tendencia de aglomeración hacia una zona que corresponde con el núcleo urbano de Murcia. Tras examinar la distribución de la muestra, con el propósito de contrastar el contagio por proximidad geográfica entre empresas fracasadas, a continuación se obtienen los tests *join-count* para todas las empresas del sector servicios sin distinguir entre subsectores de actividad. Posteriormente, se realizará el mismo análisis distinguiendo por subsectores.

Figura 4. Distribución geográfica de las empresas, en gris fracasadas y en blanco sanas



Fuente: Elaboración propia con Google My Maps.

Siguiendo la metodología *join-count* para contrastar patrones de colocación espacial de la variable fracaso empresarial, se consideran los criterios de conectividad basados en los *k*-vecinos más próximos ($k = 3, 5, 7, 9, 11, 13, 15, 17$ y 19). Partiendo de un nivel de significatividad del 10% y de la hipótesis nula (H_0 : independencia) que se rechazaría si *p*-valor fuese inferior al nivel de significación, y otra hipótesis alternativa (H_1 : existencia de patrón de colocación espacial), realizamos un análisis sobre las 2.796 empresas del sector servicios de la muestra. Los resultados obtenidos de los *join-count* (Tabla 2) indican la existencia de un patrón de colocación positivo y significativo entre empresas fracasadas para todos los valores de *k* considerados (Tabla 2. Panel A). En todos los casos, el número de pares de empresas fracasadas conectadas (J_{BB}) está por encima de lo esperado (J'_{BB}). El mayor nivel de colocación espacial lo encontramos al utilizar el criterio de las trece empresas

localizadas más próximas donde el p-valor es el menor (0,008) y el número de pares de empresas conectadas (584 conexiones) es superior al esperado (532,6 conexiones). Este resultado muestra que existe una alta probabilidad de que una empresa fracasada esté rodeada de empresas fracasadas. El test J_{BW} también permite analizar patrones de colocación entre las empresas de diferente categoría (Tabla 2. Panel B). En este caso, los resultados muestran valores del estadístico por debajo de sus valores esperados aunque las diferencias no son significativas mostrando un patrón de distribución aleatorio entre empresas Sana-Fracaso. Por último, el estadístico J_{WW} muestra una disminución significativa respecto a su valor esperado indicando que la probabilidad de encontrar una empresa sana próxima a otra sana, es menor de la esperada bajo la hipótesis de aleatoriedad⁵.

Tabla 2. Resultados del análisis *join-count* para el fracaso empresarial en 2013

TEST *JOIN-COUNT* DE AUTOCORRELACIÓN ESPACIAL
(aproximación normal)

k-empresas vecinas más próximas	Panel A			Panel B			Panel C		
	<i>join-count</i> BB test (Fracaso-Fracaso)			<i>join-count</i> BW test (Fracaso-Sana)			<i>join-count</i> WW test (Sana-Sana)		
	J_{BB}	J'_{BB}	p-valor ¹	J_{BW}	J'_{BW}	p-valor ¹	J_{WW}	J'_{WW}	p-valor ¹
3	141	122,9	0,042**	1156	1191,1	0,891	2834	2878,4	0,109
5	233	204,8	0,016**	1929	1985,3	0,937	4663	4797,3	0,062*
7	320	286,8	0,018**	2714	2779,4	0,935	6533	6716,3	0,065*
9	406	368,7	0,018**	3499	3573,5	0,936	8432	8635,3	0,064*
11	496	450,7	0,011**	4278	4367,6	0,950	10313	10554,2	0,050**
13	584	532,6	0,008***	5059	5161,7	0,957	12180	12473,1	0,043**
15	669	614,6	0,009***	5847	5955,8	0,954	14121	14392,1	0,046**
17	747	696,5	0,021**	6649	6749,9	0,928	16089	16311,1	0,072*
19	815	778,5	0,082*	7471	7543,9	0,840	18050	18230,0	0,159

¹ p-valor asociado al estadístico Z

*** significativo al 1%, ** significativo al 5% y * significativo al 10%.

Fuente: Elaboración propia mediante el programa SpaceStat.

La heterogeneidad de la muestra, en la que se consideran distintos subsectores de forma simultánea, puede inducir cierto sesgo en los resultados obtenidos en un análisis global principalmente debido al distinto impacto del fracaso empresarial en los distintos subsectores. Siguiendo este razonamiento, a continuación se desarrolla el análisis de los estadísticos *join-count* desagregando el sector servicios en aquellos subsectores para los cuales la tasa de fracaso es mayor que la media muestral del 7,6%. De este modo, nos aseguramos disponer de suficiente número de empresas fracasadas como para realizar nuestro análisis. Por tanto,

⁵ Notemos que se trata de un sistema de suma cero ya que el número total de uniones es fijo. En este caso un incremento significativo de las uniones tipo BB induce una disminución significativa de uniones WW.

centramos nuestro análisis en cuatro subsectores: Comercio (G), Hostelería (I), Actividades Inmobiliarias (L) y Actividades profesionales, científicas y técnicas (M).

Análisis del subsector Comercio

De acuerdo con la Tabla 3, los resultados del test *join-count* J_{BB} en el subsector comercio indican que entre empresas fracasadas no existe un patrón de distribución espacial (Tabla 3. Panel A) y, por tanto, estas empresas se localizan de forma aleatoria. Si nos centramos en las empresas de diferente categoría Sana-Fracaso (WB) conectadas (Tabla 3. Panel B), observamos que hay varios niveles de vecindad donde resulta un patrón de colocalización ($k = 3, 5, 11, 13$ y 19) asociado con un modelo de autocorrelación negativo ($J_{WB} > J'_{WB}$), de tal forma que la probabilidad de que una empresa sana este rodeada de empresas fracasadas es significativamente superior a la esperada bajo la hipótesis de aleatoriedad. Finalmente, cuando analizamos las empresas sanas conectadas (Tabla 3. Panel C) encontramos que para $k=13$ existen indicios de una estructura espacial, aunque el patrón de colocalización es apenas significativo (p -valor=0,094).

Tabla 3. Resultados del análisis *join-count* del subsector Comercio

TEST *JOIN-COUNT* DE AUTOCORRELACIÓN ESPACIAL
(aproximación normal)

k-empresas vecinas más próximas	Panel A			Panel B			Panel C		
	<i>join-count</i> BB test (Fracaso-Fracaso)			<i>join-count</i> WB test (Sana-Fracaso)			<i>join-count</i> WW test (Sana-Sana)		
	J_{BB}	J'_{BB}	p-valor ¹	J_{BW}	J'_{BW}	p-valor ¹	J_{WW}	J'_{WW}	p-valor ¹
3	29	32,6	0,223	424	393,7	0,008 ***	1164	1178,6	0,101
5	55	54,4	0,493	694	656,2	0,011 **	1946	1964,4	0,111
7	70	76,1	0,185	924	918,7	0,395	2748	2750,1	0,443
9	91	97,9	0,192	1204	1181,2	0,160	3525	3535,9	0,299
11	115	119,6	0,293	1498	1443,7	0,017 **	4295	4321,6	0,132
13	137	141,4	0,318	1777	1706,2	0,006 ***	5072	5107,4	0,094 *
15	155	163,1	0,219	2016	1968,7	0,061 *	5870	5893,2	0,210
17	175	184,9	0,203	2267	2231,2	0,138	6661	6678,9	0,286
19	200	206,7	0,286	2576	2493,6	0,009 ***	7424	7464,7	0,113

¹ p-valor asociado al estadístico Z

*** significativo al 1%, ** significativo al 5% y * significativo al 10%.

Fuente: Elaboración propia mediante el programa SpaceStat.

Análisis subsector Hostelería

Los resultados obtenidos en el subsector de Hostelería no presentan evidencias de patrones de colocalización significativos de empresas fracasadas con fracasadas (Tabla 4, Panel A), y

tampoco de sanas con sanas (Tabla 4, Panel C). El único resultado destacable es el estadístico J_{WB} que aparece significativo para algunos valores de k (3, 7, 13 y 15). Notemos que en todos estos casos los valores observados están por debajo de los esperados. Analizando conjuntamente los tres estadísticos *joint-count* no se pueden obtener claras conclusiones sobre los patrones de colocalización.

Tabla 4. Resultados del análisis *join-count* del subsector Hostelería

TEST *JOIN-COUNT* DE AUTOCORRELACIÓN ESPACIAL
(aproximación normal)

k-empresas vecinas más próximas	Panel A			Panel B			Panel C		
	<i>join-count</i> BB test (Fracaso-Fracaso)			<i>join-count</i> WB test (Sana-Fracaso)			<i>join-count</i> WW test (Sana-Sana)		
	J_{BB}	J'_{BB}	p-valor ¹	J_{WB}	J'_{WB}	p-valor ¹	J_{WW}	J'_{WW}	p-valor ¹
3	23	21,9	0,440	119	129,1	0,063 *	127	131,5	0,168
5	37	36,6	0,512	189	181,7	0,194	216	219,1	0,288
7	49	51,3	0,702	238	254,4	0,005 **	315	306,8	0,143
9	63	65,9	0,692	320	327,1	0,262	398	394,4	0,337
11	77	80,6	0,734	390	395,2	0,189	488	491,5	0,231
13	90	95,2	0,770	455	472,9	0,094 *	579	569,7	0,212
15	103	109,9	0,826	525	545,2	0,082 *	668	657,4	0,220
17	121	124,6	0,682	608	617,8	0,263	750	745,1	0,370
19	132	139,2	0,791	676	690,5	0,190	840	832,7	0,331

¹ p-valor asociado al estadístico Z

*** significativo al 1%, ** significativo al 5% y * significativo al 10%.

Fuente: Elaboración propia mediante el programa SpaceStat.

Análisis subsector de Actividades profesionales, científicas y técnicas

Por lo que respecta a los resultados del subsector de Actividades profesionales, científicas y técnicas, los datos obtenidos señalan que entre empresas conectadas en situación de fracaso (BB) (Tabla 5. Panel A) existe un patrón de colocalización espacial positivo, mostrando una relación directa entre la situación de fracaso de las empresas y la proximidad geográfica entre ellas. En otras palabras, las empresas de Actividades profesionales, científicas y técnicas que están en situación de fracaso presentan una probabilidad alta de estar rodeadas de empresas en la misma situación de fracaso (con respecto a una situación de distribución aleatoria). En cuanto a las empresas sanas conectadas (WW) (Tabla 5. Panel C), encontramos un comportamiento idéntico, donde las empresas sanas tienden a localizarse cerca a otras empresas sanas, no obstante, esto solo ocurre en ciertos niveles de vecindad (k = 3, 7, 9, 19). Finalmente, si examinamos la estructura Sana-Fracaso (Tabla 5. Panel B), nos encontramos con un esquema de colocalización espacial negativo, indicando que existe una mayor

probabilidad de encontrar el patrón Sana-Sana o Fracaso-Fracaso en la distribución de la variable como se apuntaba anteriormente.

Tabla 5. Resultados del análisis *join-count* del subsector Actividades profesionales, científicas y técnicas

TEST *JOIN-COUNT* DE AUTOCORRELACIÓN ESPACIAL
(aproximación normal)

k-empresas vecinas más próximas	Panel A			Panel B			Panel C		
	<i>join-count</i> BB test (Fracaso-Fracaso)			<i>join-count</i> WB test (Sana-Fracaso)			<i>join-count</i> WW test (Sana-Sana)		
	J _{BB}	J' _{BB}	p-valor ¹	J _{BW}	J' _{BW}	p-valor ¹	J _{WW}	J' _{WW}	p-valor ¹
3	27	18,1	0,004 ***	178	191,8	0,049 **	508	501,1	0,170
5	41	30,1	0,011 **	286	319,7	0,001 ***	852	835,1	0,034 **
7	52	42,2	0,041 **	447	447,6	0,000 ***	1202	1169,2	0,002 ***
9	66	54,2	0,026 **	521	575,4	0,001 ***	1531	1503,2	0,016 **
11	77	66,2	0,065 *	668	703,4	0,015 ***	1855	1837,3	0,112
13	94	78,3	0,019 **	796	831,3	0,027 **	2189	2171,3	0,141
15	106	90,3	0,024 **	915	959,2	0,013 **	2528	2505,4	0,111
17	118	102,4	0,037 **	1045	1087,1	0,025 **	2861	2839,4	0,144
19	125	114,4	0,121	1130	1215,1	0,001 ***	3216	3216,5	0,027 **

¹ p-valor asociado al estadístico Z

*** significativo al 1%, ** significativo al 5% y * significativo al 10%.

Fuente: Elaboración propia mediante el programa SpaceStat.

Análisis subsector de Actividades inmobiliarias

Observando los resultados de la Tabla 6, en el subsector de Actividades inmobiliarias detectamos un patrón de colocación espacial positivo en el conjunto de pares de empresas fracasadas (p-valores < 0,10 y valores de los *joins* BB observados mayores que los esperados) desde las 3 hasta las 13 empresas más próximas (Tabla 6. Panel A). Por tanto, la probabilidad de que una empresa inmobiliaria fracasada esté rodeada de otras fracasadas es más alta que la que se obtendría en el supuesto de distribución espacial aleatoria. A partir de las 15 empresas más próximas, el patrón de distribución espacial es aleatorio.

Por lo que respecta a la relación espacial entre empresas sanas (Tabla 6. Panel C), para k=11 los resultados indican que existe un esquema de colocación de signo positivo (p-valor < 0,10 y $J_{WW} > J'_{WW}$) aunque muy leve. En el resto de los casos la situación sana de una empresa no está interrelacionado espacialmente con la situación sana de otras empresas de su entorno. Finalmente, el patrón Sana-Fracaso (Tabla 6. Panel B) presenta un patrón de colocación de tipo negativo, identificando menos uniones tipo BW de las que cabría esperar.

Tabla 6. Resultados del análisis *join-count* del subsector de Actividades inmobiliarias

TEST JOIN-COUNT DE AUTOCORRELACIÓN ESPACIAL
(aproximación normal)

k-empresas vecinas más próximas	Panel A			Panel B			Panel C		
	join-count BB test (Fracaso-Fracaso)			join-count BW test (Sana-Fracaso)			join-count WW test (Sana-Sana)		
	J _{BB}	J' _{BB}	p-valor ¹	J _{BW}	J' _{BW}	p-valor ¹	J _{WW}	J' _{WW}	p-valor ¹
3	27	19,7	0,025 **	142	155,5	0,034 **	303	300,2	0,352
5	44	32,9	0,009 ***	238	259,1	0,013 **	499	500,4	0,400
7	57	46,1	0,024 **	342	352,8	0,031 **	705	700,6	0,302
9	73	59,2	0,010 **	439	466,5	0,014 **	911	900,7	0,140
11	84	72,4	0,045 **	548	570,2	0,055 *	1117	1100,9	0,067 *
13	96	85,6	0,081 **	654	673,8	0,096 *	1313	1301,1	0,163
15	107	98,7	0,141	761	777,5	0,162	1512	1501,2	0,233
17	119	111,9	0,211	868	881,2	0,234	1707	1701,4	0,374
19	120	125,1	0,206	976	984,8	0,324	1904	1901,6	0,445

¹ p-valor asociado al estadístico Z.

*** significativo al 1%, ** significativo al 5% y * significativo al 10%.

Fuente: Elaboración propia mediante el programa SpaceStat.

Discusión

Los resultados de este análisis confirman la existencia de un patrón de colocación espacial significativo en el fracaso empresarial entre las empresas del sector servicios. Ahora bien, este resultado global muestra indicios de cierta heterogeneidad que controlamos diferenciando por subsectores de actividad. La Tabla 7 muestra una síntesis de los resultados que hemos obtenido para los distintos subsectores de forma conjunta.

Tabla 7. Patrones de colocación espacial por principales subsectores de actividad del sector servicios

		Total	Comercio	Hostelería	Actividades profesionales, científicas y técnicas	Actividades inmobiliarias
Estructura de auto-correlación espacial	Estructura	F-F y S-S	S-F	S-F	F-F, S-S y S-F	F-F, S-S y S-F
	Signo	Positivo	Negativo	--	Positivo	Positivo
	# Vecinos	F-F (3, 5, 7, 9, 11, 13, 15, 19) S-S (3, 5, 7, 9, 11, 13, 15)	3, 5, 11, 13, 15, 19	3, 7, 13, 15	F-F (3, 7, 9, 19) S-S (5, 7, 9) y S-F (3, 7, 9, 11, 13, 15, 17, 19)	F-F (3, 5, 7, 9, 11, 13) S-S (11) y S-F (3, 5, 7, 9, 11)

El subsector de Comercio presenta una estructura de colocación espacial de tipo negativo con una mayor frecuencia de uniones tipo Sana-Fracaso (S-F) cuando consideramos distintos niveles de vecindad. Este resultado puede venir explicado a través del efecto sustitución entre empresas del mismo subsector, según el cual, las empresas supervivientes

sanas se verían favorecidas por la situación de fracaso de sus empresas vecinas. Al tratarse de un subsector con un alto efecto de sustitución en el mercado, las empresas que sobreviven absorben la clientela de las empresas fracasadas del subsector. En el caso de las empresas del subsector de Hostelería las evidencias son muy débiles y solo es posible identificar una estructura de colocación espacial positiva y significativa en el esquema S-F fruto de un menor número de este tipo de uniones que el esperado. En esta línea, en la literatura nos encontramos trabajos que sostienen que empresas pertenecientes a este subsector son más susceptibles al fracaso empresarial asociados a las características de sus mercados que las empresas de otros sectores (Thornhill y Amit, 2003). Además, Fritsch *et al.* (2006) investigan factores regionales que afectan a la supervivencia empresarial determinando que uno de los sectores más afectados por el entorno está relacionado con actividades relacionadas con “hoteles y restaurantes”. En el caso del subsector asociado a Actividades profesionales, científicas y técnicas, detectamos un patrón de colocación espacial positivo para distintos órdenes de vecindad en los esquemas S-S y F-F. Este resultado está directamente relacionado con las características de este tipo de empresas cuyas decisiones de localización en zonas de concentración de empresas del mismo subsector están motivadas por los beneficios obtenidos por la proximidad de este tipo de empresas generando estructuras de interdependencia entre sí que pueden dar mayor accesibilidad a los clientes (Keeble y Nachum, 2002). Por tanto, ante una situación de fracaso en una de ellas, las empresas próximas a ella se verán influenciadas de forma negativa aumentando también su probabilidad de fracaso. Finalmente, las empresas del subsector de Actividades Inmobiliarias presentan esta misma estructura de colocación espacial positiva, identificando un mayor número de uniones de tipo F-F y un menor número de S-F que el esperado. Al igual que el caso de las empresas de Actividades profesionales, científicas y técnicas, las decisiones de localización y las infraestructuras de interdependencia generadas entre ellas pueden dar lugar a este resultado.

Para concluir, debemos notar que no se observa una clara tendencia en los patrones de dependencia espacial al aumentar o disminuir el número (k) de vecinos, aunque en varios casos la significatividad se reduce al incrementar el valor de k indicando que las sinergias surgen a distancias cortas. Este efecto está presente en la mayoría de subsectores aunque las evidencias encontradas en el análisis de esta muestra no son sólidas.

5. Conclusiones

Este trabajo pone en evidencia la necesidad de considerar los efectos de interacción espacial derivados de la proximidad geográfica entre empresas del mismo sector como otro elemento a

considerar en la literatura del fracaso empresarial. Este resultado es de interés para los diferentes agentes relacionados con la empresa. Así, los empresarios, conociendo el efecto de esta variable, pueden anticiparse a la situación de fracaso empresarial de sus vecinos más próximos desarrollando políticas financieras con un carácter más conservador, por ejemplo disminuyendo la deuda o incrementando la liquidez de la empresa como medidas para prever las posibles tensiones financieras a las que tendrán que hacer frente en el futuro. En esta misma línea, los gerentes de las distintas empresas pueden formular estrategias de actuación partiendo del análisis de las oportunidades y amenazas del entorno y las fortalezas y debilidades de la empresa. Por otro lado, para las entidades financieras, tener en cuenta esta estructura de contagio empresarial es fundamental a la hora de gestionar la demanda de crédito. Finalmente, los investigadores que tienen como línea de investigación la predicción del fracaso empresarial pueden mejorar la capacidad predictiva de sus modelos incluyendo el componente espacial como una variable explicativa adicional.

Varias limitaciones tiene este estudio que merecen ser destacadas. En primer lugar, señalar que este análisis tiene un carácter exploratorio. Por tanto, la estimación de los patrones de colocación espacial entre empresas usando los estadísticos *join-count* como herramienta para detectar el contagio del fracaso puede no ser suficiente para confirmar la intensidad de sus efectos. Por tanto, en futuros análisis será necesario considerar información adicional de carácter financiero y no financiero de las empresas, así como del entorno para mejorar la calidad de nuestros resultados utilizando distintos métodos de estimación de modelos. En segundo lugar, el estudio se circunscribe a un territorio muy concreto y para un periodo temporal específico. Sería necesario extender el estudio tanto espacial como temporalmente para evaluar la robustez de los resultados. Esperamos que futuras investigaciones den luz sobre estas cuestiones.

Agradecimientos

La investigación ha sido soportada por el Proyecto 19884/GERM/15 del Programa de Ayudas a Grupos de Excelencia de la Región de Murcia, de la Fundación Séneca, Agencia de Ciencia y Tecnología de la Región de Murcia.

Referencias

Addoum, J., Kumar, A. y Le, N. (2014). “Contagious Negative Sentiment and Corporate Policies: Evidence from Local Bankruptcy Filings”. Finance Meeting EUROFIDAI-AFFI Paper.

- Balcaen, S. y Ooghe, H. (2006). "35 years of studies on business failure: an overview of the classic statistical methodologies and their related problems". *The British Accounting Review*, 38(1), 63-93.
- Battiston, S., Delli, D., Gallegati, M., Greenwald, B. y Stiglitz, J. (2007). "Credit chains and bankruptcy propagation in production networks". *Journal of Economic Dynamics and Control*, 31(6), 2061-2084.
- Campillo, J., Serer, G. y Ferrer, E. (2013). "Validez de la información financiera en los procesos de insolvencia. Un estudio de la pequeña empresa española". *Cuadernos de Economía y Dirección de la Empresa*, 16(1), 29-40.
- Carter, R. y Auken, H. (2006). "Small Firm Bankruptcy". *Journal of Small Business Management*, 44(4), 493-512.
- Cinca, C., Molinero, C., y Larraz, J. (2005). "Country and size effects in financial ratios: A European perspective". *Global Finance Journal*, 16(1), 26-47.
- Comisión Europea (2007). *The observatory of European SMEs*. Flash EB Series, 196.
- De Silva, D. y McComb, R. (2012). "Geographic concentration and high tech firm survival". *Regional Science and Urban Economics*, 42(4), 691-701.
- Fernández, M. y Gutiérrez, F. (2012). "Variables y modelos para la identificación y predicción del fracaso empresarial: revisión de la investigación empírica reciente". *Revista de Contabilidad*, 15(1), 7-58
- Fritsch, M., Brixy, U. y Falck, O. (2006). "The effect of industry, region, and time on new business survival—a multi-dimensional analysis". *Review of Industrial Organization*, 28(3), 285-306.
- García, A. y Mures, M. (2012). "Business failure prediction models: finding the connection between their results and the sampling method". *Economic Computation and Economic Cybernetics Studies and Research*, 46(3), 157-168.
- Gepp, A., Kumar, K., y Bhattacharya, S. (2010). "Business Failure Prediction using Decision Trees". *Journal of Forecasting*, 29(6), 536-555.
- Giesecke, K., y Weber, S. (2004). "Cyclical correlations, credit contagion, and portfolio losses". *Journal of Banking & Finance*, 28(12), 3009-3036.
- Gil, B. y Giner, B. (2013). "Predicción del fracaso empresarial en los sectores de construcción e inmobiliario: Modelos generales versus específicos". *Universia Business Review*, 39, 118-131.
- Hertzel, M., Li, Z., Officer, M. y Rodgers, K. (2008). "Inter-firm linkages and the wealth effects of financial distress along the supply chain". *Journal of Financial Economics*, 87(2), 374-387.
- Hertzel, M. y Officer, M. (2012). "Industry contagion in loan spreads". *Journal of Financial Economics*, 103(3), 493-506.
- Honjo, Y. (2000). "Business failure of new firms: an empirical analysis using a multiplicative hazards model". *International Journal of Industrial Organization*, 18 (4), 557-574.
- Iqbal, Z. (2002). "The effects of bankruptcy filings on the competitors' earnings". *International Review of Economics & Finance*, 11(1), 85-99.

- Iriondo, J., Torres, E. y Escudero, A. (2008). “Métodos para analizar variables continuas y categóricas: correlogramas y estadísticos ‘join count’”. En F.T. Maestre Gil, A. Escudero Alcántara y A. Bonet Jornet (eds.). *Introducción al análisis espacial de datos en ecología y ciencias ambientales: métodos y aplicaciones*. Dykinson, pp. 183-213.
- Keeble, D. y Nachum, L. (2002). “Why do business service firms cluster? Small consultancies, clustering and decentralization in London and southern England”. *Transactions of the Institute of British Geographers*, 27(1), 67-90.
- Kolay, M., Lemmon, M. y Tashjian, E. (2015). “Spreading the Misery? Sources of Bankruptcy Spillover in the Supply Chain”. *Journal of Financial and Quantitative Analysis*, 51(6), 1955-1990.
- Lang, L. y Stulz, R. (1992). “Contagion and competitive intra-industry effects of bankruptcy announcements”. *Journal of Financial Economics*, 32(1), 45-60.
- Le, N. (2012). “Spillover effects of intra-industry bankruptcy filings on firms cash holding policy”. Working paper.
- Li, H. y Sun, J. (2011). “Principal component case-based reasoning ensemble for business failure Prediction”. *Information and Management*, 48(6), 220-227.
- Li, H. y Xiong, T. (2012). “Predicting business risk using combined case-based reasoning in Euclidean space”. World Automation Congress (WAC), 1-6.
- Maté, M., García, D. y López, F. (2009). “Spatial effects in the productivity convergence of Spanish industrial SME’s”. *Spanish Journal of Finance and Accounting*, 38(141), 13-36.
- Maté, M., Hernández, G., Sánchez, J. y Mínguez A. (2013). “¿Hay efectos de interacción regional en el comportamiento financiero de las pyme?”. *Trimestre Económico*, 80(320), 841-867.
- Mateos, A. y López A. (2011). “Developing a business failure prediction model for cooperatives: Results of an empirical study in Spain”. *African Journal of Business Management*, 5(26), 10565-10576.
- Misas, M. (2008). “Análisis del fracaso empresarial en Andalucía. Especial referencia a la edad de la empresa”. *Cuadernos de Ciencias Económicas y Empresariales*, 1(54), 35-56.
- Ooghe, H. y Balcaen, S. (2007). “Are failure prediction models widely usable? An empirical study using a Belgian dataset”. *Multinational Finance Journal*, 11(1/2), 33-76.
- Pervan, I., Pervan, M., y Vukoja, B. (2011). “Prediction of company bankruptcy using statistical techniques – Case of Croatia”. *Croatian Operational Research Review*, 2(1), 158-167.
- Puagwatana, S. y Gunawardana, K. (2005). “Business failure prediction model: A case study of technology industry in Thailand”. In *Proceedings of IEEE International Engineering Management Conference*, 11-13.
- Pušnik, K. y Tajnikar, M. (2008). “Technical and cost efficiencies as determinants of business failures of small firms - The case of Slovenia”. *Eastern European Economics*, 46(1), 43-62.

- Quintana, M., Gallego, A. y Pascual, M. (2012). “Análisis del fracaso empresarial por sectores: factores diferenciadores”. *Pecunia: revista de la Facultad de Ciencias Económicas y Empresariales*, 1, 53-83.
- Sánchez, J. y García, D. (2003). “Influencia del tamaño y la antigüedad de la empresa sobre la rentabilidad: un estudio empírico”. *Revista de Contabilidad*, 6(12), 169-206.
- Smaranda, C. (2014). “Scoring functions and bankruptcy prediction models – case study for Romanian companies”. *Procedia Economics and Finance*, 10, 217-226.
- Staber, U. (2001). “Spatial Proximity and Firm Survival in a Declining Industrial District: The Case of Knitwear Firms in Baden-Württemberg”. *Regional Studies*, 35(4), 329-341.
- Thornhill, S. y Amit, R. (2003). “Learning about failure: Bankruptcy, firm age, and the resource-based view”. *Organization Science*, 14(5), 497-509.
- Turetsky, H. y McEwen, R. (2001). “An Empirical Investigation of Firm Longevity: A Model of the Ex Ante Predictors of Financial Distress”. *Review of Quantitative Finance and Accounting*, 16(4), 323-343.
- Vivel, M., Lado, R. y Otero, L. (2015). “¿Por qué quiebran los hoteles españoles?: un estudio de sus determinantes”. *Tourism & Management Studies*, 11(2), 25-30.
- Xu, W., Xiao, Z., Dang, X., Yang, D., y Yang, X. (2014). “Financial ratio selection for business failure prediction using soft set theory”. *Knowledge-Based Systems*, 63, 59-67.
- Youn, H. y Gu, Z. (2010). “Predicting Korean lodging firm failures: An artificial neural network model along with a logistic regression model”. *International Journal of Hospitality Management*, 29(1), 120-127.



Pronosticando la inflación mensual en Colombia un paso hacia delante: una aproximación “de abajo hacia arriba”

ALONSO, JULIO CÉSAR

Centro de Estudios en Economía y Finanzas (CIENFI)

Departamento de Economía

Universidad Icesi, Cali (Colombia)

Correo electrónico: jcalonso@icesi.edu.co

RIVERA, ANDRÉS FELIPE

Centro de Estudios en Economía y Finanzas (CIENFI)

Universidad Icesi, Cali (Colombia)

Correo electrónico: afrivera@icesi.edu.co

RESUMEN

La estructura jerárquica del Índice de Precios al Consumidor (IPC) de Colombia permite calcular la inflación como una combinación lineal de sus subcomponentes. Nuestra aproximación implica emplear modelos SARIMA para pronosticar cada componente del IPC y crear un pronóstico de la inflación como una combinación lineal de los pronósticos individuales; es decir, una aproximación “de abajo hacia arriba”. Se evalúa el desempeño fuera de muestra de los pronósticos para el siguiente mes de 12 métodos que emplean una aproximación “de abajo hacia arriba”. Estos métodos son comparados con un pronóstico agregado de la inflación empleando un modelo SARIMA para el IPC total. Nuestros resultados muestran que emplear un método “de abajo hacia arriba” para pronosticar la inflación del siguiente mes tiene un mejor comportamiento que emplear un modelo SARIMA agregado.

Palabras claves: IPC; inflación; pronósticos; de abajo hacia arriba; Colombia.

Clasificación JEL: C32; C53; E31; E37.

MSC2010: 62P20; 91B82; 91B84; 62M10.

Forecasting the Colombian Monthly Inflation One Step Ahead: A “Bottom to Top” Approach

ABSTRACT

The hierarchical structure of the Colombian Consumer Price Index (CPI) makes possible to calculate inflation as a linear combination of its subcomponents. We use SARIMA models to forecast each component of CPI and construct an forecast of inflation using a lineal combination of the forecasts of these components, i.e. a “bottom to top” approach. In this paper, we asses the out-of-sample performance of the one-step-ahead forecast of 12 “bottom to top” methodologies. These methods are compared with an aggregate forecast using a SARIMA model. Our results show that a “bottom to top” method to forecast inflation outperforms an aggregate approach for the case of monthly inflation in Colombia.

Keywords: CPI; inflation; forecasts; “bottom to top”; Colombia.

JEL classification: C32; C53; E31; E37.

MSC2010: 62P20; 91B82; 91B84; 62M10.



1. INTRODUCCIÓN

Pronosticar la inflación es una de las tareas rutinarias de los Bancos Centrales así como de buena parte de los actores del mercado financiero. Para los banqueros centrales estos pronósticos son indispensables para la implementación de una estrategia de política monetaria como es el caso de la “inflación objetivo” que maneja el Banco de la República en Colombia (Svensson, 1997). Por otro lado, los actores del mercado financiero tienen la posibilidad de obtener grandes rendimientos si están en la capacidad de generar un pronóstico acertado de la inflación para el siguiente mes (Nelson, 1976). No solamente los hacedores de política monetaria y los actores del mercado financiero están interesados en contar con un buen pronóstico de la inflación mensual, los agentes del sector real también toman decisiones económicas basadas en la información disponible sobre el comportamiento futuro de los precios (Dornbusch, 1976; Firth, 1979; Miller *et al.*, 1976; Solnik, 1983). En fin, existen numerosas razones que justifican contar con un buen modelo para predecir la inflación del próximo mes.

La inflación, y más específicamente el índice de precios del cual se deriva su cálculo, surgen de una combinación lineal de índices de precios que representan el comportamiento de canastas o subcanastas para diferentes ciudades y niveles de ingreso. Por ejemplo, en Colombia el Índice de Precios al Consumidor (IPC) nacional es el que se emplea para calcular la inflación de referencia para todos los actores. Este IPC se construye a partir de una combinación lineal de índices de precios asociados a categorías de productos o servicios, niveles de ingreso (bajo, medio y alto) y un conjunto de 24 ciudades.

En general, si una variable (a la cual denominaremos general o nivel agregado) resulta de una combinación lineal de otras variables (a las cuales denominaremos subcomponentes o nivel desagregado) la primera se podría pronosticar en dos pasos. Primero, se pronostica el comportamiento de los distintos subcomponentes. Y segundo, a partir del pronóstico de los subcomponentes, se construye una combinación lineal de estos últimos. Este pronóstico se conoce como el pronóstico de la variable general (Grunfeld y Griliches, 1960). A este procedimiento lo denominaremos como una aproximación “de abajo hacia arriba” (“*bottom to top*”, en inglés). La aproximación “de abajo hacia arriba” contrasta con una aproximación agregada (conocida como “enfoque agregado”) que implica pronosticar la variable general a partir de información agregada sin tener en cuenta cada uno de los subcomponentes de la variable general.

Existen dos argumentos a favor de una aproximación “de abajo hacia arriba”. Primero, pronosticar la variable general utilizando únicamente su información hace que se descarte otro tipo información relevante que podría permitir llegar a unos mejores resultados. Al considerar los distintos subcomponentes, la aproximación “de abajo hacia arriba” abarca un conjunto más amplio de información y un mayor rango de

heterogeneidad inherente a los distintos subcomponentes que podría generar un pronóstico de la variable general sistemáticamente más acertado. Segundo, en el caso hipotético de un choque aleatorio que afectaría a todos los subcomponentes de la variable general, los errores de pronóstico de estos podrían ser positivos o negativos. Sin embargo, al agregar los pronósticos, estos errores podrían cancelarse entre sí y de esta forma generar un pronóstico acertado de la variable general (Hendry y Clements, 2004).

Sin embargo, existen críticas a los anteriores argumentos. No necesariamente una mayor cantidad de información y heterogeneidad implican mejores pronósticos. Dado que el pronóstico de cada subcomponente requiere un modelo de estimación que describa su proceso generador de datos, es necesario conocer el verdadero proceso generador de datos de cada subcomponente. En muchas ocasiones esta condición no se cumple porque lo que el investigador tiene a la mano es una aproximación al verdadero proceso generador de datos. Al presentarse un error de especificación, es posible que la combinación lineal de los pronósticos de los distintos subcomponentes este expuesto a un error de pronóstico lo suficientemente grande que no permita lograr unos resultados satisfactorios. Respecto a los choques aleatorios, es posible que estos afecten a los subcomponentes de una misma forma, por lo cual estos errores de pronóstico no se cancelarían entre sí, generando así que el pronóstico de la variable general no sea tan acertado como podría esperarse (Theil, 1954).

El objetivo de este artículo es evaluar la conveniencia o no de emplear una aproximación “de abajo hacia arriba” para pronosticar la inflación en Colombia en el corto plazo. Considerando los argumentos que en general sugieren la aproximación “de abajo hacia arriba”, es posible esperar que el pronóstico obtenido a partir del nivel de desagregación más amplio (Grupo) muestre mejores resultados debido a que se está incorporando mayor cantidad de información de la dinámica de los componentes de la inflación. De esta manera, si distintos *shocks* económicos afectan los precios de forma asimétrica, la estimación desagregada puede incorporar en el pronóstico esta asimetría oculta en la información agregada y, por tanto, se estaría teniendo en cuenta la heterogeneidad inherente al pronóstico de los precios de los distintos bienes y servicios.

Para lograr nuestro objetivo aprovechamos la estructura con la que es construido el IPC en Colombia, el cual implica un nivel de desagregación de la canasta de productos y servicios por grupos de bienes, niveles de ingresos y por una distribución geográfica. En ese orden de ideas, el objetivo es comparar los pronósticos de la inflación en Colombia para el siguiente mes empleando una aproximación agregada con la combinación de los pronósticos asociados a los subcomponentes del IPC. Para lograr nuestro objetivo, estimamos modelos SARIMA para cada uno de los subcomponentes del IPC y se construye un pronóstico “de abajo hacia arriba” de la inflación. Este pronóstico es comparado con el obtenido de un modelo SARIMA para

el IPC total nacional. Posteriormente, comparamos el desempeño de todos los pronósticos empleando distintos criterios de evaluación, con la aproximación agregada.

Este trabajo se encuentra organizado de la siguiente manera. Después de esta introducción, en la segunda sección se presenta una revisión de la literatura sobre el tema, especialmente en la zona Euro, donde más se ha trabajado el tópico. En la tercera sección se describen los datos usados en el presente trabajo. La cuarta sección muestra la estructura metodológica a partir de la cual se construye el IPC (y la inflación) en Colombia. Además, se ilustra la estrategia empírica adoptada y los criterios tenidos en cuenta para evaluar los resultados. La quinta sección describe y analiza los resultados encontrados. Finalmente, la última sección presenta unos comentarios finales del ejercicio.

2. REVISIÓN DE LA LITERATURA

Para la zona Euro existe una amplia literatura que evalúa cuál es la metodología más acertada para pronosticar la inflación mensual. En general, en la literatura se encuentran tres tipos de ejercicios. El primer conjunto de ejercicios es el de modelar los subcomponentes del IPC nacional a través de un modelo VAR y construir una combinación lineal de los pronósticos de estos subcomponentes para lograr un pronóstico del IPC nacional. El segundo tipo de ejercicio es modelar los subcomponentes del IPC nacional y el IPC nacional en un modelo VAR para pronosticar el IPC nacional directamente. Se esperaría que la mayor información contenida en el VAR implique en mejores pronósticos del IPC nacional agregado. Finalmente, el tercer tipo de ejercicio es modelar cada subcomponente del IPC con una especificación ARIMA y construir el IPC nacional a partir de la combinación lineal de los pronósticos de estos modelos ARIMA. El desempeño de estas tres metodologías se compara con los pronósticos generados por un modelo ARIMA que utiliza únicamente información agregada del IPC nacional. Cabe resaltar el hecho que la desagregación que reportan en los trabajos respecto a la zona Euro está en función a la forma en que la Unión Monetaria reporta la información, es decir, la desagregación del IPC se da por sectores económicos y por regiones geográficas.

Espasa *et al.* (2002), empleando modelos VAR para datos de la zona Euro, encuentran que no hay una evidencia clara de cointegración entre las series de precios, según sector económico y región geográfica. La desagregación del IPC por sectores y su posterior pronóstico genera unos resultados más precisos y menos sesgados para periodos de pronóstico mayores a un mes, en comparación a los pronósticos provenientes del modelo ARIMA que usa información agregada. Pino *et al.* (2013) realizan un estudio similar para España. Empleando la existencia de una relación de largo plazo (cointegración) entre el índice de precio de un sector económico de una región y el índice de precio del mismo sector económico de una región vecina, los

autores encuentran que los pronósticos del IPC nacional a partir de la combinación lineal de los pronósticos de sus subcomponentes a través de un modelo VAR son más precisos en comparación a lo evidenciado con los pronósticos de un modelo ARIMA que usa información agregada del IPC, para periodos de pronóstico mayores a un mes. En este caso, la desagregación de los precios según el sector económico evidencia mejores resultados a los expuestos por los precios desagregados según región económica.

Fritzer *et al.* (2002), Benalal *et al.* (2004) y Harvey y Cushing (2014) consideran las tres metodologías mencionadas arriba. El primer estudio es realizado para Austria, el segundo para el caso de la zona Euro y el tercero para el caso de Ghana. Fritzer *et al.* (2002) comparan un modelo VAR (considerando todos los subcomponentes del IPC y el IPC nacional) y un modelo ARIMA (que solo considera el IPC nacional) para verificar la conveniencia de pronosticar directamente el IPC nacional a partir de la información agregada. Los autores encuentran que los pronósticos del modelo VAR son más acertados para periodos de 8 y 12 meses. A continuación, comparan el desempeño de los mismos modelos basados en un enfoque desagregado. Para este caso, el enfoque desagregado reporta un mejor desempeño a partir de los modelos ARIMA. Benalal *et al.* (2004) encuentran que los pronósticos producidos por modelos ARIMA que usan información agregada del IPC nacional son superiores a los pronósticos construidos a partir de una combinación lineal de los subcomponentes del IPC (ya sean a partir de los modelos VAR o ARIMA), esto para periodos de pronóstico de 12 y 18 meses. Para el caso de periodos de pronósticos inferiores a 12 meses, las conclusiones son mixtas. De otro lado, Harvey y Cushing (2014) concluyen algo similar. Estos autores encuentran que el pronóstico directo del IPC nacional a través de un modelo ARIMA tiene mejor desempeño a lo que resulta de los pronósticos basados en la combinación lineal de los pronósticos de los precios de los componentes del IPC.

Hubrich (2005), partiendo de las tres metodologías descritas al inicio de esta sección, encuentra que las diferencias en los pronósticos son más evidentes a medida que los periodos de pronóstico son mayores. Para periodos mayores a un mes, los pronósticos provenientes del modelo ARIMA con información agregada son superiores a los pronósticos de los demás modelos. De otro lado, el enfoque desagregado a partir del modelo VAR se desempeña mejor en comparación al enfoque desagregado a partir de los modelos ARIMA. Además de esto, también encuentra que al incluir información desagregada en un modelo VAR y pronosticar el IPC nacional resulta en mejores pronósticos en comparación al enfoque desagregado a partir del modelo ARIMA.

Con la información de los precios en Estados Unidos, Hendry y Hubrich (2011) basan sus análisis en modelos ARMA y VAR para poder concluir sobre el desempeño comparativo las tres metodologías de pronóstico. Al igual que Hubrich (2005), Hendry y Hubrich (2011) encuentran que un modelo ARIMA que utiliza información agregada,

a un modelo VAR con información de los subcomponentes del IPC pronostican mejor la inflación nacional en comparación al pronóstico de la inflación mensual que resulta de combinar los pronósticos de inflación de los subcomponentes del IPC. Es decir, incluir información desagregada resulta ser más conveniente para pronosticar la inflación mensual que combinar pronósticos desagregados. Además, incluir información desagregada en el pronóstico de la inflación mensual demuestra mejores resultados en comparación a pronosticar la inflación solo con su información (ARIMA con un enfoque agregado)

Demers y De Champlain (2005), adicional a la aproximación tradicional de los modelos AR y VAR, consideran los modelos de rezagos parsimoniosos más largos (PLL, por sus siglas en inglés). Encuentran que la aplicación de los modelos PLL, asociados con el enfoque desagregado, muestra mejores resultados. Sin embargo, estos autores resaltan la necesidad de considerar evidencia más robusta en contra de los métodos de pronóstico basados en modelos univariados que utilizan información agregada.

En general, existe evidencia mixta sobre la conveniencia de pronosticar la inflación mensual a partir de una combinación lineal de los componentes del IPC.

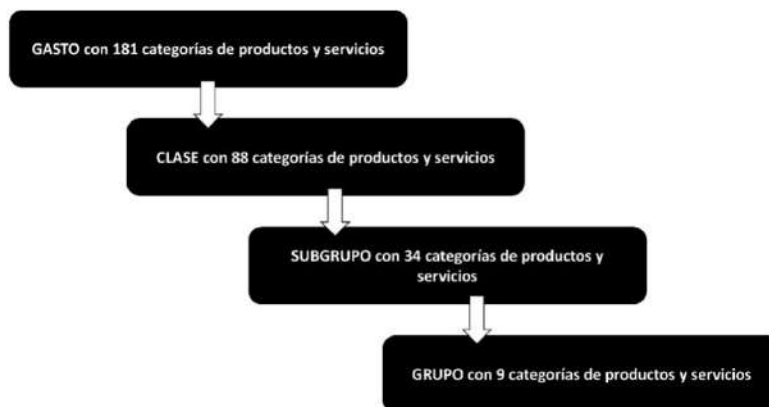
3. METODOLOGÍA

3.1. Estructura del IPC en Colombia

El Índice de Precios al Consumidor (IPC) permite seguir el comportamiento de los precios al por menor de una canasta de bienes y servicios, y su crecimiento porcentual corresponde a la medida más empleada para la inflación. En Colombia, el IPC se calcula para diferentes niveles de desagregación. Estos niveles de desagregación están dados por categorías de productos o servicios, niveles de ingreso (bajo, medio y alto) y ciudades. El primer nivel de desagregación corresponde al tipo de producto o servicio. Este criterio de desagregación implica cuatro niveles de categorización: Gasto, Clase, Subgrupo y Grupo. El nivel de categorización más amplio (más bajo) es el de Gasto, el cual cuenta con 181 productos y servicios. El siguiente nivel de categorización es la Clase con 88 subconjuntos de productos y servicios. Por ejemplo, “Frutas frescas” es un producto de la categoría de Clase la cual agrega productos como “Naranjas”, “Bananos” y “Tomate de árbol”, pertenecientes al nivel de categorización de Gasto. El tercer nivel de categorización es el de Subgrupo con 34 subconjuntos de productos y servicios. Por ejemplo, “Frutas frescas” es agrupada por “Frutas” la cual pertenece a la categorización de Subgrupo. Finalmente, el Grupo es el último nivel de categorización y abarca 9 conjuntos de productos y servicios. Siguiendo con el ejemplo anterior, el subgrupo “Frutas” está agrupada en “Alimentos”, siendo esta última uno de los nueve conjuntos de productos y servicios del nivel de categorización de Grupo. La Figura 1

muestra la estructura de desagregación descrita.

Figura 1. Esquema de la estructura jerárquica del IPC en Colombia

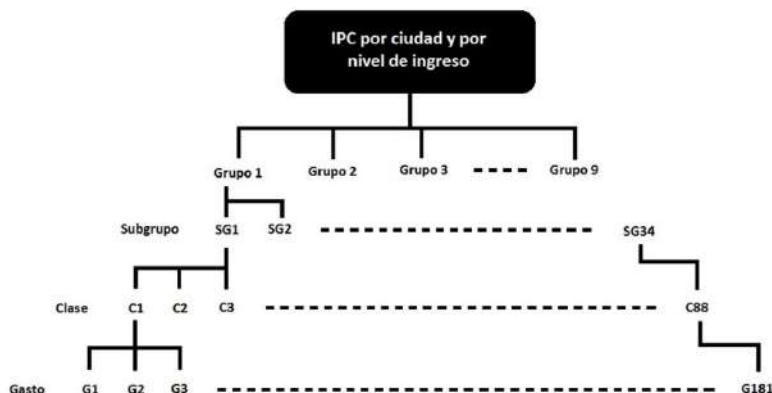


Fuente: Elaboración propia.

La estructura de clasificación es aplicada en los tres niveles de ingreso para 24 ciudades del país en las que se hace el seguimiento de los precios. Estas ciudades son: Armenia, Barranquilla, Bogotá, Bucaramanga, Cali, Cartagena, Cúcuta, Florencia, Ibagué, Manizales, Medellín, Montería, Neiva, Pasto, Pereira, Popayán, Quibdó, Riohacha, San Andrés, Santa Marta, Sincelejo, Tunja, Valledupar, y Villavicencio. De acuerdo a la estructura con la que se construye el IPC nacional, el IPC de cada nivel de ingreso por ciudad resulta de una combinación lineal de los índices de precios de cualquiera de los cuatro niveles de desagregación mencionados anteriormente. Es decir, empleando los 181 índices de precios para los productos y servicios del nivel de desagregación de Gasto para una misma ciudad y un mismo nivel de ingresos se puede obtener el índice de precios de la respectiva ciudad y nivel de ingresos. El mismo resultado se puede obtener si se emplea una combinación lineal de los 88 índices de precios correspondientes al nivel de desagregación de Clase. De manera similar se puede recuperar el índice de precios para una ciudad y nivel de ingresos a partir de los 34 subgrupos o de los 9 grupos.

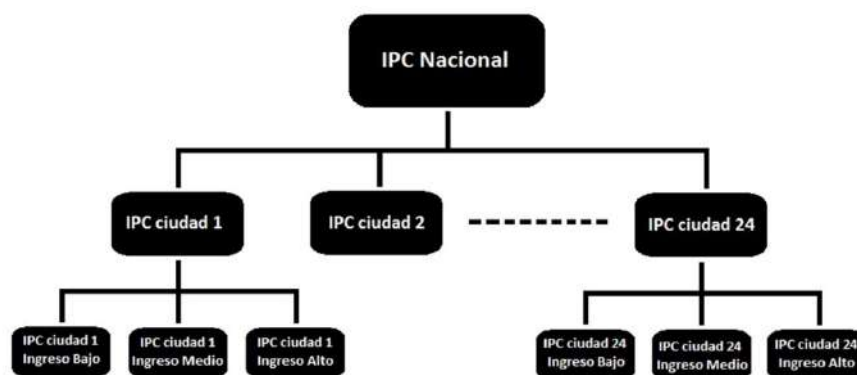
El IPC total para una ciudad resulta de la combinación lineal de los IPC de los tres niveles de ingreso de la respectiva ciudad. Finalmente, el IPC nacional es la combinación lineal del IPC de las 24 ciudades del país.

Figura 2. Esquema de la estructura del IPC para una ciudad y un nivel de ingresos en Colombia



Fuente: Elaboración propia.

Figura 3. Esquema de la construcción del IPC nacional a partir de los IPC de las 24 ciudades.



Fuente: Elaboración propia.

3.2. Representación matricial de la construcción del IPC de Colombia

Es importante adoptar una notación para la representación matricial de la construcción del IPC de Colombia. Emplearemos r para denotar los cuatro niveles de desagregación de los productos y servicios ($r \in \{Gasto, Clase, Subgrupo, Grupo\}$). j representará

el elemento posible en el correspondiente nivel de categorización r . Es decir, el número del elemento en el conjunto r . Para el nivel de Gasto existen 181 elementos, en nuestra notación esto implicará que si $r = \text{Gasto}$, entonces $j = 1, 2, \dots, 181$. Para el nivel de desagregación de Clase se tienen 88 posibles elementos de clasificación. En nuestra notación, si $r = \text{Clase}$, entonces $j = 1, 2, \dots, 88$. Para el nivel de Subgrupo existen 34 posibles elementos (si $r = \text{Subgrupo}$, entonces $j = 1, 2, \dots, 34$). Y para el nivel de Grupo se presentan 9 posibles grupos (si $r = \text{Grupo}$, entonces $j = 1, 2, \dots, 9$). i denotará el nivel del ingreso ($i \in \{\text{Bajo}, \text{Medio}, \text{Alto}\}$). Finalmente, c representará una de las 24¹ ciudades del país consideradas en la construcción del IPC.

En la construcción del IPC, cada bien o servicio o subconjunto de ellos tiene asociado un índice de precio y una ponderación, de acuerdo al nivel de desagregación. Sea $w_{i,c}^r$ un vector columna de ponderaciones de dimensión $(\text{Max}(j) \times 1)$ para la ciudad c , el nivel de ingresos i y el nivel de agregación r y sea $y_{i,c}^r$ un vector columna de dimensión $(\text{Max}(j) \times 1)$ que recoge los respectivos índices de precios para la ciudad c , el nivel de ingresos i y el nivel de agregación r . En otras palabras, $y_{i,c}^r$ es el vector de índices de precios de los j productos o conjuntos pertenecientes al nivel de desagregación r , en el nivel de ingreso i de la ciudad c . Formalmente²

$$w_{i,c}^r = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_{\text{Max}(j)} \end{bmatrix} \quad y_{i,c}^r = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\text{Max}(j)} \end{bmatrix} \quad (1)$$

El índice de precio de la ciudad c para el nivel de ingreso i y el nivel de agregación r se calcula como

$$IPC_{i,c} = (w_{i,c}^r)^T y_{i,c}^r \quad (2)$$

Con los índices de precio de los tres niveles de ingreso de la ciudad c se construye un vector columna \mathbf{I}_c el cual contiene estos índices y un vector columna \mathbf{p}_c que representa las ponderaciones de los índices anteriores para la ciudad c , tal que

¹ Es decir, $c \in \{\text{Armenia}, \text{Barranquilla}, \text{Bogotá}, \text{Bucaramanga}, \text{Cali}, \text{Cartagena}, \text{Cúcuta}, \text{Florencia}, \text{Ibagué}, \text{Manizales}, \text{Medellín}, \text{Montería}, \text{Neiva}, \text{Pasto}, \text{Pereira}, \text{Popayán}, \text{Quibdó}, \text{Riohacha}, \text{San Andrés}, \text{Santa Marta}, \text{Sincelejo}, \text{Tunja}, \text{Valledupar}, \text{Villavicencio}\}$.

² Para simplificar la notación se omitirá el subíndice correspondiente al tiempo.

$$\mathbf{I}_c = \begin{bmatrix} IPC_{bajo,c} \\ IPC_{medio,c} \\ IPC_{alto,c} \end{bmatrix} \quad \mathbf{p}_c = \begin{bmatrix} p_{bajo,c} \\ p_{medio,c} \\ p_{alto,c} \end{bmatrix} \quad (3)$$

El índice de precio en la ciudad c se construye como

$$IPC_c = (\mathbf{I}_c)^T \mathbf{p}_c \quad (4)$$

El IPC nacional se calcula como una combinación lineal de los índices de precios de 24 ciudades del país. Sea \mathbf{IPC} y \mathbf{s} vectores columna que recogen los índices de precios de cada ciudad (IPC_c) y las ponderaciones de las 24 ciudades (s_c), respectivamente; es decir

$$\mathbf{IPC} = \begin{bmatrix} IPC_1 \\ IPC_2 \\ \vdots \\ IPC_{24} \end{bmatrix} \quad \mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_{24} \end{bmatrix} \quad (5)$$

El IPC nacional es

$$IPC = (\mathbf{IPC})^T \cdot \mathbf{s} \quad (6)$$

Finalmente es importante mencionar el hecho que, por razones de reserva estadística, el Departamento Administrativo Nacional de Estadística (DANE) no reporta información desagregada del IPC para 11 ciudades del país. Esto impide construir de manera desagregada el IPC de esas 11 ciudades³.

De esta manera, nuestro ejercicio de pronóstico “de abajo hacia arriba” puede realizarse desde diferentes niveles de desagregación. Se puede construir el IPC empleando como base (nivel de máxima desagregación) cualquiera de los cuatro criterios: los 9 Grupos, los 34 Subgrupos, las 88 Clases o las 181 categorías de Gasto. En este orden de ideas, tendremos cuatro aproximaciones para construir nuestro pronóstico “de abajo hacia arriba”. La primera aproximación implicará pronosticar para cada uno de los tres niveles de ingresos cada uno de los 181 índices de precios de las categorías de Gasto en las 13 ciudades para las cuales se reporta la información desagregada⁴. A partir de estos pronósticos se construye el IPC de estas ciudades.

³ Estas ciudades son: Armenia, Florencia, Ibagué, Popayán, Quibdó, Riohacha, San Andrés, Santa Marta, Sincelejo, Tunja y Valledupar.

⁴ Estas ciudades son: Medellín, Barranquilla, Bogotá, Cartagena, Manizales, Montería, Neiva, Villavicencio, Pasto, Cúcuta, Pereira, Bucaramanga y Cali.

Posteriormente, construimos pronósticos agregados para las restantes 11 ciudades. Finalmente, se construye el pronóstico “de abajo hacia arriba” para el IPC nacional por medio de la combinación lineal de los 24 pronósticos de las ciudades. La segunda aproximación corresponde a generar pronósticos para los índices de precios de los 88 productos y servicios del nivel de desagregación de Clase para cada nivel de ingresos de las 13 ciudades para las cuales existe información desagregada. De manera similar a la primera aproximación, se construye el IPC nacional empleando los pronósticos del IPC de las 24 ciudades. La tercera aproximación se diferencia de las dos anteriores en que se generan pronósticos para los 34 elementos del nivel de desagregación de Subgrupo para cada nivel de ingresos de las 13 ciudades para las cuales existe información desagregada. La última aproximación genera pronósticos para los 9 Grupos para cada nivel de ingresos de las 13 ciudades para las cuales existe información desagregada. Nuestro pronóstico de inflación mensual será la variación porcentual de los IPC pronosticados. A continuación se describe como se estiman los pronósticos para cada una de estas aproximaciones.

3.3. Estrategia de estimación y construcción de los pronósticos

Como se mencionó anteriormente, para la construcción del pronóstico del IPC será necesario generar pronósticos univariados de cada una de las variables que conforman la “base” de las cuatro aproximaciones descritas. Para cumplir nuestro objetivo de evaluar el desempeño de corto plazo de la aproximación “de abajo hacia arriba”, estimamos un modelo SARIMA para cada producto y servicio perteneciente a los niveles de desagregación descritos arriba. Para evaluar el desempeño de los pronósticos se “guardan” los últimos 24 meses de la muestra. Es decir, se encuentra el modelo SARIMA sin incluir los últimos 24 meses de la muestra. Los pronósticos desagregados de los bienes y servicios obtenidos a través de la metodología del IPC son combinados de la forma descrita por las ecuaciones (1) a (6) para así estimar un pronóstico mensual del IPC nacional.

Para determinar el orden de los modelos SARIMA se emplearán tres criterios de información: el AIC (el criterio de información de Akaike), el AICC (el mismo AIC corregido por muestras finitas) y el BIC (criterio bayesiano). Una vez se encuentra el mejor modelo de acuerdo al respectivo criterio de información, se genera un pronóstico para el siguiente mes de manera recursiva. Es decir, se genera un pronóstico para el siguiente mes, se reestima los parámetros del modelo SARIMA incluyendo el último mes de información (pero manteniendo su estructura autoregresiva y de media móvil) y se vuelve a realizar un pronóstico para el siguiente mes. Esto se hace hasta contar con 12 pronósticos para el índice respectivo del siguiente mes (*one-step-ahead*). El pronóstico para el IPC nacional se construye como se mencionó en la sección anterior; es decir, agregando estos pronósticos individuales.

Lo anterior significa que tendremos para cada una de las cuatro aproximaciones “de abajo hacia arriba” tres conjuntos de pronósticos del IPC nacional: uno para el criterio AIC, otro para el AICC y otro para el BIC. En otras palabras, tendremos 12 métodos de pronósticos que combinan las cuatro niveles de desagregación para la aproximación “de abajo hacia arriba” y los tres criterios de información de los modelos univariados.

Por otro lado, se generan pronósticos agregados del IPC nacional empleando modelos SARIMA y los tres criterios de información mencionados anteriormente. Así, los 12 métodos de pronóstico que implican una aproximación “de abajo hacia arriba” serán comparados con 3 métodos de pronóstico agregados. Dado que nuestro objetivo es pronosticar la inflación, se comparan la inflación asociada al pronóstico del IPC nacional de los 15 métodos. Esta inflación se construye como la variación porcentual mensual entre el último IPC observado y el IPC pronosticado.

3.4. Criterios de evaluación de los métodos de pronósticos

Para evaluar el desempeño de los 12 métodos de pronóstico “de abajo hacia arriba” y de los 3 métodos que emplean la aproximación agregada, haremos una combinación de prueba estadística y métricas tradicionales que se aplicarán a todos los 12 pronósticos recursivos (*one-step-ahead*) para la inflación del siguiente mes. Primero emplearemos la prueba de Diebold y Mariano (1995). Esta prueba compara la precisión de los diferentes métodos de pronóstico. La hipótesis nula de esta prueba es que, al comparar dos métodos, ambos tengan la misma precisión estadística. Estas comparaciones se llevarán a cabo entre los 26 métodos de pronóstico.

Después de considerar la prueba de Diebold y Mariano (1995), utilizaremos las siguientes medidas de bondad de ajuste para evaluar el desempeño de los errores de pronóstico⁵: la raíz del error cuadrático promedio (RECP) y el error absoluto medio (EAM). El RECP computa el promedio de la desviación cuadrada entre la inflación pronosticada y la inflación observada:

$$RECP = \sqrt{\frac{1}{N} \sum_{i=1}^N (\%IPC_{pronosticado} - \%IPC_{observado})^2} \quad (7)$$

⁵ El error de pronóstico será la diferencia entre la inflación pronosticada y la inflación observada.

donde N es el número de pronósticos (12 en este caso). Por otro lado, el EAM mide la diferencia absoluta de la desviación entre la inflación pronosticada y la inflación observada:

$$EAM = \frac{1}{N} \sum_{i=1}^N |\%IPC_{\text{pronosticado}} - \%IPC_{\text{observado}}| \quad (8)$$

Además, también se reportarán otras medidas de dispersión como la media, mediana y los valores máximos y mínimos, todo para los errores de pronóstico. Respecto a la conveniencia de utilizar ciertas medidas de ajuste, es necesario mencionar lo planteado por Cuitiño *et al* (2010). En el presente trabajo no emplearemos medidas de ajuste porcentuales debido a que para estas, el valor en el denominador podría ser muy cercano a cero. De cierta forma la medida se magnifica llevando a resultados que pueden ser erróneos. Finalmente, es necesario tener presente que una medida cuadrática como la RECP podría resultar afectada por desviaciones extremas en los pronósticos de la inflación mensual. Sin embargo, al calcular el error absoluto medio se está tratando de atenuar ese sesgo.

3.5. Datos

Para realizar nuestro ejercicio emplearemos información mensual del IPC de Colombia publicada por el DANE en su página web⁶. La muestra bajo estudio abarca el periodo desde enero de 2009 a diciembre de 2015. Durante todo este periodo, el DANE mantiene fija las ponderaciones de los índices. No emplearemos información previa al año 2009 porque en ese año se dio un cambio metodológico que reconfiguró la estructura de clasificación del IPC. La información de los índices de precios y de la inflación es generada para los tres niveles de ingreso y 13 ciudades del país⁷. Para las demás 11 ciudades únicamente existe información desde el nivel de desagregación de Grupo, para los tres niveles de ingreso⁸.

Para determinar el correspondiente modelo SARIMA, se emplea una submuestra que va desde enero de 2009 a diciembre de 2013. Las observaciones para el año 2014 y 2015 son empleadas para evaluar los pronósticos recursivos (fuera de muestra) generados por los 15 métodos de pronóstico descritos anteriormente⁹.

⁶ La información fue descargada en el siguiente link: <http://cubindices.dane.gov.co:8084/Dane/inicioIPC.jsf>.

⁷ Estas ciudades son: Medellín, Barranquilla, Bogotá, Cartagena, Manizales, Montería, Neiva, Villavicencio, Pasto, Cúcuta, Pereira, Bucaramanga y Cali.

⁸ Estas ciudades son: Armenia, Florencia, Ibagué, Popayán, Quibdó, Riohacha, San Andrés, Santa Marta, Sincelejo, Tunja y Valledupar.

⁹ Los cálculos fueron realizados empleando el lenguaje estadístico R (R Core Team, 2016) partiendo del paquete “forecast” (Hyndman y Khandakar, 2008).

4. RESULTADOS DE LAS ESTIMACIONES

Para facilitar el análisis de los resultados, emplearemos la siguiente notación para denominar los resultados de los 12 métodos de pronóstico “de abajo hacia arriba”:

$$r.\text{criterio} \quad (10)$$

donde $r \in \{Gasto, Clase, Subgrupo, Grupo\}$ y $criterio \in \{AIC, AICC, BIC\}$. Por ejemplo, el método de pronóstico que parte del nivel de desagregación de Gasto y que emplea el criterio BIC para encontrar el mejor modelo SARIMA de los índices de precio se identifica como *Gasto.BIC*.

En el caso de la aproximación agregada, encontramos que los tres criterios de información seleccionan el mismo modelo SARIMA, por lo que se mostrarán los resultados de un solo criterio de información. Este se denotará como *Agreg.* Así, solo comparamos los 12 métodos de pronóstico “de abajo hacia arriba” y 1 método que emplea la aproximación agregada.

Por encima de la diagonal principal de la Tabla 1 se muestran los estadísticos de la prueba de Diebold y Mariano (1995) (y sus niveles de significancia se expresa con asteriscos) que resultan de la comparación de los métodos de pronóstico asociados a los 12 métodos mencionados arriba. Rechazar la hipótesis nula de la prueba implica que los dos métodos que se comparan tienen una precisión estadística distinta. De otro lado, la diagonal inferior de la Tabla 1 muestra la significancia estadística de la prueba de Diebold y Mariano (1995) cuando partimos de la hipótesis nula de que los pronósticos del modelo 2 (especificados en las columnas de la tabla) son más precisos que los pronósticos del modelo 1 (especificados en las filas de la tabla). Si los estadísticos que se encuentran por encima de la diagonal principal indican que los pronósticos de los dos modelos que se están comparando son estadísticamente distintos, la significancia estadística que se muestra debajo de la diagonal principal definirá cuál de los modelos es más preciso en su pronóstico. De la Tabla 1 se puede concluir lo siguiente:

- i) Existe una diferencia entre los niveles de los pronósticos que generan los diferentes niveles de agregación. Por ejemplo, los métodos que agregan al nivel de Gasto (sin importar el método de selección del orden del modelo SARIMA) generan pronósticos que son estadísticamente diferentes de los que generan los niveles de agregación de Clase, Subgrupo y Grupo (pero solo cuando se selecciona el número de rezagos con el criterio AIC).
- ii) A excepción del caso del nivel de agregación de Grupo, no existe diferencia estadística entre los pronósticos que genera cada nivel de agregación cuando se selecciona el orden del modelo SARIMA con diferentes métodos.

- iii) Los modelos “de abajo hacia arriba” que generan pronósticos estadísticamente diferentes al método agregado son aquellos que emplean el nivel de agregación de Subgrupo y Grupo (pero solo cuando se selecciona el número de rezagos con el criterio AIC). Sin embargo, vale la pena resaltar que los pronósticos del método desagregado a nivel de Clase (para los tres criterios) son estadísticamente diferentes a los pronósticos de los demás métodos¹⁰, excepto cuando se comparan con los pronósticos al nivel de Grupo. Igualmente, los pronósticos al nivel de Subgrupo son estadísticamente diferentes a los pronósticos de los demás métodos¹¹, excepto, de nuevo, cuando se comparan con los pronósticos a nivel de Grupo.

Por otro lado, es interesante anotar que los resultados obtenidos para los diferentes criterios de información son robustos para los diferentes niveles de agregación. En otras palabras, los diferentes criterios de información empleados no generan pronósticos estadísticamente diferentes. La diferencia estadística de los pronósticos proviene del nivel de desagregación de los productos y servicios. Por ejemplo, al comparar los pronósticos de los métodos *Grupo.AIC* y *Clase.AIC* (diagonal superior) la prueba de Diebold y Mariano (1995) concluye que los pronósticos de ambos métodos son estadísticamente distintos. Sin embargo, cuando se comparan los pronósticos de los métodos *Grupo.AIC* y *Grupo.AICC*, la prueba concluye que no existe una diferencia estadísticamente significativa. Lo mismo es cierto si comparamos los pronósticos de los métodos *Clase.AIC* y *Clase.AICC*.

En la misma Tabla 1, por debajo de la diagonal principal, se reportan los valores p de la prueba de una cola cuya hipótesis nula es que los pronósticos del método reportado en la fila genera pronósticos menos precisos o iguales en comparación a los generados por el método reportado en la columna. La alterna de esta prueba implica que el método de la fila es más preciso que el de la columna. Se somborean aquellos valores p que son inferiores a 0.1. A partir de estos resultados podemos concluir que: i) los métodos que parten del nivel de desagregación de Clase (independientemente del criterio de selección) muestran los pronósticos más precisos que el método agregado, y los métodos “de abajo hacia arriba” que emplean los niveles de agregación de Gasto, Subgrupo y Grupo para los criterios de selección AICC y BIC; ii) los métodos de pronóstico que parten del nivel de desagregación de Gasto son más precisos que los métodos de pronósticos al nivel de desagregación de Subgrupo (para los tres criterios de selección) y al nivel de Grupo para los criterios de selección AICC y BIC; iii) los

¹⁰ Aceptando la hipótesis alterna de que los pronósticos asociados al nivel de Clase son estadísticamente más acertados.

¹¹ Aceptando la hipótesis alterna de que los pronósticos asociados al nivel de Subgrupo son estadísticamente menos acertados.

métodos de pronóstico al nivel desagregación de Subgrupo solo son más precisos que los métodos de pronóstico a nivel de Grupo para los criterios de selección AICC y BIC; iv) los pronósticos del enfoque agregado solo son más precisos cuando se comparan a los pronósticos asociados al método que parte del nivel de desagregación de Subgrupo; y v) el método de pronóstico de *Grupo.AIC* muestra los pronósticos más precisos que todos los otros métodos “de abajo hacia arriba” considerados y el agregado. Una situación importante que puede destacarse de la Tabla 1 es el hecho que tener más información no es algo necesariamente conveniente. A pesar que es aparente que los métodos desagregados funcionan mejor al momento de pronosticar la inflación mensual un paso delante, relativo a lo que puede hacer el método agregado, no es cierto que entre más desagregada sea la lista de productos y servicios considerados, más acertado sea el pronósticos obtenido.

Según la Tabla 1, las pruebas de Diebold y Mariano (1995) brindan evidencia de que los pronósticos del método *Grupo.AIC* presentan el mejor desempeño entre los 13 métodos de pronóstico que consideramos en el presente análisis. La Tabla 2 muestra las medidas de bondad de ajuste de los pronósticos de los 13 métodos evaluados, además de la media, la mediana y los valores máximos y mínimos de los errores de pronóstico. Cuando comparamos el desempeño de estos 13 métodos de pronóstico, las tres medidas de bondad de ajuste favorecen a los métodos basados en la aproximación “de abajo hacia arriba”. Para el RECP, el método *Grupo.AIC* reporta el menor valor. Por otro lado, el método *Clase.AICC* reporta los menores valores para EAM.

Tabla 1. Pruebas Diebold-Mariano entre métodos de pronóstico del IPC mensual para los métodos que emplean las series de índices mensuales.

	<i>Gasto.AIC</i>	<i>Gasto.AICC</i>	<i>Gasto.BIC</i>	<i>Clase.AIC</i>	<i>Clase.AICC</i>	<i>Clase.BIC</i>	<i>Subgrupo.AIC</i>	<i>Subgrupo.AICC</i>	<i>Subgrupo.BIC</i>	<i>Grupo.AIC</i>	<i>Grupo.AICC</i>	<i>Grupo.BIC</i>	<i>Agreg</i>
<i>Gasto.AIC</i>		-0.894	-0.173	3.0029***	3.1667***	1.8057*	-5.0054***	-5.3978***	-4.4758***	2.2565**	-0.8803	-0.8973	0.7296
<i>Gasto.AICC</i>	0.8097		-0.0139	3.136***	3.2989***	1.9472*	-4.9435***	-5.3157***	-4.3457***	2.2949**	-0.8545	-0.8725	0.7694
<i>Gasto.BIC</i>	0.5679	0.5055		2.1383**	2.313**	2.2946**	-4.6103***	-5.0492***	-5.4141***	2.336**	-0.8304	-0.8512	0.7087
<i>Clase.AIC</i>	0.0032	0.0023	0.0217		1.0771	-0.4124	-7.6732***	-8.2562***	-6.6447***	1.7768*	-1.3144	-1.3303	0.1403
<i>Clase.AICC</i>	0.0022	0.0016	0.0150	0.1463		-0.6449	-7.722***	-8.3599***	-6.9169***	1.7439*	-1.3421	-1.3578	0.1035
<i>Clase.BIC</i>	0.0420	0.0319	0.0156	0.6581	0.7373		-6.6444***	-7.1965***	-7.9685***	1.8931*	-1.2532	-1.2799	0.21
<i>Subgrupo.AIC</i>	1.0000	1.0000	0.9999	1.0000	1.0000	1.0000		-1.4242	0.0064	3.4478***	0.0487	0.0245	2.2557**
<i>Subgrupo.AICC</i>	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9161		0.294	3.4978***	0.0826	0.0582	2.3054**
<i>Subgrupo.BIC</i>	0.9999	0.9999	1.0000	1.0000	1.0000	1.0000	0.4975	0.3857		3.4391***	0.0478	0.0237	2.1436**
<i>Grupo.AIC</i>	0.0169	0.0156	0.0143	0.0444	0.0473	0.0355	0.0011	0.0010	0.0011		-1.8033*	-1.8321*	-1.9467*
<i>Grupo.AICC</i>	0.8061	0.7992	0.7926	0.8992	0.9037	0.8886	0.4808	0.4674	0.4812	0.9578		-0.1947	1.0675
<i>Grupo.BIC</i>	0.8106	0.8040	0.7983	0.9018	0.9061	0.8933	0.4903	0.4771	0.4907	0.9600	0.5763		1.0824
<i>Agreg</i>	0.2365	0.2247	0.2428	0.4448	0.4592	0.4178	0.0170	0.0153	0.0214	0.9681	0.1484	0.1452	

Fuente: elaboración propia.

Tabla 2. Métricas de bondad de ajuste de los 13 métodos de pronóstico.

	Media	Mediana	Máximo	Mínimo	Rango	RECP	EAM
<i>Gasto.AIC</i>	-0.2127%	-0.2022%	0.0590%	-0.8225%	0.8814%	0.2879%	0.2205
<i>Gasto.AICC</i>	-0.2128%	-0.2107%	0.0586%	-0.8184%	0.8770%	0.2877%	0.2215
<i>Gasto.BIC</i>	-0.2159%	-0.2340%	0.0358%	-0.7903%	0.8261%	0.2855%	0.2216
<i>Clase.AIC</i>	-0.1974%	-0.1855%	0.0774%	-0.8752%	0.9526%	0.2831%	0.2046
<i>Clase.AICC</i>	-0.1967%	-0.1850%	0.0664%	-0.8782%	0.9447%	0.2824%	0.2037
<i>Clase.BIC</i>	-0.1985%	-0.2099%	0.0685%	-0.8734%	0.9419%	0.2820%	0.2067
<i>Subgrupo.AIC</i>	-0.2606%	-0.2737%	0.0405%	-0.8784%	0.9189%	0.3271%	0.6642
<i>Subgrupo.AICC</i>	-0.2618%	-0.2689%	0.0451%	-0.8772%	0.9223%	0.3279%	0.2655
<i>Subgrupo.BIC</i>	-0.2612%	-0.2781%	0.0323%	-0.8752%	0.9074%	0.3258%	0.2639
<i>Grupo.AIC</i>	-0.0804%	-0.0560%	0.1856%	-0.4241%	0.6097%	0.1821%	0.2622
<i>Grupo.AICC</i>	-0.2273%	-0.2144%	0.1677%	-1.9068%	2.0745%	0.4531%	0.2617
<i>Grupo.BIC</i>	-0.2284%	-0.2331%	0.1883%	-1.9136%	2.1018%	0.4531%	0.2628
<i>Agreg</i>	0.0986%	0.1111%	0.6914%	-0.2623%	0.9537%	0.2306%	0.5429

Por otro lado, si se considera el rango de los errores de pronóstico, se encuentra que el que tiene menor rango es el método *Grupo.AIC*. Este método también presenta el error de pronóstico medio y mediano más pequeño (en valor absoluto). El rango del error de pronóstico del método *Grupo.AIC* es aproximadamente un 55% menor que el del método *Clase.AICC*. La media del método *Grupo.AIC* es 145% menor que el del método *Clase.AICC* y la mediana es 230% menor. Es decir empleando criterios que solo tienen en cuenta la magnitud absoluta del error de pronóstico el mejor método de estimación corresponde a *Grupo.AIC*.

5. COMENTARIOS FINALES

Este trabajo permite evidenciar que para el caso de la inflación mensual en Colombia, y el período de estudio seleccionado, emplear un método “de abajo hacia arriba” para pronosticar la inflación del siguiente mes tiene un mejor comportamiento que emplear un modelo SARIMA agregado. En especial encontramos que partiendo de los índices de precios desagregados al nivel de Grupo¹² por ciudad y por nivel de ingresos se puede construir un mejor pronóstico para la inflación mensual durante el 2014 y 2015, en lugar de emplear un modelo univariado para el IPC nacional total.

El método “de abajo hacia arriba” que mejor se comportó implica estimar modelos SARIMA para 9 conjuntos de productos y servicios para los niveles de ingresos bajo, medio y alto de las 13 ciudades para las cuales se reporta la información desagregada. Estas ciudades son: Medellín, Barranquilla, Bogotá, Cartagena, Manizales, Montería, Neiva, Villavicencio, Pasto, Cúcuta, Pereira, Bucaramanga y Cali. Esto implica 351 modelos SARIMA. Para las restantes 11 ciudades que cubren el IPC en Colombia se construye un pronóstico agregado por ciudad. Empleando estos 362

¹² Y empleando el AIC para seleccionar el orden de los modelos SARIMA.

pronósticos se construye una combinación lineal de estos que produce el pronóstico de la inflación mensual para el total nacional. Según nuestros resultados, este pronóstico es estadísticamente mejor que emplear un método agregado para pronosticar la inflación mensual durante 2014 y 2015.

De esta manera, nuestros resultados muestran que, para pronosticar la inflación en el corto plazo, es útil emplear información desagregada que incorpora de forma diferente los choques macroeconómicos y microeconómicos que se presentan en la economía colombiana. Estos choques pueden ser enmascarados en el comportamiento agregado. Así un método como el empleado puede incorporar mejor toda la información disponible.

No obstante estos resultados sugerirían que los métodos “de abajo hacia arriba” son mejores que una aproximación agregada, se deberá evaluar en futuras investigaciones el comportamiento de pronóstico de los métodos “de abajo hacia arriba” para horizontes de tiempo más amplios. Por otro lado, nuestro ejercicio solo se concentró en el comportamiento de los pronósticos puntuales y en ningún momento considera el comportamiento de los intervalos de predicción que pueden generar las diferentes aproximaciones. Futuras investigaciones tendrán que evaluar el comportamiento de los intervalos de predicción de las aproximaciones “de abajo hacia arriba”.

REFERENCIAS

- Benalal, N.; del Hoyo, J. L.; Roma, M.; Landau, B. & Skudelny, F. (2004). "To aggregate or not to aggregate? Euro Area Inflation Forecasting". ECB Working Paper No. 374. European Central Bank.
- Cuitiño, F.; Ganón, E.; Tiscordio, I. & Vicente, L. (2010). "Modelos univariados de series de tiempo para predecir la inflación a corto plazo". Documento de trabajo No. 008-2010. Banco Central de Uruguay.
- Demers, F. & De Champlain, A. (2005). "Forecasting Core Inflation in Canada: Should We Forecast the Aggregate or the Components?". Working paper 2005-44. Bank of Canada.
- Diebold, F. X. & Mariano, R. S. (1995). "Comparing predictive accuracy", *Journal of Business & Economic Statistics*, 13, 253–265.
- Dornbusch, R. (1976). "Expectations and exchange rate dynamics", *Journal of Political Economy*, 84(6), 1161–1176.
- Espasa, A., Senra, E. & Albacete, R. (2002). "Forecasting inflation in the European Monetary Union: A disaggregated approach by countries and by sectors", *The European Journal of Finance*, 8(4), 402–421.

- Firth, M. (1979). "The Relationship Between Stock Market Returns and Rates of Inflation", *The Journal of Finance*, 34(3), 743–749.
- Fritzer, F., Moser, G. & Scharler, J. (2002). "Forecasting Austrian HCPI and its Components using VAR and ARIMA Models". Working Paper 73, Oesterreichische Nationalbank.
- Grunfeld, Y. & Griliches, Z. (1960). "Is aggregation necessarily bad?", *The Review of Economics and Statistics*, 42(1), 1–13.
- Harvey, S. K. & Cushing, M. J. (2014). "Does Using Disaggregate Components Help in Producing Better Forecasts for Aggregate Inflation?", *Journal of Economics and Development Studies*, 2(2), 527–546.
- Hendry, D. F. & Clements, M. P. (2004). "Pooling of forecasts", *Econometrics Journal*, 5, 1–31.
- Hendry, D. F. & Hubrich, K. (2011). "Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate", *Journal of Business & Economic Statistics*, 29(2), 216–227.
- Hubrich, K. (2005). "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?", *International Journal of Forecasting*, 21(1), 119–136.
- Hyndman, R. J., Khandakar, Y. (2008). "Automatic Time Series Forecasting: The forecast Package for R", *Journal of Statistical Software*, 27(3), 1-22.
- Miller, K. D., Jeffrey, F. J. & Mandelker, G. (1976). "The 'Fisher effect' for risky assets: An empirical investigation", *The Journal of Finance*, 31(2), 447–458.
- Nelson, C. R. (1976). "Inflation And Rates Of Return On Common Stocks", *The Journal of Finance*, 31(2), 471–483.
- Pino, G., Tena, J. D., & Espasa, A. (2013). "Forecasting disaggregates by sectors and regions: the case of inflation in the euro area and Spain". Working Paper 13-08. Statistics and Econometrics Series 07. Universidad Carlos III de Madrid.
- R Core Team (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.
- Solnik, B. (1983). "The relation between stock prices and inflationary expectations: The international evidence", *The Journal of Finance*, 38(1), 35–48.
- Svensson, L. E. O. (1997). "Inflation forecast targeting: Implementing and monitoring inflation targets", *European Economic Review*, 41(6), 1111–1146.
- Theil, H. (1954). *Linear aggregation of economic relations*. North-Holland Pub. Co.



Estimating the Price Elasticity of Demand for Cannabis: A Geographical and Crowdsourced Approach

HALCOUSSIS, DENNIS

Department of Economics

California State University, Northridge (United States)

E-mail: dhalcoussis@csun.edu

LOWENBERG, ANTON D.

Department of Economics

California State University, Northridge (United States)

E-mail: anton.lowenberg@csun.edu

ROOF, ZACH

Department of Economics

California State University, Northridge (United States)

E-mail: zroof85@gmail.com

ABSTRACT

Recent legalizations of cannabis at the state level in the United States have given rise to renewed interest in the price elasticity of demand for cannabis and implications for likely state excise and sales tax revenues. We use crowdsourced data on prices, qualities, and consumption of cannabis across regional markets in the United States to estimate the price elasticity of demand for cannabis. We use distances from growing areas to the major markets as a proxy for cost variations. We obtain an estimate of the price elasticity of demand for cannabis of -0.418 , which suggests that legalizing and taxing recreational cannabis use would be lucrative for government treasuries.

Keywords: Cannabis; price elasticity; demand.

JEL classification: D12; H71; E26.

MSC2010: 91C99; 62P20.

La estimación de la elasticidad en los precios de la demanda de cannabis: un enfoque geográfico y basado en la colaboración abierta distribuida

RESUMEN

Las recientes legalizaciones del cannabis a nivel estatal en los Estados Unidos han dado lugar a un renovado interés en la elasticidad en los precios de la demanda de cannabis y sus implicaciones en los ingresos por impuestos estatales y sobre ingresos por ventas. Utilizamos datos en colaboración abierta sobre precios, calidades y consumo de cannabis en los mercados regionales de los Estados Unidos para estimar la elasticidad en los precios de la demanda de cannabis. Utilizamos distancias desde las zonas de cultivo hasta los principales mercados como sustituto de las variaciones de los costos. Se obtiene una estimación de la elasticidad en los precios de la demanda de cannabis de -0,418, lo que sugiere que la legalización y el gravamen del cannabis para uso recreativo sería lucrativo para los erarios públicos.

Palabras claves: cannabis; elasticidad en los precios; demanda.

Clasificación JEL: D12; H71; E26.

MSC2010: 91C99; 62P20.



1. Introduction

Changes in the U.S. legal regime governing the production and sale of cannabis have gained momentum in recent years. Starting in 1996, the sale of cannabis for medical use was legalized in California and thereafter in 22 other states and in the District of Columbia, with the number of states expected to reach almost 30 in the next year (*Economist*, 2016, p. 18). Possession of cannabis in small quantities has been decriminalized in several states, starting with Oregon in 1973 and now extending to 16 states including California and New York (Norml.org, 2013). Decriminalization means no prison time or criminal record for first-time possession of a small amount of cannabis for personal consumption (Norml.org, 2013). Although the sale and possession of cannabis for recreational purposes is still illegal under federal law, ballot measures providing for legalization passed in Colorado and Washington state in November 2012. Colorado voters approved Amendment 64 to the state constitution, which allows persons aged twenty-one and older to purchase up to one ounce of cannabis at specially regulated retail stores. Possession is legal, but not public use. Adults are allowed to grow up to six cannabis plants in their homes (Gurman, 2012). Similarly, Washington state's Initiative 502 removes criminal sanctions for anyone aged twenty-one and older possessing one ounce or less of cannabis, although it also does not permit public consumption (Myers, 2012). Colorado and Washington commenced legal sales in 2014; Oregon followed suit in October 2015 and Alaska will soon join them (*Economist*, 2016, p. 18).

The Obama administration has indicated that as long as “strong and effective regulatory and enforcement systems” are maintained at the state level, the U.S. Department of Justice will not seek to overturn state laws or enforce federal prohibition (Dennis, 2013; Wing, 2013; *Economist*, 2013, p. 61). In the meantime, efforts are under way to place legalization measures on the ballot in several additional states, including Arizona and California (Wing, 2013). Canada is expected, in 2017, to become the first G7 country to legalize recreational cannabis use (*Economist*, 2016, p. 18).

The impetus for legalization stems in large part from the expected tax revenue generated from the legal sale of cannabis, especially in states confronting severe budget pressures, and from the widespread belief that legalization will mitigate drug-related crime and relieve crowding of state penal systems overburdened by minor drug offenders. Some Democratic

strategists might be motivated to place cannabis legalization measures on state ballots by the hope that such measures will mobilize electoral turnout of young and liberal voters (Wing, 2013). There is also considerable public support for legalization. According to a Gallup poll, in 2013, for the first time, a majority of Americans – 58 percent – favored legalization (NBC News, 2013).

The Washington state initiative legalizing cannabis applies a 25 percent excise tax and, according to official estimates, could generate as much as \$1.9 billion in state tax revenue over five years (State of Washington, 2015). In November 2013, voters in Colorado approved a 15 percent excise tax and a 10 percent sales tax on cannabis, which are in addition to a 2.9 percent sales tax applied to cannabis stores (9news.com, 2013). Colorado collected \$44 million in cannabis taxes in 2014 and \$72.5 million in the first eight months of 2015, with a total of \$109 million anticipated for the year (*Economist*, 2015, p. 34). Nationally, the tax base is potentially significant: \$1.43 billion of legal cannabis is estimated to have been sold in the United States in 2013 and this figure was expected to grow by 64 percent, to \$2.34 billion, in 2014. Fourteen more states are predicted to legalize cannabis for adult recreational use in the next five years, creating a potential \$10.2 billion cannabis market by 2018 (ArcView, 2013; Schwartz, 2013). One study estimates that nationwide legalization of cannabis could ultimately generate \$8.7 billion of tax revenue annually (Miron and Waldock, 2010). An earlier estimate is \$2.55 billion to \$9.09 billion (Caputo and Ostrom, 1994). Legalization in California alone is expected to produce \$1 billion per year in tax revenue (*Economist*, 2016, p. 19).

The trend toward decriminalization and legalization has raised important questions about the market for cannabis and specifically the determinants of demand for the drug and the price elasticity of demand. Estimates of price elasticity may be expected to inform the debate about the magnitude of government tax receipts in markets in which cannabis has been legalized (Caputo and Ostrom, 1994). Several efforts have recently been made in the economics literature to estimate the price elasticity of demand for cannabis. Such efforts are necessarily premised on an empirical identification of the demand equation for cannabis, which can be difficult in the case of a product that, until recently, has been traded only in illegal markets. We approach this problem by using a hedonic method of estimating the relationship between prices and quantities in the market which gives us an estimate of the price elasticity of demand for cannabis. Our method utilizes a novel dataset consisting of distances from the primary Mexican growing areas to all of

the major markets in the United States. These distances are then used as proxies to capture variation in costs of production and distribution across markets. We estimate a price elasticity of demand for cannabis of -0.418. This estimate is sufficiently inelastic that it is reasonable to believe that the tax revenues from an excise tax on legalized cannabis would be substantial.

In the next section we outline the various approaches that have been used in the literature to identify the demand for cannabis and other drugs, and to estimate the price elasticity of demand. Section 3 describes the data and model used in our study, with empirical results presented in section 4 and implications and conclusions in section 5.

2. Studies of the Market for Cannabis

Empirical estimation of demand and supply equations for illegal goods, or for goods traded in informal markets, confront unique challenges, not least of which being a lack of reliable data on prices, quantities, costs, and consumption (Bohme and Thiele, 2012). Thus for example, due to the unavailability of cannabis price data, many studies have focused on estimating cross-price elasticities of demand, rather than own-price elasticities for cannabis (Pacula *et al.*, 2000). Much of the existing literature on cannabis, cocaine, heroin and other illegal drugs uses interview or survey data (Chaloupka *et al.*, 1998; Saffer and Chaloupka, 1999a, 1999b; Dave, 2008; Pacula *et al.*, 2010) or enforcement data such as arrests and fines (Dave, 2008; Pacula *et al.*, 2010). A notable exception is Olmstead *et al.* (2015), who use actual daily price and quantity data for a sample of heroin users, as well as experimental data obtained by confronting the same heroin users with alternative hypothetical prices in a laboratory setting. Both empirical strategies yield an estimated conditional price elasticity of demand for heroin of approximately -0.80.

One of the problems of relying on individual consumer-level data is that, particularly with addictive drugs such as cocaine and cannabis, tastes may be endogenous in the sense that current consumption is partly determined by past consumption (van Ours, 1995). This endogeneity becomes especially problematic in estimating the demand equation, which requires tracing out the effects of supply shifts and resulting price changes on the quantity demanded. Habitual drug users' consumption will not vary significantly in response to such shifts whereas casual users are likely to be more sensitive to supply-induced price changes (Pacula *et al.*, 2010). The dependence of present consumption of addictive drugs on past consumption may produce

nonlinear estimates of price elasticity, with high elasticity at both low and high prices and low elasticity in the middle range of prices (Bretteville-Jensen and Biørn, 2004).

A major parameter that is used in much of the literature to identify demand for cannabis is decriminalization or the passage of so-called “medical marijuana” laws (Saffer and Chaloupka, 1999a, 1999b). Such policy changes may be associated with lower costs or risks of acquiring the drug, a supply-side shift which can be used to identify the demand equation. Thus for example, Saffer and Chaloupka (1999a, 1999b) find that cannabis decriminalization increased the probability of participation in cannabis use among high school seniors by roughly 8 percent. However, Chaloupka *et al.* (1998, p. 20) show that the elasticity of youths’ cannabis use with respect to fines for possession is very small, in the range of -0.007 to -0.008. Wen *et al.* (2014) examine the effects of medical marijuana laws in seven states between 2004 and 2011, and find that the adoption of such laws increased the probability of cannabis use by 16 percent among those aged twenty-one or older and by 5 to 6 percent among those aged twelve to twenty years.¹ A contrary result is that of Anderson *et al.* (2014) who use national and state Youth Risk Behavior Surveys, the 1997 National Longitudinal Survey of Youth and the Treatment Episode Data Set to show that legalization of medical marijuana does not in fact produce an increase in cannabis consumption among teenagers. But in a study of cannabis possession arrests across cities from 1988 to 2008, Chu (2014) finds that the enactment of medical marijuana laws increased arrests among adult males by about 15 to 20 percent and increased treatment admissions to rehabilitation facilities by 10 to 20 percent.

Pacula *et al.* (2000) use survey data on high school seniors’ cannabis participation to estimate time-series and cross-sectional models of demand for the drug. In order to account for the endogeneity of the price variable and identify the demand equation in the time-series model, these authors use variations in the amount of resources allocated to apprehension and conviction of cannabis dealers as an instrument for exogenous shifts of the supply curve. The supply function is assumed to be infinitely elastic, since high school seniors are price takers if they represent a small fraction of the cannabis market. The study finds that the estimated price elasticity of annual cannabis participation ranges from -0.06 to -0.47.

Davis *et al.* (2016) study the cannabis demand function in a cross-section of U.S. states, cities and counties. They use electricity prices, the legal status of medical marijuana and

¹ For expanded findings, see Wen *et al.* (2015).

distance to Mexico City as instrumental variables to identify the cannabis demand function, these three variables being treated as exogenous supply-side factors which affect the price of cannabis. The authors argue that, although medical marijuana legislation might be expected to reduce the risk of use and increase social acceptance, thereby influencing demand, this variable is likely to have a stronger impact on the supply side, as it allows for home cultivation. Electricity prices are argued to be the largest single component of production cost for domestic producers, since most domestic production takes place indoors and requires electricity for high-intensity lighting, dehumidification, space heating, and carbon dioxide generation. The authors use distance to Mexico City to capture the cost of imported cannabis (Davis *et al.*, 2016). Using these variables as instruments, Davis *et al.* (2016) estimate that the price elasticity of demand for cannabis ranges from -0.67 to -0.79. These findings are broadly consistent with earlier estimates of the price elasticity, ranging from -0.40 to -1.51, cited by Caputo and Ostrom (1994, p. 486).

Like Davis *et al.* (2016), we use crowdsourced transaction data on prices and quantities for high-, medium- and low-quality grades of cannabis in a cross-section of market areas, obtained from the priceofweed.com website, in order to estimate the price elasticity of demand. The validity of crowdsourced data is attested to in a study by Lutz (2016) which shows that crowdsourced samples produce results similar to those in the laboratory. Also, following the method established by Davis *et al.* (2016), we incorporate a measure of spatial distances from primary growing areas to the major markets in the United States. The distance from growing area to market is a meaningful indicator of cost differences across markets and therefore can be used as a proxy for production costs in a hedonic model. However, whereas these latter authors use the distance from each state's geographic center to Mexico City, we take the average of the distance from the center of each county in the United States to the three states in Mexico which account for the majority of Mexican cannabis production. According to the 2009 United Nations World Drug Report, the Mexican states of Sinaloa, Chihuahua, and Durango accounted for 76 percent of cannabis crop eradication in Mexico. This implies that these states, located in the north of the country, are the major growers of cannabis.² In the next section we discuss the data and model used in our study in more detail.

² The 2009 United Nations World Drug Report can be found at: <http://www.unodc.org/unodc/en/data-and-analysis/WDR-2009.html>.

3. Data and Model

In order to evaluate the potential for state governments to raise revenue by legalizing and taxing recreational cannabis use, it would be useful to have an estimate of the price elasticity of demand for cannabis in the United States, but transactions data for illegal markets can be difficult to acquire. We address this potential stumbling block by using data for individual market transactions that were collected by [priceofweed.com](http://www.priceofweed.com), a site that describes its function in the following manner: “We crowdsource the street value of marijuana from the most accurate source possible: you, the consumer. Help by anonymously submitting data on the latest transaction you’ve made.”³ Each transaction entered by a consumer contains the city in which the transaction occurred, the total price of the transaction, the weight of the cannabis involved, the quality of the product (low, medium, or high), and the date. There were 29,472 transactions that were entered onto the site from September 2010 through March 2012. County level data were also used to allow additional control variables and instruments to be implemented that were not available at the city level.

We use Full Information Maximum Likelihood (FIML) to estimate a semi-log model. The logarithm of price and logarithm of quantity of cannabis bought and sold are endogenous variables. The logarithms are appropriate to capture the non-linear aspect of cannabis demand as well as facilitating the interpretation of slope estimates as elasticities. Gallet and List (2003, p. 825), in a metastudy, report that out of 523 estimated cigarette price elasticities, 292 come from a double or semi log model. Exogenous demand parameters include the total number of arrests for possession or sale of cannabis in each county, median household income in each county, and dummies measuring the quality of cannabis. To represent supply costs, we use a measure of the distance that imported cannabis was transported to each of the market areas in our dataset. These variables are given more precise definitions below.

The model is estimated using each transaction in the dataset as a separate observation, giving 29,461 observations (eleven of the 29,472 observations that were originally collected from the [priceofweed.com](http://www.priceofweed.com) site were either incomplete or were omitted because the price was recorded as zero, so that 29,461 observations are available).

³ See the front page of www.priceofweed.com.

The model to be estimated with full information maximum likelihood is as follows:

$$\ln(Q_i) = B_0 + B_1 \cdot \ln(P_i) + B_2 \cdot \text{QUALITY1}_i + B_3 \cdot \text{QUALITY3}_i + B_4 \cdot \ln(\text{INCOME}_i) + B_5 \cdot \text{ARRESTS}_i + B_6 \cdot \ln(\text{DISTANCE}_i) + e_i$$

where

$\ln(Q_i)$ = the natural logarithm of the quantity, in grams, of cannabis for transaction i . Each quantity observation is rounded to the nearest whole integer. The data are from priceofweed.com for 29,461 transactions submitted to the website from September 2010 through March 2012.

$\ln(P_i)$ = the natural logarithm of the price, in U.S. dollars, per gram of cannabis for transaction i . As in the case of the quantity data, the price data are from priceofweed.com for 29,461 transactions submitted to the website from September 2010 through March 2012.

$\text{QUALITY1}_i = 1$ if the quality of the cannabis in the transaction is rated as low (out of low, medium, and high); $\text{QUALITY1}_i = 0$ otherwise. Quality ratings are submitted by the consumers and correspond to the same 29,461 transactions submitted to priceofweed.com from September 2010 through March 2012 mentioned above for the price and quantity data.

$\text{QUALITY3}_i = 1$ if the quality of the cannabis in the transaction is rated as high; $\text{QUALITY3}_i = 0$ otherwise. As above, quality ratings are submitted by the consumers and correspond to the same 29,461 transactions submitted to priceofweed.com from September 2010 through March 2012. Note that the middle quality rating is represented by the intercept in the model.

$\ln(\text{INCOME}_i)$ = the natural logarithm of median household income, in thousands of dollars, for the county in which transaction i occurred. Data are from the United States Census Bureau, "Small Area Income and Poverty Estimates for 2011". They can be found at <http://www.census.gov//did/www/saipe/> (accessed October 2, 2015).

ARRESTS_i = the number of arrests for sale or possession of cannabis per 100,000 residents for the county in which transaction i occurred. These data are from the

Federal Bureau of Investigation's "Uniform Crime Reports", http://www.drugscience.org/States/US/US_home.htm (accessed October 2, 2015). Data are for 2007, except for Wyoming counties and the Missouri counties of Adair, Audrain, and Boone, which are for 2006. There are 12 counties in total for which there were no arrests, so the natural logarithm cannot be used in those cases.

$\ln(\text{DISTANCE}_i)$ = the natural logarithm of the average distances, in miles, between the county of transaction i and the Mexican states of Sinaloa, Chihuahua, and Durango, which are the major North American growing regions for cannabis. For each U.S. county, three distances are calculated, namely, the straight line distance from the center of the county to the geographic center of each of the three Mexican states. These three distances are then averaged for each U.S. county in the dataset. It should be noted that this measure does not consider differences in the ease of crossing borders at different points or in the characteristics of distribution networks.

4. Empirical Results

Table 1 shows descriptive statistics for Q, P, QUALITY1, QUALITY3, INCOME, ARRESTS, and DISTANCE. Table 2 gives the FIML results. As part of the FIML estimation process, an additional instrument was used. In addition to DISTANCE as defined above, a domestic version of DISTANCE was used as an instrument where weighted distances from all other states were calculated, using cannabis crop eradication data as the weights. The eradication data were used to represent the prominence of each state as a growing area. These eradication data are from a United States Department of Justice website which is no longer active; however, the data can be obtained from the authors on request. Note that this variable is used as an additional instrument in the FIML process to represent supply costs along with DISTANCE; it is not a variable in the regression.

The main variable of interest, the logarithm of P, has a slope estimate that is statistically significant at a 1% level. The -0.418 coefficient estimate for the log of P indicates a relatively inelastic demand for cannabis in the United States. This estimate falls within the range of values for the price elasticity of demand for cannabis found in previous work (see Caputo and Ostrom,

1994, and Davis *et al.*, 2016), and bodes well for state governments that hope to raise revenue by legalizing and taxing recreational cannabis use.

Table 1. Descriptive Statistics for Variables Used in FIML Model, for 29,461 transactions conducted from September 2010 to March 2012.

Variable	Mean (Standard Deviation)	Minimum	Maximum
Q (Quantity in grams)	14.70 (10.97)	4.0	28.0
P (Price in dollars)	10.81 (8.76)	0.036	249.75
QUALITY1 (Low quality rating)	0.067 (0.250)	0	1
QUALITY3 (High quality rating)	0.347 (0.476)	0	1
INCOME (Median household income in thousands of U.S. dollars, by county, for 2011)	54.403 (13.596)	21.865	119.525
ARRESTS (For sale or possession of cannabis per 100,000 residents by county)	297.25 (175.38)	0	3411
DISTANCE (Average distance in miles between county and major growing areas of Sinaloa, Chihuahua, and Durango)	1506.99 (497.99)	283.17	3375.46

Table 2. FIML Results for Cannabis Market for 29,461 transactions conducted from September 2010 to March 2012.

Left-Hand-Side Variable: Ln(Q) (Natural logarithm of quantity in grams)			
Variable	Coefficient (Standard Error)	Z-statistic	p-value
Intercept	3.190 (0.223)	14.33	0.000
Ln(P) (Natural logarithm of price in dollars)	-0.418 (0.007)	-60.12	0.000
QUALITY1 (Low quality rating)	0.133 (0.021)	6.281	0.000
QUALITY3 (High quality rating)	0.187 (0.011)	17.43	0.000
Ln(INCOME) (Natural logarithm of median household income in thousands of U.S. dollars, by county, for 2011)	-0.012 (0.021)	-0.600	0.548
ARRESTS (For sale or possession of cannabis per 100,000 residents by county)	0.0000683 (0.0000270)	2.53	0.011
Ln(DISTANCE) (Natural logarithm of the average distance in miles between county and major growing areas of Sinaloa, Chihuahua, and Durango)	0.0000239 (0.0000095)	2.51	0.012
R-squared: 0.161			
Adjusted R-squared: 0.161			

The coefficient estimate for QUALITY1 is positive and significant at a 1% level, indicating that, holding other independent variables constant, lower quality cannabis (which presumably would cost less) is bought in larger quantities than “medium” quality cannabis (note,

however, that consumers themselves determine the quality of the cannabis when they submit the transaction data). Likewise, *QUALITY3* has a positive coefficient estimate that is significant at 1%, indicating that high quality cannabis is also purchased in larger quantities than medium quality cannabis (even though it would likely cost more). The coefficient estimate for the logarithm of *INCOME* is insignificant. The coefficient estimates for *ARRESTS* and the logarithm of *DISTANCE* are also positive and significant, in this case at 5%, suggesting that increased levels of law enforcement and increased distances from growing areas are both positively correlated with quantity consumed, although the effect is not large. The positive relationship between quantity and *DISTANCE* might reflect the Alchian-Allen effect, according to which the higher quality product, which, as we have seen, is purchased in larger quantities than medium quality, is also more likely to be shipped over longer distances. This is because the addition of a per-unit transportation cost lowers the relative price of the higher quality product (see Borcharding and Silberberg, 1978).

The adjusted R-squared of 0.161 indicates a relatively weak fit for the model. This is a symptom of the illegality of the market, as a result of which information is not as reliable and transaction costs are higher than in a comparable legal market, such as that for cigarettes. An examination of the correlation matrices for the data used in Table 2 does not suggest any evidence of multicollinearity.

5. Implications and Conclusions

Although the market for cannabis has existed in the United States for longer than anyone alive can remember, a lack of good data concerning this market makes it difficult for empirical economists to contribute to the current debate concerning the legalization of cannabis. The relatively recent phenomenon of “crowdsourcing”, or obtaining data from a large number of market participants through the internet, gives rise to the hope that researchers will no longer be dependent on traditional sources of information that ignore illegal or underground markets. Our study uses transaction data collected from individual purchasers of cannabis, reported at priceofweed.com, to estimate the price elasticity of demand for cannabis. Of course, as cannabis is legalized in more states, and probably eventually nationwide, more reliable sales and price data will ultimately become available.

Following Davis *et al.* (2016), we use the spatial distance from major growing areas of imported cannabis to markets within the United States as a measure of variation across markets in the cost of supplying cannabis. In our approach, the distance measure focuses on the actual areas within Mexico where most of the cannabis is grown. For each U.S. county, we calculate the average of the distance from the county's geographic center to each of the three major cannabis growing states in Mexico. This distance variable enters in our regression model as a production cost parameter.

The market for cigarettes is a reasonable proxy for that of a legalized and taxed cannabis market. In a metastudy, Gallet and List (2003) report that the median short-run price elasticity of demand for cigarettes is -0.40 (out of 368 estimates) and that the long-run price elasticity of demand is -0.44 (out of 155 estimates). Recent research also indicates that the absolute value of the price elasticity of demand for cigarettes increases monotonically with price (Tauras *et al.*, 2016). The price elasticities reported in the cigarette literature are comparable to our estimate of the price elasticity of demand for cannabis of -0.418. Goolsbee *et al.* (2010) estimate excise tax rate elasticities for cigarettes to study the effect of internet cigarettes sales on tax elasticities. They find that internet cigarette sales have increased excise tax elasticities for cigarettes, however, the elasticity values are still low, as their results yield a range of excise tax rate elasticities from -0.189 to -0.267, even with internet sales accounted for.

State and local governments collected over \$4 billion in tax revenues from cigarettes in 2013, with state and local governments in California collecting over \$280 million. These data are from the United States Census Bureau, "Annual Survey of State and Local Government Finances," Government Finances, Volume 4, and the Census of Governments (1977-2011), "State and Local Government Finance Data Query System," <http://slfdqs.taxpolicycenter.org/pages.cfm> (accessed January 2, 2016), and can be found at the Tax Policy Center, <http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=403> (accessed January 2, 2016). If the price elasticities of demand for cannabis and cigarettes are both in the same range, as suggested by our results for cannabis, then there is every reason to expect that legalizing and taxing recreational cannabis use will yield similarly large state excise tax revenues to those collected from the taxation of cigarettes.

Acknowledgment

The authors are grateful to Michael Pistone for research assistance. All errors remain the responsibility of the authors alone. We thank the David Nazarian College of Business and Economics at California State University, Northridge for support in the form of a sabbatical grant for one of the authors.

References

- Anderson, D. M.; Hansen, B. and Rees, D. I., “Medical Marijuana Laws and Teen Marijuana Use”, NBER Working Paper No. 20332, National Bureau of Economic Research, Cambridge, MA: 2014.
- ArcView, *The State of Legal Marijuana Markets*, Second Edition, San Francisco: ArcView Market Research, <http://www.arcviewmarketresearch.com/order/state-of-the-legal-marijuana-markets-2nd-edition>, accessed December 10, 2013.
- Bohme, M. and Thiele, R., “Is the Informal Sector Constrained from the Demand Side? Evidence for Six West African Capitals”, *World Development*, 40: 1369-1381 (2012).
- Borcherding, T. E. and Silberberg, E., “Shipping the Good Apples Out: The Alchian and Allen Theorem Reconsidered”, *Journal of Political Economy*, 86: 131-138 (1978).
- Bretteville-Jensen, A. L. and Biørn, E., “Do Prices Count? A Micro-Econometric Study of Illicit Drug Consumption Based on Self-Reported Data”, *Empirical Economics*, 29: 673-695 (2004).
- Caputo, M. R. and Ostrom, B. J., “Potential Tax Revenue from a Regulated Marijuana Market: A Meaningful Revenue Source”, *American Journal of Economics and Sociology*, 53: 475-490 (1994).
- Chaloupka, F. J.; Grossman, M. and Tauras, J. A., “The Demand for Cocaine and Marijuana by Youth”, NBER Working Paper No. 6411, National Bureau of Economic Research, Cambridge, MA: 1998.
- Chu, Y-W. L., “The Effects of Medical Marijuana Laws on Illegal Marijuana Use”, *Journal of Health Economics*, 38: 43-61 (2014).
- Dave, D., “Illicit Drug Use among Arrestees, Prices and Policy”, *Journal of Urban Economics*, 63: 694-714 (2008).
- Davis, A. J.; Geisler, K. R. and Nichols, M. W., “The Price Elasticity of Marijuana Demand: Evidence from Crowd-Sourced Transaction Data”, *Empirical Economics*, 50: 1171-1192 (2016).

- Dennis, B., “Obama Administration Will Not Block State Marijuana Laws, If Distribution Is Regulated”, *Washington Post*, August 29, 2013, available at http://www.washingtonpost.com/national/health-science/obama-administration-will-not-preempt-state-marijuana-laws--for-now/2013/08/29/b725bfd8-10bd-11e3-8cdd-bcdc09410972_story.html?wpisrc=al_comboPN.
- Economist*, September 28, 2013, p. 61; October 17, 2015, p. 34; February 13, 2016, pp. 18-21.
- Federal Bureau of Investigation, “Uniform Crime Reports”, available at http://www.drugscience.org/States/US/US_home.htm, accessed October 2, 2015.
- Gallet, C. A. and List, J. A., “Cigarette Demand: A Meta-Analysis of Elasticities”, *Health Economics*, 12: 821-835 (2003).
- Goolsbee, A., Lovenheim, M. F. and Slemrod, J., “Playing With Fire: Cigarettes, Taxes, and Competition from the Internet”, *American Economic Journal: Economic Policy*, 2: 131-154 (2010).
- Gurman, S., “Coloradans Say Yes to Recreational Use of Marijuana”, *The Denver Post*, November 6, 2012, available at http://www.denverpost.com/breakingnews/ci_21941918/nation-watches-colorados-marijuana-legalization-vote.
- Lutz, J., “The Validity of Crowdsourcing Data in Studying Anger and Aggressive Behavior: A Comparison of Online and Laboratory Data”, *Social Psychology*, 47: 38-51 (2016).
- Miron, J. A. and Waldo, K., *The Budgetary Impact of Ending Drug Prohibition*, Washington, D. C.: Cato Institute, 2010.
- Myers, L. L., “Washington Marijuana Law: Prosecutors Warn Pot Smoking In Public Still Outlawed”, *Huffington Post*, December 5, 2012, available at http://www.huffingtonpost.com/2012/12/05/washington-marijuana-law_n_2248778.html.
- NBC News, “Poll: Majority of Americans in Favor of Legalizing Marijuana”, October 22, 2013, available at <http://www.nbcnews.com/video/nightly-news/53347727/#53347727>.
- Norml.org, “States That Have Decriminalized”, available at <http://norml.org/aboutmarijuana/item/states-that-have-decriminalized>, accessed December 8, 2015.
- Olmstead, T. A.; Alessi, S. M.; Kline, B.; Pacula, R. L. and Petry, N. M., “The Price Elasticity of Demand for Heroin: Matched Longitudinal and Experimental Evidence”, *Journal of Health Economics*, 41: 59-71 (2015).
- van Ours, J. C., “The Price Elasticity of Hard Drugs: The Case of Opium in the Dutch East Indies, 1923-1938”, *Journal of Political Economy*, 103: 261-279 (1995).

Pacula, R. L.; Grossman, M.; Chaloupka, F. J.; O'Malley, P. M.; Johnston, L. D. and Farrelly, M. C., "Marijuana and Youth", NBER Working Paper No. 7703, National Bureau of Economic Research, Cambridge, MA: 2000.

Pacula, R. L.; Kilmer, B.; Grossman, M. and Chaloupka, F. J., "Risks and Prices: The Role of User Sanctions in Marijuana Markets", *The B.E. Journal of Economic Analysis and Policy*, 10, DOI: [10.2202/1935-1682.1992](https://doi.org/10.2202/1935-1682.1992) (2010).

priceofweed.com, accessed various dates.

Saffer, H. and Chaloupka, F. J., "The Demand for Illicit Drugs", *Economic Inquiry*, 37: 401-411 (1999a).

Saffer, H. and Chaloupka, F. J., "Demographic Differentials in the Demand for Alcohol and Illicit Drugs", in F. J. Chaloupka, M. Grossman, W. K. Bickel and H. Saffer (eds.), *The Econometric Analysis of Substance Use and Abuse: An Integration of Econometric and Behavioral Economics Approaches*, Chicago: University of Chicago Press, 1999b, pp. 187-211.

Schwartz, C., "Marijuana Market Poised To Grow Faster Than Smartphones", *Huffington Post*, November 4, 2013, available at http://www.huffingtonpost.com/2013/11/04/marijuana-market_n_4209874.html.

State of Washington, "I-502 - Fiscal Impact Statement", <http://vote.wa.gov/guides/2012/I-502-Fiscal-Impact.html>, accessed December 14, 2015.

Tauras, J. A.; Pesko, M. F.; Huang, J.; Chaloupka, F. J. and Farrelly, M. C. "The Effect of Cigarette Prices on Cigarette Sales: Exploring Heterogeneity in Price Elasticities at High and Low Prices", NBER Working Paper No. 22251, National Bureau of Economic Research, Cambridge, MA: 2016.

Tax Policy Center, <http://www.taxpolicycenter.org/taxfacts/displayafact.cfm?Docid=403>, accessed January 2, 2016.

United Nations Office on Drugs and Crime, "World Drug Report, 2009", available at <http://www.unodc.org/unodc/en/data-and-analysis/WDR-2009.html>, accessed October 2, 2015.

United States Census Bureau, "Small Area Income and Poverty Estimates for 2011", available at <http://www.census.gov//did/www/saipe/>, accessed October 2, 2015.

United States Census Bureau, "Annual Survey of State and Local Government Finances", *Government Finances*, Volume 4, available at <http://slfdqs.taxpolicycenter.org/pages.cfm>, accessed January 2, 2016.

- United States Census Bureau, Census of Governments (1977-2011), “State and Local Government Finance Data Query System”, available at <http://sldqs.taxpolicycenter.org/pages.cfm>, accessed January 2, 2016.
- Wen, H.; Hockenberry, J. M. and Cummings, J. R., “The Effect of Medical Marijuana Laws on Marijuana, Alcohol, and Hard Drug Use”, NBER Working Paper No. 20085, National Bureau of Economic Research, Cambridge, MA: 2014.
- Wen, H.; Hockenberry, J. M. and Cummings, J. R., “The Effect of Medical Marijuana Laws on Adolescent and Adult Use of Marijuana, Alcohol, and Other Substances”, *Journal of Health Economics*, 42: 64-80 (2015).
- Wing, N., “These States Are Most Likely To Legalize Pot Next”, *The Huffington Post*, August 31, 2013, available at http://www.huffingtonpost.com/2013/08/30/marijuana-legalization-states_n_3838866.html.
- 9news.com, “Voters Pass Pot Measure, Reject School Tax Hike”, November 6, 2013, available at <http://www.9news.com/news/article/363083/339/Voters-pass-pot-measure-reject-school-tax-hike>.



Proceso de reestructuración bancaria: análisis de las entidades participantes, fusionadas y nacionalizadas

JIMBER DEL RÍO, JUAN ANTONIO

Departamento de Estadística, Econometría, Investigación Operativa,
Organización de Empresas y Economía Aplicada
Universidad de Córdoba (España)
Correo electrónico: jjimber@uco.es

CARIDAD Y OCERIN, JOSÉ MARÍA

Departamento de Estadística, Econometría, Investigación Operativa,
Organización de Empresas y Economía Aplicada
Universidad de Córdoba (España)
Correo electrónico: ccjm@uco.es

CEULAR VILLAMANDOS, NURIA

Departamento de Estadística, Econometría, Investigación Operativa,
Organización de Empresas y Economía Aplicada
Universidad de Córdoba (España)
Correo-e: nuria.ceular@uco.es

RESUMEN

En la década de los noventa en el sector financiero se produjo un proceso sin precedentes de fusiones y adquisiciones en toda Europa. Esta tendencia se inició en Estados Unidos en los ochenta, extendiéndose rápidamente al continente europeo. El estudio realiza un análisis de las variables que tienen un efecto marginal sobre la probabilidad por parte de una entidad financiera de participar en un proceso de reestructuración financiera como entidad fusionada o nacionalizada. Para ello se usan modelos con variable endógena categórica para representar las distintas situaciones de absorción consideradas. En el proceso de modelización se han utilizado las técnicas econométricas usuales de selección de variable y la interpretabilidad financiera de éstas.

Palabras claves: fusiones; nacionalizadas; efectos marginales; Basilea III.

Clasificación JEL: C50; G21.

MSC2010: 62P20; 91G70; 91G99.

Bank Restructuring: Analysis of Participating, Merged and Nationalized Entities

ABSTRACT

In the nineties in the financial sector, there was an unprecedented process of mergers and acquisitions in Europe. This trend started in the United States in the eighties, quickly spreading to Europe. The study analyzes the variables that have a marginal effect on the probability by a financial institution to participate in a process of financial restructuring as a merged or nationalized entity. For this, categorical models with endogenous variable are used to represent the different situations of absorption being considered. In the modeling process we have used the usual econometric techniques for variable selection and their financial interpretability.

Keywords: Mergers; nationalized; marginal effects; Basel III.

JEL classification: C50; G21.

MSC2010: 62P20; 91G70; 91G99.



1. INTRODUCCIÓN

El sector financiero en la década de los noventa se vió inmerso en un proceso sin precedentes de fusiones y adquisiciones en toda Europa. Estados Unidos fue la punta de lanza en los ochenta, poco después la tendencia iniciada por EEUU se extendió por Europa. Las economías de escala son las principales causas de las fusiones y adquisiciones. Dichas economías se derivan de la revolución tecnológica, la eliminación de normas y la liberalización del sector financiero, disminuyendo las dificultades para implementar fusiones. El incremento exponencial del mercado financiero español viene determinado por la liberación del sector; dicho incremento ha llevado a perder de vista la rentabilidad de la entidad financiera y una competencia derivada del incremento de oficinas que ha llevado a la concesión de préstamos sin las garantías necesarias y se ha reducido las exigencias del *scoring* de las entidades financieras a cotas poco recomendables. Para prevenir las pérdidas en las entidades financieras y la posterior intervención del Banco Central de cada país encargado de supervisarla, se desarrolla una nueva regulación internacional. La exigencia de nuevos ratios de capital sobre el total de activos ponderados al riesgo (RWA) desde el 2013 al 2020, deriva, a causa del no cumplimiento de dicha exigencia, en la intervención de las entidades por parte del Estado (nacionalización), o las fusiones en igualdad entre entidades similares (SIP) y la adquisición de entidades (fusión por absorción). Se pone de manifiesto ante la crisis la necesidad de un sistema financiero saneado, competitivo y sólido, que facilite la finalidad original de las entidades financieras: la intermediación entre los poseedores de recursos financieros y quienes tienen necesidad de financiación.

Estas deficiencias del mercado financiero internacional se están tratando de solucionar en dos ambitos: primero, medidas que prevengan nuevas crisis del sistema financiero en el conjunto de la economía, la mejora de la estructura financiera internacional, la supervisión financiera de calidad y la exigencia de mayores requerimientos de capital; segundo, el apoyo al sistema financiero con las medidas para recuperar la confianza que los Estados están adoptando y reducir las presiones de liquidez. Se hace necesario el desarrollo de nuevas medidas que lleven a buen fin los avances realizados en las cajas de ahorro españolas, con medidas como la creación del FROB (Fondo de Reestructuración Ordenada Bancaria) en el verano de 2009 y la publicación de los resultados de las pruebas de resistencia, la Reforma de la Ley de Cajas en julio de 2010, así como las de reforzamiento de la solvencia de las entidades

financieras de febrero de 2011, el saneamiento del sector financiero, el saneamiento y venta de los activos inmobiliarios en 2012 y la reestructuración y resolución de entidades de crédito en 2013.

2. ANÁLISIS EMPÍRICO

2.1. ANTECEDENTES Y OBJETIVOS

El mayor tamaño de las entidades financieras produce sinergias operativas, fenómeno común en las operaciones de fusión (Sutton, 1980; Wheelock y Wilson, 2001). La reducción del coste unitario es uno de los factores que determina las fusiones de empresas, aprovechando de este modo las economías de escala, procedentes de costes fijos en los servicios centrales y en la comercialización, costes que son independientes del nivel de producción (Fanjul y Maraval, 1985; Vives, 1988; Berger y Ofek, 1995). Las entidades más pequeñas que se vean inmersos en un proceso de fusión podrán beneficiarse de ellas debido a que la curva de costes medios de entidades financieras tiene forma de U aplanada (Berger *et al.*, 1987; McAllister y McManus, 1993). Por esta razón las entidades medianas y grandes tendrían menores costes unitarios que las entidades pequeñas.

Desde 1970 en adelante, el papel que juegan las entidades financieras ha sufrido un cambio sustancial: la regulación financiera se ha mantenido estable durante mucho tiempo. La competencia entre entidades financieras era escasa. En la década de 1980, en los países más desarrollados, los cambios regulatorios y la innovación quebraron las limitaciones de los mercados financieros, eliminando la segmentación tan elevada existente en el momento. La década de 1990 continúa con la liberalización del sector iniciada en los ochenta; tienen que afrontar nuevos riesgos en la gestión debidos a la innovación tecnológica, que se abre paso en las entidades financieras y a una mayor competencia, especialmente en los mercados mayoristas, donde la globalización ha erosionado aún más las barreras. Las entidades financieras llevan su negocio fuera de balance, centrando su actividad de negocio de ingresos no financieros. Como consecuencia la gestión del riesgo se ha ampliado para incluir la gestión de los derivados fuera de balance.

Altunbas *et al.* (2001) ponen de manifiesto la destacable disminución media anual de los costes en las entidades financieras alrededor de un 3% en los noventa como factor de mejora derivada del progreso tecnológico. Por su parte, Cavallo y Rossi

(2001), en su estudio de 442 entidades financieras de seis países europeos durante el período entre 1992 y 1997, concluyen que en todos los países analizados las economías de escala son significativas, en especial en las entidades financieras de tamaño inferior. De forma similar, en un análisis de diez países de la Unión Europea durante el período 1993–1996, Maudos *et al.* (2002) demuestran que las entidades financieras europeas de tamaño medio son más eficientes, por lo que las entidades de menor tamaño tienen la posibilidad de reducir sus costes incrementando su dimensión.

Por último, Carbó y Humphrey (2004) encuentran economías de escala para las cajas de ahorro españolas en los diferentes grupos de tamaño analizados, utilizando diferentes especificaciones de funciones de costes.

Ante esta situación, las entidades financieras se pueden plantear la posibilidad de obtener una mayor economía de escala a través del crecimiento interno o del externo. El crecimiento interno permite controlar los problemas de gestión originados en el proceso de crecimiento, pero resulta demasiado lento como para aprovechar, a corto plazo, las ventajas que ofrece tamaño mayor.

Por el contrario, las fusiones y adquisiciones al constituirse en una forma de crecimiento externo, conducen a un importante, a la vez que rápido, aumento del tamaño, lo que toma relevancia especial en entornos poco estables y con modificaciones de estructura importantes, tales como los sufridos por el sector financiero en los últimos años (Mitchell y Mulherin, 1996).

Toda la regulación internacional desde 1988, la normativa derivada de Basilea I, II y III, los nuevos requerimientos de capital, la calificación por el BCE de las entidades de riesgo sistémico, las reservas de capital y las nuevas reservas anticíclicas, así como la adaptación normativa en España, como detalla Fernández (2011), con el objeto de reestructurar y sanear el sistema financiero español, ha originado a una disminución del número de entidades financieras en el país, bien por fusión por absorción, por fusión en condiciones de igualdad, o por nacionalización de entidades financieras por parte del Estado, y posteriormente venta a otra que la absorbe.

Las entidades de mayor tamaño son las que disponen de una mayor capacidad para realizar una adquisición o abonar una elevada prima (Cyree *et al.*, 2000; Worthington, 2004), siendo, de cualquier modo, las reducciones de costes unitarios por economías de escala mayores cuando el adquirido se integra en una entidad de gran tamaño, la cual podría facilitarle, por ejemplo, acceso a mejores canales de distribución o a tecnología superior.

En el caso de las fusiones, cuando las entidades financieras participantes son de menor tamaño, las economías de escala obtenidas en el proceso de integración son significativas (Sutton, 1980). De hecho, las motivaciones de integraciones entre entidades financieras de mayor tamaño se basarían más en reducir ineficiencias, alcanzar una mejor posición competitiva o el cierre de oficinas solapadas geográficamente (Vander Venet, 1999).

El objetivo del presente análisis trata de estudiar de forma cuantitativa las distintas hipótesis, referidas al mercado financiero nacional y estimar modelos que nos permitan calcular los efectos marginales de las variables explicativas para fusionarse, fusionarse en igualdad y ser intervenida/nacionalizada.

Las hipótesis a contrastar son:

1. Las entidades responden con agilidad al cambio normativo nacional e internacional.
2. Las entidades con mayor ratio de capital son propensas a participar fusionándose en igualdad.
3. Las entidades con mayor retorno del capital y de los activos medios son propensas a participar fusionándose en igualdad.
4. Las entidades con mayores préstamos deteriorados son propensas a ser nacionalizadas.
5. Las entidades que son más eficientes son propensas a participar fusionándose en condiciones de igualdad.
6. Las entidades de cuyo porcentaje de activo fijo sobre el activo total es de mayor tamaño son propensas a ser nacionalizadas.
7. Las entidades de tamaño menor y con una carga elevada de personal y de oficinas son propensas a participar en condiciones de igualdad.
8. Las entidades con mayor importe préstamos deteriorados son propensas a ser nacionalizadas.

2.2. DATOS Y SELECCIÓN DE LA MUESTRA

El estudio estadístico se realiza para una muestra de entidades financieras de España en el período de tiempo comprendido entre 2004 y 2010. La información contable necesaria para este análisis procede de la base de datos Bankscope. Partiendo de las 309 entidades seleccionadas, se ha llevado a cabo una depuración eliminando aquellas para las que no

se dispone información suficiente, y separando las entidades activas de las entidades disueltas. Finalmente se dispone de una muestra de 71 entidades financieras, incluyendo las más representativas del sector.

La base de datos dispone de 482 variables correspondientes al balance y a la cuenta de pérdidas y ganancias, así como diversos coeficientes de rentabilidad, morosidad, porcentajes de activos financieros, etc. Para los modelos de entidades que participan en un proceso de reestructuración financiera, participar en una fusión en condiciones de igualdad y ser nacionalizada se han realizado pruebas con 93 variables de las cuales forman parte definitivamente de los modelos 24 variables:

- **Préstamos brutos:** préstamos brutos concedidos a terceros incluyendo las reservas por deterioro de préstamos. Expresado en millones de euros.
- **Tiempo:** se asigna el valor de uno a siete desde el año 2004 al 2010. Expresado en años.
- **Préstamos:** son los préstamos concedidos a terceros. Expresado en millones de euros.
- **Préstamos deteriorados:** comprende los préstamos cuya probabilidad de cobro es escasa debido a que la calidad crediticia del prestatario ha disminuido. Expresado en millones de euros.
- **Total depósitos de la clientela:** comprende los depósitos a largo y corto plazo de los clientes. Expresado en millones de euros.
- **Capital Tier 1:** se compone del capital de mayor calidad y de los elementos adicionales de Tier 1 (instrumentos híbridos, como algunas participaciones preferentes); son capaces de absorber pérdidas cuando la entidad está en funcionamiento. Expresado en millones de euros.
- **Resultado antes de impuestos:** es el resultado bruto de la entidad financiera antes del impuesto de sociedades. Expresados en millones de euros.
- **Margen de intermediación:** comprende el diferencial de precio de los depósitos recibidos frente a los préstamos concedidos. Expresado en millones de euros.
- **Total de Gastos de intereses:** son los importes por los intereses y otros costes en los que la entidad financiera incurre y que están relacionados con los fondos que ha prestado. Expresado en millones de euros.

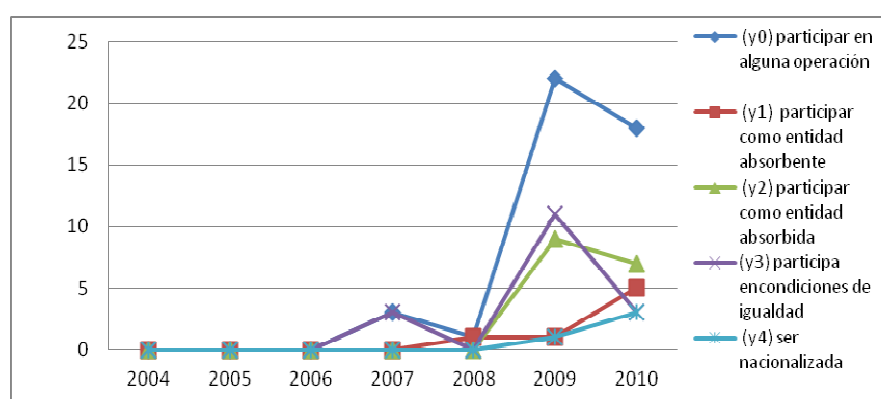
- **Ingresos/Activos ponderados por riesgo %:** es el porcentaje de los ingresos en relación a los activos ponderados por riesgo.
- **BENEFICIO/Activos ponderados por riesgo %:** es el porcentaje del beneficio en relación a los activos ponderados por riesgo.
- **Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados %:** es el porcentaje de la suma préstamos deteriorados cuya cartera está vencida y los activos adjudicados de préstamos impagados sobre la suma de los préstamos brutos y los activos adjudicados.
- **Préstamos deteriorados/Préstamos Brutos %:** es el porcentaje de préstamos deteriorados cuya cartera está vencida y se encuentra en situación de impago sobre el total de préstamos brutos de la entidad.
- **Ratio de capital:** es el porcentaje de la suma de los conceptos considerados por el Banco de España como capital sobre los activos ponderados al riesgo.
- **Retorno de la media de capital (ROAE):** es la rentabilidad del capital promedio invertido. La unidad es en porcentaje.
- **Retorno sobre activos promedio (ROAA):** son los resultados del ejercicio después de impuestos dividido por los activos promediados. La unidad es en porcentaje.
- **AF/AT:** es la relación que existe entre el activo fijo sobre el activo total de la entidad. La unidad es en tanto por uno.
- **Resultado/Nº Empleados:** es el resultado por empleado. La unidad es en millones de euros.
- **Ratio de eficiencia:** es un indicador que relaciona el conjunto de ingresos que obtiene una entidad financiera con los costes de explotación que ha asumido. La unidad es en porcentaje.
- **Gastos personal/B^{io}:** es la relación del gastos de personal sobre el beneficio. La unidad es en tanto por uno.
- **Volumen de negocio/Nº empleados:** es la cantidad de negocio que aporta a la entidad cada empleado. Expresado en millones de euros.

- **Resultado/Nº Sucursales:** es el resultado por sucursal. La unidad es en millones de euros.

Para identificar las integraciones, fusiones, absorciones y nacionalizaciones que suceden durante el periodo temporal 2004–2010, se recurre a la información facilitada por el Banco de España y al registro de hechos relevantes de la Comisión Nacional del Mercado de Valores. Del análisis de esta información, se identifican 44 entidades que participan en una operación de reestructuración financiera, de las cuales 7 entidades son absorbentes, 16 absorbidas, 17 fusionadas en condición de igualdad y 4 nacionalizadas por el Estado.

El número de operaciones de reestructuración financiera se incrementa notablemente en los años 2008 y 2010.

Gráfico 1: Evolución de operaciones de reestructuración financiera



2.3. ESPECIFICACIÓN DE LOS MODELOS Y VARIABLES SELECCIONADAS PARA EL ANÁLISIS

Para el análisis y contraste de las hipótesis planteadas, se evalúa en qué medida las entidades financieras son propensas a participar en operaciones de reestructuración, de participar en un proceso de fusión en igualdad de condiciones y de ser nacionalizada. Para ello se usan modelos con variable endógena categórica para representar las distintas situaciones de absorción, fusión o reestructuración consideradas.

Siguiendo la demostración empírica más representativa (Thompson, 1997; Worthington, 2004), la probabilidad de que una entidad financiera participe en un proceso de reestructuración financiera, ya sea como entidad absorbente, absorbida o nacionalizada, depende de una serie de variables, entre las que se incluye el

cumplimiento de los requisitos de capital, el beneficio por empleado, el beneficio por oficina o el tamaño.

2.3.1. Variables endógenas

Para el análisis de las entidades financieras que pueden participar en un proceso de reestructuración financiera, se ha definido una variable binaria (Y_0) que identifica si una entidad participa (1) o no (0). Esta variable se ha modelado en función de diversos factores causales, y el modelo permite evaluar la probabilidad, para una entidad concreta, de pertenecer a cada uno de estos dos grupos.

Del mismo modo, se define la variable artificial binaria (Y_3), que determina si una entidad se ha fusionado en condiciones de igualdad (1) o no (0).

Para finalizar, se define la variable artificial binaria (Y_4), que determina si la entidad ha sido nacionalizada por el Estado o no, tomando igualmente los valores (0) y (1).

2.3.2. Variables exógenas

El proceso de selección de variables explicativas se ha desarrollado a partir de la información disponible en la base de datos utilizada. En el proceso de modelización se han utilizado las técnicas econométricas usuales de selección de variable, así como la interpretación financiera de las mismas.

2.3.3. Resultados obtenidos

Para el proceso de cálculo de las variables regresoras sobre la variable regresada, se han tomado todos los valores i -ésimos de la variable exógena de la cual se quiere calcular el efecto marginal y se ha mantenido constante el resto de variables explicativas. De este modo, se ha calculado variable a variable el efecto marginal sobre la variable endógena en cada uno de los modelos.¹

Para el cálculo del Ratio ODDS, se ha calculado e^{β_i} .

¹ Se han realizado más de 70 modelos combinando 94 variables exógenas.

2.3.3.1. Modelo de probabilidad de participar en alguna operación de restructuración bancaria (Y_{0i})

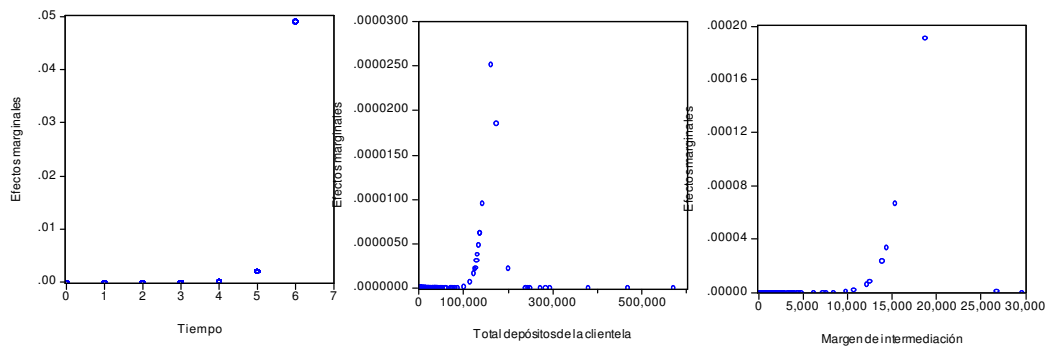
El modelo predice correctamente el 99,28% de los (0), es decir, la probabilidad de no participar en un proceso de restructuración bancaria; y un 95,24% de los (1), es decir, la probabilidad de participar en algún tipo de restructuración bancaria.

a) Efectos marginales

A continuación se analizan los efectos marginales de las variables exógenas una a una sobre la variable endógena de cada modelo. Para el modelo Y_{0i} obtenemos los siguientes datos:

- **Tiempo:** conforme nos vamos acercando a los años en los que ha habido mayor número de operaciones de restructuración financiera, es mayor el efecto marginal de ser una entidad que participa en operaciones de restructuración bancaria.
- **Total depósitos de la clientela:** para importes inferiores a 100.000 millones de euros y superiores a 250.000 millones de euros, el efecto marginal es cero. Para valores intermedios el efecto marginal es positivo.
- **Margen de intermediación:** a mayor margen de intermediación, mayor es el efecto marginal. Para valores inferiores a 10.000 millones de euros y superior a 26.500 millones de euros el efecto marginal es cero. Para valores intermedios el efecto marginal es positivo.

Gráfico 2: Análisis de los efectos marginales respecto de las variables Tiempo, Total depósitos de la clientela y Margen de intermediación

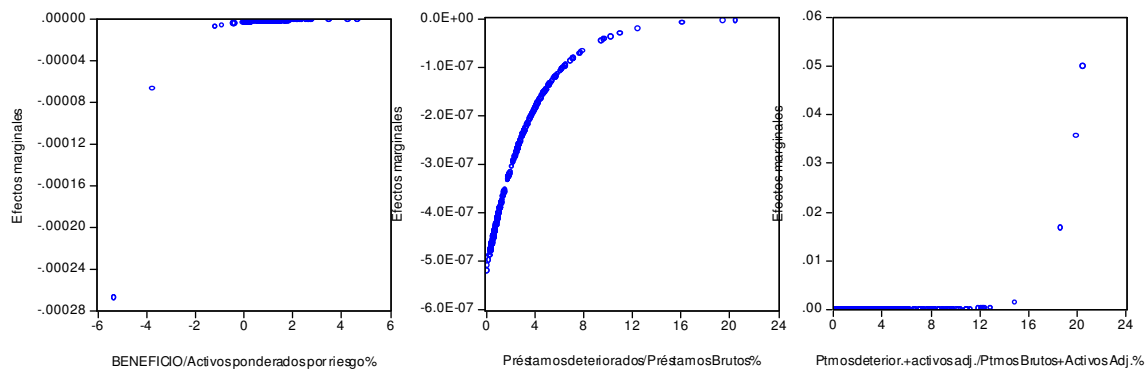


- **BENEFICIO/Activos ponderados por riesgo%:** a menor porcentaje de beneficio sobre los activos ponderados por riesgo, mayor es el efecto marginal.

Los valores están en un intervalo de -6% a 6% . Para valores cuyo porcentaje es menor que 4% el efecto marginal es de mayor valor negativo, para valores mayores a 4% el efecto marginal es cero.

- **Préstamos deteriorados (cartera vencida)/Préstamos Brutos²**: a mayor porcentaje de préstamos deteriorados sobre la cartera bruta, menor es el efecto marginal. Para valores cercanos a cero en el porcentaje de préstamos deteriorados sobre los préstamos brutos el efecto marginal es negativo y de mayor valor. Para valores mayores al 16% de efecto marginal es cercano a cero.
- **Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados²**: a mayor porcentaje, mayor es el efecto marginal. Para valores menores al 12% , el efecto marginal es cero; y para valores superiores a 12% , el efecto marginal es creciente.

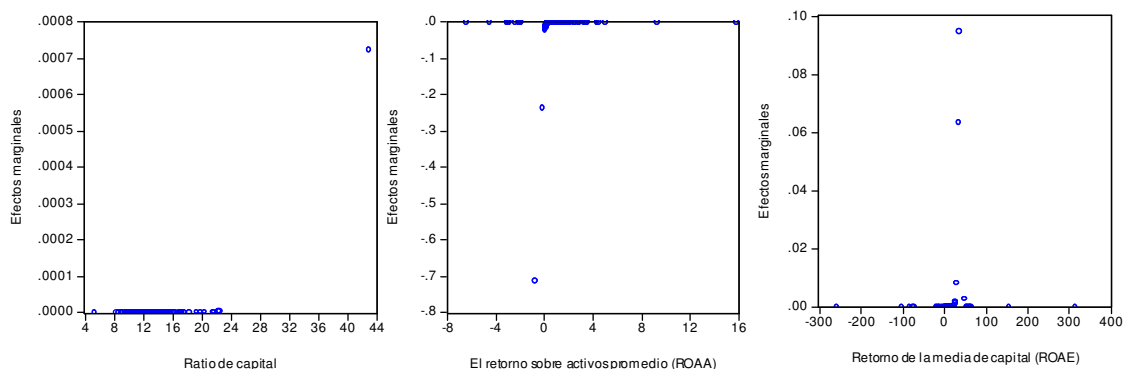
Gráfico 3: Análisis de los efectos marginales respecto de las variables BENEFICIO/Activos ponderados por riesgo%, Préstamos deteriorados/Préstamos Brutos% y Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados%



- **Ratio de capital**: para la mayoría de los valores del ratio de capital, el efecto marginal de participar en un proceso de reestructuración bancaria es cero. Para un valor de 40% de ratio de capital, el efecto marginal es positivo.
- **Retorno sobre activos promedio (ROAA)**: los valores cercanos a cero tienen efectos marginales negativos, para el resto de valores el efecto marginal es cero.
- **Retorno de la media de capital (ROAE)**: para valores comprendidos en el intervalo 25% y 52% , el efecto marginal es positivo, para el resto de valores es cero.

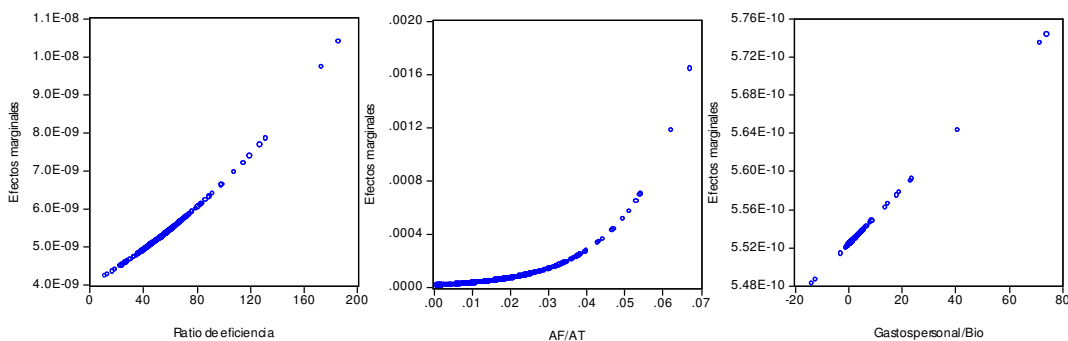
²Se ha representado la gráfica de la totalidad de efectos marginales, que por motivos de espacio no se muestran.

Gráfico 4: Análisis de los efectos marginales respecto de las variables Ratio de capital, Retorno sobre activos promedio (ROAA) y Retorno de la media capital (ROAE)



- **Ratio de eficiencia:** a mayor porcentaje de Ratio de eficiencia, mayor es el efecto marginal, estando la mayoría de las observaciones entre 40% y 80%.
- **AT/AF:** a mayor porcentaje, mayor estructura de activo fijo sobre el activo total mayor efecto marginal, estando la mayoría de las observaciones entre 0% y 0,04%
- **Gastos personal/Beneficio:** para valores negativos debido a las pérdidas, los efectos marginales son menores y negativos; a medida que la relación de beneficios sobre gastos de personal aumenta, el efecto marginal es mayor y positivo.

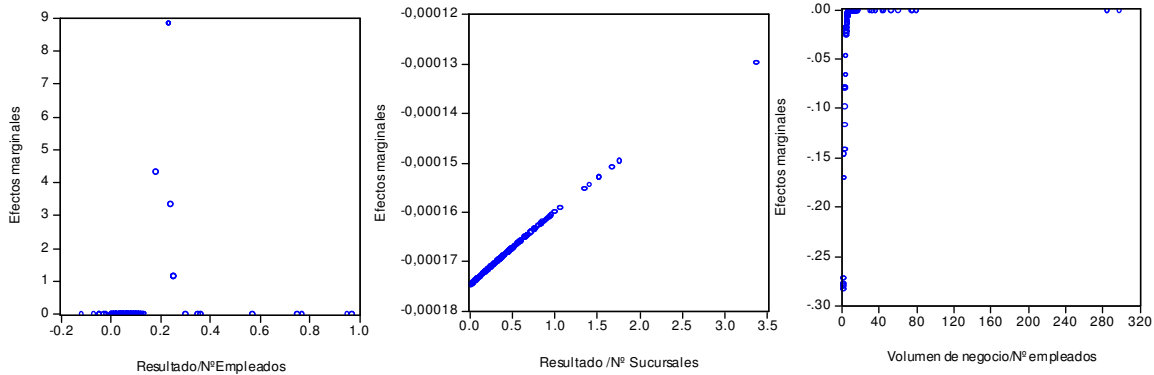
Gráfico 5: Análisis de los efectos marginales respecto de las variables Ratio de eficiencia, AF/AT y Gastos personal/Beneficio



- **Resultado/Nº Empleados:** para valores cercanos a 0,2 millones de euros por empleado, el efecto marginal es positivo; para el resto de valores es cero.
- **Resultado/Nº Sucursales:** a menor resultado por sucursal, mayor efecto marginal negativo.

- **Volumen de negocio/Nº empleados:** a menor volumen de negocio por sucursal, mayor efecto marginal negativo. Para valores superiores a 11 millones de euros de volumen de negocio por empleado, el efecto marginal es cero.

Gráfico 6: Análisis de los efectos marginales respecto de las variables Resultado/Nº Empleados, Resultado/Nº Sucursales y Volumen de negocio/Nº empleados



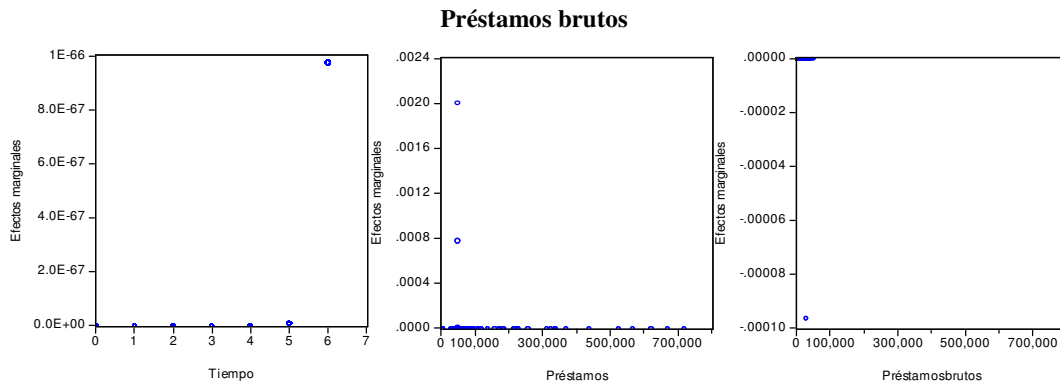
2.3.3.2. Modelo de probabilidad de participar en alguna reestructuración bancaria en condiciones de igualdad (Y_{3i})

El modelo predice correctamente el 98,65% de los ceros, es decir, la probabilidad de no participar en un proceso de reestructuración bancaria fusionándose en condiciones de igualdad; y un 90,00% de unos, es decir, la probabilidad de participar en condiciones de igualdad.

a) Efectos marginales de las variables

- **Tiempo:** conforme nos vamos acercando a los años en los que ha habido mayor número de operaciones de reestructuración financiera, mayor es el efecto marginal.
- **Préstamos:** para valores cercanos a 40 millones de euros, existe mayor probabilidad de participar en un proceso de reestructuración bancaria en condiciones de igualdad; para el resto de valores, el efecto marginal es cero.
- **Préstamos brutos:** para la mayoría de los valores, el efecto marginal es cero.

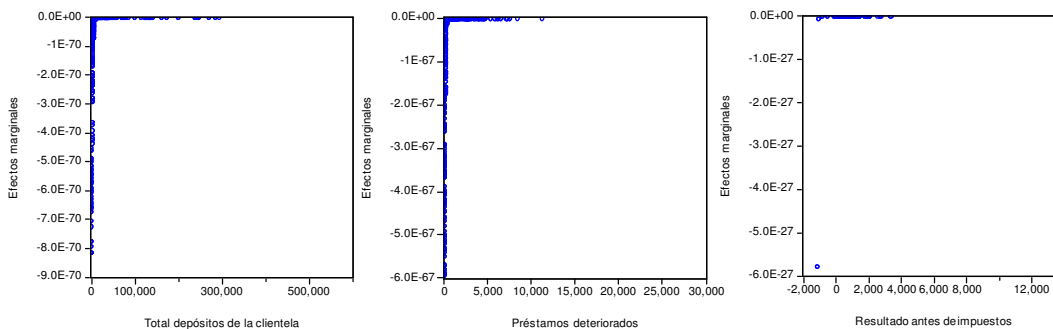
Gráfico 7: Análisis de los efectos marginales respecto de las variables Tiempo, Préstamos y



- **Total depósitos de la clientela:** para valores inferiores a 20.000 millones de euros, el efecto marginal es creciente y negativo; para valores superiores, el efecto marginal es cero.
- **Préstamos deteriorados:** para valores inferiores a 339 millones de euros, el efecto marginal es mayor y negativo; para valores superiores, el efecto marginal es cero.
- **Resultado antes de impuestos:** a mayor valor de resultados antes de impuestos, menor es el efecto marginal. Para valores inferiores a 71 millones de euros, el efecto marginal es mayor y negativo; para valores superiores, el efecto marginal es cero.

Gráfico 8: Análisis de los efectos marginales respecto de las variables Total depósitos de la clientela,

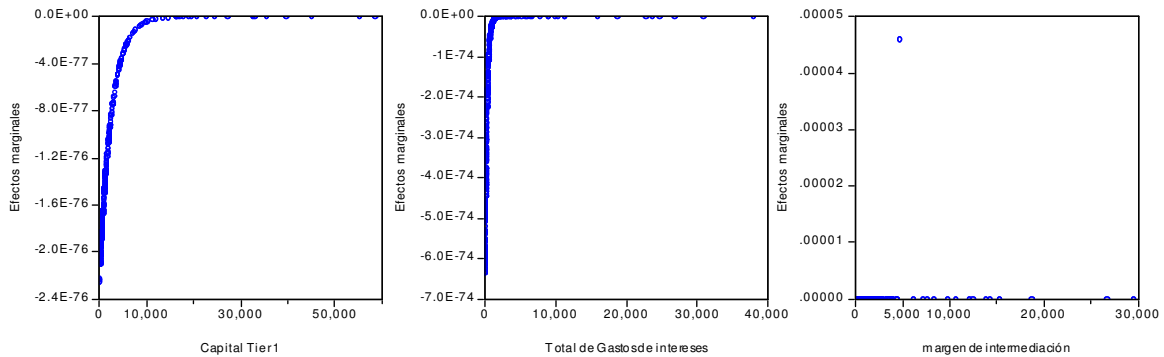
Préstamos deteriorados y Resultados antes de impuestos



- **Capital Tier 1:** a mayor importe de Tier 1, menor es el efecto marginal. Para valores superiores a 15.000 millones de euros, el efecto marginal es cero. Para valores inferiores, se va incrementando el efecto marginal negativo.

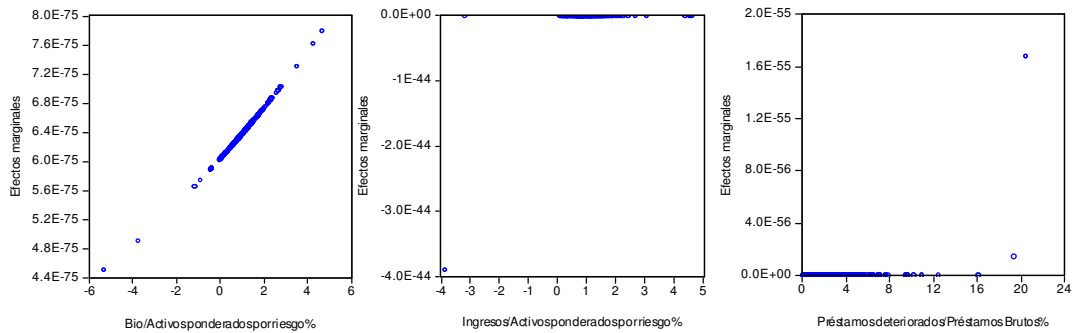
- **Total de Gastos de Intereses:** a menor gasto por intereses, mayor es el efecto marginal. Para valores superiores a 1.400 millones de euros, el efecto marginal es cero; para valores inferiores, se va incrementado el efecto marginal negativo.
- **Margen de intermediación:** para la mayoría de los valores, el efecto marginal es cero. Para valores entre 300 y 400 millones de euros, el efecto marginal es positivo.

Gráfico 9: Análisis de los efectos marginales respecto de las variables Capital Tier 1, Total de Gastos de Intereses y Margen de intermediación



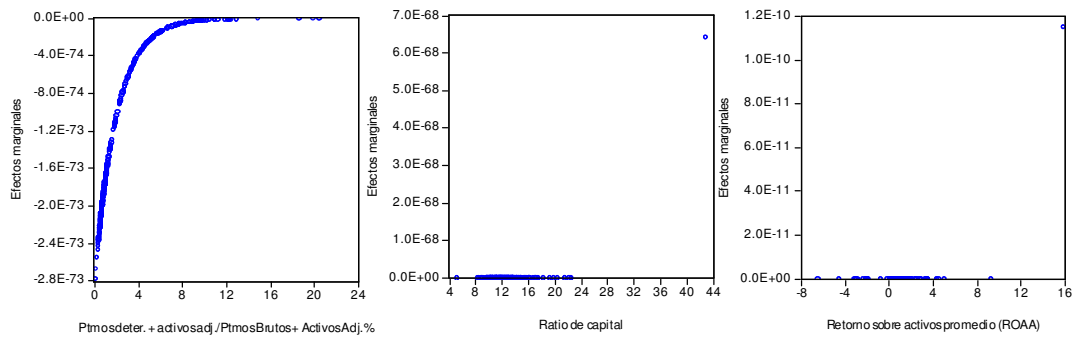
- **BENEFICIO/Activos ponderados por riesgo%:** a mayor porcentaje de beneficio sobre los activos ponderados por riesgo, mayor es el efecto marginal.
- **Ingresos/Activos ponderados por riesgo%:** para valores negativos, el efecto marginal es negativo; para valores positivos, el efecto marginal es cero.
- **Préstamos deteriorados/Préstamos Brutos%:** a mayor porcentaje de préstamos deteriorados sobre la cartera bruta, mayor es el efecto marginal.

Gráfico 10: Análisis de los efectos marginales respecto de las variables BENEFICIO/Activos ponderados por riesgo%, Ingresos/Activos ponderados por riesgo% y Préstamos deteriorados/Préstamos Brutos%



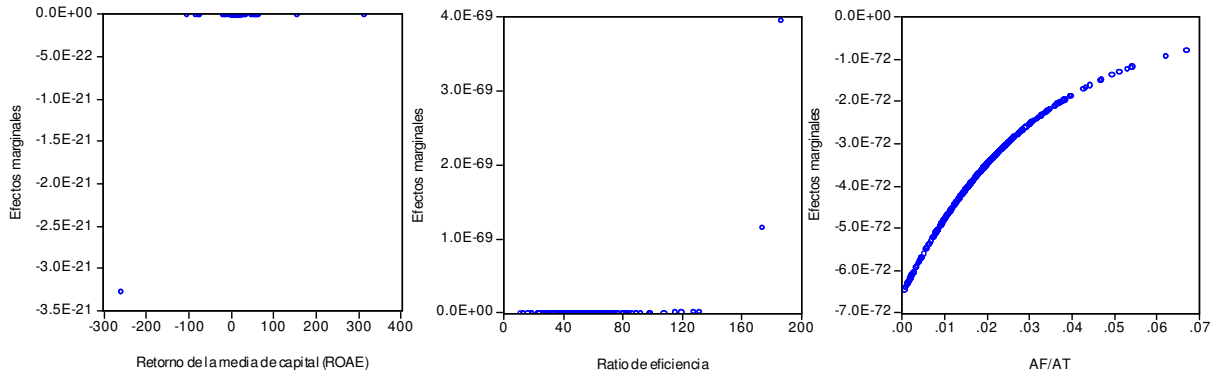
- **Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados%:** a menor porcentaje, mayor efecto marginal. Para valores superiores al 9%, el efecto marginal es cero.
- **Ratio de capital:** a mayor porcentaje de ratio de capital, mayor es el efecto marginal. Para valores inferiores al 8%, el efecto marginal es cero.
- **Retorno sobre activos promedio (ROAA):** a mayor porcentaje de retorno sobre activos promedio, mayor es el efecto marginal. Para valores menores al 0,8%, el efecto marginal es cero.

Gráfico 11: Análisis de los efectos marginales respecto de las variables Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados%, Ratio de capital y Retorno sobre activos promedio (ROAA)



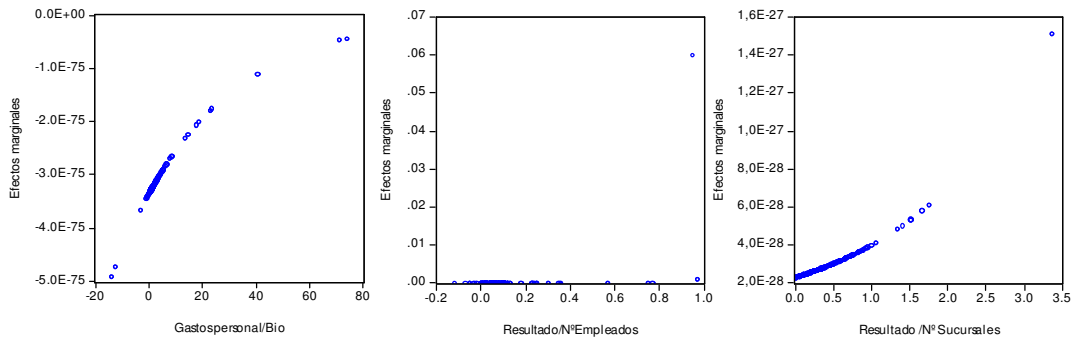
- **Retorno de la media de capital (ROAE):** a menor retorno de la media del capital, mayor es el efecto marginal negativo. Para valores superiores al 12%, el efecto marginal es cero.
- **Ratio de eficiencia:** a mayor porcentaje de ratio de eficiencia, mayor es el efecto marginal.
- **AF/AT,** a mayor valor de activo fijo sobre el activo total mayor es el efecto marginal.

Gráfico 12: Análisis de los efectos marginales respecto de las variables Retorno de la media de capital (ROAE), Ratio de eficiencia y AF/AT



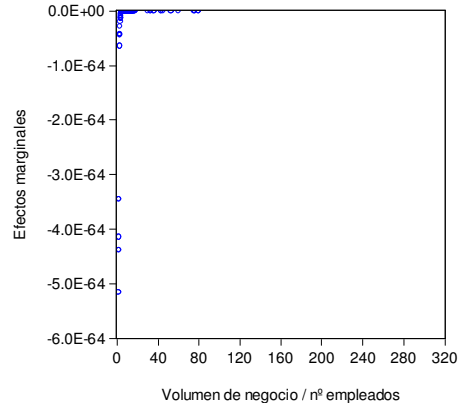
- **Gastos personal/B^{io}**: a menor valor de gasto de personal por unidad de beneficio, mayor es el efecto marginal negativo.
- **Resultado/N^o Empleados**: a mayor valor de resultado por empleado, mayor efecto marginal.
- **Resultado/N^o Sucursales**, el efecto marginal es poco significativo.

Gráfico 13: Análisis de los efectos marginales respecto de las variables Gastos personal/B^{io}, Resultado/N^o Empleados y Resultado/N^o Sucursales



- **Volumen de negocio/n^o empleados**: a menor volumen de negocio por empleado, mayor efecto marginal negativo.

Gráfico 14: Análisis de los efectos marginales respecto de la variable Volumen de negocio/nº empleados



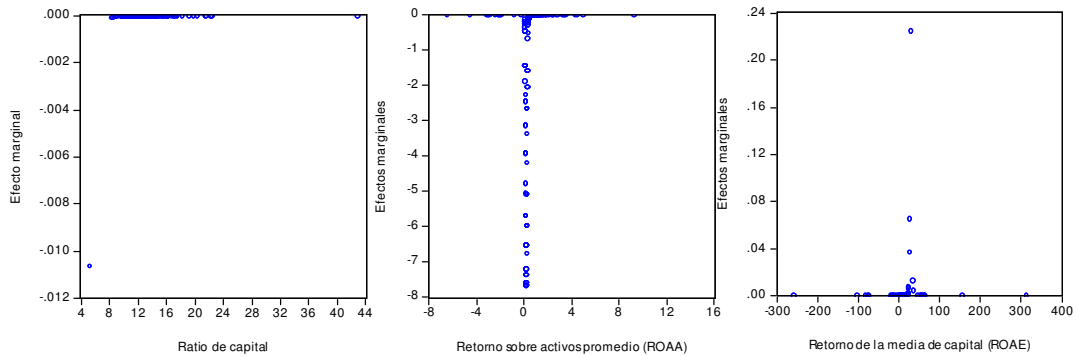
2.3.3.3. Modelo de probabilidad de que participe en alguna reestructuración bancaria como entidad nacionalizada (Y_{4i})

El modelo predice correctamente el 100,0% de los (0), es decir, la probabilidad de no participar en un proceso de reestructuración bancaria siendo nacionalizada; y un 66,67% de los (1), es decir, la probabilidad de participar siendo nacionalizada.

a) Efectos marginales de las variables

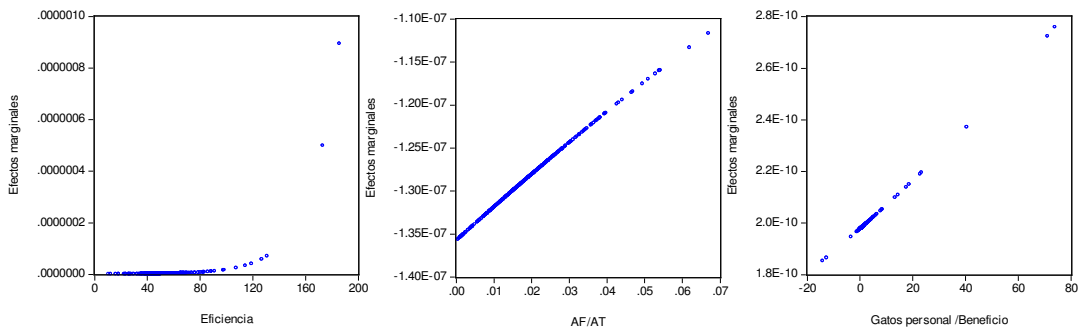
- **Ratio de capital:** a mayor porcentaje de ratio de capital, menor es el efecto marginal. Para valores superiores al 12%, el efecto marginal es cero.
- **Retorno sobre activos promedio (ROAA):** para valores cercanos a cero, el efecto marginal es mayor y negativo. Para los valores restantes, el efecto marginal es cero.
- **Retorno de la media de capital (ROAE):** para valores cercanos a cero, es mayor el efecto marginal. Para el resto de valores, el efecto marginal es cero.

Gráfico 15: Análisis de los efectos marginales respecto de las variables Ratio de capital, Retorno sobre activos promedio (ROAA) y Retorno de la media capital (ROAE)



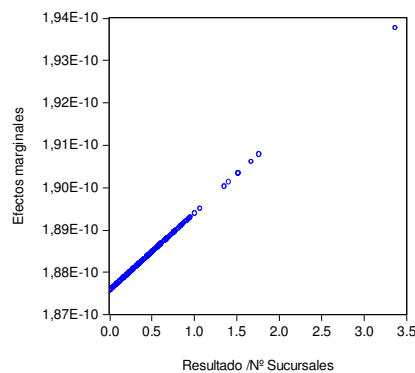
- **Ratio de eficiencia:** a mayor porcentaje de ratio de eficiencia, mayor es el efecto marginal negativo.
- **AF/AT:** mayor estructura de activo fijo sobre el activo total y, por tanto, mayor efecto marginal.
- **Gastos personal/B^{io}:** a mayor gasto de personal por unidad de beneficio, mayor efecto marginal.

Gráfico 16: Análisis de los efectos marginales respecto de las variables Ratio de eficiencia, AF/AT y Gastos personal/B^{io}



- **Resultado/Nº Sucursales:** el efecto marginal para valores menores a 0.5%, que representa 66.67% de las observaciones, es positivo.

Gráfico 17: Análisis de los efectos marginales respecto de la variable Resultados/Nº Sucursales



2.3.4. ODDS Ratio

Una forma de interpretar los coeficientes del modelo es mediante el coeficiente de ventajas o el cálculo de ODDS RATIO, lo calculamos como el exponente del coeficiente de la variable exógena sobre la que queremos calcular el ODDS Ratio³.

³ Cálculos realizados con Eviews8.

Tabla 1: ODDS Ratios

	Variable	Y_0 ,participar en un proceso de reestructuración bancaria	Y_3 ,participar en un proceso de reestructuración bancaria en condiciones de igualdad	Y_4 ,participar en un proceso de reestructuración bancaria como entidad nacionalizada
X_{16i}	Temporal	24,91347165	425,08911	
X_{17i}	Préstamos		1,0230295	
X_{18i}	Préstamos Brutos		0,97814929	
X_{24i}	Total de Depósitos de la clientela	1,000101136	0,99934501	
X_{26i}	Préstamos deteriorados		0,98088362	
X_{30i}	Resultado antes de impuestos		0,86352497	
X_{31i}	Impuesto de sociedades			
X_{34i}	Capital Tier 1		1,0015039	
X_{38i}	Total de Gastos de intereses		0,99499763	
X_{39i}	Margen de intermediación	1,000806887	1,06633775	
X_{65i}	BENEFICIO / Activos ponderados por riesgo%	0,409596799	1,56905189	
X_{66i}	Ingresos / activos ponderados por riesgo neto%		2,13E-06	
X_{67i}	Préstamos deteriorados / Préstamos Brutos%	0,769014033	25,3129263	
X_{72i}	Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados%	1,930798062	0,34341778	
X_{83i}	Ratio de capital	1,297225251	2,04335809	0,19523054
X_{84i}	El retorno sobre activos promedio (ROAA)	28.752.250.584,15	25.907,76	4,90E-14
X_{85i}	Retorno de la media de capital (ROAE)	1,670047833	0,40945009	2,47145602
X_{87i}	Ratio de eficiencia	1,005119153	1,11479944	1,0476031
X_{89i}	AF/AT	2,35E+43	1,23E-34	0,05365888
X_{90i}	Gatos personal /Beneficio	1,000529272	0,88976113	0,00453108
X_{91i}	Resultado /Nº Empleados	3,06E+62	1,08E+103	
X_{92i}	Resultado /Nº Sucursales	0,914827659	1,7567747	1,00969172
X_{93i}	Volumen de negocio / nº empleados	0,319801406	0,16293117	

3. CONCLUSIONES

Los efectos marginales y los ODDS ratio calculados, llevan a realizar las siguientes afirmaciones:

1. La variable tiempo influye en los modelos de participar en un proceso de reestructuración financiera y en participar en una fusión en condiciones de igualdad. Conforme nos acercamos temporalmente al desarrollo normativo, tanto nacional como internacional, mayor es la probabilidad de participar en un proceso de reestructuración bancaria y de fusionarse en condiciones de igualdad.

Lo que nos indica que este tipo de entidades son proactivas ante cambios tales como la creación del FROB en el año 2009, la publicación de los resultados de las pruebas de resistencia, la Reforma de la Ley de Cajas en 2010, así como las de reforzamiento de la solvencia de las entidades financieras en 2011, el saneamiento del sector financiero, el saneamiento y venta de los activos inmobiliarios del sector financiero y la reestructuración y resolución de entidades de crédito en 2012.

2. Ratio de capital tiene efecto marginal para todos los valores, siendo positivo para participar en un proceso de reestructuración financiera y para participar en una fusión en condiciones de igualdad, y negativo para participar siendo una entidad nacionalizada por el Estado. Los mayores efectos marginales positivos son para valores superiores al 14%, que representa el 30,53% de las observaciones. Los mayores efectos marginales negativos son para valores inferiores al 12%, que representa el 48,85% de las observaciones.
3. El retorno de sobre los activos promedio y de la media del capital (ROAE–ROAA) tienen efecto marginales inversos:
 - a. Para valores cercanos a cero, el efecto marginal es negativo para el retorno de los activos y positivo para valores en el intervalo de 18% a 60% de participar en un proceso en general de reestructuración financiera.
 - b. Para valores mayores a 0,65%, el efecto marginal es positivo para el retorno de los activos y negativo para valores inferiores al 10% de retorno del capital de participar en una fusión en condiciones de igualdad.
 - c. Para valores cercanos a cero de retorno de los activos, el efecto marginal es negativo; y para valores comprendidos en el intervalo 20% y 47% del retorno del capital, el efecto marginal es positivo.
4. Préstamos deteriorados / Préstamos Brutos% y Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados% tienen efectos marginales inversos:
 - a. Participar en procesos de reestructuración financiera en general: los Préstamos deteriorados/Préstamos Brutos% tiene efectos marginales

negativos para valores cercanos a cero; para valores mayores, el efecto marginal es cercano a cero. Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados% tiene efectos marginales positivos para valores superiores al 13%, que representan el 2,96% de las observaciones.

- b. Participar en fusiones en condiciones de igualdad: los Préstamos deteriorados/Préstamos Brutos% tienen efectos marginales positivos para valores superiores al 16%, que representan el 0,99% de las observaciones. Los Préstamos deteriorados + activos adjudicados / Préstamos Brutos + Activos Adjudicados% tienen efectos marginales negativos para valores menores al 8%, que representan el 85,20 % de las observaciones; para valores mayores al 8%, los efectos marginales son cercanos a cero, que representan el 14,80% de las observaciones.
5. El ratio de eficiencia: para participar en un proceso de reestructuración en general, todos los valores tienen efecto marginal positivo; para participar en una fusión en condiciones de igualdad, tiene efecto positivo para valores superiores al 73%, de los valores observados el 8,02% son superiores al 70%, lo que indica que el efecto marginal es poco representativo de las entidades observadas; para ser nacionalizada, tiene efectos marginales positivos para valores superiores al 39% de ratio de eficiencia, de los valores observados el 86,53% tienen valores entre 39% y 70%. El efecto marginal es mayor en las entidades propensas a ser nacionalizadas.
 6. La proporción de activo fijo sobre activo total: para participar en procesos de reestructuración financiera en general, el efecto marginal es positivo para todos los valores, siendo mayor el efecto marginal para valores superiores a 0.04; para fusiones en condiciones de igualdad, el efecto marginal es negativo, siendo para los valores más elevado el efecto marginal negativo y para valores cercanos a cero mayores efectos marginales negativos; para entidades nacionalizadas, el efecto marginal es negativo para todos sus valores.
 7. Gastos personal /Beneficio: el efecto marginal es creciente para todos los valores para ser nacionalizadas y negativo para participar en fusiones en condiciones de igualdad; es positivo para valores cercanos a cero para ser nacionalizada.

8. Resultado /Nº Sucursales: es negativo para participar en un proceso de reestructuración financiera en general en todos sus valores; es positivo con efectos marginales muy reducidos para participar en una fusión en condiciones de igualdad; y es positivo para ser nacionalizada, siendo para la mayoría de observaciones cercano a cero.

REFERENCIAS BIBLIOGRÁFICAS

- Altunbas, Y.; Gardener, E. P. M.; Molyneux, P. y Moore, B. (2001): Efficiency in European Banking. *European Economic Review*, 45:10, pp. 1931–1955.
- Basel Committee on Banking Supervision (2010): *Basel III: A global regulatory framework for more resilient banks and banking systems*. Disponible en <http://www.bis.org/publ/bcbs189.pdf>.
- Berger, A. N.; Hanweck, G. y Humphrey, D. B. (1987): Competitive Viability in Banking, Scale, Scope, and Product Mix Economies. *Journal of Monetary Economics*, 20, pp. 501–520.
- Berger, P. G. y Ofek, E. (1995): Diversification's Effect on Firm Value. *Journal of Financial Economics*, 37:1, pp. 39–65.
- Carbó, S. y Humphrey, D. B. (2004): Predicted and Actual Costs from Individual Bank Mergers. *Journal of Economics and Business*, 56:2, pp. 137–157.
- Cavallo, L. y Rossi, S. P. S. (2001): Scale and Scope Economies in the European Banking Systems. *Journal of Multinational Financial Management*, 11, pp. 515–531.
- Cyree, K. B.; Wansley, J. W. y Black, H. A. (2000): Determinants of Bank Growth Choices. *Journal of Banking and Finance*, 24, pp. 709–734.
- Comité de Supervisión Bancaria de Basilea, Comunicado de prensa (2010): *El Grupo de Gobernadores y Jefes de Supervisión alcanzan un amplio consenso en torno al paquete de reformas del Comité de Basilea sobre el capital y liquidez*. Disponible en http://www.bis.org/press/p100912_es.pdf.

- Fanjul, O. y Maraval, F. (1985): *La eficiencia del sistema bancario español*. Alianza, Madrid.
- Fernández, M. (2011): *La reestructuración del sector bancario español y el Real Decreto-ley para el reforzamiento del sistema*. Banco de España, Eurosistema. Disponible en <http://www.bde.es/webbde/es/secciones/prensa/intervenpub/gobernador/mfo210211.pdf>.
- Maudos, J.; Pastor, J. M.; Pérez, F. y Quesada, J. (2002): Cost and Profit Efficiency in European Banks. *Journal of International Financial Markets, Institutions and Money*, 12:1, pp. 33–58.
- McAllister, P. H. y McManus, D. (1993): Resolving the Scale Efficiency Puzzle in Banking. *Journal of Banking and Finance*, 17:2–3, pp. 389–405.
- Ministerio de Asuntos Exteriores y de Cooperación (2012): *Memorando de entendimiento sobre condiciones de política sectorial financiera*. Disponible en <https://www.boe.es/boe/dias/2012/12/10/pdfs/BOE-A-2012-14946.pdf>.
- Mitchell, M. L. y Mulherin, J. H. (1996): The Impact of Industry Shocks on Takeover and Restructuring Activity. *Journal of Financial Economics*, 41:2, pp. 193–229.00
- Real Decreto-ley 9/2009, de 26 de junio, sobre reestructuración bancaria y reforzamiento de los recursos propios de las entidades de crédito. Disponible en <http://www.boe.es/boe/dias/2009/06/27/pdfs/BOE-A-2009-10575.pdf>.
- Real Decreto-ley 11/2010, de 9 de julio, de órganos de gobierno y otros aspectos del régimen jurídico de las Cajas de Ahorros. Disponible en <https://www.boe.es/boe/dias/2010/07/13/pdfs/BOE-A-2010-11086.pdf>.
- Real Decreto-ley 6/2010, de 9 de abril, de medidas para el impulso de la recuperación económica y el empleo. Disponible en <https://www.boe.es/boe/dias/2010/04/13/pdfs/BOE-A-2010-5879.pdf>.

- Real Decreto–ley 2/2011, de 18 de febrero, para el reforzamiento financiero. Disponible en <http://www.boe.es/boe/dias/2011/02/19/pdfs/BOE-A-2011-3254.pdf>.
- Real Decreto–ley 16/2011, de 14 de octubre, por el que se crea el Fondo de Garantía de Depósitos de Entidades de Crédito. Disponible en www.boe.es/boe/dias/2011/10/15/pdfs/BOE-A-2011-16173.pdf.
- Real Decreto–ley 2/2012, de 3 de febrero, de saneamiento del sector financiero. Disponible en www.boe.es/boe/dias/2012/02/04/pdfs/BOE-A-2012-1674.pdf.
- Real Decreto–ley 24/2012, de 31 de agosto, de reestructuración y resolución de entidades de crédito. Disponible en <http://www.boe.es/boe/dias/2012/08/31/pdfs/BOE-A-2012-11247.pdf>.
- FROB Fondo de Reestructuración Ordenada Bancaria (2012): *Reestructuración y Recapitalización del Sector Bancario: la Sociedad de Gestión de Activos (Sareb)*. Disponible en http://www.frob.es/es/Lists/Contenidos/Attachments/342/20121029_Presentacion_FROB_SGA_es_%20prot.pdf.
- Banco de España (2012): *Resultados de la evaluación independiente del sector bancario español*. Disponible en http://www.bde.es/f/webbde/SSICOM/20120928presentacion2809_12.pdf.
- Sutton, C. J. (1980): *Economics and Corporate Strategy*. Cambridge University Press, Nueva York.
- Thompson, S. (1997): Takeover Activity Among Financial Mutuals: An Analysis of Target Characteristics. *Journal of Banking and Finance*, 21:1, pp. 37–53.
- Vander Venet, R. (1999): Causes and Consequences of EU Bank Takeovers. En Eijffinger, S.; Koedijk, K.; Pagano, M. y Portes, R. (eds.): *The Changing European Financial Landscape*, London: CEPR, pp. 45-61.
- Vives, X. (1988): Concentración bancaria y competitividad. *Papeles de Economía Española*, 36, pp. 62–75.

- Wheelock, D. C. y Wilson, P. W. (2001): New Evidence on Returns to Scale and Product Mix Among US Commercial Banks. *Journal of Monetary Economics*, 47:3, pp. 653–674.
- Worthington, A. C. (2004): Determinants of Merger and Acquisition Activity in Australian Cooperative Deposit– Taking Institutions. *Journal of Business Research*, 57:1, pp. 47–57.



La técnica de escalamiento lineal por intervalos: una propuesta de estandarización aplicada a la medición de niveles de bienestar social

ACTIS DI PASQUALE, EUGENIO

Grupo Estudios del Trabajo, Facultad de Ciencias Económicas y Sociales

Universidad Nacional de Mar del Plata (Argentina)

Correo electrónico: edipasq@mdp.edu.ar

BALSA, JAVIER

Centro de Investigaciones sobre Economía y Sociedad en la Argentina Contemporánea

Universidad Nacional de Quilmes (Argentina)

Consejo Nacional de Investigaciones Científicas y Técnicas (Argentina)

Correo electrónico: jjbalsa@unq.edu.ar

RESUMEN

En el presente estudio realizamos una propuesta de estandarización de indicadores basada en la técnica de escalamiento lineal (LST por sus siglas en inglés) que denominamos “LST por intervalos”. Los métodos tradicionales poseen ciertas limitaciones diferenciales en cuanto a invariancia, robustez, posibilidad de analizar series temporales, pérdida de la variabilidad original de los datos y capacidad de generar escalas ordinales que representen niveles —bajo, medio, alto— del fenómeno social estudiado. La LST por intervalos consigue superar esas limitaciones a través de dos etapas: 1) la fijación de niveles de correspondencia intermedios a través de un consenso intersubjetivo entre especialistas, y 2) la estandarización lineal entre los valores de referencia.

El método propuesto representa un doble avance. Por un lado, permite construir índices parciales (correspondientes a cada dimensión de un concepto) que satisfacen todas las condiciones deseadas. Por otro lado, generar una escala ordinal equiparable entre los distintos componentes que están listos para ser agregados por algún método particular que no elimine esa condición.

Para ilustrar estas ventajas se presenta un ejercicio de transformación aplicado a tres indicadores de bienestar social con datos de Argentina. No obstante, el método puede ser aplicado para la estandarización de otros conceptos teóricos.

Palabras claves: estandarización; índices sintéticos; bienestar social.

Clasificación JEL: C10; C43.

MSC2010: 91B82; 91B15; 62P10.

Artículo recibido el 14 de junio de 2015 y aceptado el 22 de mayo de 2017.

Interval Linear Scaling Technique: Proposal for Standardization Applied to the Measurement of Social Well-Being Levels

ABSTRACT

In the present study, we propose a standardization of indicators based on the Linear Scaling Technique (LST) that is called “interval LST”. Traditional methods have certain limitations in terms of robustness, capability to analyze time series, loss of original variability for data and capability to generate ordinal scales for representing levels —low, medium, and high— of the social phenomenon to be studied. This technique manages to overcome these limitations through two stages: 1) establishing intermediate levels of correspondence through an inter-subjective consensus among specialists, and 2) the linear standardization among the reference values.

The proposed method represents a double advance. On the one hand, it allows us to build partial indices (one for each dimension of a concept) that meet all the required conditions. On the other hand, the method can generate a comparable ordinal scale between the various components that are available to be added by any other particular method that preserves such a condition.

To illustrate these advantages, we present an exercise of modification which is applied to three social well-being indicators using data from Argentina. However, the method can be applied to the standardization of other theoretical concepts.

Keywords: Standardization; Synthetic Indices; Social Well-Being.

JEL classification: C10; C43.

MSC2010: 91B82; 91B15; 62P10.



1. INTRODUCCION

El proceso de operacionalización en Ciencias Sociales implica la descomposición del concepto teórico hasta llegar a indicadores empíricos. Es decir, cuando partimos desde un enfoque metodológico-conceptual para analizar la realidad compleja, se comienza por elaborar un constructo teórico. Luego, se conforman las dimensiones que lo componen y en cada una de éstas, los indicadores que dan cuenta de la extensión que alcanza esa dimensión. En este sentido, al final de este proceso, en la medida que continuamos pensando la realidad en términos de un concepto, y no como una serie de dimensiones autónomas, vamos a requerir algún tipo de operación que resuma ese conjunto de indicadores en una única medida: el índice sintético.

No obstante, la elaboración de índices sintéticos no es sencilla, ya que nos enfrentamos a una serie de problemas que intentan ser resueltos en cada una de las siguientes etapas: 1) selección de las dimensiones y variables; 2) estandarización; 3) ponderación; y 4) agregación. La primera etapa dependerá de la problemática específica de cada estudio y, por ende, no la abordaremos en el presente trabajo.

La segunda etapa busca homogeneizar un conjunto de indicadores con unidades de medida heterogéneas y de distinta naturaleza para que puedan ser agregados con algún método específico. A diferencia de las mediciones estrictamente económicas (como el cálculo del producto nacional en el que todos sus componentes se miden en unidades monetarias), los indicadores sociales se pueden medir en años, personas por metros cuadrados, distancias entre la vivienda y el centro prestador de servicios, porcentajes o tasas de ocurrencia de un determinado fenómeno o característica del hogar o de los individuos por cada 10.000 o 100.000 habitantes y también, en unidades monetarias destinadas a determinados servicios, entre otros. Esta complejidad de cómo agregar variables medidas en formas tan disímiles, intenta ser resuelto a través de la estandarización. El meollo estaría dado en la elección de una técnica que resulte válida para el objetivo de medición, dado que los resultados obtenidos variarán con cada tipo de procedimiento. En este sentido, se puede preferir una técnica que reúna ciertas bondades o condiciones deseadas.

Sin embargo, la propia selección de método de estandarización va a generar otro inconveniente: el de la ponderación implícita. Dado que el resultado obtenido dependerá de la técnica utilizada, aquella que sobreestime ciertos parámetros estadísticos de la distribución del indicador entre las unidades de análisis, le estará asignando más peso implícitamente en la medida final (por ejemplo, considérense dos indicadores de igual rango entre sus unidades de análisis que se estandaricen a través de la inversa de la desviación estándar σ , si σ es más bajo en el indicador 1 que en el 2, entonces al estandarizar el índice 1 será mayor en una gran cantidad de unidades de análisis; al agregarlos en la medida final, se pueden obtener resultados

con sesgos indeseados). Por este motivo, resulta perentorio analizar las ventajas y limitaciones de cada técnica de estandarización.

La tercera y la cuarta etapa se pueden efectuar en forma conjunta o separada, dependiendo de si el método utilizado las integra en un solo paso o no. Respecto a la ponderación (explícita), se debe determinar el peso de cada índice parcial en la medida sintética que se intenta elaborar. Independientemente que las unidades de medida sean iguales y las variaciones sean proporcionales, algunas dimensiones pueden tener más incidencia en el valor global que otras, ya sea porque así lo dictamina la teoría, o las estimaciones a través de métodos estadísticos, o bien por una asignación subjetiva basada en la validez o calidad de los datos. Así, por ejemplo, podría considerarse en base a alguno de esos criterios que una tasa de analfabetismo del 4% tendría que tener más incidencia sobre un índice sintético de bienestar social que una tasa de no acceso a cloacas del 4%.

En cuanto a la agregación, se busca conjugar a los índices parciales en un índice sintético a través de algún procedimiento estadístico, matemático o lógico. Resulta deseable que el valor que arroje la medida final represente niveles factibles de ser interpretados en una escala ordinal – alto, medio, bajo o nulo –. Esto último dependerá de si la técnica de estandarización utilizada cumplió con la condición de conseguir escalas equiparables entre los diferentes índices parciales. Entonces, estamos en presencia del problema de encontrar equivalencias en los valores de diferentes dimensiones, todo lo cual requiere de un trabajo de reflexión conceptual y, a la vez, técnico que construya la forma teóricamente más adecuada de resolver esta cuestión.

Estos tipos de razonamientos son los que, lamentablemente, muchas veces se obvian en los procesos de estandarización y construcción de índices sintéticos y le restan validez teórica a estas operaciones. En este sentido, si bien mantendremos la denominación de “estandarización de los indicadores”, no compartimos cierto halo semántico tecnicista que tiende a asociarse con esta noción, ya que esta etapa incluye la toma de una serie de decisiones de tipo teórico que, explícita o implícitamente, se realizan al escoger entre una u otra técnica de estandarización.

En este trabajo proponemos un proceso de estandarización basado en una reformulación de la técnica denominada técnica de escalamiento lineal (en inglés *Linear Scaling Technique*, con las siglas *LST*) con el fin de obtener índices sintéticos que brinden (sin perder, en cierto sentido, la medición de intervalo) una categorización ordinal que represente niveles equiparables entre los distintos indicadores que componen la operacionalización de un concepto.

Para ello, en la segunda sección, basándonos en cinco condiciones deseables, analizamos las características propias de cada una de las técnicas (incluyendo la *LST* estándar), de modo que se pueda explicitar cuál nos parece el procedimiento más adecuado. Luego, en la tercera sección, presentamos una propuesta original que permite satisfacer el objetivo enunciado y que denominamos “*LST* por intervalos”. Para ilustrar las ventajas que posee, la aplicamos a

tres indicadores de bienestar social que permiten generar la escala de niveles equiparables ya mencionada.

Consideramos que la técnica propuesta representa un doble avance, dado que no sólo permite construir índices parciales que satisfacen todas las condiciones deseadas, sino también porque consigue generar una escala ordinal equiparable entre los distintos componentes que están listos para ser agregados por algún método particular que no elimine esa condición. En este sentido, las bondades de la estandarización conseguida se verán reflejadas en la agregación y, por ende, en el índice sintético.

2. ANALISIS COMPARATIVO DE LAS TÉCNICAS DE ESTANDARIZACIÓN

2. 1. Las propiedades deseables

Para poder analizar las ventajas y limitaciones de cada técnica de estandarización, tendremos en cuenta cuatro propiedades que la literatura especializada ha señalado como claves para poder medir la validez de estas técnicas: la robustez, el cumplimiento de la propiedad de invariancia, la capacidad de permitir el análisis de series temporales y el grado de conservación de la variabilidad original. Asimismo, añadimos una quinta propiedad que para nosotros es significativa debido a que metodológicamente facilita la interpretación y valoración conceptual de los resultados luego de la agregación: la posibilidad de atender a la categorización ordinal de los valores de los índices parciales. Dado que todas las técnicas presentan una cierta simplicidad en su cálculo, no consideramos al nivel de complejidad como una condición relevante para discriminar procedimientos, aunque sí destacamos cuáles son los más sencillos. El significado de cada una de estas características se detalla a continuación:

- a) La robustez se refiere a que la estandarización se vea poco afectada por las observaciones atípicas. No obstante, si la técnica de estandarización empleada no logra despojarse del efecto de los valores atípicos, se pueden efectuar transformaciones que se aplican a menudo a los datos antes de la estandarización. Una de ellas es el uso de valores ordinales (con la consiguiente pérdida de toda información de la medición por intervalos) y otra es el truncamiento, en el cual se recortan las colas de las distribuciones de los indicadores con el objetivo de evitar que los valores extremos dominen demasiado el resultado y, en parte, para corregir problemas de calidad de datos. Sin embargo, el truncamiento no resulta del todo aconsejable pues está dejando de lado casos reales cuyo desconocimiento debería poder justificarse teóricamente. Otra es la transformación funcional, que se aplica a los datos en bruto para representar el significado de los cambios marginales (Molpeceres Abella, 2008: 16).

- b) La invariancia es una propiedad que considera que el indicador estandarizado debe ser invariante respecto a un cambio de origen y/o de escala en las unidades en que vengan expresados los valores de los componentes. Es decir, se puede suponer una transformación lineal del indicador X como $X' = a + bX$, donde a y b son dos parámetros – b positivo –. A partir de allí, se analiza si el indicador estandarizado sin transformación lineal es igual al indicador estandarizado con esa transformación (Zarzosa Espina, 1996).
- c) La posibilidad de análisis de series temporales implica que la técnica no sólo genere un *ranking* con valores que sirvan para analizar un momento determinado en el tiempo, sino que a su vez permita análisis diacrónicos. Una escala que no sea comparable a través del tiempo anula la viabilidad de medir la evolución de las distintas unidades de análisis.
- d) La variabilidad original intra- e inter-indicadores¹ se verá afectada a través del proceso de estandarización, y su magnitud dependerá no sólo de la fórmula que se utilice sino también de los parámetros que se fijen en la misma. Algunas técnicas realizan el cálculo a través de la inversa de la desviación estándar entre las unidades de análisis y, por ende, a mayor desviación estándar, menor será la variabilidad del índice parcial y viceversa. En cambio, otras técnicas ajustan al índice parcial en un rango de máximos y mínimos (teóricos o empíricos), con lo cual la variabilidad conseguida dependerá del tamaño de ese rango. Por lo tanto, la variabilidad original siempre se ve modificada, aunque se puede llegar a mantener parcialmente o bien perder totalmente. Si la magnitud del cambio es significativa, se generan dos problemas: uno de interpretación conceptual de ese valor numérico y otro al momento de la agregación, dado que se estaría agrupando variabilidad desconocida.
- e) Cuando proponemos la categorización ordinal de los valores de los índices parciales hacemos referencia a la posibilidad de representar niveles – alto, medio, bajo o nulo – de cualquier fenómeno social o concepto que se esté estudiando. Lo que se pretende conseguir es que un determinado valor, por ejemplo 0,60, represente en todos los indicadores estandarizados la misma situación (por ejemplo, un nivel medio-bajo). Para llegar a ello, la estandarización de los indicadores, en una escala entre 0 y 1 para cada dimensión del concepto, debería darse a través de puntos de correspondencia intermedios que implican tanto una discusión teórico-conceptual como también el empleo de técnicas específicas.

¹ La variabilidad intra-indicadores se refiere a la desviación estándar y el rango de un mismo indicador calculado con los valores de todas las unidades de análisis. En cambio, la variabilidad inter-indicadores alude a la comparación de esas medidas estadísticas entre dos o más indicadores que componen un concepto.

A partir de estas cinco cuestiones o propiedades deseables es que procederemos a comparar las principales técnicas de estandarización.

2. 2. Ventajas y desventajas de cada una de las técnicas

Con el fin de estudiar la bondad de distintos procedimientos, efectuamos un análisis comparativo de las seis técnicas más empleadas en la literatura y que son también las más representativas, dado que de éstas se derivan otras similares. Estas técnicas de estandarización son: el recuento de Borda o método ordinal, la estandarización a un valor de referencia o año base, el ratio o proporción respecto a la media, la tipificación, la distancia relativa a la desviación estándar σ y la técnica de escalamiento o transformación lineal (o LST por ser las siglas en inglés de *Linear Scaling Technique*); véase Booyesen (2002), Saisana y Tarántola (2002), Salzman (2003) y Schuschny y Soto (2009).

La notación general que utilizaremos es la siguiente:

- X_i : indicador X sin estandarizar del territorio geográfico i ;
- X_* : valor de referencia;
- X_t : indicador X en el momento t ;
- X_{t-1} : indicador X en el momento $t-1$;
- μ : media del indicador X entre todos los territorios geográficos analizados;
- σ : dispersión del indicador X entre todos los territorios geográficos analizados;
- I_i : índice parcial del territorio geográfico i (que es el indicador X estandarizado).

En los casos que el método posea una notación específica se la definirá para ese caso particular.

2.2.1. Recuento de Borda o método ordinal

Consiste en ordenar las observaciones conforme a cada componente o indicador simple y así clasificar en un *ranking* a las regiones o países. A partir de allí se le asigna una puntuación categórica que puede ser ordinal – como, por ejemplo: $n, n-1, n-2, \dots, n-(n-1)$ – o bien, que esté basada en los percentiles de la distribución del valor del indicador. Para los estudios de series temporales, la clasificación se lleva a cabo en cada punto en el tiempo y, por ende, solamente se puede seguir el desempeño del país en términos de la posición relativa.

Este método tiene varias ventajas, como la simplicidad, la independencia a los valores atípicos y la invariancia, ya que ante una transformación lineal de igual magnitud para todas las

unidades de análisis, el orden no se modifica y se cumple esa propiedad. Una limitación que reviste gravedad es la pérdida total de información de la variabilidad original de los indicadores y, por ello, no se pueden obtener conclusiones sobre el nivel alcanzado por cada unidad de análisis, ni tampoco la posibilidad de efectuar un seguimiento de la evolución a través del tiempo.

Por estas limitaciones, esta técnica de estandarización no es muy empleada para construir índices parciales con las características buscadas y, por ende, resulta poco recomendable.

2.2.2. Estandarización a un valor de referencia o año base

Este método divide el valor del indicador para una región o país dado en un punto determinado en el tiempo con un valor de referencia. En los casos en los que se pretende estudiar la evolución de las series temporales, ese valor puede ser el de una región o país en un año base. Una alternativa del método es que en lugar de utilizar la simple relación entre el valor del indicador y una referencia, primero se podría restar el valor de referencia desde el dato en bruto y luego dividir por ese mismo valor. Es decir:

$$I_i = \frac{X_t}{X_{t-1}} \text{ o bien, } I_i = \frac{X_t - X_{t-1}}{X_{t-1}}$$

$$I_i = \frac{X_t}{X_*}, \text{ o bien: } I_i = \frac{X_t - X_*}{X_*}$$

El método es sencillo en su cálculo y permite que la variabilidad original se reduzca a términos proporcionales respecto al valor de referencia, con lo cual se pierde parcialmente. No obstante, es poco robusto cuando hay valores atípicos dado que no logra suavizarlos; tampoco permite comparar niveles, solo tasas de crecimiento y no cumple con la propiedad de invariancia:

$$\frac{X_t}{X_*} \neq \frac{a + bX_t}{a + bX_*}$$

$$\frac{X_t - X_{t-1}}{X_{t-1}} \neq \frac{b(X_t - X_{t-1})}{a + bX_{t-1}}$$

En definitiva, no resulta recomendable para el objetivo propuesto, dado que sólo cumple con la propiedad de analizar series temporales

2.2.3. *Ratio o proporción respecto a la media*

Este método se calcula como una proporción respecto a la media para cada indicador:

$$I_i = \frac{X_i}{\mu}$$

Como el método anterior, presenta la ventaja de simplicidad. Sin embargo, es menos robusto cuando hay valores atípicos dado que la media es un estadístico muy sensible a la presencia de datos extremos. Por otra parte, la variabilidad original se pierde dado que el índice parcial resulta una proporción respecto al valor de la media del indicador original. En virtud de ello, las medias de los índices parciales convergen a 1. Asimismo, los valores obtenidos no se pueden categorizar en una escala ordinal equiparable. La posibilidad de su utilización en series temporales estaría acotada al empleo de una media de un año base, lo que distorsiona el objetivo del método. Por otra parte, no cumple con la propiedad de invariancia:

$$\frac{X_i}{\mu} \neq \frac{a + bX_i}{a + b\mu}$$

En este sentido, dado que el método no satisface ninguna condición no resulta recomendable para calcular índices con las características buscadas.

2.2.4. *Tipificación, estandarización normal o puntuación Z*

Ésta es una de las técnicas de estandarización más difundidas, hasta el punto de comprenderse casi como “la” estandarización². En particular, es uno de los métodos más utilizados en estudios de bienestar social (Diener, 1995; Estes, 2003; Distaso, 2007) y se calcula como el cociente de la diferencia entre el valor del indicador X y la media μ dividida por la desviación estándar σ :

$$Z_i = \frac{X_i - \mu}{\sigma}$$

Esto lo convierte en una escala común con media de cero y desviación estándar de uno, lo que posee el significado de factor de escala.

La ventaja principal de la normalización en relación a los métodos anteriores es que resulta más robusto cuando se trata con valores atípicos (Saisana y Tarántola, 2002: 10-11). No obstante, la transformación continúa da un peso mayor a los países o regiones con valores extremos, lo cual podría ser una propiedad deseable sólo si se quiere premiar el comportamiento excepcional. Asimismo, cumple con la propiedad de invariancia:

² De hecho, es común considerar normalización como sinónimo de estandarización.

$$\frac{b(X_i - \mu)}{b\sigma} = \frac{X_i - \mu}{\sigma}$$

Las limitaciones principales que posee son dos: la falta de validez como medida que arroje una escala ordinal equiparable y la pérdida de la variabilidad original de los indicadores. Respecto a lo primero, el resultado no asegura que los valores más bajos de Z_i para cualquier indicador signifiquen niveles bajos de algún fenómeno. De hecho, la tipificación genera valores con signo negativo y positivo respecto a la diferencia con el promedio. Y la media no necesariamente tiene que ser un parámetro de un nivel estándar para delimitar los valores bajos y altos. Por ende, Z_i solamente indica la distancia a la media como una relación inversa a la desviación típica.

Esto último nos lleva a la segunda desventaja respecto a la pérdida de variabilidad original intra- e inter-indicadores, dado que se verá modificada respecto a la distancia con la media y en relación inversa al valor de la desviación estándar. Un ejemplo puede ser ilustrativo de esta limitación: supongamos dos indicadores X_A y X_B , pertenecientes a dos dimensiones de un fenómeno social que están expresadas en la misma unidad de medida; en este caso, porcentajes (para simplificar didácticamente el problema). Como vemos en el Cuadro 1, el indicador X_A tiene un rango mayor y una variabilidad más grande que el indicador X_B (con σ tomando valores 27,39 y 11,01 y rangos 80 y 37, respectivamente). Al aplicar la estandarización normal, los valores de la media y de la desviación σ de los índices parciales (Z_A y Z_B) convergen a 0 y 1, y el mayor rango pasa a ser el del indicador B estandarizado. De este modo, se perdió la variabilidad intra- e inter-indicadores que estaba representada por la mayor variabilidad de la dimensión A. Por otra parte, al estudiar los casos concretos, un mismo valor de 40% (que podría estar significando lo mismo si fueran dimensiones similares), en un indicador normalizado obtendría el valor de -0,37, mientras que en el otro sería de -1,18. Similares valores estandarizados estarían dando cuenta de valores originales muy diferentes (así, -1,10 en la dimensión A es la estandarización de 20%). Esto indica que, cuando la variabilidad inter-indicadores cambia significativamente, se generan limitaciones en la interpretación de los resultados y en su posterior agregación.

Cuadro 1 - Ejemplo de pérdida de variabilidad original intra- e inter-indicadores empleando la estandarización normal

Países	Indicadores sin normalizar		Indicadores normalizados	
	X_A	X_B	Z_A	Z_B
A	10%	50%	-1,46	-0,27
B	20%	45%	-1,10	-0,73
C	30%	50%	-0,73	-0,27
D	40%	55%	-0,37	0,18
E	50%	40%	0,00	-1,18
F	60%	60%	0,37	0,64
G	70%	44%	0,73	-0,82
H	80%	56%	1,10	0,27
I	90%	77%	1,46	2,18
σ	27,39	11,01	1,00	1,00
μ	50	53	0	0
Rango	80	37	2,92	3,36

Fuente: elaboración propia

Una tercera limitación se da al momento de estudiar las series temporales dado que modificaciones en la dispersión de datos entre países provocarían alteraciones en la estandarización, sin que los umbrales del concepto operacionalizado se hayan visto modificados. Una alternativa sería calcular la estandarización para cada año, utilizando los valores de la media y desviación estándar de un año de referencia. Nuevamente, dependerá de esos valores el resultado de la normalización.

Finalmente, la existencia de valores con signo negativo limita su uso en índices sintéticos de distancia, dado que una de las condiciones exigidas en el espacio métrico es la no negatividad.

En definitiva, la tipificación, aunque es la técnica más conocida de estandarización, no es recomendable para construir índices sintéticos de las características buscadas dado que solamente satisface la invariancia y robustez.

2.2.5. Distancia relativa a σ (D_i)

Esta técnica logra superar la presencia de valores negativos tras aplicar la estandarización al incorporar el valor absoluto en el numerador. Su cálculo se realiza a través de la siguiente fórmula:

$$D_i = \frac{|X_i - X_*|}{\sigma} = \frac{d_i}{\sigma}$$

donde el valor que se resta al indicador X_i es un valor de referencia u objetivo X_* (Ivanovic, 1973; Pena Trapero, 1977).

Esta estandarización, al igual que la anterior, cumple con la condición de invariancia:

$$\frac{b|X_i - X_*|}{b\sigma} = \frac{d_i}{\sigma}$$

Asimismo, también resulta robusto cuando se tratan valores atípicos. Sin embargo, al seguir utilizando σ en el denominador, no permite escapar de las limitaciones antes mencionadas respecto a la pérdida de la variabilidad original que genera esta relación inversa. Es decir, al darle menor importancia a las distancias correspondientes a los componentes cuyos valores presentan mayor σ , termina otorgando una ponderación implícita que puede llegar a ser indeseada en el resultado del índice sintético. Con lo cual σ no es un buen relativizador para efectuar mediciones que arrojen resultados equiparables a categorías ordinales.

Al usar el valor absoluto, no logra obtener una medida que represente una escala ordinal equiparable dado que su resultado indica la distancia al valor de referencia como una relación inversa a la desviación típica, lo cual no es posible interpretar conceptualmente.

Finalmente, comparte con la tipificación las limitaciones al estudiar series temporales, con lo que la única alternativa sería calcular la estandarización para cada año, utilizando los valores de la desviación estándar de un año de referencia.

Por ende, al satisfacer las mismas condiciones que la tipificación, tampoco es recomendable para el cálculo de índices sintéticos que posean las características que escogimos al inicio. No obstante, ambas técnicas podrían ser coherentes con otra forma de configurar un concepto complejo.

2.2.6. Técnica de escalamiento lineal (LST)

Esta técnica fue definida por Drewnowski y Scott (1966) y es una de las más utilizadas en la construcción de numerosos índices sintéticos sociales y económicos (Morris, 1979; Zárate Martín, 1988; PNUD, 1990-2011; Velázquez y Gómez Lende, 2005; Velázquez, 2008). El cálculo se diferencia de los anteriores dado que utiliza los valores máximos (X_{max}) y mínimos (X_{min}) de los indicadores y el rango en lugar de la media y/o desviación estándar. Estos valores pueden ser empíricos, históricos o bien ideales, dependiendo del objetivo de la medición.

El índice parcial se calcula como el cociente de la diferencia entre el valor del indicador X_i y el valor mínimo X_{min} dividido por el rango absoluto ($X_{max} - X_{min}$)

$$I_i = \frac{(X_i - X_{min})}{(X_{max} - X_{min})}$$

En el caso de que exista una relación inversa entre el indicador y el constructo teórico – por ejemplo, la tasa de analfabetismo y el bienestar–, el cálculo será el siguiente, que incluye un cambio direccional:

$$I_i = \frac{(X_{\max} - X_i)}{(X_{\max} - X_{\min})}$$

En ambos casos, el resultado queda comprendido entre cero y uno, siendo cero la peor situación y uno la mejor. De esta manera, la estandarización se realiza en base a la distancia al valor mínimo –o máximo– en relación al recorrido total, sin importar la dispersión de los datos³. Por lo tanto, lo podemos considerar como una estandarización en términos de distancia respecto al valor objetivo de cada indicador

Entre sus ventajas, es robusto, permite el análisis de series temporales a través de la fijación de máximos y mínimos históricos o ideales, y cumple con la propiedad de invariancia:

$$\frac{b(X_i - X_{\min})}{b(X_{\max} - X_{\min})} = \frac{(X_i - X_{\min})}{(X_{\max} - X_{\min})}$$

Sin embargo, resulta parcialmente válida para establecer categorías ordinales que expresen niveles debido a estar supeditada a la fijación de los extremos de referencia sin considerar los valores intermedios. Es decir, que un valor de 0,70 solamente nos indica que la distancia respecto al mínimo representa el 70% del rango establecido y no podemos asegurar que represente una situación media o media-alta del fenómeno estudiado.

Asimismo, la variabilidad original se ve afectada parcialmente. Si bien la respectiva a una unidad de análisis no se pierde al modificarse la dispersión del total de los territorios estudiados, la variabilidad dependerá de la selección del rango de máximos y mínimos. En aquellos casos en los que los valores del indicador están situados en un rango pequeño de máximo y mínimo, la aplicación de LST aumenta el tamaño del rango y, por consiguiente, pequeñas variaciones en los datos originales se traducen en grandes variaciones en los datos estandarizados, lo que puede producir fuertes sesgos en el cálculo del índice sintético. Es decir, la variabilidad de los indicadores va a resultar sobrestimada en los casos que el indicador posea un rango pequeño para los territorios estudiados y lo contrario ocurre cuando el rango es muy amplio.

Un claro ejemplo de estos problemas fue el recálculo del Índice de Desarrollo Humano (IDH) para las provincias argentinas realizado en el informe del PNUD Argentina (2002). En

³En su versión original, en el cálculo interviene como multiplicando el factor $e = 1-k$, siendo k el coeficiente de concentración –deducido de la curva de Lorenz– que se introduce para tener en cuenta la desigualdad de la distribución.

líneas generales, los indicadores de las tres dimensiones que miden el IDH eran buenas en todas las jurisdicciones de la Argentina, por lo cual generaban poco poder discriminante entre las mismas –en un extremo se ubicaba Formosa con un valor de IDH de 0,745 y, en el otro, la Ciudad de Buenos Aires con 0,845–. Recuérdese que, en el cálculo de este índice, se emplearon como valores mínimos y máximos las realidades presentadas por los mejores y peores países para cada dimensión.

Por lo tanto, para reflejar mejor la disparidad de condiciones de vida en las provincias se incorporaron nuevos indicadores a cada una de las dimensiones (longevidad, conocimientos y estándar de vida), construyendo lo que se denominó el Índice de Desarrollo Humano Ampliado (IDHA)⁴. Pero, además, se recalcularon los índices parciales a partir de los valores mínimos y máximos que presentaban las provincias argentinas y no el conjunto mundial de países. De modo que el IDHA para Formosa fue de 0,156, y el de la Ciudad de Buenos Aires de 0,867. El problema es que estos resultados fueron incomparables con los valores internacionales e incluso podían dar lugar a fuertes equívocos interpretativos, ya que algunas provincias presentaban valores muy bajos que podían ser entendidos como equivalentes a algunos países que tenían un IDH mucho peor que estas provincias.

En virtud de las ventajas señaladas, si se definen valores adecuados de máximo y mínimo, la LST surge como el único método que satisface la mayor parte de las características deseadas: invariancia, robustez y análisis de series temporales sin mayores complicaciones. Sin embargo, la generación de categorías ordinales se realiza con posterioridad a la estandarización y, por ende, no necesariamente refleja niveles bajos, medios y altos de ese indicador (tal como se verá en el próximo ejemplo). La variabilidad original intra- e inter-indicadores se mantiene parcialmente, dado que tiene una relación inversa con el rango.

A modo de resumen, en el Cuadro 2 se recogen las principales características de cada tipo de estandarización. Por un lado, encontramos tres métodos muy sencillos pero que no resultan adecuados para nuestro objetivo. El recuento de Borda, al desentenderse de los valores absolutos (sólo queda un *ranking*) satisface la invariancia y robustez, pero, por ese mismo motivo, no cumple con el resto de las propiedades. La estandarización a un valor de referencia y la proporción respecto a la media son las técnicas con menos propiedades deseables, a tal punto que esta última no cumple con ninguna condición.

Por otro lado, las técnicas más utilizadas (la tipificación, la distancia relativa y la LST) poseen algunas ventajas dado que cumplen tanto con la invariancia como con la robustez y, con ciertas precauciones, permiten analizar series temporales. Sin embargo, la tipificación y la distancia relativa poseen limitaciones muy importantes debido a que en su cálculo interviene la desviación estándar. En rigor, se pierde totalmente la variabilidad original, a tal punto que

⁴ Véase PNUD Argentina (2002).

resulta en una convergencia inter-indicadores y el resultado no permite la generación de categorías ordinales equiparables que representen niveles de un fenómeno social estudiado. Por ende, la LST se posiciona como la técnica más recomendable dado que, si bien no se genera una escala ordinal equiparable, la variabilidad original se mantiene parcialmente.

Cuadro 2 - Comparación de métodos de estandarización

MÉTODO DE ESTANDARIZACIÓN	CARACTERÍSTICAS				
	Invariancia	Robustez	Análisis de series temporales	Variabilidad original	Categorías ordinales equiparables
1- Recuento de Borda	SI	SI. Independencia a los valores atípicos	NO, dado que permite estudiar sólo cambios en el <i>ranking</i> .	Se pierde totalmente.	NO. Pérdida de información en los niveles absolutos.
2- Estandarización a un valor de referencia o año base	NO	Poco robusto	SI. Sólo sirve para analizar series temporales	Se pierde parcialmente.	NO
3-Ratio o proporción respecto a la media	NO	Poco robusto	NO	Se pierde parcialmente.	NO
4- Tipificación o estandarización normal	SI	SI	Solamente si se considera μ y σ de un año base.	Se pierde totalmente. Converge inter-indicadores.	NO. Debido a la relatividad respecto a σ .
5- Distancia relativa	SI	SI	Solamente si se considera σ de un año base.	Se pierde totalmente. Converge inter-indicadores.	NO. Debido a la relatividad respecto a σ .
6- Técnica de Escalamiento Lineal (LST)	SI	SI	SI (con valores max y min históricos o ideales).	Se mantiene parcialmente.	NO.

Fuente: elaboración propia

A continuación presentamos un ejemplo numérico con el objeto de ilustrar las diferencias de resultados, sus implicancias y limitaciones, que se presentan al estandarizar un mismo indicador en dos momentos de tiempo con las tres técnicas más utilizadas: la tipificación, la distancia relativa y la LST.

2. 3. Un ejemplo ilustrativo

Supongamos dos indicadores $-X_A$ y X_B - pertenecientes a dos dimensiones de un fenómeno social que están expresadas en la misma unidad de medida; en este caso, porcentajes y donde el valor 10 representa niveles de malestar (X_{min}) y el valor 90 niveles de bienestar (X_{max}), que pueden estar teórica y/o empíricamente definidos. En ambos indicadores, si bien la media es la misma ($\mu_A = \mu_B = 50$), algunos países presentan situaciones diferentes, a tal punto que la desviación estándar es: $\sigma_A = 27,39$ y $\sigma_B = 20,98$; véase Cuadro 3.

Cuadro 3 - Ejemplo de dos indicadores de bienestar

Países	Indicador A	Indicador B
A	10	10
B	20	45
C	30	50
D	40	40
E	50	55
F	60	44
G	70	60
H	80	56
I	90	90
□	27,39	20,98
□	50	50

Fuente: elaboración propia

El país D es uno de los que tiene el mismo valor en ambos indicadores: 40. Si los parámetros de bienestar/malestar son idénticos, uno esperaría que los valores de los índices parciales para esa unidad de análisis sean los mismos.

Ahora comparemos los resultados de estandarizar el valor 40 en ambos indicadores con las tres técnicas que resultaron superadoras en nuestro análisis crítico:

Cuadro 4 - Diferencias al aplicar tipificación Z, distancia relativa o LST al País D.

Indicador	Método de estandarización		
	Tipificación Z	Distancia relativa*	LST
A	$Z_t = \frac{40-50}{27,39} = -0,365$	$D_t = \frac{ 40-50 }{27,39} = 1,095$	$I_t = \frac{40-10}{90-10} = 0,375$
B	$Z_t = \frac{40-50}{20,98} = -0,477$	$D_t = \frac{ 40-50 }{20,98} = 1,430$	$I_t = \frac{40-10}{90-10} = 0,375$

Fuente: elaboración propia.

Nota: * se consideró un valor de referencia X_* de 50, que coincide con la media, para que sea comparable con la tipificación Z.

Con la tipificación Z, el valor es menor en B dado que, al relativizar un valor negativo con una desviación estándar menor, este se hace aún más negativo. Con la distancia relativa, la estandarización arroja un valor mayor ya que el resultado es inversamente proporcional a la varianza. En cambio, mediante la aplicación de la LST, el valor estandarizado es el mismo, aunque la desviación estándar de cada indicador sea diferente. En este ejemplo se puede apreciar cómo la estandarización cumple con la función de ponderación implícita de los componentes.

En definitiva, la LST es el que más se ajustaría al objetivo propuesto de conseguir un índice parcial que arroje una categorización ordinal válida dado que, si una determinada unidad de análisis posee dos indicadores (pertenecientes a distintas dimensiones de un mismo concepto)

con el mismo valor (40), los índices parciales resultantes de esa estandarización serán los mismos (0,375), sin importar la variabilidad del resto de las unidades de análisis.

Entonces el éxito de la técnica depende de la fijación de esos valores extremos, lo cual significa un problema. Si en el ejemplo se hubieran alterado, la ventaja se hubiera disipado. Por lo general, estos valores pueden mantenerse fijos al asociarlos a máximos y mínimos teóricos, históricos o ideales.

Sin embargo, habría un segundo problema ya que intuitivamente es común asociar los valores de la estandarización a una escala ordinal y lo cierto es que no necesariamente tiene que haber una correspondencia entre el resultado y el significado del mismo. En el ejemplo, el valor 0,375 podría no representar realmente niveles bajos del fenómeno estudiado. Lo que indica es que la distancia respecto al mínimo representa el 37,5% del rango establecido.

Por ende, el desafío estaría dado en elaborar una transformación que permita soslayar las debilidades de la LST. En rigor, conseguir medidas equiparables que permitan la ordenación de niveles y controlar la pérdida de variabilidad original del indicador. En estos términos se enmarca nuestra propuesta de la siguiente sección.

3. PROPUESTA DE ESTANDARIZACIÓN: LST POR INTERVALOS

Para superar la limitación principal de la LST proponemos una transformación lineal pero por intervalos, con lo que se reducirían los sesgos indeseados que pudiera generar la selección de máximos y mínimos. Es decir, esta concepción va un paso más allá que la desarrollada para el *United Nations Research Institute for Social Development* (UNRISD) por Drewnowski y Scott (1966) y que considera solamente los niveles máximos y mínimos que pueden ser representativos de valores extremos ideales, mundiales, nacionales o bien del rango correspondiente a los datos analizados. Nuestra propuesta la denominamos “LST por intervalos” y, a través de su aplicación, se logran soslayar las limitaciones enunciadas ya que la transformación se realiza entre valores de correspondencia intermedios y no entre los extremos. Para ello, hay que proceder en dos etapas:

1. Conceptualización y fijación de valores de referencia;
2. Cálculo de los valores intermedios al interior de los intervalos.

En la primera etapa se debe elaborar conceptualmente una escala de conversión con límites que oscilan entre 0 y 1. Los valores de referencia a fijar pueden ser los siguientes: a) 0 (cero) que implica la peor situación; b) 0,40 que representa un nivel bajo (tal vez un “mínimo aceptable”); c) 0,70 que indica un nivel medio (tal vez calificable como una “situación buena”); y d) 1 constituye un nivel alto, donde pueden convivir una diversidad de situaciones. Al conceptualizar de esta manera el valor más alto del indicador, se consigue eliminar el sesgo que

provocaría un valor atípico. Por ejemplo, se pueden fijar como 1 los niveles de ingresos monetarios por encima de un determinado valor que ya es considerado como saturador del nivel de bienestar⁵.

La asimilación de un determinado valor del indicador sin estandarizar a un valor de referencia debe ser a través de un proceso de revisión de estudios previos específicos a cada componente, el análisis de estándares y las recomendaciones de organismos internacionales, la exploración de legislación nacional (en los casos que corresponda) y, fundamentalmente, la participación de especialistas o informantes cualificados. De esta manera, a través de un *consenso intersubjetivo*, se obtiene una medida válida desde el punto de vista de las “voces cualificadas”. Esta es la validez de expertos, que se consideró por muchos años como parte de la validez interna, aunque ya se concibe como un tipo de evidencia distinta (Bostwick y Kyte, 2005).

En la segunda etapa, luego de establecer los valores de referencia, se calculan los valores intermedios a través de una modificación del método LST. La fórmula utilizada difiere según estemos en presencia de indicadores positivos o negativos. En términos generales, para los primeros, la fórmula utilizada para cada intervalo de referencia es:

$$I_{\text{interm}} = \left(\frac{X_i - \text{MIN}_j}{\text{MAX}_j - \text{MIN}_j} \times \text{rango}_j \right) + \text{min}_j$$

mientras que para los indicadores negativos, la fórmula resulta ser:

$$I_{\text{interm}} = \left(\frac{\text{MAX}_j - X_i}{\text{MAX}_j - \text{MIN}_j} \times \text{rango}_j \right) + \text{min}_j$$

donde:

- I_{interm} : es el valor intermedio del índice parcial;
- X_i : es el valor del indicador sin estandarizar del territorio geográfico i ;
- MAX_j : límite superior del intervalo j de ese indicador sin estandarizar;
- MIN_j : límite inferior del intervalo j de ese indicador sin estandarizar;
- rango_j : es la diferencia entre max y min del intervalo j de valores de referencia;
- min_j : límite inferior del intervalo j de los valores de referencia;
- max_j : límite superior del intervalo j de los valores de referencia.

⁵ Esta idea de colocar un nivel de saturación a partir de un valor de un componente del bienestar, podría conceptualizarse en términos de satisfacción de niveles de “buen vivir” (Acosta, 2010; Kowii Maldonado, 2011) como un estado de consumo compatible con ciertos equilibrios ecológicos y de igualdad social, diferenciados de metas de consumo que podrían incrementarse a niveles incompatibles con estos objetivos.

Para ilustrar las ventajas del LST por intervalos, respecto a sus implicancias en la variabilidad original de los datos y la medición de niveles ordinales, presentamos a continuación tres casos de índices parciales de bienestar social elaborados con esta propuesta metodológica. En todos ellos vamos a considerar la fijación de los valores de referencia: 0 (cero) la peor situación; 0,40 como límite entre el nivel bajo y medio-bajo; 0,70 como límite del nivel medio-bajo y medio-alto; y 1 nivel alto. En los casos que amerite, se fijaran valores de referencia intermedios a éstos. Los datos utilizados corresponden a la Argentina para el año 2005.

- A) *Índice de ingreso relativo a la canasta básica.* Este es un indicador que mide indirectamente el nivel de bienestar. A diferencia de los indicadores de pobreza de ingresos, tiene en cuenta los niveles de ingreso de toda la población. Se calcula dividiendo el ingreso total familiar por el valor de la canasta básica ajustado por la escala de equivalencias. El resultado representa cantidades de canastas que pueden comprar en ese hogar de acuerdo a las necesidades de cada edad y sexo. A partir de esto, se fijan las escalas intermedias mencionadas al valor de cada persona que habita en ese hogar, lo que permite efectuar una ordenación correcta de los valores de los tramos asociados a la medición de niveles de bienestar. Para su cálculo es necesario una base de microdatos.
- B) *Índice de personas por dormitorio.* Este índice considera la situación habitacional de todas las personas, no sólo aquellas que se encuentren en situación de hacinamiento. Por lo tanto, no se corre el riesgo de dejar afuera a aquellos que se encuentran muy cerca del umbral definido. Nuevamente, para su cálculo es necesario una base de microdatos.
- C) *Índice de homicidios culposos en accidentes de tránsito.* Este indicador tiene dos particularidades. Por un lado, da cuenta de la peor situación que puede ocurrir como consecuencia de la inseguridad (dado que no incluye cifras sobre delitos a la integridad personal, robos o hurtos de bienes, entre otros), con lo cual es un indicador a nivel agregado que no permite caracterizar los distintos niveles de inseguridad del total de la población. Por otro lado, y en relación a lo anterior, los valores de referencia para la estandarización por tramos corresponden a distintos niveles de homicidios culposos en accidentes de tránsito. En este tipo de problemáticas, que involucra sólo a las muertes, resulta más difícil encontrar consenso en cuáles son los niveles intermedios aceptables para dar cuenta de un gradiente de inseguridad.

3.1. El Índice de ingreso relativo a la canasta básica (INGREL)

Los valores de referencia se ajustan a la cantidad de canastas básicas alimentarias (CBA) o de canastas básicas totales (CBT) a las que puede acceder el hogar. La CBA fija el límite de la indigencia, en 0,20 puntos, y la CBT delimita la pobreza en 0,40 puntos. El resto de los valores de correspondencia está basado en estudios empíricos y conceptuales estadísticos.

Murmis y Feldman (1997: 57) aluden a que, en contextos económicos inflacionarios, hasta un valor de 1,50 CBT se incluyen segmentos próximos a la línea de pobreza "que se encuentran en una situación de particular incertidumbre en la que pueden caer más o menos rápidamente por debajo de la línea de pobreza".

Otro grupo de trabajos arriba a un valor similar, entendiendo que el límite debiera colocarse en torno a las 3,3 CAB (o 1,52 CBT aproximadamente) para que la satisfacción de las necesidades humanas logre una vida digna y sostenible. A ese punto lo denominan línea de dignidad; y lo reconocen no como óptimo, pero sí como intermedio y factible de alcanzar en pocos años⁶ (Schatan, 1999; Wautiez, 2000; Aedo y Larraín, 2002).

Tomando como base estas propuestas, podría considerarse que un valor de 2 CBT resulta atinado para el nivel aceptable de 0,70 dado que se encuentra a una distancia prudente con el valor que consideramos pobreza. Y por último, para mantener una brecha similar entre los valores de referencia, asignamos el máximo puntaje a partir de las 4 CBT, nivel que se encuentra en el percentil 87 del total de aglomerados⁷ y que replica el máximo valor de análisis de los estudios empíricos citados. Este valor daría cuenta de un nivel de ingresos que permite no solo la satisfacción de necesidades cubiertas por la CBT, sino también de participación, esparcimiento y demás actividades que hacen al pleno desarrollo de las personas como seres sociales. Asimismo, existen dos razones más que apoyan esta elección:

- Porque a partir de las 4 CBT comienza un cambio de pendiente en la serie de ingreso relativo. De esta manera, logramos reducir el impacto por peso adicional y la distancia con respecto a los mayores valores, que se ubican entre 11,3 y 84,47CBT en el último percentil⁸;
- Porque existe una cuasi-relación geométrica entre cada uno de los niveles de referencia, dado que el nivel 0,40 de 1 CBT es igual a 1 CBA x 2 aproximadamente –de acuerdo al coeficiente de Engel que varía en el período en 2,15 y 2,17–; el nivel de 0,70 es igual a 1 CBT x 2; y el nivel de 1 es igual al cálculo anterior por 2.

⁶ Las 3,3 CBA corresponden a un monto monetario equivalente al ingreso promedio de los veintiles 9 y 10 de la población de Chile de fines de la década de 1990; es decir, el de la población con ingresos medios. El nivel escogido se acerca bastante al que algunos países europeos utilizan para determinar la frontera entre pobreza y no pobreza. Allí se consideran pobres todos aquellos que tienen un ingreso inferior a la mitad del promedio nacional.

⁷ Según análisis de la base de microdatos de la EPH-INDEC para el IV trimestre de 2005.

⁸ Según datos del IV Trimestre de 2005 EPH-INDEC.

El cálculo de los intervalos entre cada uno de los valores de referencia lo realizamos a través de la LST por intervals.

Cuadro 5 – Índice de nivel de ingreso relativo (INGREL)

CONDICIÓN	CANTIDAD DE CANASTAS A LAS QUE ACCEDE EL HOGAR	INTERVALOS DE INGREL	NIVEL DE BIENESTAR
Indigente	Sin ingresos	0	Nulo
	Menos de 1 CBA	(0; 0,20)	
Pobre no indigente	1 CBA	0,20	Bajo
	Más de 1 CBA y menos de 1 CBT	(0,20; 0,40)	
No pobre	1 CBT	0,40	Medio-Bajo
	Más de 1 CBT y menos de 2 CBT	(0,40; 0,70)	
	2 CBT	0,70	Medio-Alto
	Más de 2 CBT y menos de 4 CBT	(0,70-1)	
4 o más CBT	1	Alto	

Fuente: elaboración propia

Este indicador resulta una alternativa superadora con respecto a las que se centran en una parte de la población –indicadores de pobreza de ingresos– y de aquellos indicadores que dan cuenta de la distribución de ingresos percibidos sin considerar alguna medida normativa de referencia. Es decir, la medida elaborada resulta válida ya que permite analizar la totalidad de la población por medio de un análisis percentílico y, de esta manera, se pueden ubicar los valores cercanos a la línea de pobreza sin necesidad de realizar cálculos alternativos.

Asimismo, en la literatura existe un debate acerca de la sensibilidad de los indicadores de ingreso monetario para medir el bienestar o la pobreza dado que, ante pequeñas alteraciones en la línea de pobreza, estos indicadores se ven modificados (Feres y Mancero, 2001). El INGREL cumple con una serie de propiedades relacionadas con la monotonicidad y transferencia a saber:

1- Si aumenta el valor de la CBT, las posiciones percentílicas de las personas no cambian, solamente la cantidad de personas en cada nivel –más personas en los grupos más bajos– y, por ende, del INGREL. Es decir: si aumenta la CBT, disminuye INGREL.

2- Una disminución del ingreso de cualquier persona –u hogar–, incluso si se produce solamente entre los que ya son indigentes o pobres, provoca una disminución del INGREL. Esto demuestra una total sensibilidad del indicador ante cambios del ingreso. La disminución de INGREL será mayor si la persona es pobre que no pobre debido a la transformación escalar que realizamos. Es decir, ante una variación de 1 CBT, INGREL varía según se indica en Cuadro 6.

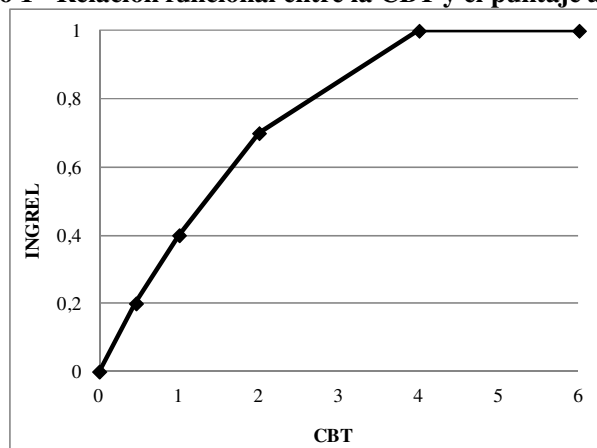
Cuadro 6 -Pendiente de la relación funcional INGREL

INTERVALOS de INGREL	Δ INGREL
[0; 0,20)	0,431
[0,20; 0,40)	0,371
[0,40; 0,70)	0,30
[0,70; 1]	0,15

Fuente: elaboración propia

Esto queda reflejado en la pendiente de la relación funcional entre la CBT y el índice obtenido, donde se observa el menor impacto por peso adicional de ingreso a medida que aumentamos de tramo (véase Gráfico 1). Cabe aclarar que esta pendiente no está relacionada con la utilidad, sino con la posibilidad de satisfacción de las necesidades fundamentales.

Gráfico 1 - Relación funcional entre la CBT y el puntaje asignado



Fuente: elaboración propia

3- Una transferencia de ingresos de una persona de mayores ingresos a una de menores ingresos, siempre que ambas estén en el mismo intervalo (indigentes, pobres o no pobres) y que la variación sea menor o igual a la diferencia entre ambos, no modifica el valor promedio del INGREL, pero sí disminuye la desigualdad de ingresos.

4- Ante una transferencia de ingresos de una persona no pobre a una pobre o indigente, el valor promedio del INGREL aumenta dado que la caída del indicador en el intervalo (0,40, 1) es menor al aumento en el intervalo por debajo de 0,40, tal y como se detalla en el Cuadro 6. Asimismo, mejora la distribución de ingresos.

Estas cuatro propiedades dan cuenta de la bondad del indicador elaborado para dar cuenta de los niveles de bienestar.

3.2. El Índice de personas por dormitorio (IPD)

Para definir los valores de referencia, tuvimos en cuenta tanto aspectos teóricos como empíricos. Los umbrales de hacinamiento crítico y semicrítico sirven para definir los niveles menores a 0,40. La mayor parte de los países, incluida Argentina, lo establecen en más de 3 personas por dormitorio⁹.

⁹Respecto a la noción de hacinamiento, es un concepto que varía según el nivel de desarrollo de las sociedades, el momento histórico y las particularidades culturales, lo cual implica que no existe un umbral universal de hacinamiento. Algunos autores lo establecen partir de las tres personas por habitación (UN-HABITAT, 2009: 9; Rodríguez, 1999: 225; Szalachman, 2000: 23) y se lo suele denominar como hacinamiento semicrítico; mientras que

Debido a que el tamaño del hogar y el número de dormitorios son variables relativamente acotadas, la definición del umbral para el nivel de hacinamiento puede afectar fuertemente los resultados. En nuestra propuesta, un hogar constituido por un matrimonio y dos hijos que cuente con dos dormitorios, sería una situación adecuada –2 personas por dormitorio–. Si luego llega un tercer hijo a ese hogar y se cuenta con la misma cantidad de habitaciones, las condiciones pasarían a estar en el umbral de lo aceptable –2,5 personas por dormitorio–. Un cuarto hijo con la misma infraestructura ubicaría a esta familia en el umbral de lo mínimo –3 personas por dormitorio–. Más allá de este valor, aparece el hacinamiento en sus diversas formas –semicrítico y crítico–.

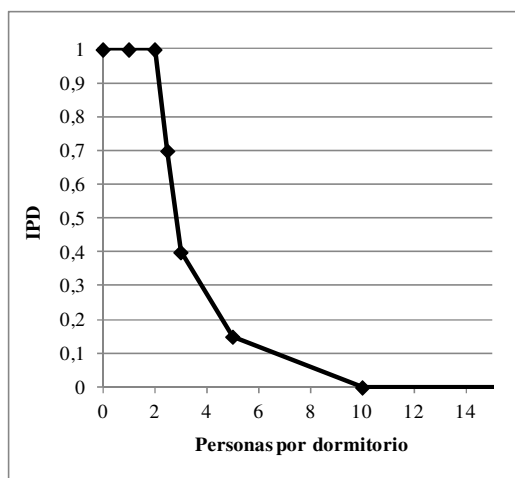
Cuadro 7- Índice de personas por dormitorio (IPD)

INTERVALOS DE PERSONAS POR DORMITORIO	INTERVALOS DEL IPD	NIVEL DE BIENESTAR
10 o más	0	Nulo
(5; 10)	(0; 0,15)	Bajo
5	0,15	
(3; 5)	(0,15;0,40)	
3	0,40	Medio-Bajo
(2,5;3)	(0,40; 0,70)	Medio-Alto
2,5	0,70	
(2; 2,5)	(0,70; 1)	Alto
2 o menos	1	

Fuente: elaboración propia

El cálculo para los niveles intermedios se realizó con la fórmula de la LST por intervalos, lo que arrojó una relación funcional como la representada en el Gráfico 2, muy sensible a la incorporación de un nuevo habitante en el primer intervalo –hasta 3 personas por cuarto–, relativamente sensible entre 3 y 5 personas por dormitorio y poco sensible a partir de las 5 personas por dormitorio.

Gráfico 2 - Relación funcional entre la cantidad de personas por dormitorio y el IPD



Fuente: elaboración propia.

otros consideran a más de tres personas por habitación como hacinamiento crítico (Lentini y Palero, 1997: 26). Sobre las diferencias de umbrales entre países, se recomienda visitar el sitio de la CEPAL: <http://www.eclac.cl/esalc/>

No obstante, asumida la sensibilidad ante pequeños cambios de los niveles establecidos, esta forma de elaboración del índice permite que no se pierda información cercana a los umbrales a diferencia del tradicional indicador porcentual que muestra la proporción de población por encima de cierto valor. En este sentido, el análisis percentílico permite observar la distribución de todos los niveles de personas por dormitorio.

3.3. El Índice de homicidios culposos en accidentes de tránsito (IHCAT)

Este caso es un ejemplo de un indicador en el cual resulta muy difícil encontrar un consenso sobre qué nivel de homicidios se corresponde con cada valor de referencia intermedio dado que, por ejemplo, estaríamos ponderando entre los expertos consultados hasta que nivel de homicidios en accidentes de tránsito resulta aceptable o adecuado para hablar de bienestar social medio.

En el nivel de bienestar alto, al valor de referencia 1 le asignamos una tasa de 0/100.000, que indica la ausencia de muertes. También se podría considerar como situación de bienestar alto la existencia de un número muy bajo de muertes por inseguridad. En cambio, en el nivel más bajo, el índice 0 (cero) se corresponde con los valores máximos mundiales de cada indicador. Los valores de 0,40 y 0,70 los fijamos de acuerdo a los intervalos que elabora la United Nations Office on Drugs and Crime (UNODC, 2011) para comparar las tasas de homicidios dolosos de los distintos países del mundo. Así por ejemplo, cuando el nivel de homicidios se encuentra entre 5 y 10 muertes por cada 100.000 habitantes de acuerdo a la escala de la UNODC, el nivel de bienestar medio-bajo (entre 0,40 y 0,70) en nuestra propuesta. De esta manera, asignamos los valores de referencia a los límites de cada una y luego aplicamos la fórmula de la LST por intervalos (véase Cuadro 8).

Cuadro 8 -Índice de homicidios culposos en accidentes de tránsito (IHCAT).

INTERVALOS DE LA THCAT (Muertes cada 100.000 habitantes)	INTERVALOS DEL IHCAT	NIVEL DE BIENESTAR
La THCAT más alta del mundo	0	Nulo
20 o más	(0; 0,15)	Bajo
20	0,15	
(10; 20)	(0,15; 0,40)	Medio-Bajo
10	0,40	
(5; 10)	(0,4; 0,70)	Medio-Alto
5	0,70	
(2; 5)	(0,70; 0,9)	
2	0,90	Alto
(0; 2)	(0,90; 1)	
0	1	

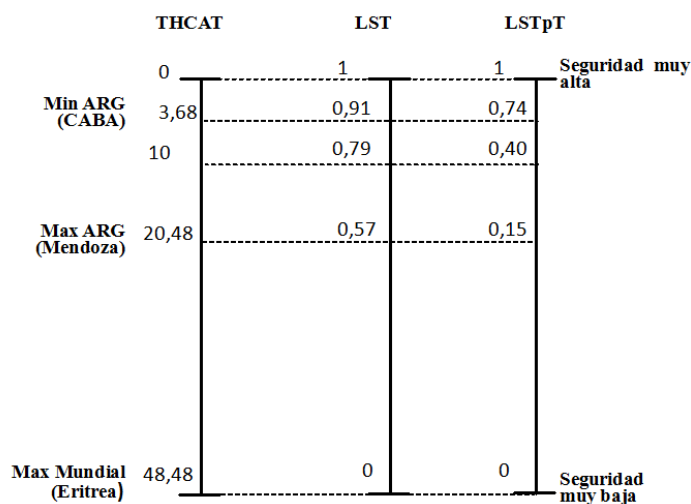
Fuente: elaboración propia en base a escala ordinal UNODC (2011).

Analicemos ahora la diferencia entre estandarizar la tasa de homicidios culposos en accidentes de tránsito (THCAT) a través de la LST y de la LST por intervalos para observar las

implicaciones en la variabilidad original de los datos y la medición de niveles ordinales. En Argentina en el año 2005, la THCAT tuvo un valor máximo para la provincia de Mendoza de 20,48 muertes por cada 100.000 habitantes y un valor mínimo de 3,68 para la Ciudad Autónoma de Buenos Aires¹⁰. En el *ranking* mundial de países, el peor valor es el de 48,4 que corresponde al país africano de Eritrea¹¹. Como mejor valor consideramos la inexistencia de muertes; es decir, una tasa de homicidios igual a cero. Veamos la diferencia entre la aplicación de la LST y nuestra propuesta de LST por intervalos:

- Si estandarizamos los valores de las provincias argentinas a través de la LST con el máximo y mínimo mundial, la peor situación quedaría en 0,57 y la mejor en 0,91 (véase LST en Esquema 1).
- En cambio, si asignamos los valores de referencia a cada intervalo de nivel de seguridad, entonces la aplicación de la LST por intervalos (véase LSTpT en Esquema 1) permite obtener un índice que da cuenta de los niveles de bienestar. De esta manera el peor valor de Argentina resulta un nivel muy bajo (0,15) y el mejor valor es alto (0,74).

Esquema 1– Comparación de estandarizar THCAT con LST y LST por intervalos



Fuente: elaboración propia en base a datos suministrados por la Dirección Nacional de Política Criminal del Ministerio de Justicia y Derechos Humanos

En este sentido, la variabilidad del índice obtenido se encuentra relacionada directamente con la noción de bienestar que representa y, de esta manera, se evitan las transformaciones lineales asociadas a valores extremos arbitrarios. La utilización de valores de referencia conceptualmente establecidos convierten a la tasa de homicidios en accidentes de

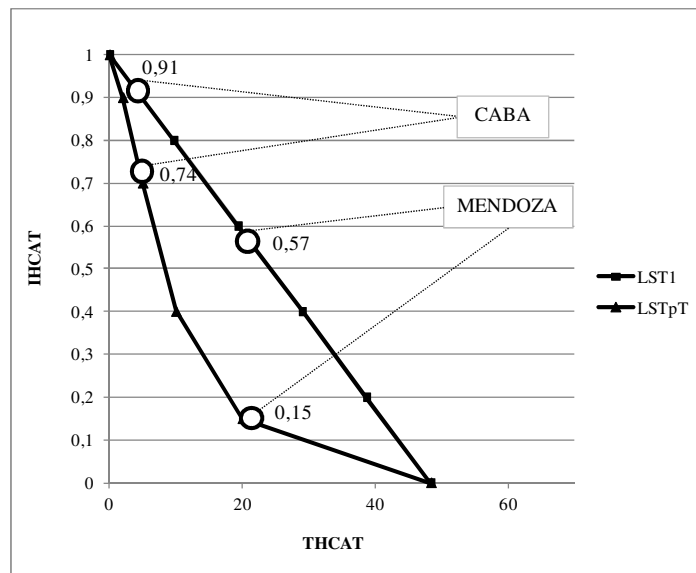
¹⁰ En base a datos suministrados por la Dirección Nacional de Política Criminal del Ministerio de Justicia y Derechos Humanos.

¹¹ En el año 2006, la tasa más alta fue de 48,4 (Eritrea-África) y la más baja 1,7 (Islas Marshall). Fuentes: THD: UNODC - Homicide statistics - Data series; THCAT: WHO - Global Health Observatory Data Repository.

tránsito de la Ciudad Autónoma de Buenos Aires de un valor que parecía casi ideal por su proximidad con el valor nulo (0,91 según la LST estándar) en un valor alto pero apenas por encima de lo que hemos fijado como “medio” (0,74 en la LST por intervalos). Al mismo tiempo, la tasa de Mendoza desciende de un valor que parecía aceptable (0,57), por estar muy lejos de lo que sucede en el peor país del mundo en accidentes de tránsito, a un valor claramente muy bajo (0,15) por encontrarse en una posición intermedia entre dicho país y el valor que hemos fijado como “bajo” (véase Gráfico 3).

También se puede apreciar que la relación se encuentra suavizada y la pendiente de la relación funcional es diferente de acuerdo al nivel de seguridad que se posea. En otras palabras, en los valores de seguridad más alto, un aumento porcentual de la tasa de homicidios se traduce en una disminución más que proporcional del índice obtenido. En cambio, en los niveles de seguridad más bajos, un aumento porcentual de la tasa de homicidios se traduce en una disminución menos que proporcional del índice estimado.

Gráfico 3 – Comparación de las relaciones funcionales entre el indicador original y el índice parcial a través de LST y LST por intervalos



Fuente: elaboración propia

3.4. Equiparación de los índices parciales

El análisis de cada índice por separado demuestra que la LST por intervalos logra superar las limitaciones de la LST estándar en cuanto a los cambios en la variabilidad original y la posibilidad de generar una escala ordinal que represente niveles, en este caso, de bienestar.

No obstante, las ventajas no terminan ahí dado que, si conceptualmente los valores de referencia (0; 0,4; 0,7; y 1) representan las mismas situaciones de malestar/bienestar, entonces la

validez de nuestra propuesta se extiende a la equiparación de las escalas de los tres índices, lo que permite agregarlos mediante algún procedimiento que no elimine esta condición de equiparabilidad y, por ende, el índice sintético permitirá ser interpretado en los mismos términos.

En el Cuadro 9 se presenta esta relación de correspondencia con las limitaciones que representa la tasa de homicidios culposos en accidentes de tránsito, que es el único indicador agregado que no surge de bases de microdatos sino de registros administrativos agregados. Así por ejemplo, un nivel de bienestar medio (0,70) incluiría que, en el territorio geográfico, la THCAT sea de 5/100.000, que las personas habiten en un hogar con 2,50 personas por dormitorio y con un nivel de ingresos correspondiente a 2 CBT.

Cuadro 9 – Equiparación de los índices parciales.

NIVEL de BIENESTAR	Intervalos de los Índices	Intervalos de ingreso relativo (CBT)	Intervalos de personas por dormitorio	Intervalos de la THCAT (Muertes cada 100.000 habitantes)
Nulo	0	Sin ingresos	10 o más	Máximo valor mundial
Bajo	(0; 0,40)	(0; 1)	(3; 10)	Más de 10 y menos del máximo mundial
Medio-Bajo	[0,40; 0,70)	[1; 2)	(2,50; 3]	(5; 10]
Medio-Alto	[0,70; 1)	[2; 4)	(2; 2,50]	(0; 5]
Alto	1	4 o más	2 o menos	0

Fuente: elaboración propia

Sin lugar a dudas, estos tres índices dan cuenta de tan solo una parte de las dimensiones del bienestar social dado que también habría que incluir medidas sobre salud, educación, seguridad ciudadana, trabajo y vivienda, entre otros. Por ende, la conclusión sobre la posible agregación de los índices parciales resulta incompleta. Sin embargo, en un estudio realizado con trece índices equiparables de bienestar social para Argentina, la agregación mediante la media aritmética ponderada arrojó resultados que pueden ser interpretados en términos de la misma escala ordinal de bienestar (Actis di Pasquale, 2013).

4. CONCLUSIONES

Como hemos visto el “escalamiento lineal por intervalos” permite transformar variables de distinta naturaleza en índices equiparables que representan niveles de algún fenómeno social estudiado. De esta manera, se logra combinar intervalos con niveles ordinales que poseen una interpretación conceptual. La validez quedó demostrada en el ejercicio de aplicación en tres casos de indicadores de bienestar social: a) el índice de ingreso relativo, basado en una variable

cuantitativa con información individual proveniente de bases de microdatos; b) el índice de personas por dormitorio, que cuenta con similares características; y c) el índice de homicidios culposos en accidentes de tránsito, calculado a partir de un indicador agregado que da cuenta de una parte de la situación estudiada (solamente las muertes). Por ende, es posible aplicar este método incluso con variables de límites teóricos mucho menos claros y con niveles intermedios muy arbitrarios, como es este último caso.

De esta manera, el LST por intervalos consigue un doble progreso. Por un lado, logra superar las limitaciones del LST estándar dado que la variabilidad del indicador cambia, pero no se pierde, al reasignarse a través de los puntos de correspondencia que se definan. De esta manera, se puede obtener una medición más precisa del objeto de estudio y, por ende, se reduce la brecha entre teoría y empiria.

Por otro lado, y como consecuencia de esa reasignación, se obtienen medidas equiparables que no sólo permiten la interpretación de niveles previamente acordados a través de un *consenso intersubjetivo*, sino que se encuentran en condiciones para ser agregados por algún método particular que no elimine esas propiedades.

Este alto grado de subjetividad, que permite la participación de especialistas en cada disciplina, es el ingrediente principal que lo diferencia de otro tipo de técnicas de mayor rigidez, como por ejemplo, la tipificación y la distancia relativa. Y ello no resulta en perjuicio de la medición, sino todo lo contrario, ya que permite despojarse del isomorfismo que caracteriza a los métodos más duros, que parten implícitamente de una correspondencia entre la estructura lógico-matemática y la estructura de la realidad.

5. BIBLIOGRAFIA

- Actis di Pasquale, E. (2013): *Bienestar Social, resignificación del concepto y de su operacionalización. Un aporte metodológico aplicado al caso argentino (1988-2005)*. Tesis doctoral. Universidad Nacional de Quilmes.
- Acosta, A. (2010): *El Buen Vivir en el camino del post-desarrollo. Una lectura desde la Constitución de Montecristi*, Policy Paper 9, Fundación Friedrich Ebert, Quito.
- Aedo, P. y Larraín, S. (2002): *Línea de dignidad: Desafíos sociales para la sustentabilidad*. Programa Cono Sur Sustentable, Santiago de Chile.
- Booyesen, F. (2002): “An overview and evaluation of composite indices of development”, *Social Indicators Research*, 59:2, pp. 115–151.
- Bostwick, G. J., y. Kyte, N. S (2005): *Measurement Social Work: Research and Evaluation Quantitative and Qualitative Approach*. 7ª ed. Oxford University Press, New York.

- Diener, E. (1995): “A value based index for measuring national quality of life”, *Social Indicators Research*, 36, pp 107–127.
- Distaso, A. (2007): “Well-being and/or quality of life in EU countries through a multidimensional index of sustainability”. *Ecological Economics*, 64:1, pp. 163–180.
- Drewnowski, J y Scott, (1966): *The level of living index*. Report N° 4. UNRISD. Ginebra.
- Estes, R.J. (2003): “European social development trends: development challenges of the New Europe”, en Voguel, J. (ed.): *Good times and hard times in Sweden during 1990's*, Living Conditions Series, Report 100. Statistics Sweden, Estocolmo, pp. 435–468.
- Feres, J.C. y Mancero, X. (2001): “Enfoques para la medición de la pobreza. Breve revisión de la literatura”. *Estudios Estadísticos y Prospectivos 4*. Naciones Unidas, Santiago de Chile.
- Ivanovic, B. (1973): *Comment établir une liste des indicateurs de développement*. Etude XXVI. UNESCO, París.
- Kowii Maldonado, W.A. (2011): “El Sumak Kawsay”. *Aportes Andinos*, 28, pp. 1–5.
- Lentini, M. y Palero, D. (1997): El hacinamiento: la dimensión no visible del déficit habitacional. *Revista invi*, 12:31, pp. 23–32.
- Molpeceres Abella, M. de las M. (2008): *Métodos de aproximación a la medición del bienestar: una panorámica*. Documento de trabajo para el Informe FOESSA 2008, España.
- Morris, D. (1979): "Measuring the Condition of the World's Poor: the Physical Quality of Life Index". *Pergamon Policies Studies*, 42. Pergamon, New. York.
- Pena Trapero, B. (1977): *Problemas de medición del bienestar y conceptos afines. Una aplicación al caso español*. INE, Madrid.
- PNUD (1990 a 2011): *Informe sobre Desarrollo Humano*. Programa de Naciones Unidas para el Desarrollo. Mundiprensa Libros, Madrid / Fondo de Cultura Económica, México.
- PNUD Argentina (2002): *Aportes para el Desarrollo Humano de la Argentina / 2002. Desigualdad y Pobreza*. PNUD, Buenos Aires.
- Rodríguez, J. (1999): “Información censal relevante para la medición del déficit habitacional”, en *América Latina: aspectos conceptuales de los Censos del 2000 (Seminario Censos 2000: diseño conceptual y temas a investigar en América Latina)*, Serie Manuales 1. Comisión Económica para América Latina y el Caribe, Santiago de Chile.
- Saisana, M. y Tarántola, S. (2002): *State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development*. Report EUR 20408 EN. Joint Research Centre. European Commission, Ispra.

- Salzman, J. (2003): *Methodological Choices Encountered in the Construction of Composite Indices of Economic and Social Well-Being*. Center for the Study of Living Standards Ottawa, Ontario.
- Schatan W. J. (1999): *Pobreza, Distribución del Ingreso y Línea de Dignidad*. Documento elaborado para el Programa de Economía Ecológica del Instituto de Ecología Política, Santiago de Chile.
- Schuschny, A. y Soto, H. (2009): *Guía metodológica. Diseño de indicadores compuestos de desarrollo sostenible*. Documento de proyecto. DDSAH/CEPAL, Santiago de Chile.
- Szalachman, R. (2000): “Perfil del déficit y políticas de vivienda de interés social: situación de algunos países de la región en los noventa”, *Serie Financiamiento y Desarrollo* 103. Unidad de Financiamiento para el Desarrollo, División de Comercio Internacional y Financiamiento Habitacional, Santiago de Chile.
- UNODC -United Nations Office on Drugs and Crime (2011): *2011 Global Study on Homicide. Trends, contexts, data*. UNODC, Viena.
- UN-HABITAT (2009): *Urban Indicators Guidelines. 'Better Information, Better Cities': Monitoring the Habitat Agenda and the Millenium Development Goals-Slums Target*. UN-HABITAT.
- Velázquez, G.A. (2008): “Bienestar y jerarquía urbana. Un análisis regional en la Argentina (2001)”, en Lucero, P. (ed.): *Territorio y calidad de vida, una mirada desde la geografía local: Mar del Plata y Partido de General Pueyrredón*. EUDEM. Mar del Plata, pp. 41–72.
- Velázquez, G.A. y Gómez Lende, S. (2005): “Población y calidad de vida en la Argentina (1991-2001). La fragmentación de la sociedad y el territorio”, en Velázquez, G.A. y Gómez Lende, S. (comp.): *Desigualdad y calidad de vida en la Argentina (1991-2001) Aportes empíricos y metodológicos*, UNCEN-REUN-CIG, Tandil, pp. 63–86.
- Wautiez, F. (2000): “La equidad socio-ambiental en Chile: una tarea pendiente”. Programa Chile Sustentable, Santiago de Chile.
- Zárata Martín, M.A. (1988): “Bienestar Social y diferenciación interna del espacio urbano. Vitoria-Gasteiz”, *Revista de la Facultad de Geografía e Historia*, 1, pp. 163–178.
- Zarzosa Espina, P. (1996): *Aproximación a la medición del Bienestar Social*. Universidad de Valladolid, Secretariado de Publicaciones, Valladolid.



Links between the Eurozone Stock Markets: A New Perspective, Considering the Capitalization Level

GABRIEL, VÍTOR

UDI - Research Unit for Inland Development

Polytechnic Institute of Guarda (Portugal)

E-mail: vigab@ipg.pt

SARAIVA, HELENA

UDI - Research Unit for Inland Development

Polytechnic Institute of Guarda (Portugal)

E-mail: helenasaraiva@ipg.pt

ABSTRACT

This paper examines short-term and long-term linkages among stock markets within EMU, taking into account the business capitalization. According to this objective, we have analysed four capitalization segments, corresponding to the Micro, Small, Mid and Large Caps indices, in the period between November 2007 and December 2013.

In order to identify the existence of interdependencies and short-term links between the European indices, we have used a vector autoregressive error-correction model, the concept of Granger causality and the impulse-response functions. We have concluded that the Large Cap described relatively autonomous movements and contained information that helped to explain the changes in other indices.

With regard to the existence of long-term connections, the usual cointegration tests were used, which showed that the segment index of the largest capitalizations described a different route compared to the indices of the two segments with smaller capitalizations. This proves to be particularly important for an international portfolio diversification strategy.

Keywords: European stock markets; cap segments; vector autoregressive; cointegration.

JEL classification: G11; G15.

MSC2010: 91G10; 91G50.

Relación entre los índices bursátiles europeos: una nueva perspectiva a partir de los niveles de capitalización

RESUMEN

Este artículo analiza las relaciones y las interdependencias a corto y largo plazo entre los mercados de valores de la eurozona, teniendo en cuenta el nivel de capitalización. De acuerdo con este objetivo, se analizaron cuatro segmentos de capitalización correspondientes a los índices de micro, pequeñas, medianas y grandes capitalizaciones en el período comprendido entre noviembre de 2007 y diciembre de 2013.

Con el fin de identificar la existencia de interdependencias y relaciones a corto plazo entre los índices europeos, se ha recurrido a un vector autorregresivo con mecanismo corrector de errores, al concepto de causalidad de Granger y a funciones de impulso-respuesta. Se concluyó que el índice *Large Cap* describe movimientos relativamente autónomos y que contiene información que ayuda a explicar los cambios en otros índices.

En cuanto a la existencia de relaciones a largo plazo, se utilizó las habituales pruebas de cointegración, lo que ha permitido entender que el índice para el segmento de las grandes capitalizaciones describe un comportamiento diferente en comparación con los índices de los dos segmentos de capitalización más pequeños. Este resultado es particularmente importante para una posible estrategia de diversificación de la cartera de inversiones en una perspectiva internacional.

Palabras claves: mercados bursátiles europeos; segmentos de capitalización; vectores autorregresivos; cointegración.

Clasificación JEL: G11; G15.

MSC2010: 91G10; 91G50.



1. INTRODUCTION

The links between international stock markets have inspired a large number of studies, partly encouraged by the pioneering work of Grubel (1968). Early works on this subject did not always produce overlapping conclusions. Ripley (1973) concluded that there was a certain degree of interdependence between markets open to investments and foreign capital. On the other hand, Granger and Morgenstern (1970), Agmon (1972) and Branch (1974) found no significant evidence of lead-lag relationships between international stock markets. For his part, Bertoneche (1979), when studying the links between the stock markets of the USA, Germany, Belgium, the UK, the Netherlands, France and Italy in the period between 1969 and 1976, identified a high degree of segmentation between the seven markets studied by him, which favoured the adoption of international diversification strategies. In the same vein, Roll (1988) concludes that the links between international stock markets were weak.

Bibliography on the theme of stock market links highlights the stock market crash of October 1987 as a key element in bringing markets together. In this context, we highlight the works of Eun and Shim (1989), Lau and McInish (1993) and Arshanapalli *et al.* (1995), among others.

Eun and Shim (1989) identified significant interdependencies between nine large international stock caps. Jeon and Von-Furstenberg (1990) found a significant increase in international stock market co-movements, as a consequence of the 1987 crash. Similarly, Lau and McInish (1993), using daily information on the period between 1986 and 1989, identified changes in the structure of the lead-lag co-movements in international stock markets, after the occurrence of the crash. Furthermore, Arshanapalli *et al.* (1995) concluded that the crash helped to create more similar and nearby responses among global stock markets. The statistical confirmation of lead-lag type effects, such as those seen in the above-mentioned studies, can be understood as a violation of the efficient market hypothesis since, according to Fama (1970), price behaviour is equivalent to a random walk process, implying the non-predictability of asset prices, which is partly counteracted with confirmation of such effects.

Later studies have emphasized a gradual strengthening of the links between international stock markets and highlighted the role of the main market capitalization of the US market, by demonstrating the ability to help explaining movements in other markets.

Hassan and Atsuyuki (1996) resorted to daily information on the period from 1984 to 1991, in order to study the dynamic linkages between the US, Japan, the UK and Germany, and they concluded that there are significant relationships between these markets. In the short term, the US market proved the most exogenous before and after the 1987 crash. In the long run, it is identified a cointegration relationship between the four markets, with implications in terms of international diversification opportunities. Peiro *et al.* (1998) studied the links between the stock markets of New York, Tokyo and

Frankfurt in the time-lapse from 1990 to 1993 and concluded that the former is the most influential stock market and the Japanese market may prove more sensitive to international shocks. Similarly, Ozdemir and Cakan (2007) studied the US, Japan, France and the UK markets between 1990 and 2006, using the concept of Granger causality, and concluded that the US market leads the remaining markets.

Goetzmann *et al.* (2005) showed that the correlation structure of the world stock markets has varied considerably over the past 150 years, and was particularly strong in periods of financial and economic integration. Similarly, Bekaert *et al.* (2007) emphasized the importance of market integration in the global context due to the processes of liberalization of capital markets, stock markets and the respective banking systems, facilitating the further development of international linkages between markets.

In Europe, the introduction of the single currency produced empirical evidence that the correlations between stock market indexes, representative for the Eurozone, have increased considerably. They were, however, unstable in time and higher in "bear market" periods, with implications in terms of diversification reduction within the investment portfolio - see Cappiello *et al.* (2006), for example.

More recently, other scientific studies highlighted the role of the recent global financial crisis (GC) as a determining element in bringing the stock markets together. Using Granger causality tests and impulse-response functions, Tudor (2011) concluded that the links between stock markets of Central and Eastern Europe and the US market have increased with the emergence of this crisis. Similarly, Mandigma (2014) concluded that there was a strengthening of dynamic links between Southeast Asian countries and the US market.

In order to study the possible existence of interdependencies between markets and to seek diversification benefits, the previously mentioned studies considered some of the most representative international indexes. But in either case, these indexes basically represent the largest companies in their respective markets, ignoring companies with lower capitalization levels.

With regard to investment diversification, individual and institutional investors search for expanding investment alternatives at international level but also those arising from a possible advantage provided by companies of different capitalization levels. According to portfolio theory, based on binomial risk-return, if Small Caps yields do not correlate perfectly with those of Large Caps, investors can obtain benefits if they always adopt a diversification strategy based on company size. The benefits are a better relationship between risk and profitability; i.e., reducing the risk associated with the investment portfolio, without penalizing its profitability.

According to Petrella (2005) and Eun *et al.* (2006), company size is an important factor in the behaviour of market yields. If Large Cap yields are fundamentally explained by factors of global nature, Small Cap yields are mainly explained by idiosyncratic factors. According to these authors, this difference in the generating mechanism of yields can be explained by the fact that large companies (multinationals in some cases) develop activity in several countries and have an international shareholder base. On the other hand, small businesses, with a more locally focused activity, have a more limited international exposure. Consequently, it must be assumed that the yields of the smallest and the largest capitalizations are not perfectly correlated, which brings advantages to investors who adopt a strategy based on "size diversification". However, this line of argument is only valid if the time structure of market links, reflected in the correlation between Small and Large Cap yields, is stable and does not show asymmetrical behaviour; i.e., in periods of bear market the correlation between yields is not higher than the one in bull market periods.

The main purpose of this study involves the analysis of short-term and long-term links between daily yields of different size company stock indexes in the space defined by the Eurozone. The overwhelming majority of the studies has tended to focus successively on the use of indexes dominated by large companies, so our job is to fill a lack in terms of empirical research and study the degree of connection between the yields of companies with different levels of market capitalization, using representative indexes from four cap segments: micro, small, mid, and large caps.

In terms of structure, this research continues in Section 2 with a description of the data and methodology. The third sections deals with the presentation of empirical results and 4 with the presentation of the summary and key findings.

2. DATA AND METHODOLOGY

2.1 Data

Our analysis is based on daily data, for a period of time of about six years, and uses four European stock indexes corresponding to four cap segments: *Micro Cap*, *Small Cap*, *Mid Cap* and *Large Cap*. The data used in this study were provided by Morgan Stanley Capital International and encompass the sample period of 30/11/2007 to 05/12/2013, translated in 1543 everyday observations.

The series of index closing values were transformed into yield logarithmic series, instant or continuously compounded, through the following expression:

$$r_t = \ln P_t - \ln P_{t-1} \quad (1)$$

in which r_t is the return rate in day t , and P_t and P_{t-1} are the series closing values in the moments t and $t-1$, respectively.

2.2 Methodology

In order to study the short-term connections between the selected stock exchange indexes and the direction of influence between these, we resorted to the concept of Granger causality, introduced by Granger (1969) and later made popular by Sims (1980), specifically to the *VAR Granger Causality/Block Exogeneity Wald Tests*, developed from a vector autoregression, whose theoretical basis is presented below.

In order to further study the short-term connections between the stock markets, we have also considered generalized impulse-response functions, introduced by Koop *et al.* (1996) and Pesaran and Shin (1998), to improve the understanding of the results of the chain reactions that occurred in the indexes. This approach has advantages over the traditional orthogonalized impulse-response analysis since the results do not depend on the ordering of variables in the VAR model.

The study of long-term connections between the indexes involves consideration of the concept of cointegration, introduced by Granger (1981). According to this author, if two or more sets are not stationary individually, but a linear combination of these is stationary, the series is said to be cointegrated. A cointegration relationship between series may happen in the long run, even if in the short run these series do not have a convergent behaviour, but they recover the same trend in the long run.

In general, two alternatives are considered for testing cointegration: One encompasses the Engle-Granger test, which is used to study simple series; the other uses the Johansen approach, which is especially useful in the multivariate case.

After testing series integration, using, for example, unit root tests, you can apply the Johansen approach.

Considering Y_t as the index vector, and assuming that it has the following VAR representation:

$$Y_t = z + \sum_{i=1}^p \Pi_i Y_{t-i} + u_t \quad (2)$$

where:

z is a vector of endogenous variables ($n \times 1$);

u is a white noise error vector ($n \times 1$);

and Π_i is a coefficient matrix ($n \times n$).

The VAR model for the error correction mechanism is given by:

$$\Delta Y_t = z + \sum_{i=1}^{p-1} B_i \Delta Y_{t-i} + \Pi Y_{t-1} + u_t \quad (3)$$

in which Y_t is a column vector, of parameters $N \times 1$, of dependent variables, integrated of order one; μ is a deterministic parameter vector; and u_t is an innovation column vector, of $N \times 1$ size.

Before applying the Johansen test, the VAR models should be transformed in a vector error correction (VEC) in the form:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{n-1} \Gamma_i \Delta Y_{t-i} + \mu + u_t \quad (4)$$

where Y_t is an integrated variables vector, $I(1)$, ΔY_t are variables $I(0)$, Δ is the first difference operator, B is a matrix of coefficients ($n \times n$) and Π is a matrix ($n \times n$), in which r determines the number of cointegrating vectors between variables.

The Johansen cointegration test is based on Π matrix analysis. If Π is defined by $r = n$, the variables are stationary in levels; if $r = 0$, there is no cointegration between variables.

The Johansen cointegration approach uses the trace (λ_{trace}) and maximum eigenvalue (λ_{max}) tests, based on the likelihood ratio. The statistics are expressed as follows:

$$\lambda_{trace} = -T \sum_{i=r+1}^k \log(1 - \lambda_i) \quad (5)$$

The trace test tests the null hypothesis that the number of cointegrating vectors is r against the alternative hypothesis of being $r + 1$.

The maximum eigenvalue test is given by:

$$\lambda_{max} = -T \log(1 - \lambda_{r+1}) \quad (6)$$

The maximum eigenvalue tests the hypothesis that the number of cointegrating vectors is larger than r , against the alternative hypothesis of being $r + 1$.

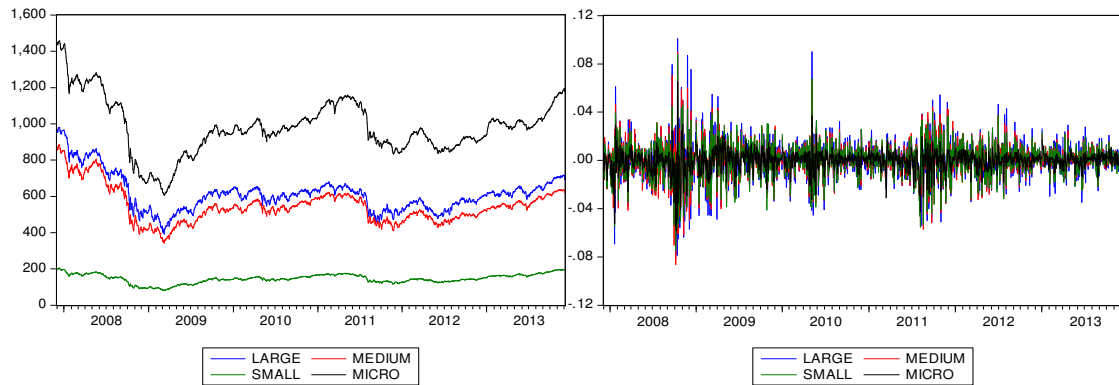
In both tests, T sets the number of observations and λ_i the eigenvalues, estimated and ordered in decreasing order.

When there is a discrepancy in the results produced by the two statistics, we follow the recommendations of Johansen (1991), Kasa (1992), Serletis and King (1997) and Hubrich *et al.* (2001), which promotes the importance of trace statistics, given the maximum eigenvalue, as it is more reliable.

3. EMPIRICAL RESULTS

Let us appreciate the index series evolution in levels and respective yields. Figure 1 shows the evolution of the four stock market indexes in levels (left) and of the return rate (right), in the sample period between November 30, 2007 and December 5, 2013.

Figure 1: Index evolution.



The graphical analysis of the indexes shows that they have very similar behaviour patterns over the sample period, especially the Large and Mid Caps indexes. The behaviour patterns were strongly marked by the emergence of the global financial crisis, in particular after the fourth quarter of 2008, which started a bear market phase, coupled with a sharp increase in the volatility of European markets. The values of the indexes suffered significant variations over the studied period. However, despite variations, yields have average stationary features, giving a first indication that these may be stationary.

The main yield descriptive statistics of the four indexes are shown in Table 1. The analysis of descriptive statistics supports the conclusion that all indexes showed a negative daily average profitability.

Table 1: Yield descriptive statistics

	<i>Large</i>	<i>Mid</i>	<i>Small</i>	<i>Micro</i>
Average	-0.00022	-0.00022	-0.00004	-0.00013
Median	-0.00011	0.00028	0.00079	0.00094
Maximum	0.10083	0.08973	0.08757	0.06501
Minimum	-0.08113	-0.08621	-0.07424	-0.05817
Standard Deviation	0.01617	0.01471	0.01371	0.00818
Asymmetry	0.04194	-0.23670	-0.30629	-1.07280
Kurtosis	7.71691	6.85468	6.31291	12.54646
Jarque-Bera	1430.89400	969.69150	729.75260	6155.19200
Probability	(0.00000)	(0.00000)	(0.00000)	(0.00000)
ADF (levels)	(0.14511)	(0.20771)	(0.33961)	(0.32594)
ADF (returns)	(0.00000)	(0.00000)	(0.00000)	(0.00000)

All yield series showed signs of deviation due to the normality hypothesis, given the skewness and kurtosis coefficients, different from zero and three, respectively. The analysed series are leptokurtic and feature asymmetric tails.

In order to know the adjustment of the normal distribution to the empirical distributions of the four series, the Jarque-Bera test was also applied, and its statistical values can be seen in Table 1. Taking into account the test values and their respective probabilities (equal to zero), we concluded that all series are statistically significant at 1%, so the hypothesis of their normality is clearly rejected.

In order to determine the series stationarity (levels and yields), the traditional Augmented Dickey-Fuller tests (ADF; see Dickey and Fuller, 1979) were applied. The null hypothesis (H_0) of this test stipulates that the series has a unit root, that is, that the series is integrated of order 1, I (1), given the alternative hypothesis (H_a) that the series does not have a unit root or is I (0). The results of the series stationary tests are shown in Table 1. The yield series evidenced immobility, I (0), at the significance level of 1%, whereas the series in levels were shown to be non-stationary, or R (1), for the same level of significance.

In order to further study stock market behaviour, one starts by calculating the coefficients of Contemporary Correlations between European index yields (see Table 2).

Table 2: Contemporary Correlations between yields

	<i>Large</i>	<i>Mid</i>	<i>Small</i>	<i>Micro</i>
<i>Large</i>	1.000			
<i>Mid</i>	0.946	1.000		
	(0.000)			
<i>Small</i>	0.910	0.967	1.000	
	(0.000)	(0.000)		
<i>Micro</i>	0.782	0.842	0.890	1.000
	(0.000)	(0.000)	(0.000)	

All the correlation coefficients recorded positive values and were statistically significant at a significance level of 1%, according to the values of the respective p-values. The values of correlation coefficients were quite high. The coefficients for the smaller caps produced, in either case, strong correlations. The remaining pairs have resulted in very strong correlations exceeding 90%, which possibly means the European stock markets described, in the short term, very similar movements.

In order to analyze the existence of a causal relationship between the four indexes, we applied the VAR Granger Causality/Block Exogeneity Wald Tests procedure, whose null hypothesis states that there is no causal relationship between the indexes. This procedure assesses the significance of each

joint lagged endogenous variable in each equation of the VAR, through the statistic χ^2 , and simultaneously the significance of the joint contribution of all offset endogenous variables, present in the equation using F-statistic. For each of the equations of the estimated VAR models, we present the Wald test statistics, about the joint significance of each of the other endogenous variables, based on the selected offsets, defined by the Akaike-Schwarz information criteria.

Table 3 is a summary of the results for Granger causality tests. These results revealed some statistically significant unidirectional relations. However, we did not identify any two-way relationships, with statistical significance, and this differentiates these results from other recent research, including Tudor (2011) and Mandigma (2014), favoring the idea that yield generating mechanism in the segment indexes can differentiate the mechanism of conventional indexes that take into consideration the largest market capitalizations.

The Small Cap and Micro Cap indexes proved to be the most endogenous, as they were caused, in the Granger sense, by the Mid Cap and Large Cap indexes, also being caused by the combined effect of the remaining offset indexes at the significance level of 1%. The Large Cap index was the less endogenous, which, in the Granger sense, was not caused individually by any other indexes, or by the combined effect of the other indexes. In terms of exogeneity, this index clearly stood out from its peers, as it contained information on the movements of other indexes. These results led us to conclude that the Large Cap describes an autonomous behavior, but also that this index plays the role of the driver of the other stock indexes, as it helps to explain their movements.

Table 3: Granger Causality/Block Exogeneity Wald Tests

		Dependent Variable			
		<i>Large</i>	<i>Mid</i>	<i>Small</i>	<i>Micro</i>
Excluded variables	<i>Large</i>		6.1652	9.5698	10.7161
			(0.0458)	(0.0084)	(0.0047)
	<i>Mid</i>	1.0862		9.2792	8.6212
		(0.5809)		(0.0097)	(0.0134)
	<i>Small</i>	0.0377	0.5156		0.3774
		(0.9813)	(0.7727)		(0.8280)
	<i>Micro</i>	1.0518	2.7613	4.1728	
		(0.5910)	(0.2514)	(0.1241)	
	<i>All</i>	4.7200	9.9412	17.3152	26.5036
		(0.5802)	(0.1272)	(0.0082)	(0.0002)

The statistical significance of various causal relationships allows us to question the assumptions of the efficient market hypothesis, since the movements in a particular index are partly preceded by past movements in other indexes, which shows certain predictability in these movements and enables the development of arbitration operations.

Granger causality test helps to understand the links between markets, but does not show if the indexes have a negative or a positive effect on their peers, or lets you know if some links are stronger than others. To get these answers, we use generalized impulse response functions, amplitude corresponding to a standard deviation, in order to obtain further evidence about the short-term movements of the transmission mechanisms. The results of the impulse response functions are summarized in Table 4.

Response functions to impulses, calculated from the autoregression vector model, lets us to realize that the four European indexes showed statistically significant reactions as a result of shocks from abroad (see Table 4). A significant part of these shocks statistically significant effects for only a period, fading afterwards. All reactions displayed the expected signal; i.e., they accompanied the direction of index variation originated where the shock occurred and there were no correction reactions to the initial shock.

Table 4: Summary of the results of the impulse-response functions

	<i>Large</i>	<i>Mid</i>	<i>Small</i>	<i>Micro</i>
<i>Large</i>		1 1 +	1 1 +	1 1 +
<i>Mid</i>	1 1 +		2 1,2 +,+	1 1 +
<i>Small</i>	2 1,2 +,+	2 1,2 +,+		2 1,2 +,+
<i>Micro</i>	6 1,2,3,4,5,6 +,+,+,+,+,+	6 1,2,3,4,5,6 +,+,+,+,+,+	6 1,2,3,4,5,6 +,+,+,+,+,+	

Notes: Online market responses to impulses from the column in markets; each cell has three rows: the first row indicates the number of periods during which the online market response was statistically significant, according to the criterion of one standard deviation; the second row indicates the order number of these periods; while the third row indicates the signal of the impulse response.

With regard to the persistence of the reactions, we concluded that, in general, the indexes for the larger cap segments (Mid and Large) were the ones which triggered the less persistent responses, in most cases for a period only. Conversely, the Micro Cap index recorded the most persistent reactions to impulses by its peers (six days). This high persistence may be interpreted as evidence that information on this index has not been incorporated instantly, contradicting the efficient market hypothesis.

In order to study the possibility of long-term relationships between European indexes, bivariate Johansen cointegration tests were applied.

A necessary condition for the application of the cointegration test is that the variables have the same integration order. Taking the results shown in Table 1 into account, it is concluded that, in any

case, the ADF unit root tests showed that the indexes are at levels I (1), satisfying the prerequisite for the application of cointegration tests. If the series are cointegrated, there is statistical evidence of long-term relationship between the indexes.

Notably, the Johansen method presupposes prior estimation of a vector model. A key element in the specification of the Vector models is the lag size. In each of the bivariate analyses, the choice of the optimum number of VAR lags took into account the usual Akaike-Schwarz information criteria.

After estimating the VAR model, the Johansen cointegration test was applied, whose statistics are summarized in Table 5, namely in the two columns on the right side. The first column presents results for the trace test, while the second column shows the test results for the maximum eigenvalue. In both cases, in the above-mentioned columns, there are values for test statistics and for the probability of rejecting the null hypothesis (p-value); e.g., that there is no cointegrating vector. Taking into account that, in some cases, there was some discrepancy in the results produced by the two statistics, the trace test was preferred, in accordance with the recommendation of Johansen (1991), Kasa (1992), Serletis and King (1997) and Hubrich *et al.* (2001).

Table 5: Johansen cointegration tests

		Trace test (λ_{trace})	Maximum Eigenvalue Test (λ_{max})
Large	Mid	24.3267 (0.013)	15.7673 (0.0523)
	Small	16.2118 (0.1647)	13.9076 (0.0999)
	Micro	17.1430 (0.1273)	13.9371 (0.0989)
Mid	Small	14.8163 (0.0631)	14.4033 (0.0475)
	Micro	15.1520 (0.0563)	11.4471 (0.1319)
Small	Micro	39.7495 (0.0005)	31.7660 (0.0005)

The null hypothesis of non-cointegration between the Large Cap and the Mid Cap, but also between the Small Cap and the Micro Cap, was rejected by the trace test, a cointegrating vector being identified for the significance levels of 5% and 1%, respectively. In any case, the results revealed the existence of a balance between long-term pairs formed by these indexes.

In summary, the Johansen bivariate cointegration analysis identified some balanced relationships in the long run between pairs of European indexes. The occurrence of common movements suggests the existence of common factors that limit the independent variation between the indexes and allow a certain degree of market behaviour predictability. Verifying balanced relationships

conditions investment alternatives, in the light of a possible international diversification strategy. However, the results also show that it is possible for investors, in a long-term perspective, to adopt an investment strategy which combines assets from large/ small businesses or large /micro, because these pairs do not reveal balanced relationships in the long run and can, therefore, be viewed as alternatives for investment diversification.

According to Granger (1986) and Baillie and Bollerslev (1989), and similarly to the conclusion drawn as to the short-term connections between the indexes, the long-term connections, measured from the Johansen methodology, also contradict the assumptions of the efficiency hypothesis of the markets, as it contributes to a predictability degree in market behaviour.

4. SUMMARY, CONCLUSIONS AND LIMITATIONS

The study of links and interdependencies between stock markets has been a very popular study topic in finance. The absence of links between markets translates into an advantage in terms of diversification of investment portfolios in the global context. Several studies have reported a reduction in benefits associated with diversification, especially after the occurrence of extreme events, as are the stock market crashes.

In the present study, we considered a number of Eurozone stock market indexes, representing different cap segments, namely Micro Cap, Small Cap, Mid Cap and Large Cap. In order to ascertain the existence of short-term co-movements, coefficients of contemporary correlation were calculated, an autoregression vector was estimated, from which causality tests were applied and impulse-response functions estimated. The study of possible long-run balanced relationships between the indexes involved the application of bivariate cointegration tests.

The coefficients of contemporary correlation recorded very high values, particularly among indexes for the three largest capitalizations. The results of Granger causality tests revealed several statistically significant unidirectional links. The Large Cap index stood out, emerging as the least endogenous, but also as the most exogenous. This reflects, on the one hand, an autonomous behaviour of this index compared to others and, on the other hand, the high capacity of this index to help explain peer movements. In turn, the impulse response functions show that, in general, index reactions are not persistent, fading fast some days after the stimulus. The only exception was the Micro Cap Index, which reported statistically significant answers six days after the impulse originated in their peers.

The analysis of Johansen bivariate cointegration identified a number of balanced relationships in the long run between European index pairs. The occurrence of common movement may be the result of the existence of common factors, which limits the independent variation between the indexes and

enables a certain degree of predictability. The verification of these relationships ultimately conditions the investment alternatives in an international diversification perspective. The results suggest that it is possible for investors, in a long-term perspective, to adopt an investment strategy which combines assets from large/ small businesses or large /micro, because these pairs do not reveal balanced relationships in the long run and can, therefore, be viewed as alternatives for investment diversification. The short-term and long-term results obtained question the validity of the market efficiency hypothesis assumptions, since the prediction of the movement of some indexes can be improved if we consider the offset of the remaining movements, giving rise to arbitrage transactions, which contradicts the genesis of this theory.

In future research, we plan to further study the links between European stock markets, giving again emphasis to the use of cap segments, so that we follow two research lines. In the first, we will analyse the transmission of information between the indexes, using multivariate models of conditional heteroskedasticity, in order to accommodate the volatility clusters and the asymmetric effect. In the second, we will study alternative allocation of international investment, based on size diversification and optimization models.

REFERENCES

- Agmon, T. (1972). "The relations among equity markets: A study of share price comovements in the United States, United Kingdom, Germany and Japan". *Journal of Finance*, 27, pp. 839-855.
- Arshanapalli, B.; Doukas, J. and Lang, L. (1995). "Pre and post-October 1987 stock market linkages between U.S. and Asian markets". *Pacific-basin Finance Journal*, 3, pp. 57-73.
- Baillie, R. and Bollerslev, T. (1989). "The Message in Daily Exchange Rates: A Conditional- Variance Tale". *Journal of Business and Economic Statistics*, 7(3), pp. 297-305.
- Bekaert, G.; Harvey, C.; Lundblad, C. and Siegel, S. (2007). "Global Growth Opportunities and Market Integration". *Journal of Finance*, 62(3), pp. 1081-1137.
- Bertoneche, M. (1979). "An empirical analysis of the interrelationships among equity markets under changing exchange rate systems". *Journal of Banking and Finance*, 3(4), pp. 397-405.
- Branch, B. (1974). "Common stock performance and inflation: an international comparison". *Journal of Business*, 47, pp. 48-52.
- Cappiello, L.; Engle, R. and Sheppard, K., (2006). "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns". *Journal of Financial Econometrics*, 4, pp. 537-572.

- Dickey, D. e Fuller, W. (1979). “Distribution of the Estimators for Time Series Regressions with a Unit Root”. *Journal of the American Statistical Association*, 74, pp. 427-31.
- Eun, C. e Shim, S. (1989). “International Transmission of Stock Market Movements”. *Journal of Financial and Quantitative Analysis*, 24, pp. 241-256.
- Eun, C.; Huang, W. and Lai, S. (2006). “International Diversification with Large and Small-Caps”. *Journal of Financial and Quantitative Analysis*, 43:2, pp. 489-523.
- Fama, E. (1970). “Efficient capital markets: A review of empirical work”. *Journal of Finance*, 25, pp. 383-417.
- Goetzmann, W.; Li, L. and Rouwenhorst, K. (2005). “Long-Term Global Market Correlations”. *Journal of Business*, 78(1), pp. 1-38.
- Granger, C. (1969). “Investigating causal relationships by econometric models and cross-spectral methods”. *Econometrica*, 37(3), pp. 424-438.
- Granger, C. (1981). “Some Properties of Time Series Data and Their Use in Econometric Model Specification”. *Journal of Econometrics*, 16:1, pp. 121–130.
- Granger, C. (1986). “Developments in the study of cointegrated economic variables”. *Oxford Bulletin of Economics and Statistics*, 48, 213-228.
- Granger, C. and Morgenstern, O. (1970). *Predictability of stock market prices*. Lexington Books.
- Grubel, H. (1968). “Internationally diversified portfolios: welfare gains and capital flows”. *American Economic Review*, 58, pp. 1299-1314.
- Hassan, M. and Atsuyuki, N. (1996). “Short-run and long-run dynamic linkages among international stock markets”. *International Review of Economics and Finance*, 5:4, pp. 387-405.
- Hubrich, K.; Lütkepohl, H. and Saikkonen, P. (2001). “A Review of Systems Cointegration Tests”. *Econometric Reviews*, 20, 247-318.
- Jeon, B. and Von-Furstenberg, G. (1990). “Growing international co-movement in stock price indexes”. *Quartely Review of Economics and Business*, 30:3, pp. 15-30.
- Johansen, S. (1991). “Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”. *Econometrica*, 59: 1551-1580.
- Kasa, K. (1992). “Common stochastic trends in international stock markets”. *Journal of Monetary Economics*, 29, 95-124.

- Lau, S. and McInish, T. (1993). “Comovements of international equity returns: A comparison of pre- and post-October 19, 1987, periods”. *Global Finance Journal*, 4, pp. 1-19.
- Koop, G.; Pesaran, M. and Potter, S. (1996). “Impulse response analysis in non-linear multivariate models”. *Journal of Econometrics*, 74, pp. 119-147.
- Mandigma, M. (2014). “Stock Market Linkages among the ASEAN 5+3 Countries and US: Further Evidence”. *Management and Administrative Sciences Review*, 3(1), pp. 53-68.
- Ozdemir, Z. and Cakan, E. (2007). “Non-linear Dynamic Linkages in the International Stock Markets”. *Physica A: Statistical Mechanics and its Applications*, 377, pp. 173–180.
- Peiro, A.; Quesada, J. and Ezequiel, U. (1998). “Transmission of movements in stock markets”. *The European Journal of Finance*, 4, pp. 331- 343.
- Pesaran, H. and Shin, Y. (1998). “Generalized impulse response analysis in linear multivariate models”. *Economics Letter*, Elsevier. 58(1), pp. 17-29.
- Petrella, G. (2005). “Are Euro Area Small Cap Stocks an Asset Class? Evidence from Mean-Variance Spanning Tests”. *European Financial Management*, 11, pp. 229-253.
- Ripley, D. (1973). “Systematic Elements in the Linkage of National Stock Market Indices”. *Review of Economics and Statistics*, 55(3), pp. 356-361.
- Roll, R. (1988). “The International Crash of October 1987”. *Financial Analyst Journal*, 44, pp. 19-35.
- Serletis, A. and King, M. (1997). “Common stochastic trends and convergence of European Union stock markets”. *The Manchester School*, 65: 44-57.
- Sims, C. (1980). “Macroeconomics and reality”. *Econometrica*, 48, pp. 1-48.
- Tudor, C. (2011). “Changes in Stock Markets Interdependencies as a Result of the Global Financial Crisis: Empirical Investigation on the CEE Region”. *Panoeconomicus*, 4, pp. 525-543.



La tasa de variación del PIB en un modelo simple de determinación de la renta

ESCARTIN, EDUARDO

Grupo SEJ442: Métodos Cualitativos y Optimización en Sistemas Dinámicos
Económicos

Universidad de Sevilla (España)

Correo electrónico: escartin@us.es

VELASCO, FRANCISCO

Grupo SEJ442: Métodos Cualitativos y Optimización en Sistemas Dinámicos
Económicos

Universidad de Sevilla (España)

Correo electrónico: velasco@us.es

GONZÁLEZ-ABRIL, LUIS

Grupo SEJ442: Métodos Cualitativos y Optimización en Sistemas Dinámicos
Económicos

Universidad de Sevilla (España)

Correo-e: luisgon@us.es

RESUMEN

En este artículo se demuestra que en un modelo simple de determinación de la renta, cuando se tienen en cuenta funciones de consumo lineales, la tasa de variación del producto interior bruto (PIB) es igual al cociente entre el incremento de la inversión y la suma del consumo exógeno más la inversión inicial. Esto significa que la pendiente de la función de consumo no influye en la tasa de variación del PIB y que lo relevante para esa tasa de variación es el consumo exógeno y la inversión inicial.

Palabras claves: macroeconomía; producto interior bruto (PIB); modelo de determinación de la renta; tasa de crecimiento.

Clasificación JEL: E10.

MSC2010: 37N40; 91B02; 91B55.

The Variation of GDP in a Simple Model of Income Determination

ABSTRACT

In this paper, we prove that in a simple model of income determination, when taking into account linear functions, the change rate for the gross domestic product (GDP) is equal to the quotient of the increase in investment and the sum of exogenous consumption plus the initial investment. This means that the slope of the consumption function does not influence the change rate for the GDP and what matters for this variation rate is the exogenous consumption and the initial investment.

Keywords: Macroeconomics; gross domestic product (GDP); model of income determination; growth rate.

JEL classification: E10.

MSC2010: 37N40; 91B02; 91B55.



1 Introducción

La tasa de variación del producto interior bruto (PIB), aunque es un concepto de por sí claro, plantea la siguiente pregunta: ¿a qué equivale la tasa de variación del PIB, o $\Delta Y/Y$?

En este artículo se expone una solución a esta cuestión, derivándola de un modelo simple de análisis de la renta; esto es una aportación original, porque, según nuestro conocimiento, no hay ningún manual donde se aborde este tema. Permítasenos resaltar que, al tratarse de una demostración personal sin referencias previas, no hay bibliografía a la que acudir para tratar sobre esta demostración, pues, evidentemente, no sería una novedad. Por tal motivo las referencias bibliográficas son necesariamente mínimas.

El enunciado a demostrar es el siguiente:

La tasa de variación del producto interior bruto (PIB), en un modelo simple de determinación de la renta (Y), teniendo en cuenta funciones de consumo lineales con pendiente positiva y menor que la unidad, es igual al cociente entre el incremento de la inversión y la suma del consumo exógeno (C_0) más la inversión inicial (I); es decir, su fórmula es:

$$\frac{\Delta Y}{Y} = \frac{\Delta I}{C_0 + I}$$

Aunque la inspiración de la pregunta proviene de la *Teoría General* de Keynes (1989: 126)¹, se señala que la enunciación y la demostración son totalmente ajenas a su teoría y que no se pretende entrar en ella ni establecer cuestión ni crítica sobre la misma. Se presenta una demostración y se extrae de ella determinadas conclusiones útiles para la práctica política del crecimiento económico y del empleo. Así pues, el artículo pone su foco en la demostración y se ruega al lector que se centre en la exposición y permanezca al margen de lo que dijo o pretendió decir Keynes².

¹ En la nota 2 a pie de la página 126 de *The General Theory*, Keynes indica: «la razón entre el cambio proporcional de la demanda total respecto al cambio proporcional de la inversión

$$= \frac{\Delta Y}{Y} : \frac{\Delta I}{I} = \frac{\Delta Y}{Y} \cdot \frac{Y - C}{\Delta Y - \Delta C} = \frac{1 - \frac{C}{Y}}{1 - \frac{\Delta C}{\Delta Y}} \gg$$

² Keynes ya tuvo tempranamente inquietudes para analizar la mejoría del empleo en Keynes y Henderson (1929: 112-116) y en Keynes (1933: 340-341).

2 Demostración analítica

Se considera el modelo tradicional:

Sea una función de demanda agregada (D) que, en equilibrio, es igual a la oferta agregada, es decir, al producto agregado (PIB); a su vez el PIB es igual a la renta (Y), o retribución de los factores de la producción, porque éstos son remunerados del valor del producto agregado. Así pues: $D=Y$.

La función de demanda agregada es la suma de dos componentes: el consumo (C) y la inversión (I); por lo que: $D=C+I$. Pero como en el equilibrio la demanda agregada debe ser igual a la oferta agregada (es decir, $D=Y$), resulta que $Y=C+I$.

Adicionalmente, se considera que el consumo es función de la renta, $C = f(Y)$. Esta función tiene pendiente, dC/dY , positiva, constante y menor que la unidad. A su vez, se consideran funciones lineales, $C = C_0 + cY$; es decir, dicha función consta de dos elementos: un consumo exógeno constante³, C_0 , y un consumo inducido por la renta, cY , siendo c una constante que actúa de factor de escala sobre la renta e indica una cierta propensión al consumo de esa renta.

La inversión es una variable exógena, I.

Sustituyendo y operando en todas estas ecuaciones, una vez logrado el equilibrio, se sigue:

$$Y = C_0 + cY + I$$

$$Y - cY = C_0 + I$$

$$(1 - c)Y = C_0 + I$$

$$Y = \frac{1}{1 - c}(C_0 + I)$$

$$Y = m(C_0 + I)$$

si se usa la notación $m = \frac{1}{1 - c}$.

Los montos de estos componentes, alcanzados al final del periodo anterior, son los que se consideran los iniciales para el siguiente periodo, que es el que se va a analizar.

Así pues, manteniendo el valor de estos componentes iniciales ($C_0 + I$), se añade exógenamente una inversión nueva (ΔI), cuyos efectos son los que se pretenden analizar.

³ Exógeno significa, en este modelo, que no proviene de la renta del periodo.

De las igualdades anteriores y tras incrementar sólo la inversión, se deduce que⁴:

$$\Delta Y = m \Delta I$$

La tasa de variación de la renta al final de esa adición es $\Delta Y/Y$, cuyo valor es:

$$\frac{\Delta Y}{Y} = \frac{m \Delta I}{Y} = \frac{m \Delta I}{m(C_0 + I)} = \frac{\Delta I}{C_0 + I}$$

que es lo que se quería demostrar.

3 Ilustración gráfica y demostración geométrica

La fórmula (o solución) recién deducida en la sección anterior se ilustra geoméricamente mediante la Figura 1.

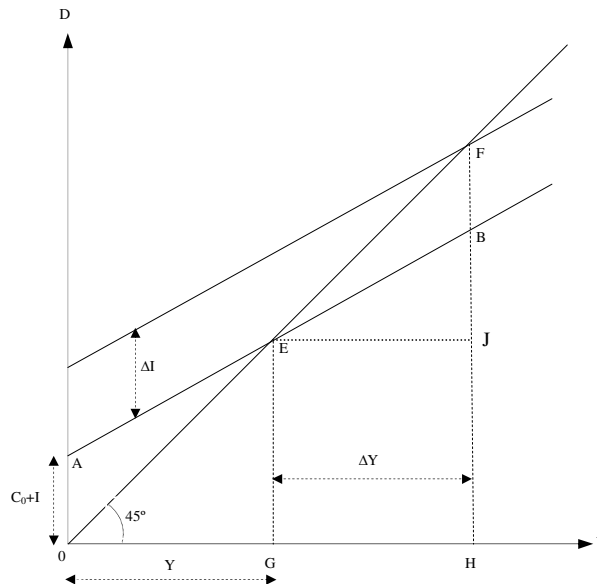


Figura 1: Incremento de la inversión y nuevo equilibrio, resultando que $\frac{\Delta Y}{Y} = \frac{\Delta I}{C_0 + I}$

Se denota por «E» el punto que marca la situación inicial y representa el último nivel de renta alcanzado, cuya condición de equilibrio es $Y=D$. Por tanto, se encuentra en la bisectriz del primer cuadrante: $OG=GE$.

Se traza por «E» cualquier recta de pendiente positiva y menor que la unidad (la cual representa la demanda agregada), de modo que el punto de corte con el eje de ordenadas es

⁴ Esta expresión coincide con la dada por Keynes (1989: 115).

«A». La pendiente de tal recta es «c». OA representa el nivel de consumo exógeno más la inversión exógena inicial (C_0+I) que permite que la renta, $Y=OG$, alcance en el equilibrio el volumen de la demanda agregada GE ($OG=GE$). Por lo tanto, la ecuación de tal recta es $Y=C_0+cY+I$.

El punto «A» es de lo más importante en esta demostración geométrica. Determinados «A» y «E» se conoce implícitamente la pendiente «c» de la recta que representa la demanda agregada (y de la de consumo). No obstante, esta «c», que se tomó arbitrariamente, es por completo irrelevante para el resultado de la demostración. Así pues, lo único decisivo es tener determinado el punto inicial de partida «A», pero para efectuar la demostración se necesita el apoyo de cualquier función lineal de demanda agregada.

Supongamos que, sobre la situación inicial, se origina un incremento exógeno de la inversión cuyo monto viene representado por el segmento BF, del cual resulta al final el incremento de la renta GH; de modo que $GH=EJ=JF$.

Debido a la proporcionalidad de los triángulos semejantes OAE y EBF se puede establecer la siguiente relación:

$$\frac{BF}{OA} = \frac{EF}{OE}$$

Los triángulos OGE y EJF también son semejantes, por lo que:

$$\frac{EF}{OE} = \frac{EJ}{OG} = \frac{GH}{OG}$$

En ambas parejas de triángulos semejantes se obtiene el mismo valor de proporcionalidad: $\frac{EF}{OE}$. Así pues:

$$\frac{GH}{OG} = \frac{BF}{OA}$$

Puesto que GH representa a ΔY , OG es el segmento que muestra el nivel de renta Y, BF es el representativo del incremento de la inversión (es decir, ΔI) y OA es la suma del consumo exógeno y la inversión inicial (esto es, C_0+I), la igualdad se convierte en la misma fórmula obtenida en la demostración analítica:

$$\frac{\Delta Y}{Y} = \frac{\Delta I}{C_0 + I}$$

A continuación, en la Figura 2 se muestra que esta solución no está supeditada a la pendiente de la función de demanda agregada, pese a haber ya quedado de manifiesto.

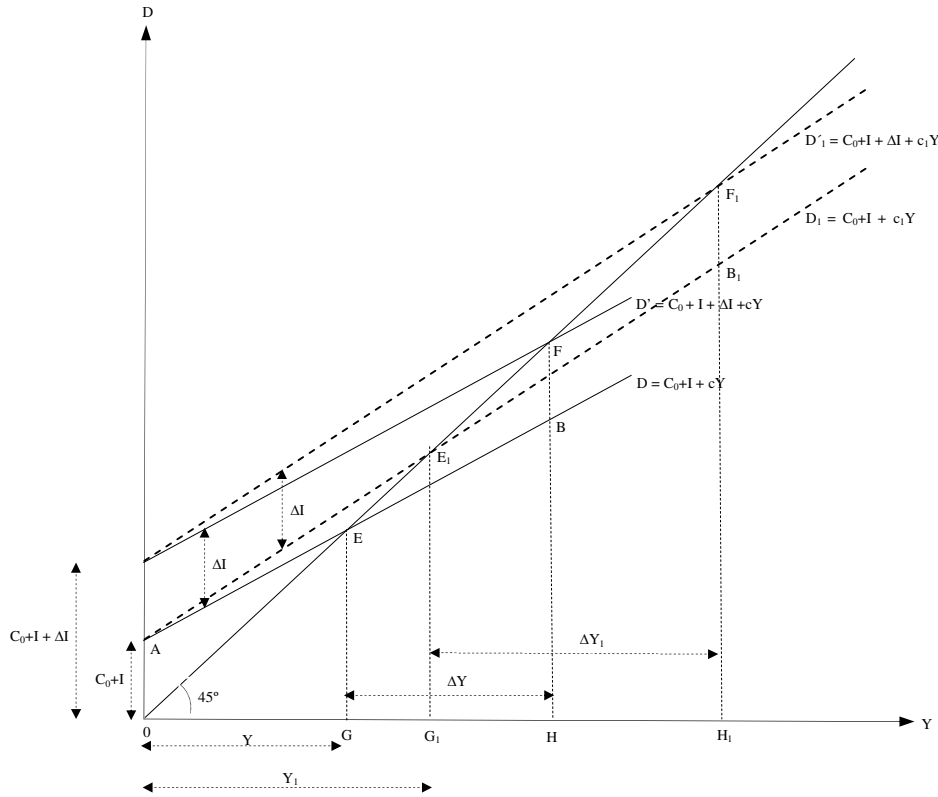


Figura 2: Independencia de la pendiente de la Función de Demanda, ya que $\frac{\Delta Y}{Y} = \frac{\Delta I}{C_0 + I} = \frac{\Delta Y_1}{Y_1}$

Para ello se consideran dos funciones que pasan por el punto «A» (que representa al valor C_0+I): la una con poca pendiente, $D=C_0+I+c_1Y$, y la otra con mayor pendiente que la anterior (mas siempre menor que la unidad), $D_1=C_0+I+c_1Y$, a las que se aplica el mismo incremento de la inversión (ΔI), y, en ambos casos, el resultado es el mismo.

En efecto: la primera de estas funciones conlleva un equilibrio inicial a un bajo nivel de PIB (OG) en comparación con la segunda, que implica un PIB más cuantioso (OG₁). Al incrementarse la inversión en idéntico monto para ambos supuestos (BF=B₁F₁), en el primer caso crece poco el PIB (GH), habiendo sido bajo su PIB inicial; mientras que en el segundo caso aumenta mucho el PIB (G₁H₁), siendo alto su PIB inicial. Al dividir sus respectivos incrementos de PIB por sus respectivos PIB iniciales, resulta que sus incrementos proporcionales son exactamente iguales (recuérdese que BF=B₁F₁):

$$\frac{GH}{OG} = \frac{BF}{OA} = \frac{G_1H_1}{OG_1} = \frac{B_1F_1}{OA}$$

En resumidas cuentas, la solución, no depende de la pendiente de la función de demanda agregada que, pudiendo ser cualquiera, no afecta a la solución, como ya se ha tenido

ocasión de apreciar en su fórmula:

$$\frac{\Delta Y}{Y} = \frac{\Delta I}{C_0 + I} = \frac{\Delta Y_1}{Y_1}$$

4 Interpretación

Si esta solución (o fórmula) puede parecer extraña es debido al rechazo inicial de la mente, acomodada a ideas preconcebidas, que suele bloquear lo desconocido (para lo que se requiere adaptación). Pero la ciencia matemática es una de las más lógicas y sus derivaciones son totalmente ciertas en la medida que se acepten las hipótesis de partida y no haya errores en el proceso deductivo. Si se asume que hay gastos exógenos, en consumo y en inversión⁵ (lo que parece tan evidente que no precisa demostración), y se acepta que de un periodo al siguiente una aproximación lineal es suficientemente satisfactoria, lo demás es resultado de la lógica matemática y sólo requiere interpretación.

La interpretación más inmediata de esta demostración es que lo relevante para la tasa de variación del PIB es el gasto exógeno agregado inicial ($C_0 + I$ en este modelo simple). Lo cual no desentona con lo dicho por Skidelsky (2009: 38): si «el PIB se espera que caiga un 5%, el Gobierno inyectará un 5% de gasto extraordinario en la economía». El fomento del gasto extraordinario debe ser tanto en consumo exógeno (ΔC_0) como en inversión exógena (ΔI); lo cual no significa que se niegue la existencia de un consumo e inversión inducidos por la renta generada en el periodo⁶. Lo que resulta notable, tras la demostración matemática, es que el gasto inducido no afecta al porcentaje de variación del PIB al considerar funciones lineales.

Otra interpretación concierne a los términos de la fórmula de la demostración. Es preciso recordar que C_0 e I quedaron prefijados al principio del periodo a considerar. Aunque siempre con carácter de prefijados, para su estudio se pueden contemplar valores altos o bajos de C_0 e I y comprobar su repercusión en la fórmula; pero, en el modelo, sólo se considera la

⁵Se aprovecha la introducción del término gasto para indicar que se entiende por «demanda» lo que se pide, es decir, una potencialidad o expectativa, y por «gasto» lo que se ha realizado. En la medida en que la expectativa se realice, ambos términos coinciden. Dando esto por supuesto, a veces se usan como sinónimos.

⁶ Hace años, hubo autores seguidores de Keynes, como Samuelson (1969: 272) y Lipsey (1968: 614-615), que tuvieron en cuenta la inversión inducida por la renta; pero en ediciones posteriores de sus manuales la suprimieron, dejando sólo la inversión exógena.

variación de I en el periodo, la cual está recogida en ΔI .

Por otra parte, se debe tener en cuenta que el resultado está vinculado a la condición de equilibrio. Por cierto, el efecto de la Contabilidad Nacional, que está diseñada bajo la perspectiva de la igualdad entre producto interior bruto y gasto interior bruto, es que en ella siempre se da la condición de equilibrio $D=Y$ (los errores y demás descuadres son corregidos en la rúbrica de variación de existencias). Su consecuencia es que siempre, en cualquier periodo y para cualquier cuantía, el cálculo estadístico del PIB da un punto, en su representación gráfica, que se halla sobre la bisectriz del primer cuadrante, y en él ha de cortar a la bisectriz cualquier recta representativa de la función de demanda agregada ($C+I$) y, por encima de ella, ha de figurar la nueva función de demanda agregada que, incorporando la adición del nuevo gasto exógeno en inversión (ΔI), proporciona un nuevo punto de equilibrio sobre la bisectriz.

Además, conviene resaltar las siguientes consideraciones:

- 1) El porcentaje de variación de la renta siempre será inferior a la variación porcentual de la inversión, ya que el denominador de la fórmula satisface $C_0+I > I$; lo que quiere decir que, en este modelo, la elasticidad de la renta respecto a la inversión ($\epsilon_{Y/I}$) siempre es rígida, o menor que la unidad: $\epsilon_{Y/I} = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta I}{I}} < 1$ (cuya demostración es inmediata a partir de la fórmula que se ha demostrado).

Esto significa que la sensibilidad de la renta ante variaciones de la inversión es moderada; es decir, para obtener un determinado porcentaje de aumento del PIB, se requiere incrementar la inversión en un porcentaje bastante más elevado que el deseado del PIB. Pero si consideráramos variables los gastos exógenos, tanto en consumo como en inversión, en $Y = C+I$, resultaría que: $\Delta Y = \Delta C_0 + \Delta I$ (siempre y cuando no haya dependencia funcional entre C_0 e I , cuya independencia parece plausible dado que quienes deciden consumir no suelen ser los mismos que deciden invertir, y, si en algún caso lo fueran, sus decisiones en ambos campos no se vinculan entre sí). Dividiendo ésta última ecuación por la anterior igualdad se tendrá:

$$\frac{\Delta Y}{Y} = \frac{\Delta C_0 + \Delta I}{C+I}.$$

De ahí que, si a la vez se consiguen subir todos los gastos exógenos (consumo e inversión, en este modelo simplificado), entonces el PIB ascendería en idéntica

proporción a la que lo hicieron todos los gastos exógenos. Así pues, en la medida que el aumento del empleo dependa del porcentaje de incremento del PIB, los elementos exógenos del gasto agregado son los que cobran importancia.

Acerca de esto conviene señalar que, en modelos que contemplaran otros gastos, pueden suscitarse incrementos de cualquiera de los componentes del gasto agregado sin que haya que limitarse exclusivamente a la inversión; es decir, tan importante es fomentar el consumo exógeno como la inversión, como las exportaciones, como la reducción de las importaciones.

- 2) El consumo exógeno puede jugar un papel moderador en las variaciones de la renta nacional ante alteraciones exógenas de la inversión, porque cuanto más alto sea el consumo exógeno tanto menor será la elasticidad de la renta respecto a la inversión y, por lo tanto, el porcentaje de variación de la renta nacional respecto a la variación porcentual de la inversión será tanto menor (tanto al alza como a la baja). Según esto, aplicando este principio ahora al caso de una recesión económica, cuando merma la inversión, nuevas prestaciones sociales cuantiosas y el fomento de más consumos exógenos (sanidad, educación, pensiones, seguros sociales, etc.) ayudan a moderar la intensidad de la crisis. También es verdad que al invertirse el ciclo se modera la magnitud del crecimiento económico medido en porcentaje de aumento del PIB (por eso, entre otras cosas, resulta dificultoso remontar una recesión). La norma, insistimos, es que, en todo caso, tanto en crisis como en auge, las fluctuaciones de la renta se moderan con un consumo exógeno elevado.
- 3) El razonamiento anterior es extensible, con idénticos resultados, a la inversión, cuyos valores de partida pueden ser altos o bajos. Está claro, que niveles altos de inversión también moderan las fluctuaciones de la renta nacional (al alza y a la baja), como así suele suceder en los países desarrollados (excepto cuando en una profunda crisis, la inversión se recorta drásticamente).
- 4) Del modelo simplificado y sus equivalencias ampliamente aceptadas y recogidas por la Contabilidad Nacional (CN), es posible extraer otra consideración. En ese modelo quedó establecida, mediante la condición de equilibrio, que la producción de un periodo es la suma del consumo más la inversión en dicho periodo, implantándose con nitidez un modelo económico de demanda (o de gasto agregado, pues, según lo que antes dijimos del efecto de la CN, ambos conceptos son siempre

iguales en la CN): $Y=D= C_1+I$. Por otra parte, también se considera que de la renta del periodo, lo que no se consume se ahorra, es decir, $Y= C_2+S$. Pero hay que dudar que verdaderamente al final de cada periodo (arbitrario de por sí), la economía esté en equilibrio y de que en esas dos igualdades sean equivalentes los respectivos consumos; antes bien se puede suponer que son distintos: $C_1 \neq C_2$, porque en C_1 intervienen partidas no provenientes de la renta del periodo, mientras que todo C_2 , por definición, está generado por dicha renta. Por eso, pudiendo ser uno mayor que el otro, no hay que descartar una insuficiencia de demanda para llegar al pretendido equilibrio, o a cualquier nivel de ocupación prefijado (o un exceso de la demanda que se resuelva con una inflación); y, en el caso de una insuficiencia, ésta es debida tanto al consumo como a la inversión (y también al tándem de las exportaciones e importaciones que también conviene introducir ampliando el modelo a estos componentes).

- 5) Además, es preciso recordar que del valor del PIB se retribuyen los factores de la producción en forma de salarios (W), beneficios del capital (B) y rentas de la propiedad (R), de modo que⁷: $Y=W+B+R$.

De esta equivalencia resulta evidente que $W < \text{PIB}$, por lo que, para superar la insuficiencia de la demanda, en aras de lograr un aumento del empleo, o sencillamente remontar una recesión económica aumentando el PIB, el peso del incremento del gasto agregado no debe esperarse de la elevación del consumo de los asalariados, porque los asalariados de un país jamás podrán adquirir todo lo producido en ese periodo con sus retribuciones del periodo y mucho menos consumir más para alcanzar el objetivo de elevar el PIB. Forzosamente, para retirar todo el PIB se necesita inyectar en el sistema económico íntegramente un montante equivalente al valor de $B+R$, y más si se precisa un crecimiento económico. Y resulta que esto es harto difícil de lograr, porque en cuanto alguien ahorre un poco sin dedicarlo a promover trabajo productivo, o llevando su ahorro a un paraíso fiscal, o donándolo en el extranjero, o consumiendo productos foráneos, prácticas habituales con dineros

⁷ Smith decía: «El precio total o el valor en cambio de aquel producto anual no puede por menos de resolverse necesariamente en esas tres partes, y distribuirse entre los habitantes del país, como salarios del trabajo, o como beneficios del capital, o como renta de la tierra» (Smith, 1994: 259). No obstante, algunos impuestos, como los indirectos (IVA), no forman parte de la retribución a esos tres factores de la producción; de ahí la distinción de PIB a coste de los factores y PIB a precios de mercado.

procedentes de los beneficios y las rentas (y los tributos, detraídos de la remuneración de los factores) o reduciendo los salarios, ya se ponen las bases para que la economía entre en crisis por insuficiencia de la demanda, que es sistémica y solo se palia temporalmente (hasta su estallido) mediante el consumo a crédito y el déficit público, los cuales causan inflación y ahorro forzoso (que merma el consumo de los perceptores de rentas fijas o cuya variación sea con apreciable retardo). El déficit público es un arma de doble filo, pues mientras por un lado fomenta la demanda (en consumo e inversión), a la vez, por el otro lado, suscita la inflación, que merma el gasto real (es decir, provoca un injusto ahorro forzoso⁸).

En definitiva, parece ser que el fomento de la iniciativa privada, fallando ésta, para promover el dinamismo económico, aumentando los gastos exógenos agregados (en consumo, inversión, sustitución de importaciones, o promoción de programas de investigación y desarrollo), debe provenir del sector público que ha de aplicar a ello todos los recursos posibles con eficaz gestión y controlando al máximo el déficit.

5 Conclusión

En lo que respecta a la tasa de variación del PIB, cuando se contempla un modelo lineal de determinación de la renta, lo único influyente es el gasto exógeno inicial, $C_0 + I$, y el incremento de los gastos autónomos, mientras que es absolutamente irrelevante la pendiente de la función de consumo.

Referencias

Keynes, J. M. (1933). *The Means to Prosperity*. Londres: Macmillan. Traducido al español bajo el título *Medios para la prosperidad*, en Keynes, J. M. (1997): *Ensayos de Persuasión*. Vol. II, Barcelona: Ediciones Folio, pp. 337-366.

Keynes, J. M. (1989 [1936]). *The General Theory of Employment, Interest and Money*. Londres: Macmillan.

Keynes, J. M. y Henderson, H. D. (1929). *Can Lloyd George Do it? The Pledge Examined*. Londres:

⁸ Según dijo Thornton (2000: 232): «Pero este ahorro, como todo ahorro adicional proveniente de un desfaldo de los ingresos de los miembros de la sociedad, irá acompañado de una privación e injusticia proporcionales».

- The Nation and Athenaeum. Traducido al español bajo el título *¿Puede hacerlo Lloyd George? Examen de la promesa*, en Keynes J. M. (1997): *Ensayos de Persuasión*. Vol. I, Barcelona: Ediciones Folio, pp. 96-133.
- Lipsey, R. G. (1968). *An Introduction to Positive Economics*. 2nd Edition. Londres: Weinsenfeld and Nicholson.
- Samuelson, P. A. (1969). *Curso de Economía Moderna*. Madrid: Aguilar Ediciones.
- Skidelsky, R. (2009). *El regreso de Keynes*. Barcelona: Crítica.
- Smith, A. (1994 [1776]). *Investigación sobre la naturaleza y causas de la riqueza de las naciones*. México: Fondo de Cultura Económica.
- Thornton, H. (2000 [1802]). *Crédito papel*. Madrid: Ediciones Pirámide.

Qualitative Analysis of the Goodwin Model of the Growth Cycle

SEREBRIAKOV, VLADIMIR

Department of Economics, Faculty of Business and Management
Brno University of Technology (Czech Republic)
E-mail: serebriakov@fbm.vutbr.cz

DOHNAL, MIRKO

Department of Economics, Faculty of Business and Management
Brno University of Technology (Czech Republic)
E-mail: dedicova@lib.vutbr.cz

ABSTRACT

Goodwin's model is a set of ordinary differential equations and is a well-known model of the growth cycle. However, its four constants require an extensive numerical study of its two differential equations to identify all possible unsteady state behaviors, i.e. phase portraits, which corresponds to infinitely many combinations of numerical values of the constants. Qualitative interpretation of Goodwin's model solves these problems by replacing all numerical constants and all derivatives by trends (increasing, constant and decreasing). The model has two variables — the employment rate V , and the labour share U . A solution of the qualitative Goodwin's model is a scenario. An example of a Goodwin's scenario is — *V is increasing more and more rapidly, U is decreasing and the decrease is slowing down.* The complete set of all possible 41 Goodwin's scenarios and 168 time transitions among them are given. This result qualitatively represents all possible unsteady state Goodwin's behaviours. It is therefore possible to predict all possible future behaviours if a current behaviour is known/chosen. A prediction example is presented in details. No prior knowledge of qualitative model theory is required.

Keywords: Goodwin model; business cycle; qualitative; scenario; transition.

JEL classification: E17.

MSC2010: 34C60.

Análisis cualitativo del modelo de Goodwin de ciclos de crecimiento

RESUMEN

El modelo de Goodwin es un conjunto de ecuaciones diferenciales ordinarias y resulta un modelo bien conocido para ciclos de crecimiento. Sin embargo, sus cuatro constantes requieren de un extenso estudio numérico de sus dos ecuaciones diferenciales para identificar todos los posibles comportamientos de estado no estacionario, i.e. retratos de fase, que corresponden a infinitamente muchas combinaciones de valores numéricos de las constantes. La interpretación cualitativa del modelo de Goodwin resuelve estos problemas reemplazando todas las constantes numéricas y todas las derivadas por tendencias (creciente, constante y decreciente). El modelo consiste en dos variables: la tasa de empleabilidad V y la repartición del valor agregado U . Una solución del modelo cualitativo de Goodwin es un escenario. Un ejemplo de escenario de Goodwin es el siguiente: *V es creciente cada vez más rápidamente y U es decreciente pero el decrecimiento se está ralentizando.* Se obtiene el conjunto completo de los 41 escenarios posibles de Goodwin con las 168 transiciones temporales entre ellas. Este resultado representa cualitativamente todos los posibles comportamientos de estado no estacionario de Goodwin. Por tanto, es posible predecir todos los comportamientos futuros posibles si un comportamiento actual es conocido o elegido. Un ejemplo de predicción es presentado en detalle. No se requiere ningún conocimiento previo de la teoría de modelos cualitativos.

Palabras claves: modelo de Goodwin; ciclo empresarial; cualitativo; escenario; transición.

Clasificación JEL: E17.

MSC2010: 34C60.



1. Introduction

Goodwin (1967) states a simple but useful model of the struggle between capital and labour for shares in national income, based on the classic Volterra–Lotka predator–prey model for fish populations. Since then, the model has been extended in many directions and as such has proved to be a useful framework for combining growth and cycles in a simple non-linear model (Desai *et al.*, 2006).

Some researchers intended to add additional variables to the original model in order to make the model more generalized and realistic, see e.g. Sordi and Vercelli, (2014) or Sportelli, (1995). Other researchers studied the stability and other properties of the original model (Yoshida and Asada, 2007; Cao and Jiang, 2011; Veneziani and Mohun, 2006). Those who tried to evaluate the model empirically (Weber, 2005; Harvie, 2000; Moura Jr. and Ribeiro, 2013) and investigate the cyclical behaviour, used the data available from different sources or just solved the model numerically.

However, a numerical solution of Goodwin’s model requires knowledge of all the four numerical constants. A qualitative interpretation of such parameters is possible and is studied in this paper. The paper presents a qualitative approach to the Goodwin model. Deep knowledge items reflect undisputed elements of the corresponding theory. The law of gravity is an example. This law has no exceptions. This is a typical feature of deep-knowledge items. Soft sciences as e.g. macroeconomics, are just very rarely based on deep-knowledge items. Goodwin’s model is based partially on shallow knowledge.

A shallow-knowledge item is usually a heuristic or a result of a statistical analysis of observations and has usually many exceptions, see e.g. Oliveira and Rezende, (2013) or Ahn and Kim (2009).

2. Qualitative algebra

The following description of Qualitative Algebra is sufficient to make this paper self-contained. For additional details see, e.g. Dohnal (1991), Kuipers, (1994), Salles *et al.* (2006) or Šuc *et al.* (2004).

Volterra-Lotka model is a set of ordinary nonlinear differential equations (ONDEs). A qualitative solution of ONDEs is specified if all its n qualitative variables:

$$x_1, x_2, \dots, x_n \quad (1)$$

are described by the corresponding qualitative triplets:

$$(X_1, DX_1, DDX_1), (X_2, DX_2, DDX_2), \dots, (X_n, DX_n, DDX_n), \quad (2)$$

where X_i is the i -th variable and DX_i and DDX_i are the first and second qualitative derivatives with respect to the independent variable t (which is usually time). DX is the qualitative interpretation of a numerical value for dx/dt . The reason why the third and higher derivatives are ignored is that these derivatives are known just rarely.

A qualitative model has m qualitative solutions i.e. scenarios. The j -th qualitative state is the n -triplet:

$$(X_1, DX_1, DDX_1), (X_2, DX_2, DDX_2), \dots, (X_n, DX_n, DDX_n)_j, \quad (3)$$

where $j = 1, 2, \dots, m$.

A simple algorithm can evaluate all possible transitions among the set of one-dimensional scenarios $n = 1$ (see (3)). One-dimensional transitions are based on the list of possible one-dimensional transitions, see Table 1. Multidimensional transitions must satisfy the Table 1 for n one-dimensional scenarios. However, this table is not a dogma. If a user feels/ knows/believes that a certain transition is not possible then this transition can simply be removed from the table.

Table 1: Some transition rules

	From		To	Or	Or	Or	Or	Or	Or
			a	b	c	d	e	f	g
1	+++	→	++0						
2	++0	→	+++	+-					
3	+-	→	++0	+0-	+00				
4	+0+	→	+++						
5	+00	→	+++	+-					
6	+0-	→	+-						
7	+-+	→	+-0	+0+	+00	0-+	00+	000	0-0
8	+ - 0	→	+-+	+-	0-0				
9	+-	→	+-0	0--	0-0				

An oriented graph is commonly used to represent graphically the set of all the transitions. If it is possible to transfer the r -th solution (3) into the s -th solution, then an oriented arc represents the corresponding transition from the node r to the node s .

A qualitative addition

$$X_i + X_j = X_s \quad (4)$$

is represented by the matrix shown in Table 2.

Table 2: Qualitative Addition

+	X_j			
		+	o	-
X_i	+	+	+	?
	o	+	o	-
	-	?	-	-

It is sometimes possible to find more than one qualitative value. It is impossible to predict a sign of the result:

$$(X_i=positive) + (X_j=negative) = (X_s=?) \quad (5)$$

A qualitative derivative of a sum of qualitative variables is a sum of their qualitative derivatives.

$$\begin{aligned} DX_i + DX_j &= DX_s \\ DDX_i + DDX_j &= DDX_s \end{aligned} \quad (6)$$

A qualitative multiplication

$$X_i * X_j = X_s \quad (7)$$

is described by Table 3.

Table 3: Qualitative Multiplication

*	X_j			
		+	o	-
X_i	+	+	o	-
	o	o	o	o
	-	-	o	+

A known relation for the first qualitative derivative gives

$$X_i * DX_i + X_j * DX_j = DX_s \quad (8)$$

ONDEs are interpreted as a set of qualitative differential equations and solved using qualitative additions and multiplications.

A multiplication by a qualitative constant c is irrelevant:

$$c \times X = 2 \times X = 2000 \times X = X \quad (9)$$

where c is a numerical constant and X is variable.

3. The qualitative Goodwin model of the growth cycle

The Goodwin model can be represented as follows (Desai *et al.*, 2006; Sordi and Vercelli, 2014):

$$\begin{aligned} \frac{\dot{v}}{v} &= \frac{1-u}{\sigma} - (\alpha + \beta) \\ \frac{\dot{u}}{u} &= -(\gamma + \alpha) + \rho v \end{aligned} \quad (10)$$

where numerical constants α, δ, γ and σ are considered to be positive; v is the employment rate and u is the labour share.

A numerical solution of the set of differential equations shown in (10) requires knowledge of all the four constants. Therefore a qualitative interpretation of the model represented by (10) generates a meaningful solution based on trends only.

A qualitative interpretation of Eq. (10) is, see (9):

$$\begin{aligned} DV + V \times U &= V \\ DU + U &= V \times U \end{aligned} \quad (11)$$

Algorithms used to solve qualitative models are combinatorial tasks and are not studied in this paper; for details see e.g. Dohnal (1991). Table 4 gives 41 scenarios, see (3), of the qualitative Goodwin model described by (10).

Table 4: List of all scenarios of Goodwin model

No	V	U	No	V	U
1	+++	+++	22	+0-	+++
2	+++	++0	23	+0-	++0
3	+++	++-	24	+0-	++-
4	+++	+0+	25	+ - +	++-
5	+++	+ - +	26	+ - +	+0-
6	+++	+ - 0	27	+ - +	+ - +
7	+++	+ - -	28	+ - +	+ - 0
8	++0	+++	29	+ - +	+ - -
9	++0	++0	30	+ - 0	++-
10	++0	++-	31	+ - 0	+0-
11	++0	+0+	32	+ - 0	+ - +
12	++0	+ - +	33	+ - 0	+ - 0
13	++-	+++	34	+ - 0	+ - -
14	++-	++0	35	+ - -	+++
15	++-	++-	36	+ - -	++0
16	++-	+0+	37	+ - -	++-
17	++-	+ - +	38	+ - -	+0-
18	+0+	+ - +	39	+ - -	+ - +
19	+0+	+ - 0	40	+ - -	+ - 0
20	+0+	+ - -	41	+ - -	+ - -
21	+00	+00			

There are 168 transitions among them. As an example, a set of 20 transitions among the scenarios included in Table 4, is given in Table 5.

Table 5: Some transitions of Goodwin model

No of transition	From	To
1	1	2
2	1	8
3	1	9
4	2	1
5	2	3
6	2	8
7	2	9
8	2	10
9	3	2
16	5	11
26	8	14
41	11	8
61	14	24
78	19	5
83	21	1
94	24	38
109	28	19
136	33	28
145	34	33
156	38	34

Weber (2005) simulates the behaviour of Goodwin's model by applying randomly selected values. As a result, a graph is achieved, which represents the cyclical behaviour of the model. One-dimensional transitions, given in Table 1, are used to identify some possible transitions among the scenarios set, see Table 4.

Let us suppose that the current scenario is the scenario No. 28. The following time sequence of scenarios is possible, see Table 4:

$$\begin{aligned}
 28 \xrightarrow{109} 19 \xrightarrow{78} 5 \xrightarrow{16} 11 \xrightarrow{41} 8 \xrightarrow{26} 14 \xrightarrow{61} 24 \xrightarrow{94} & \\
 38 \xrightarrow{156} 34 \xrightarrow{145} 33 \xrightarrow{136} 28 & \quad (12)
 \end{aligned}$$

The time sequence shown in (12) corresponds to the cyclical behaviour of the model in Weber (2005). The described behaviour is not the only possible one, keeping in mind that there are 168 possible transitions among 41 scenarios.

The following examples give an interpretation of qualitative Goodwin's results, see Tables 4 and 5. Different qualitative answers are presented. The first qualitative question to be answered is the following: Is there a steady state?; i.e. is there the following scenario (see $n = 2$ in (2))?:

$$\begin{array}{cc}
 V & U \\
 + 0 0 & + 0 0
 \end{array}$$

The answer is solved just by searching through Table 4. According to this table, scenario No. 21 is the qualitative steady state as both variables V and U are positive and have the first and the second derivatives equal to zero.

Scenario No. 1 (see Table 4) has the following triplets:

$$\begin{array}{cc}
 V & U \\
 + + + & + + +
 \end{array}$$

It means that both employment rate and labour share are increasing more and more rapidly. Scenario No. 21 is the steady state. The list of transitions (see Table 5) indicates that it is possible to transfer Scenario No. 21 into Scenario No. 1. It is impossible to

transfer Scenario No.1 into Scenario No. 21. For instance, if it is necessary to reach the steady state of the system, there are many paths being available. For example:

$$28 \xrightarrow{109} 19 \xrightarrow{78} 5 \xrightarrow{16} 11 \xrightarrow{41} 8 \xrightarrow{26} 14 \xrightarrow{58} 15 \xrightarrow{65} 21 \quad (13)$$

4. Conclusion

The well-known Goodwin's model is the first model which tried to combine cyclical behaviour and economic growth. Therefore a complex system with its negative features must be studied. It means that business cycle analysis must be done under the following conditions:

- Severe shortage of information
- High level of subjectivity of available knowledge
- Inconsistencies of information items of interdisciplinary nature

Throughout the years the model has been tested both theoretically and numerically. Although it was mentioned above that constants α, δ, γ and σ of the system (10) are hard to identify numerically, so this paper attempted to evaluate the model qualitatively and apply the achieved results into practice.

We have qualitatively described the trajectory of the cyclical behaviour of the Goodwin's model, studied in Weber (2005), by a sequence of one-dimensional scenarios and found out that this sequence represents only a subset of our qualitative model.

Our qualitative Goodwin model gives all the possible developments based on two variables. No statistical data sets are needed and all the possible solutions are identified.

References

- Ahn, H. and Kim, K. (2009): Bankruptcy prediction modelling with hybrid case-based reasoning and genetic algorithms approach. *Applied Soft Computing*, 9, 599-607.
- Cao, J. and Jiang H. (2011): Stability and Hopf bifurcation analysis on Goodwin model with three delays. *Chaos, Solitons & Fractals*, 44, 613-618.

- Desai, M.; Henry, B.; Mosley, A. and Pemberton, M. (2006): A clarification of the Goodwin model of the growth cycle. *Journal of Economic Dynamics & Control*, 30, 2661-2670.
- Dohnal, M. (1991): A methodology for common-sense model development. *Computers in Industry*, 16(2), 141–158.
- Goodwin, R. M. (1967): A Growth Cycle. In Feinstein, C.H. (ed.): *Socialism, Capitalism and Economic Growth*. Cambridge University Press, Cambridge, pp. 54-58.
- Harvie, D. (2000): Testing Goodwin: growth cycles in ten OECD countries. *Cambridge Journal of Economics*, 24, 349-376.
- Kuipers, B.J. (1994): *Qualitative reasoning. Modelling and simulation with incomplete knowledge*. MIT Press, Cambridge.
- Moura, Jr. N.J. and Ribeiro, M.B. (2013): Testing the Goodwin growth-cycle macroeconomic dynamics in Brazil. *Physica A*, 392, 2088-2103.
- Oliveira, R.D.S. and Rezende, A.C. (2013): Global phase portraits of a SIS model. *Applied Mathematics and Computation*, 219, 4924-4930.
- Salles, P.; Bredeweg, B. and Araújo, S. (2006): Qualitative models about stream ecosystem recovery: Exploratory studies. *Ecological Modelling*, 194, 80–9.
- Sordi, S. and Vercelli, A. (2014): Unemployment, income distribution and debt-financed investment in a growth cycle model. *Journal of Economic Dynamics & Control*, 48, 325-348.
- Sportelli, M.C. (1995): A Kolmogoroff generalized predator–prey model of Goodwin’s growth cycle, *Zeitschrift für Nationalökonomie*, 61: 1, 35–64.
- Šuc, D.; Vladušić, D. and Bratko, I. (2004): Qualitatively faithful quantitative prediction. *Artificial Intelligence*, 158, 189–214.
- Veneziani, R. and Mohun, S. (2006): Structural stability and Goodwin’s growth cycle. *Structural Change and Economic Dynamics*, 17, 437-451.
- Weber, L. (2005): A contribution to Goodwin’s growth cycle model from a system dynamics perspective. In *Proceedings of the 23rd International System Dynamics*

Conference, System Dynamics Society, Boston, July 17-21, 28pp. Available at <http://www.systemdynamics.org/conferences/2005/proceed/papers/WEBER196.pdf>.

Yoshida, H. and Asada, T. (2007): Dynamic analysis of policy lag in a Keynes-Goodwin model: Stability, instability, cycles and chaos. *Journal of Economic Behaviour & Organization*, 62, 441-469.

Elección contable para la valoración de las inversiones inmobiliarias. Contribución de las técnicas de minería de datos para determinar patrones de decisión

DE VICENTE LAMA, MARTA

Departamento de Economía Financiera y Contabilidad
Universidad Loyola Andalucía (España)
Correo electrónico: mvicente@uloyola.es

MOLINA SÁNCHEZ, HORACIO

Departamento de Economía Financiera y Contabilidad
Universidad Loyola Andalucía (España)
Correo electrónico: hmolina@uloyola.es

RAMÍREZ SOBRINO, JESÚS N.

Departamento de Economía Financiera y Contabilidad
Universidad Loyola Andalucía (España)
Correo-e: jramirez@uloyola.es

TORRES JIMÉNEZ, MERCEDES

Departamento de Métodos Cuantitativos
Universidad Loyola Andalucía (España)
Correo-e: mtorres@uloyola.es

RESUMEN

La normativa contable internacional ofrece con la Norma Internacional de Contabilidad 40 (NIC 40) “Inversiones inmobiliarias” un caso referente para investigar la decisión que toman las empresas cuando se les ofrece el valor razonable o el coste histórico como criterios alternativos de valoración. En este trabajo aprovechamos la oportunidad que ofrece esta norma para aportar evidencia adicional en un contexto multinacional y multisectorial sobre cuáles son los motivos que explican la elección contable. Además, en este trabajo introducimos y comparamos el uso de las redes neuronales artificiales y los árboles de decisión, con el objetivo de evaluar la capacidad predictiva de estas metodologías, frente a la tradicionalmente utilizada regresión logística para la resolución de problemas de clasificación en este área. Los resultados de la clasificación indican que tanto las redes neuronales como los árboles de decisión pueden ser una alternativa interesante a los métodos clásicos estadísticos como la regresión logística. En particular, las dos metodologías mostraron una mayor capacidad predictiva frente a la regresión logística aunque no se encontraron diferencias significativas entre ambas.

Palabras claves: elección contable; valor razonable; NIIF; redes neuronales; árboles de decisión.

Clasificación JEL: M10.

MSC2010: 62P20; 90B15.

Artículo recibido el 9 de febrero 2015 y aceptado el 23 de mayo de 2017.

Accounting Choice for Measuring Investment Properties. Data Mining Techniques Contribution to Determine Decision Patterns

ABSTRACT

International Accounting Standard 40 (IAS 40 - Investment properties) offers an ideal setting for research on accounting choice as it represents a paradigmatic case choosing between the fair value and the historical cost as the measurement criteria. In this paper, we take the opportunity of this standard to provide additional evidence in a multinational and multi-context on the determinants that explain the accounting choice. Furthermore, in this paper, we introduce and compare the use of artificial neural networks and decision trees in order to assess the predictive capability of these methodologies, compared to other techniques commonly used to solve classification problems in this area such as the logistic regression. The classification results indicate that both neural networks and decision trees can be an interesting alternative to classical statistical methods such as the logistic regression. In particular, both methods outperformed the logistic regression in terms of predictive ability, although no significant differences were found between both.

Accounting choice; fair value; IFRS; neural networks; decision trees.

Clasificación JEL: M10.

MSC2010: 62P20; 90B15.



1. INTRODUCCIÓN

Con la Norma Internacional de Contabilidad (NIC) 40 “Inversiones Inmobiliarias” (IASB, 2013)¹ el *International Accounting Standards Board* (IASB) introduce, por primera vez para un activo no financiero, la posibilidad de emplear el valor razonable en la valoración posterior de las inversiones inmobiliarias con contrapartida el resultado del ejercicio. Esta elección contable supone para las empresas optar entre el reconocimiento del valor razonable de este tipo de activos en los estados financieros principales (modelo del valor razonable) o su revelación en la Memoria (modelo del coste)² lo que hace que la NIC 40 sea un caso referente para investigar patrones de decisión.

Por este motivo, varios autores han centrado su atención en determinar qué factores explican la elección contable que, de acuerdo con la NIC 40, toman las empresas para este tipo de activos (Muller *et al.*, 2011; Quagli y Avallone, 2010; Christensen y Nikolaev, 2013, De Vicente *et al.*, 2013). En cuanto a la metodología empleada, estos trabajos utilizan la regresión logística (RL) como técnica de análisis con carácter explicativo.

En la normativa estadounidense, el *Financial Accounting Standards Board* (FASB) ha tenido durante varios años en su agenda el desarrollo de una nueva norma que determine si bien se permite (de forma similar a la NIC 40) o bien se requiere a las empresas seguir el modelo del valor razonable, proyecto que no ha culminado. Por tanto, ampliar la investigación en este ámbito es de interés directo para el regulador contable estadounidense, como apunta Linsmeier (2013), para comprender los motivos que explican la opción que podrían tomar las empresas entre un criterio de valoración u otro en caso de que el FASB se decantara por la opcionalidad contable de la NIC 40.

La elección contable que ofrece la NIC 40 es una de las pocas alternativas contables libres que ofrece la normativa internacional a los preparadores de la información. En este trabajo aprovechamos la oportunidad que ofrece esta norma y aplicamos diferentes

¹ Para los grupos cotizados europeos entró en vigor el 1 de enero de 2005 con la adopción por parte de la Unión Europea de las Normas Internacionales de Información Financiera (NIIF),

² Cuando una empresa opta por el modelo del valor razonable como política contable, para lo cual se requiere que este pueda medirse de forma fiable, valorará las inversiones inmobiliarias a valor razonable entendido este como “el precio que se recibiría por vender un activo o que se pagaría por transferir un pasivo en una transacción ordenada entre participantes de mercado en la fecha de la medición” (NIC 40.5). Por el contrario, la opción por el modelo del coste implica valorar las inversiones inmobiliarias aplicando los requisitos establecidos en la NIC 16 – Propiedad, planta y equipo para este modelo, es decir, a coste histórico menos amortización acumulada menos el importe acumulado de las pérdidas por deterioro de valor.

metodologías de clasificación que ofrece la minería de datos con el objetivo de predecir la opción contable que tomarían las empresas. En particular, en este trabajo introducimos el uso de las redes neuronales artificiales, en concreto el perceptrón multicapa (MLP) y los árboles de decisión (algoritmo C4.5), con el objetivo de evaluar el rendimiento de estas metodologías, frente a la tradicionalmente utilizada RL para la resolución de problemas de clasificación en trabajos de elección contable. Adicionalmente, para garantizar la capacidad predictiva de los resultados obtenidos y a diferencia de los trabajos anteriores, validamos estos mediante muestras de test (*out-of-sample validation*). Para analizar la significación de las diferencias obtenidas en los resultados alcanzados con cada metodología (RL, MLP y C4.5) empleamos un test sobre muestras relacionadas.

Cabe añadir que, excepto en el caso de Christensen y Nikolaev (2013) y De Vicente *et al.* (2013), los trabajos anteriores directamente relacionados con este estudio, se centran únicamente en empresas del sector inmobiliario y en países europeos donde la normativa local previa a la entrada en vigor de las Normas Internacionales de Información Financiera (NIIF) no permitía el uso del valor razonable como criterio de valoración; fundamentalmente, para evitar el efecto de la experiencia pasada en la elección contable (Quagli y Avallone, 2010, p. 463). En este trabajo, empleamos una muestra de empresas pertenecientes a distintos sectores de actividad en el Reino Unido y en España, dos de los mercados inmobiliarios europeos más dinámicos a la entrada en vigor de las NIIF (Ball, 2005) y, por tanto, países en los que el impacto de la valoración a valor razonable en los estados financieros pudiera ser mayor que en otros países europeos. Adicionalmente estos países tienen una tradición contable contrapuesta para la valoración de las inversiones inmobiliarias: en el Reino Unido la normativa local anterior a la entrada en vigor de las NIIF requiere que las inversiones inmobiliarias se valoren a valor razonable mientras que en España la normativa local obliga a valorarlas a coste histórico. Este contexto nos permite aportar evidencia adicional a esta línea de investigación analizando la decisión que toman las empresas cuando la norma permite libremente optar entre dos atributos de valoración alternativos mientras que la normativa local previa requería el uso de uno de ellos. En este sentido, trabajos previos como el de Kvaal y Nobes (2010) han puesto de manifiesto que las diferencias en la normativa contable local entre distintos países han provocado la existencia de un “patrón nacional” en la adopción posterior de las NIIF. Esta es una cuestión importante a investigar y de utilidad para los reguladores contables pues la distinta aplicación de las normas, sobre todo en los casos en los que éstas permiten criterios alternativos, afecta a la comparabilidad internacional de la información financiera.

El objetivo por tanto de este trabajo es doble. En primer lugar, analizar y comparar la utilidad y rendimiento de otras técnicas de minería de datos distintas de la RL en una de las

áreas clásicas de investigación en contabilidad: las elecciones contables. En segundo lugar, el objetivo de este estudio es contribuir a la literatura contable aportando evidencia adicional en un contexto multinacional y multisectorial sobre cuáles son los motivos que explican la elección entre valor razonable y coste histórico para la valoración de las inversiones inmobiliarias.

Para lograr estos objetivos, este trabajo se ha estructurado en 4 secciones. En primer lugar, en la sección 2 se presentan los resultados obtenidos por los trabajos más representativos en la investigación sobre elección contable relacionados con el presente estudio y se revisan algunos de los métodos de clasificación que han sido empleados en este área. En la sección 3 describimos la población objeto de estudio, definimos las variables utilizadas y describimos las técnicas estadísticas que aplicamos en nuestro estudio empírico. A continuación, presentamos y discutimos los principales resultados obtenidos en la sección 4. Por último, en la sección 5 resumimos los hallazgos principales de la investigación y aportamos nuestras conclusiones finales.

2. REVISIÓN DE LA LITERATURA

Elección contable para la valoración de las inversiones inmobiliarias

La literatura que ha analizado los factores que explican las decisiones contables es muy extensa. Fields *et al.* (2001) ofrecen un amplio repaso de la investigación empírica sobre elección contable de los años 90 y clasifican en tres categorías los factores que han resultado ser explicativos de las decisiones contables que son la existencia de costes de agencia, las asimetrías informativas y el comportamiento oportunista de los directivos.

Directamente relacionados con la elección contable entre el modelo del coste y el modelo del valor razonable que ofrece la NIC 40, los trabajos de Quagli y Avallone (2010) y Muller *et al.* (2011) se centran exclusivamente en empresas del sector inmobiliario domiciliadas en países donde la normativa local anterior a la adopción de las NIIF no permitía el uso del valor razonable para la valoración de las propiedades de inversión. Los resultados obtenidos por Quagli y Avallone (2010) sugieren que todas las razones que han sido descritas por la literatura previa sobre elección contable (véase Fields *et al.*, 2001), influyen en la decisión que toman las empresas para la contabilización de las inversiones inmobiliarias. Sin embargo, no encuentran una asociación significativa entre variables como el endeudamiento y la elección contable. Por su parte, Muller *et al.* (2011) encuentran que las empresas cuyos estados financieros son auditados por una de las BIG4 y las que presentan menor dispersión geográfica de sus operaciones son más propensas a proporcionar, de forma voluntaria, el valor razonable de las

inversiones inmobiliarias bien en los estados financieros principales o bien en la memoria. No obstante, la estructura de propiedad no resulta estar asociada con la elección contable.

Por otra parte, y recientemente, Christensen y Nikolaev (2013) analizan la decisión que toman las empresas que operan en distintos sectores de actividad en Alemania y el Reino Unido y De Vicente *et al.* (2013) en España. Los primeros encuentran que la opción por el modelo del valor razonable es más frecuente para las inversiones inmobiliarias que para el resto de activos no financieros, como propiedad, planta y equipo. Sus resultados ponen de manifiesto además que las empresas del Reino Unido (Alemania), las que operan en el sector inmobiliario (otros sectores), las de menor (mayor) tamaño así como las más (menos) endeudadas son más (menos) propensas a optar por el modelo del valor razonable de la NIC 40. Por su parte, De Vicente *et al.* (2013) encuentran en España que la importancia de las inversiones inmobiliarias en el balance de situación, el impacto del valor razonable sobre el total de activos y la naturaleza del auditor son factores explicativos de la elección contable.

En este trabajo, estrechamente relacionado con el trabajo de Christensen y Nikolaev (2013), aprovechamos la oportunidad que ofrece la NIC 40 para profundizar en el estudio de los factores que influyen en las decisiones contables en un entorno multisectorial y multinacional pero incorporamos otras variables que no han resultado ser explicativas de la elección contable en trabajos anteriores y mejoramos la definición de algunas de las variables utilizadas por estos autores.

Enfoque metodológico en la investigación sobre elección contable

La mayoría de los trabajos empíricos sobre elección contable utilizan la regresión logística para explicar los motivos de la elección (Gaeremynck y Veugelers, 1999; Lin y Peasnel, 2000; Hervás Oliver, 2005; Barlev *et al.*, 2007; Missonier-Piera, 2007; Muller *et al.*, 2011; Quagli y Avallone, 2010; Christensen y Nikolaev, 2013; De Vicente *et al.*, 2013;). Otras metodologías utilizadas en menor medida son los modelos probit (Brown *et al.*, 1992; Whittred y Chan, 1992) y la regresión lineal simple (Cotter y Zimmer, 1995; Cotter, 1999). Trabajos previos como el de Stone y Rasp (1991) ponen de manifiesto que, aún en investigaciones donde el tamaño de la muestra analizada es pequeño (menor de 50), es preferible el uso de la regresión logística frente a la regresión lineal simple.

Por su parte, otras técnicas de minería de datos en el área de la inteligencia artificial, como los árboles de decisión o las redes neuronales, son frecuentemente utilizadas en el ámbito financiero para resolver problemas de clasificación, mostrando su rendimiento y superando con frecuencia el de otras técnicas clásicas como la regresión logística o el análisis discriminante.

En el ámbito contable y financiero se han empleado con frecuencia para predecir la insolvencia empresarial (véase en la revisión de la literatura realizada por Sun *et al.*, 2014), su liquidez (De Andrés *et al.*, 2010), para determinar las características que explican los rating de las empresas aseguradoras (Florez-Lopez, 2007) o para predecir el fraude empresarial (Kirkos *et al.*, 2007), y en el ámbito de la auditoría, para predecir la opinión del auditor sobre el principio de empresa en funcionamiento (Koh, 2004; Martens *et al.*, 2008) o para predecir la elección del tipo de auditor (Kirkos *et al.*, 2008). La principal ventaja de estas técnicas es que no imponen a priori restricciones sobre las características de los datos, la distribución de las variables o sus propiedades métricas y, por tanto, se adecúan más a las características de la información financiera.

En este trabajo evaluamos el rendimiento de otros métodos de clasificación (redes neuronales artificiales, en concreto perceptrón multicapa MLP y árboles de decisión, C4.5) en problemas de elección contable. Además, los estudios anteriores que investigan la elección contable para las inversiones inmobiliarias utilizan la RL, un método clásico usado para distintas aplicaciones de clasificación binaria, pero lo hacen para la totalidad de la muestra por lo que podría existir un sobreajuste del modelo que reduciría la robustez de los resultados. En este estudio validamos los resultados obtenidos en cada uno de modelos aplicados dividiendo la muestra en submuestras de entrenamiento y generalización lo que asegura la validez de los resultados e incrementa su potencial generalización.

3. METODOLOGÍA

3.1. Muestra

En nuestra investigación hemos obtenido las cuentas anuales consolidadas de todos los grupos españoles y británicos (excluyendo sector financiero y asegurador) que cotizan en el mercado continuo en 2005 y cuyo marco normativo de aplicación son, por tanto, las NIIF. Posteriormente, tras su lectura y análisis, identificamos un total de 106 empresas que mantienen registradas en su balance de situación inversiones inmobiliarias en el ejercicio 2005 (42 grupos españoles y 64 grupos británicos). Por tanto, la población objeto de estudio incluye el 100% de los grupos cotizados españoles y británicos que mantienen inversiones inmobiliarias e incluye tanto los grupos que operan en el sector inmobiliario como los que operan en otros sectores. Sin embargo, 3 de las empresas que componen la muestra (2 españolas y 1 británica) presentaban ratios de endeudamiento negativos al presentar cifras de patrimonio neto negativo, por lo que fueron excluidas del análisis. Así, la muestra final se compone de 103 empresas. Toda la

información necesaria de acuerdo con las variables definidas para el contraste de las hipótesis ha sido obtenida directamente de las cuentas anuales.

3.2. Selección y definición de variables

La variable dependiente que empleamos en este estudio es la elección del modelo contable (ELECCION) para la presentación de las inversiones inmobiliarias en las primeras cuentas anuales consolidadas emitidas por las empresas tras la entrada en vigor de las NIIF. Esta variable toma dos valores: 1 si la empresa opta por el modelo del valor razonable y 0 si opta por el modelo del coste. De acuerdo con la evidencia que proporciona la literatura previa seleccionamos las siguientes variables explicativas:

País de origen

España y el Reino Unido son países con tradiciones contables contrapuestas respecto al empleo del valor razonable en la normativa local anterior a la adopción de las NIIF. Ball (2006) identifica que las diferencias en el sistema legal, fiscal y financiero entre países explican también las diferentes políticas contables. Este trabajo se concentra únicamente en dos países por lo que no es posible distinguir los factores institucionales del efecto de la normativa contable previa en la posterior elección; no obstante, al igual que Christensen y Nikolaev (2013) esperamos que las empresas domiciliadas en el Reino Unido sean más propensas a elegir el valor razonable para la valoración de sus inversiones inmobiliarias que las empresas españolas.

Definimos la variable PAÍS que toma el valor 1 si la empresa está domiciliada en el Reino Unido y el valor 0 si está domiciliada en España.

Importancia de las inversiones inmobiliarias en el balance de situación

La proporción de las inversiones inmobiliarias sobre el total de activos de la empresa es claramente mayor en el sector inmobiliario que en otros sectores de actividad. Nuestro trabajo se lleva a cabo en un entorno multisectorial por lo que es de esperar que las empresas del sector inmobiliario sean más propensas a elegir el modelo del valor razonable que las empresas que operan en otros sectores, tal y como ha puesto de manifiesto la literatura previa (p.e. Quagli y Avallone, 2010; Christensen y Nikolaev, 2013; De Vicente *et al.*, 2013), pues este modelo proporciona información relevante a los usuarios para evaluar empresas en las que las inversiones inmobiliarias son un componente muy significativo del total de activos de la empresa.

Definimos la variable INV_TA que obtenemos como el cociente de las inversiones inmobiliarias entre la cifra total de activos al final del ejercicio, ambos importes ajustados, en su caso, por las variaciones en el valor razonable.

Endeudamiento

Los analistas de riesgos de las entidades financieras utilizan la información contable para analizar la situación financiera de las empresas y determinar el riesgo que asumen al conceder financiación. Quagli y Avallone (2010, p. 465) argumentan que las entidades financieras preferirán la opción más conservadora (modelo del coste histórico) para evitar la distribución del valor de la empresa vía dividendos. Sin embargo, sus resultados no evidencian una relación significativa entre el endeudamiento y la elección contable. Por otra parte, Christensen y Nikolaev (2013) argumentan que el reconocimiento de las inversiones inmobiliarias a valor razonable no conlleva costes incrementales elevados para la obtención de valores de mercado fiables ya que las entidades financieras requieren la tasación de los inmuebles para conceder la financiación. Por este motivo, plantean que las empresas más endeudadas serán más propensas a optar por el modelo del valor razonable. Los resultados muestran el endeudamiento como factor determinante en la elección del criterio de valor razonable para el registro de las inversiones inmobiliarias y, especialmente, el endeudamiento a corto plazo. Adicionalmente, a la entrada en vigor de las NIIF en 2005, el ciclo inmobiliario se encuentra en su fase alcista por lo que el registro de las inversiones inmobiliarias a valor razonable supone incrementar el activo y el patrimonio neto reduciendo el ratio de endeudamiento y ofreciendo, por tanto, una imagen de menor riesgo financiero. Por estos motivos, esperamos que las empresas con mayores niveles de endeudamiento serán más propensas a optar por el valor razonable.

Definimos la variable END como el cociente entre el total pasivo y el patrimonio neto ajustado este último por la variación en el valor razonable de las inversiones inmobiliarias durante el ejercicio³.

³ Quagli y Avallone (2010) utilizan como medida del endeudamiento el ratio total pasivo entre total activo sin ajustar por la variación en el valor razonable durante el ejercicio por lo que la medida puede estar influenciada por la propia decisión contable. Por otra parte, Christensen y Nikolaev (2013) incluyen la variable total pasivo entre total de activos a valor de mercado, obtenidos estos últimos como el valor contable de los pasivos más el valor de mercado del patrimonio neto. En este caso, la decisión contable no influiría en la propia medida del endeudamiento asumiendo que el mercado incorpora en su valoración el valor razonable de las inversiones inmobiliarias, tanto si la empresa ha optado por el modelo del valor razonable (reconocido) como por el modelo del coste (revelado en las notas de la Memoria).

Concentración de la propiedad

Los países en los que los derechos de los inversores externos se encuentran escasamente protegidos por el sistema legal, como los de influencia continental europea entre los que se encuentra España, cuentan con una mayor concentración de la propiedad como mecanismo de defensa de los accionistas (La Porta *et al.*, 1998). Por otra parte, La Porta *et al.* (1998) muestran que los países regidos por el derecho común (*common-law*) como el Reino Unido presentan estructuras de propiedad más dispersas.

La influencia de la estructura de propiedad en la elección contable para activos no financieros se analiza únicamente en Muller *et al.* (2011) y en su *working paper* de 2008 (Muller *et al.* 2008). Muller *et al.* (2008, p. 15) argumentan que, ante estructuras de propiedad concentradas, los accionistas preferirán una menor volatilidad de los resultados (modelo del coste histórico) frente a la relevancia que aporta el valor razonable a la información financiera porque las asimetrías de información, de existir, las pueden resolver por otras vías. Los resultados confirman la hipótesis planteada. Sin embargo, la estructura de propiedad no parece influir en la decisión de reconocer o revelar en las notas de la memoria de forma voluntaria el valor razonable de las inversiones inmobiliarias (Muller *et al.*, 2011). De acuerdo con lo anterior, esperamos que la propensión a optar por el valor razonable será menor ante estructuras de propiedad más concentradas.

Definimos las variables CONC como una variable *dummy* que toma el valor 1 cuando se produce concentración y 0, si no la hay. Siguiendo a La Porta *et al.* (1999, p. 476), consideramos que la estructura de propiedad presenta concentración si algún accionista posee directa o indirectamente el 20% o más de las acciones ordinarias.

Variables de control

De acuerdo con la literatura previa, incluimos en nuestro modelo una variable adicional que consideramos puede influir en la elección contable como es el tamaño de la empresa. Los resultados obtenidos en trabajos anteriores ponen de manifiesto que las empresas de menor tamaño son más propensas a optar por el modelo del valor razonable (Quagli y Avallone, 2010; Christensen y Nikolaev, 2013).

Definimos la variable tamaño (TAM) como el logaritmo natural del importe total de ventas en lugar de tomar la cifra de activos, como hacen Quagli y Avallone (2010), para evitar la influencia de la propia elección contable en la variable. Trabajos previos que analizan las razones por las que las empresas deciden revalorizar sus activos no financieros han utilizado el logaritmo natural de la cifra de negocios como medida de aproximación al tamaño de la

empresa (Lin y Peasnell, 2000; Hervás Oliver, 2005 y Missonier-Piera, 2007). De acuerdo con los resultados obtenidos en trabajos anteriores, esperamos una relación negativa entre el tamaño de la empresa y la opción por el modelo del valor razonable.

En nuestro modelo no consideramos algunas variables que han resultado estar asociadas con la elección contable en trabajos previos tales como el ratio *market-to-book* (Quagli y Avallone, 2010). La decisión para no incluir esta variable en nuestro modelo es la diferente composición de nuestra muestra que abarca no sólo empresas del sector inmobiliario sino también empresas que operan en otros sectores de actividad por lo que diferencias en este indicador podrían deberse a otros motivos distintos del modelo contable elegido para la contabilización de las inversiones inmobiliarias.

3.3. Técnicas de análisis

En este trabajo empleamos dos metodologías alternativas a la RL para resolver problemas de clasificación: redes neuronales artificiales y árboles de decisión. El software que se ha empleado para desarrollar los modelos ha sido WEKA 3.6.12.

Para evitar que la separación de la muestra en los grupos de entrenamiento y generalización afecte a los resultados se utilizó el método de validación cruzada de k iteraciones o, como comúnmente se denomina en inglés *k-fold cross validation*. Los datos de muestra se dividen en k subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba o generalización y el resto (k-1) como datos de entrenamiento. El proceso de validación cruzada es repetido durante k iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. Este método es muy preciso puesto que evaluamos a partir de k combinaciones de datos de entrenamiento y de prueba. En la práctica, la elección del número de iteraciones depende de la medida del conjunto de datos. En este caso utilizamos la validación cruzada de 10 iteraciones ya que es el número de iteraciones más utilizado en la práctica pues garantiza la validez de los resultados sin generar un coste computacional demasiado elevado

Regresión logística (RL)

En los modelos de regresión logística la variable dependiente es siempre una variable definida como categórica. En este trabajo, utilizamos el modelo de regresión logística binaria puesto que nuestra variable dependiente es la elección contable y puede tomar únicamente dos valores (valor 1 si la elección es por el modelo del valor razonable y 0 por el modelo del coste).

El modelo de regresión logística implica la obtención de la probabilidad de que una observación pertenezca a un grupo determinado, en función del comportamiento de las variables independientes. El modelo se formula como sigue:

$$\text{Logit}(p) = \ln(p/1-p) = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n \quad [1]$$

donde $(p/1-p)$ es la razón de oportunidades (*odd ratio*) o ventaja de que un suceso ocurra y $\ln(p/1-p)$ es la transformación logarítmica de la razón de *odd*; x_n son las variables independientes y b_n los coeficientes obtenidos por el modelo para cada una de las variables independientes.

Redes Neuronales Artificiales (MLP)

En este trabajo aplicamos las redes neuronales en su formulación clásica determinada por el perceptrón multicapa (MLP), con una sola capa oculta y funciones de activación de las neuronas de tipo sigmoide. Las neuronas de la capa de entrada estarían constituidas por las variables explicativas en el modelo de regresión logística y la capa de salida por la variable que representa la elección contable. Se seleccionó una sola capa oculta porque como ha demostrado la teoría (Cybenko, 1989; Hornik, 1991) estas redes son aproximadores universales y pueden representar cualquier función de sus entradas. Con respecto al número de neuronas a incluir en la capa oculta se determinó experimentalmente hasta encontrar el menor número de neuronas en la capa oculta que generaba buenos resultados, que resultó ser igual a 4.

Árboles de decisión (C4.5)

La principal ventaja de los árboles de decisión es que generan reglas de clasificación muy intuitivas en cuanto a la interpretación de los resultados obtenidos y que, además, indican la importancia relativa de las variables seleccionadas para la clasificación y predicción (Koh, 2004).

En este trabajo utilizamos el algoritmo de minería de datos C4.5 de Quinlan (2005) que en el software Weka se ha implementado como J48. El algoritmo C4.5 es una extensión del algoritmo ID3, el cual sólo trabaja con variables que toman valores discretos; en cambio, el algoritmo C4.5 permite trabajar con valores tanto discretos como continuos. C4.5 genera un árbol de decisión mediante particiones realizadas de manera recursiva para lo cual selecciona aquellas variables que maximizan el criterio de separación. El criterio de separación empleado por C4.5 a la hora de realizar un test es la razón de ganancia.

4. RESULTADOS

4.1. Análisis descriptivo

En la Tabla 1 se ofrecen los principales estadísticos descriptivos que describen la distribución de cada una de las variables, agrupadas en función del modelo contable elegido y para el total de la muestra. Asimismo, se presentan en la última columna los resultados obtenidos con las técnicas Chi-cuadrado de Pearson y prueba de rangos U de Mann-Whitney, utilizadas para contrastar la influencia de cada una de las variables independientes (como el país de origen, la concentración de la propiedad, el tamaño, el nivel de endeudamiento y la proporción de inversiones inmobiliarias) en la elección contable.

El análisis de frecuencias en relación con la variable PAÍS revela que 20 (31,7%) de las 63 empresas británicas que componen la muestra optan por el modelo del coste en 2005 pasando, por tanto, del modelo de revalorización requerido por la normativa local previa hacia el modelo del coste. Por otra parte, en España únicamente 4 empresas (10%) optan por el modelo del valor razonable, criterio no permitido por la normativa local. Los resultados del test de Chi-cuadrado ($\chi^2 = 33,466$ y $p = 0,000$) muestran la influencia del país de origen en la elección contable.

Respecto a la variable CONC, los resultados muestran que el 35% de las empresas que componen la muestra presentan estructuras de propiedad concentradas. De éstas, el 50% optan por el modelo del valor razonable y el otro 50% lo hacen por el modelo del coste. Los resultados de la prueba de Chi-cuadrado confirman que no existe relación entre las variables, dado el elevado nivel de significación obtenido para el estadístico χ^2 ($p = 0,514$). Por tanto, en el plano univariante, no se aprecia la influencia de la estructura de propiedad en la elección contable.

Por lo que respecta al resto de las variables, observamos que las empresas que optan por el modelo del valor razonable presentan de media una mayor proporción de las inversiones inmobiliarias sobre el total de activos (INV_TA) y son de menor tamaño (TAM). Por otra parte, sobre la base de los estadísticos descriptivos observamos que las empresas que optan por el modelo del coste están de media más endeudadas (END) que las que eligen el modelo del valor razonable (ratio de endeudamiento medio de 2,05 frente a 1,61). El test de rangos U de Mann-Whitney muestra que las variables señaladas son significativamente distintas entre las empresas que optan por el modelo del valor razonable y las que lo hacen por el modelo del coste.

Tabla 1: Análisis descriptivo (por modelo contable y para el total de la muestra)

Variables independientes	Coste (n = 56)		Valor razonable (n = 47)		Total (n = 103)		Test (Coste vs. valor razonable)	
	n = 1	n = 0	n = 1	n = 0	n = 1	n = 0	Chi-cuadrado	U Mann Whitney Pr > Z
<i>Variables dummy</i>								
PAÍS	20	36	43	4	63	40	$\chi^2 = 33,466$ (p= 0,000)	N/A
CONC	18	38	18	29	36	67	$\chi^2 = 0,426$ (p= 0,514)	N/A
<i>Variables continuas</i>								
	Media	Desv.típ.	Media	Desv.típ.	Media	Desv.típ.		
INV_TA	0,0810	0,1837	0,5119	0,3760	0,2776	0,3584	N/A	$\alpha = 0,000$
END	2,0450	1,3759	1,6140	1,4643	1,8483	1,4263	N/A	$\alpha = 0,017$
TAM	13,7825	2,1097	11,8063	1,3642	12,8808	2,0537	N/A	$\alpha = 0,000$

4.2. Predicción de la elección contable

A continuación mostramos los resultados obtenidos por cada uno de los modelos. Para evaluar la capacidad predictiva de cada una de las metodologías utilizadas mostramos la matriz de confusión con los resultados medios obtenidos en las 10 iteraciones realizadas. En esta matriz incluimos el porcentaje de casos correctamente clasificados sobre el total y sobre cada una de las clases, determinadas por la elección contable (coste o valor razonable).

4.2.1. Regresión logística

El modelo de regresión logística aplicado presenta la siguiente forma:

$$\text{Ln}\Omega (\text{ELECCION})= b_0 +b_1\text{PAIS}+b_2\text{INV_TA} +b_3\text{END} +b_4\text{CONC} +b_5\text{TAM} \quad [2]$$

Los resultados obtenidos aparecen en la Tabla 2. La medida de Hosmer y Lemeshow de ajuste global, nos permite afirmar que no hay una diferencia estadísticamente significativa entre las clasificaciones observadas y predichas por el modelo, distribuyéndose como una Chi-cuadrado con 8 grados de libertad. Esta última es una prueba especialmente adecuada para muestras pequeñas y modelos en los que se incluyen una o más variables independientes de naturaleza continua, como es nuestro caso.

Tabla 2: Resultados regresión logística: Coeficientes, estadístico de Wald y razón de *odd*

Variable	Signo	β (Wald)	e^β
Constante		2,277 (1,039)	1,3198
PAIS	+	3,854 (13,340) ***	47,2050
INV_TA	+	3,854 (6,842) ***	47,1908
END	+	0,441 (2,975) *	1,5544
CONC	-	1,169 (2,414)	3,2175
TAM	-	-0,576 (7,455) ***	1,7793
<i>Estadísticos</i>			
R ² Nagelkerke		0,673	
-2LL		69,885 ***	
Hosmer y Lemeshow		6,597	

p < 0,1, ** p < 0,05 y *** p < 0,01

Para contrastar la significación de los coeficientes, utilizamos el estadístico de Wald. Los resultados del modelo 1 confirman los resultados obtenidos en el análisis descriptivo (véase apartado 4.1). La variable PAIS⁴ muestra tener una fuerte relación directa con la elección contable. Asimismo, la proporción de las inversiones inmobiliarias sobre el total de activos (INV_TA) y el endeudamiento (END) resultan estar directamente asociadas con la opción por el modelo del valor razonable. Por otra parte, el coeficiente de la variable tamaño (TAM) es negativo y significativo. Por último, la estructura de propiedad (CONC) no parece estar asociada con la elección contable.

Por último, la precisión en la clasificación del modelo se muestra en la Tabla 3. La RL clasifica correctamente, en media, el 83,2% de las observaciones. Los resultados muestran además un error de tipo I del 17,2% y un error de tipo II del 16,1%.

Tabla 3: Matriz de confusión: regresión logística (RL)

Porcentaje correcto de clasificación: 83,2%		Grupo pronosticado		Total
		Modelo valor razonable	Modelo coste	
Grupo real	Modelo valor razonable	39 (82,8%)	8 (17,2%)	47
	Modelo coste	9 (16,1%)	47 (83,9%)	56

4.3. Redes neuronales artificiales (MLP)

El modelo de redes neuronales empleado es clásico modelo de perceptrón multicapa basado en unidades sigmoides. Los resultados obtenidos de la precisión en la clasificación del modelo se muestran en la Tabla 4. El modelo MLP clasifica correctamente el 92,1% de las empresas que optan por el valor razonable, lo que resulta en un error de tipo I del 7,9%. El error de tipo II es del 18,1%. La precisión media en la clasificación del modelo es del 87,3%, superando el rendimiento obtenido con la RL.

⁴ La interpretación de la razón de *odd* (e^{β}) estimada por el modelo para la variable PAIS que toma un valor de 42,205 (estadísticamente significativa a un nivel del 1%) es que la probabilidad de que las empresas domiciliadas en el Reino Unido opten por el modelo del valor razonable es 42 veces superior a la de las empresas domiciliadas en España. De igual modo debe procederse para interpretar el ratio de *odd* obtenido para el resto de variables (INV_TA, END y TAM); no obstante, en estos casos las variables no son de respuesta binaria por lo que su interpretación no es tan sencilla.

Tabla 4: Matriz de confusión: redes neuronales (MLP)

Porcentaje global de aciertos: 87,3%		Grupo pronosticado		Total
		Modelo valor razonable	Modelo coste	
Grupo real	Modelo valor razonable	43 (92,1%)	4 (7,9%)	47
	Modelo coste	10 (18,4%)	46 (81,6%)	56

4.4. Árboles de decisión (C4.5)

Al aplicar C4.5 se obtienen las reglas de clasificación contenidas en la Tabla 5 que son útiles para identificar los perfiles obtenidos por el algoritmo.

Tabla 5: Reglas obtenidas a partir del algoritmo C4.5: soporte, confianza e índice

Regla	Expresión	Soporte	Confianza	Índice
1	Si PAIS =1, $INV_TA \leq 0,0859$ y $CONC=0$, $ELECCION=0$	19	89,47%	164,57%
2	Si PAIS =1, $INV_TA \leq 0,0859$ y $CONC=1$, $ELECCION=1$	9	66,67%	146,10%
3	Si PAIS =1, $INV_TA > 0,0859$, $ELECCION=1$	35	100,00%	219,15%
4	Si PAIS =0, $ELECCION=0$	40	90,00%	165,54%

El análisis de la reglas de clasificación obtenidas muestra que de las empresas domiciliadas en el Reino Unido ($PAIS = 1$), con una proporción de inversiones inmobiliarias sobre el total de activos es residual ($INV_TA \leq 0,0859$) y que presentan estructuras de propiedad dispersas ($CONC=0$), el 89,5% optan por el modelo del coste (Regla 1, soporte 19). Por su parte, las empresas domiciliadas en el Reino Unido con inversiones inmobiliarias menos significativas sobre el total de activos pero con una estructura de propiedad concentrada (Regla 2, soporte 9), el 66,7% de las empresas optan por el modelo del valor razonable.

Los resultados muestran el predominio de las empresas que optan por el modelo del coste cuando existe dispersión de la propiedad (Regla 1), y el indicador del índice obtenido sugiere que la probabilidad de que una empresa con ese perfil opte por el coste ($ELECCION=0$) es un 64,6% superior respecto a una empresa aleatoria de la muestra general. En teoría, el valor razonable reduce las asimetrías informativas y estas últimas se presumen mayores cuando la estructura de propiedad es muy dispersa. No obstante, a la vista de los resultados obtenidos en la Regla 1, también es razonable pensar que a mayor dispersión de la propiedad los directivos de empresas donde la actividad patrimonialista es residual (valores inferiores en INV_TA), tenderán a evitar la volatilidad que introduce el valor razonable en la cuenta de resultados optando, por tanto, por el modelo del coste que requiere la revelación de los valores razonables

en la memoria. En este sentido, Godfrey y Jones (1999) obtienen que las empresas con estructuras de propiedad más dispersas son más proclives a emplear prácticas de alisamiento del resultado frente a aquellas donde existe una mayor concentración de la propiedad.

Por su parte, la Regla 3 (soporte 35) indica que el 100% de las empresas domiciliadas en el Reino Unido y en las que la actividad patrimonialista es más significativa ($INV_TA > 0,0859$) optan por el modelo del valor razonable.

Por último, la Regla 4 (soporte 40) indica que el 90,0% de las empresas domiciliadas en España, optan por el modelo del coste.

En resumen, el árbol de decisión creado por el algoritmo C4.5 pone de manifiesto, al igual que los resultados obtenidos en la RL, que las variables que tienen mayor peso para la clasificación de las empresas en la elección contable son las variables PAIS e INV_TA. Estos resultados nos permiten contribuir a los trabajos previos de elección sobre la NIC 40 porque sugieren que las empresas británicas que cambian de criterio contable, desde el valor razonable hacia el coste, lo hacen fundamentalmente porque esta actividad no es significativa en su modelo de negocio y, probablemente, en términos coste-beneficio, se decantan por el modelo del coste.

Por último, la variable CONC que no parecía influir en la elección contable a la luz de los resultados obtenidos en la RL, aparece como predictor de la elección contable en el grupo de empresas británicas donde la actividad patrimonialista es poco significativa.

En la Tabla 6 mostramos los resultados en cuanto a la precisión en la clasificación del modelo:

Tabla 6: Matriz de confusión: algoritmo C4.5

Porcentaje global de aciertos: 86,9%		Grupo pronosticado		Total
		Modelo valor razonable	Modelo coste	
Grupo real	Modelo valor razonable	43 (92,5%)	4 (7,5%)	47
	Modelo coste	11 (19,8%)	45 (80,2%)	56

De las 47 empresas que optan por el modelo del valor razonable, el modelo clasifica correctamente 43 de ellas (92,5%), lo que implica un error de tipo I del 7,5%. De igual forma, el error de tipo II es del 19,8%, clasificando el modelo correctamente el 80,2% de las empresas (45 sobre un total de 56) que optan por el modelo del coste. En global, el modelo clasifica de forma correcta el 86,9% de las empresas, superando el rendimiento de la RL no así el del MLP.

4.5. Comparación de la precisión en la clasificación de las tres metodologías: RL, MLP y C4.5

Para analizar la significación de las diferencias obtenidas en los resultados alcanzados con cada metodología (RL, MLP y C4.5) empleamos un t-test sobre muestras relacionadas. El MLP obtiene el mejor resultado (mejor porcentaje medio de clasificación) entre todas las metodologías empleadas. El test de comparación realizado entre los resultados medios obtenidos por cada uno de los métodos pone de manifiesto que la capacidad predictiva del MLP y C4.5 es significativamente superior a la obtenida por la RL, con un nivel de significación del 5%. Sin embargo, no se han encontrado diferencias significativas en el rendimiento relativo entre las aplicaciones de los modelos MLP y C4.5.

5. CONCLUSIONES

Las normas contables habilitan escasas opciones para presentar los recursos porque la existencia de alternativas contables conduce a que recursos y transacciones similares se presenten de forma distinta en los estados financieros. Esto puede traducirse en una menor comparabilidad de la información financiera, motivo por el cual las opciones contables han ido desapareciendo a lo largo del tiempo (Zeff, 2012).

La NIC 40 es una de las pocas opciones contables abiertas en la normativa internacional pues permite optar entre el reconocimiento del valor razonable de las inversiones inmobiliarias en el balance y en la cuenta de resultados (modelo del valor razonable) o su revelación en las notas de la memoria (modelo del coste). Por su parte, el regulador contable estadounidense, el FASB, ha mantenido durante varios años en su agenda un proyecto, que no ha culminado, en el que se plantea el desarrollo de una nueva norma que bien ofrezca a las empresas la opción contable (similar a la NIC 40) o bien requiera únicamente el modelo del valor razonable para la contabilización de las inversiones inmobiliarias. Por tanto, resulta de interés directo para los reguladores contables, ampliar la investigación en este área.

Así, con el objetivo de predecir la opción que podrían tomar las empresas entre un criterio de valoración u otro cuando se les presenta esta elección contable, en este trabajo se introducen diferentes metodologías de clasificación que ofrece la minería de datos (MLP y C4.5), distintas de las utilizadas en los trabajos anteriores (RL), y se comparan en términos de capacidad predictiva. En segundo lugar, este trabajo contribuye a la literatura contable previa pues amplía la investigación contable existente en este área en un contexto multinacional (Reino Unido y España) y multisectorial.

Los resultados obtenidos en el análisis descriptivo y en la RL ratifican la influencia de variables como el país de origen, la importancia de las inversiones inmobiliarias sobre el total de activos, el nivel de endeudamiento y el tamaño en la elección contable. Así, las empresas domiciliadas en el Reino Unido (España), aquellas en las que las inversiones inmobiliarias son más (menos) importantes sobre el total de activos, las que presentan menores (mayores) niveles de endeudamiento y son de menor (mayor) tamaño son más propensas a optar por el valor razonable (coste) como criterio de valoración.

Por su parte, MLP y C4.5 son metodologías que aunque no muestran la significación individual de cada variable exógena en el modelo, su mayor capacidad predictiva pone de manifiesto la influencia del conjunto de ellas en la decisión contable. En particular, el análisis comparativo entre los porcentajes medios de clasificación correcta demuestra que MLP y C4.5 presentan una mayor capacidad predictiva frente a la RL. Sin embargo, no se encontraron diferencias significativas respecto a dicha capacidad entre MLP y C4.5.

Este estudio pone de manifiesto el potencial de las técnicas de minería de datos para el análisis predictivo en un contexto de elección contable. La capacidad de interpretación de relaciones no lineales o interacciones entre las variables que ofrecen dichas técnicas supone un complemento a las metodologías estadísticas tradicionales.

REFERENCIAS

- Ball, M. (2005). *RICS European Housing Review*. RICS, Londres.
- Ball, R. (2006). International Financial Reporting Standards (IFRS): pros and cons for investors. *Accounting and Business Research. International Accounting Forum*, 36: 5–27.
- Barlev, B.; Fried, D.; Haddad, J.R. & Livnat, J. (2007). Reevaluation of Revaluations: A Cross-Country Examination of the Motives and Effects on Future Performance. *Journal of Business Finance and Accounting*, 34(7-8): 1025–1050.
- Brown, P.; Izan, H.Y. & Loh, A.L. (1992). Fixed Assets Revaluations and Managerial Incentives. *Abacus*, 28(1): 36–57.
- Christensen, H.B. & Nikolaev, V.V. (2013). Does fair value accounting for non-financial assets pass the market test? *Review of Accounting Studies*, 18: 734–775.
- Cotter, J. (1999). Asset Revaluations and Debt Contracting. *Abacus*, 35(3): 268–285.

- Cotter, J. & Zimmer, I. (1995). Asset Revaluation and Assessment of Borrowing Capacity. *Abacus*, 31(2): 136–151.
- Cybenko, G. (1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Systems*, 2: 303–314.
- De Andrés, J.; Landajo, M., Lorca, P. & Ordoñez, P. (2010). Assessing the Liquidity of Firms: Robust Neural Network Regression as an Alternative to the Current Ratio. En Lytras, M.D.; Ordonez De Pablos, P.; Ziderman, A.; Roulstone, A.; Maurer, H. & Imber, J.B. (eds.). *Knowledge Management, Information Systems, E-learning & Sustainability Research*, Springer, Berlín, pp. 537–544.
- De Vicente, M.; Molina, H. & Ramírez, J.N. (2013). Inversiones inmobiliarias: la elección contable valor razonable versus coste en los grupos cotizados españoles. *Cuadernos de contabilidad*, 14(34): 25–51.
- Fields, T.D.; Lyz, T.Z. & Vincent, L. (2001). Empirical research on accounting choice. *Journal of Accounting & Economics*, 31: 255–307.
- Florez-Lopez, R. (2007). Modelling of insurers' rating determinants. An application of machine learning techniques and statistical models. *European Journal of Operational Research*, 183(3): 1488–1512.
- Gaeremynck, A. & Veugelers, R. (1999). The revaluation of assets as a signalling device: a theoretical and an empirical analysis. *Accounting and Business Research*, 29(2): 123–138.
- Godfrey, J.M. & Jones, K.L. (1999). Political cost influences on income smoothing via extraordinary item classification. *Accounting and Finance*, 39: 229–254.
- Hervás Oliver, J.L. (2005). La revalorización de activos fijos. Contraste empírico de un modelo financiero de elección contable. *Investigaciones Europeas de Dirección y Economía de la Empresa*, 11(1): 31–51.
- Hornik, K. (1991). Approximation Capabilities of Multilayer Feedforward Networks. *Neural Networks*, 4(2): 251–257.
- International Accounting Standards Board (IASB) (2013). *International Accounting Standard 40, Investment Property (Part A)*. IASB, Londres.

- Kirkos, E., Spathis, C. & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32: 995–1003.
- Kirkos, E.; Spathis, C. & Manolopoulos, Y. (2008). Support vector machines, Decision Trees and Neural Networks for auditor selection. *Journal of computational Methods in Sciences and Engineering*, 8: 213–224.
- Koh, H-C. (2004). Going concern predictions using data mining techniques. *Management Auditing Journal*, 19(3): 462–476.
- Kvaal, E. & Nobes, C. (2010). International differences in IFRS policy choice: a research note. *Accounting and Business Research*, 40(2): 173–187.
- La Porta, R.; López-de-Silanes, F. & Shleifer, A. (1999). Corporate Governance Around the World. *The Journal of Finance*, vol. LIV (2): 471–515.
- La Porta; R., López de Silanes, F.; Shleiffer, A. & Vishny, R. (1998). Law and Finance. *Journal of Political Economy*, 106: 1.113–1.155.
- Lin, Y.C. & Peasnell, K.V. (2000). Fixed Asset Revaluation and Equity Depletion in the UK. *Journal of Business Finance & Accounting*, 27(3-4): 359–394.
- Linsmeier, T.J. (2013). A Standard setter’s framework for selecting between fair value and historical cost measurement attributes: a basis for discussion of “Does fair value accounting for nonfinancial assets pass the market test?”, *Review of Accounting Studies*, 18: 776–782.
- Martens, D.; Bruynseels, L.; Baesens, B.; Willekens, M. & Vanthienen, J. (2008). Predicting going concern opinion with data mining. *Decision Support Systems*, 45(4): 765–777.
- Missonier-Piera, F. (2007). Motives for fixed-asset revaluation: An empirical analysis with Swiss data. *The International Journal of Accounting*, 42: 186–205.
- Muller, K.A.; Riedl, E.J. & Sellhorn, T. (2008). *Causes and Consequences of Choosing Historical Cost versus Fair Value*. Working paper, Pennsylvania State University, Harvard Business School and Ruhr-Universität Bochum.
- Muller, K.A.; Riedl, E.J. & Sellhorn, T. (2011). Mandatory Fair Value Accounting and Information Asymmetry: Evidence from the European Real Estate Industry. *Management Science*, 57(6): 1138–1153.

- Quagli, A. & Avallone, F. (2010). Fair Value or Cost Model? Drivers of Choice for IAS 40 in the Real Estate Industry. *European Accounting Review*, 19(3): 461–493.
- Quinlan, J.R. (2005). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers, San Francisco.
- Stone, M. & Rasp, J. (1991). The Tradeoffs in the Choice between Logit and OLS in Accounting Choice Studies. *The Accounting Review*, 66(1): 170–187.
- Sun, J.; Li, H.; Huang, Q-H. & He, K-Y. (2014). Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches. *Knowledge Based Systems*, 57: 41-56.
- Whittred, G. & Chan, Y.K. (1992). Asset Revaluation and the Mitigation of Underinvestment. *Abacus*, 28(1): 58–74.
- Zeff, S.A. (2012). The Evolution of the IASC into the IASB, and the Challenges it Faces. *The Accounting Review*, 87(3): 807–837.



Forecast Intervals for US/EURO Foreign Exchange Rate

SIMIONESCU, MIHAELA

Romanian Academy (Romania)

E-mail: mihaela_mb1@yahoo.com

ABSTRACT

The main goal of this research is to construct and assess forecast intervals for monthly US/EURO foreign exchange rate. The point forecasts used to build the intervals are based on a vector autoregression (VAR model) and on a Bayesian VAR model for data starting with the first month of 1999. The forecast intervals are based on the prediction error of the previous month. All the interval predictions based on VAR model included the actual values from 2014. The probability that the intervals based on BVAR model include the registered values of exchange rate is less than 0.8, according to likelihood ratio and chi-square tests.

Keywords: Forecast intervals; exchange rate; VAR model; Bayesian VAR model.

JEL classification: C51; C53.

MSC2010: 00A71; 97M10; 62M10; 62P20.

Intervalos de pronóstico para los tipos de cambio US/EURO

RESUMEN

El objetivo principal de esta investigación es construir y evaluar intervalos mensuales previstos para los tipos de cambio US/EURO. Las previsiones puntuales usadas para construir los intervalos se basan en un modelo de vectores autorregresivos (VAR) y en un modelo VAR bayesiano para los datos a partir del primer mes de 1999. El pronóstico se basa en el error de intervalos de predicción del mes anterior. Todas las predicciones de intervalos basadas en el modelo VAR incluyen los valores reales de 2014. La probabilidad de que los intervalos basados en el modelo BVAR incluyan los valores registrados de los tipos de cambio es inferior a 0,8, según las pruebas de coeficiente de chi-cuadrado y de verosimilitud.

Palabras claves: intervalos de pronóstico; tipos de cambio; modelo VAR; modelo VAR bayesiano.

Clasificación JEL: C51; C53.

MSC2010: 00A71; 97M10; 62M10; 62P20.



1. Introduction

The treasury secretary of USA considered that a stronger Dollar is and will be always a positive aspect for US. However, there was a short period when monthly Euro was stronger than US dollar. The prediction of exchange rate is fundamental for the construction of macroeconomic policies and for making-decision process. The exchange rate predicting is essential to gauge the future evolution of export performance and price competitiveness. Most of the studies use point predictions, but fewer analyses are dedicated to forecast intervals.

The objective of this study is the construction of forecast intervals for real exchange rate, where the predictions are based on different types of econometric models (VAR and BVAR models). These forecast intervals are assessed by using likelihood ratio and chi-square tests.

After the literature review, a presentation of methods to assess the forecast intervals is made. Then, the forecasts based on the two types of models and the associated intervals. Moreover, the evaluation of prediction intervals is made and some conclusions are drawn.

2. Literature review

The real exchange rate is modeled using many types of models of fundamentals from productivity measures, commodity prices, openness, interest rate differentials to capital flows and fiscal balance. Another approach considers the exchange rate as asset price, taking into account not only the current fundamentals but also the fundamentals' expectations at a future moment. The price-asset approach allows the check of efficiency assumption in the exchange foreign markets.

Meese and Rogoff (1983) showed the superiority of random walk forecasts in terms of accuracy compared to structural models predictions. However, Wright (2008) argued the superiority of predictions based on Bayesian Model Averaging technique (BMA technique) compared to naïve forecasts. Kilian and Taylor (2003) are those who showed that the exchange rates the horizons for predictions based on economic model that includes a possible nonlinear exchange rate vary from 2 years to 3 years.

Wing and Wu (2009) attacked the Meese-Rogoff puzzle from the perspective of the interval predicting. The authors started from a group of Taylor rule models for which robust

semi-parametric interval forecasting was applied. They discovered that Taylor rule models conduct to tighter forecast intervals compared to random walk. Molodtsova and Papell (2009) showed a significant predictability of exchange rate with Taylor-rule fundamentals.

Fattouh *et al.* (2008) proved, using empirical data, that forecasts based on a linear error correction model including fundamentals outperformed the predictions based on Markov-switching error correction model.

There were large preoccupations in literature for constructing theoretical models like Band-Threshold Autoregressive (Band-TAR models) and exponential smooth-transition autoregressive models (ESTAR models) that include transaction costs. A discrete regime switching is the main characteristic of Band-TAR models, while the ESTAR model supposes smooth transition between regimes.

Aye *et al.* (2013) compared the performance of real exchange rate predictions based on linear and, respectively non-linear models in South Africa. The authors presented all the varieties of forecasts: point predictions, interval forecasts and density forecasts for linear autoregressive models, ESTAR and TAR models. The differences in performance of forecasts for these types of models are non-significant for the real exchange rate in South Africa (South Africa Rand/US dollar and South Africa Rand/British pound).

Lam *et al.* (2008) compared the performance of predictions based on Bayesian model averaging technique, Purchasing Power Parity model, Uncovered Interest Rate Parity model and combined forecasts based on these models. The reference prediction is the naïve one and the historical average return. The combined predictions gave better results in terms of accuracy than the expectations based on a single model of those mentioned above.

Mtonga (2006) explained the necessity of predicting the real exchange rate for reducing the uncertainties in policy and decision making. The deviations of the exchange rate from the long-run equilibrium are more important for the policy making.

The forecast interval is built starting from the point forecasts and prediction error and a probability is attached in accordance to the hypothesis regarding the errors distribution. In general, we assume that the random shocks follow a normal distribution $e_t \rightarrow N(0, \sigma_\varepsilon^2)$, which implies a normally distributed probability density $x_{t+h} \rightarrow N(\hat{x}_{t+h}, \sigma_h^2)$.

Olave and Miguel (2012) proposed a bootstrap method in order to construct forecast intervals starting from ARCH models of exchange rate. The main advantage of this approach is the fact that there is no assumption regarding the conditional repartition of errors.

Boero and Marrocu (2004) compared the performance of predictions based on GARCH and SETAR models for exchange rate, using as reference model a random walk. The assessment criteria refer to point and interval forecasts.

For a fixed probability (nominal coverage), the empirical coverage of forecast intervals based on an economic model is a probability that the registered exchange rates be placed inside the intervals. The difference between the upper and the lower limit of an interval is called length of the interval, being a measure of tightness. Most of evaluation methods suppose the comparison of empirical coverage across models. The best model is the one with the most accurate coverage. If we have equal coverage accuracy, we have to check if the intervals have the same length. Tighter the forecast interval is, better the forecasting model is. Wing and Wu (2009) showed that in the case of equal coverage accuracy the economic models provide tighter prediction intervals compared to random walk. The authors proposed some loss criteria to assess the quality of intervals and some statistics presented by Giacomini and White (2006).

The forecast intervals are assessed by making the comparison with the out-of-sample interval predictions from the random walk. The likelihood ratio tests for conditional, unconditional and independent coverage were used by Wallis (2003) and Aye *et al.* (2013) to compare the predictions based on linear random walk models to those based on non-linear random walk models.

Empirical forecast intervals were built by Lee and Scholtes (2014) that used the repartition of the previous prediction errors. The construction of the prediction interval is based on the distribution quantiles for the prediction errors and the point forecasts based on a certain model.

For high-frequency data, Cai and Zhang (2016) analyzed the exchange rate movement predictability. The authors used an autoregressive conditional multinomial–autoregressive conditional duration model and they found a high predictability in data. By filtering the data, the forecasts performance improved.

A Bayesian VAR and a VAR model with Dornbusch prior were used by Ca'Zorzi *et al.* (2015) to predict the exchange rate. The standard VAR predictions and the naïve forecasts were outperformed by the predictions based on the VAR model with Dornbusch prior.

3. The assessment of forecast intervals

For assessing the density or forecast intervals more tests are used. Wallis (2003) used inferences on p-values rather than the asymptotic repartition of chi-square tests. This author developed the conditional, unconditional and independent coverage proposed by Christoffersen (1998). Good interval forecasts should be independently distributed, having an acceptable coverage.

I. Likelihood ratio (LR) tests for forecast intervals

We consider time series for the forecast intervals and we fix the probability π for the value to be inside the interval. We registered time series for the results registered in reality and we fix as objective the evaluation of *ex-ante* probability correction. If n_1 results are inside the forecast intervals and n_2 outside it, the coverage probability is: $p = n_1/n$. The distribution is binomial, under the null hypothesis and the likelihood is $L(\pi) = (1-\pi)^{n_2} \pi^{n_1}$; while under the alternative hypothesis, it is $L(p) = (1-p)^{n_2} p^{n_1}$. Christoffersen (1998) established the statistic of likelihood ratio as $LR_{UC} = 2(n_2 \log \frac{1-p}{1-\pi} + n_1 \log \frac{p}{\pi}) \xrightarrow{H_0} \chi^2_1$.

This is a test of unconditional coverage that is actually unsuitable for time series. Therefore, Christoffersen (1998) proposed another test that combines the test of unconditional coverage with the independence one.

The independence test is based on matrix of transition frequencies $[n_{ij}]$, which is the number of observations that are in state i at moment $t-1$ and in state j at moment t . The maximal likelihood estimations of transition probabilities are computed as a ratio between frequencies in a cell and the total number of frequencies of a line. For a forecast interval, two cases are possible: The values are inside or outside the interval, being denoted with 1 and 0. The transition matrix of estimated probabilities is:

$$P = \begin{pmatrix} 1-p_{01} & p_{01} \\ 1-p_{11} & p_{11} \end{pmatrix} = \begin{pmatrix} \frac{n_{00}}{n_0} & \frac{n_{01}}{n_0} \\ \frac{n_{10}}{n_1} & \frac{n_{11}}{n_1} \end{pmatrix}$$

The likelihood for P is $L(P) = (1-p_{01})^{n_{00}} \cdot p_{01}^{n_{01}} \cdot (1-p_{11})^{n_{10}} \cdot p_{11}^{n_{11}}$. The null hypothesis of independence test fixes that the $t-1$ state is independent of t state, which is equivalent with $\pi_{01} = \pi_{11}$. The estimator of maximal likelihood of the common probability is $p = \frac{n_{\cdot 1}}{n}$. The likelihood under the null hypothesis assessed at p is: $L(p) = (1-p)^{n_0} \cdot p^{n_1}$. The LR test statistic is: $LR_{ind} = -2 \log \frac{L(p)}{L(P)} \xrightarrow{H_0} \chi_1^2$. The test proposed by Christoffersen (1998) that combine the unconditional coverage test with the independence one has the statistic: $LR_{CC} = -2 \log \frac{L(\pi)}{L(P)} \xrightarrow{H_0} \chi_2^2$.

If the first observation is ignored, then: $LR_{CC} = LR_{UC} + LR_{ind}$.

II. Chi-square (χ^2) tests for forecast intervals

Stuart *et al.* (1999) demonstrated that likelihood ratio tests are equivalent with Pearson's goodness-of-fit tests. Wallis (2003) used them for the first time for density and forecast intervals. Chi-square test for unconditional coverage uses the statistic $X^2 = \frac{n(p-\pi)^2}{\pi(1-\pi)}$. If we consider the matrix of observed frequencies,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$X^2 = \frac{n(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}$$

The conditional coverage test combined with the independence test uses the contingency table of the observed frequencies with expected frequencies under the null hypothesis of independent lines and using the coverage probability π .

The matrix of expected frequencies is

$$\begin{bmatrix} (1-\pi)(a+b) & \pi(a+b) \\ (1-\pi)(c+d) & \pi(c+d) \end{bmatrix}.$$

The proportions column is considered under the test hypothesis; the tests has 2 degrees of freedom. The statistic is computed as a sum of square normal standard statistics of the sample of proportions, a proportion for each row of the table. For low-volume samples the additive relationship of LR statistics cannot be transposed exactly in chi-square test terms.

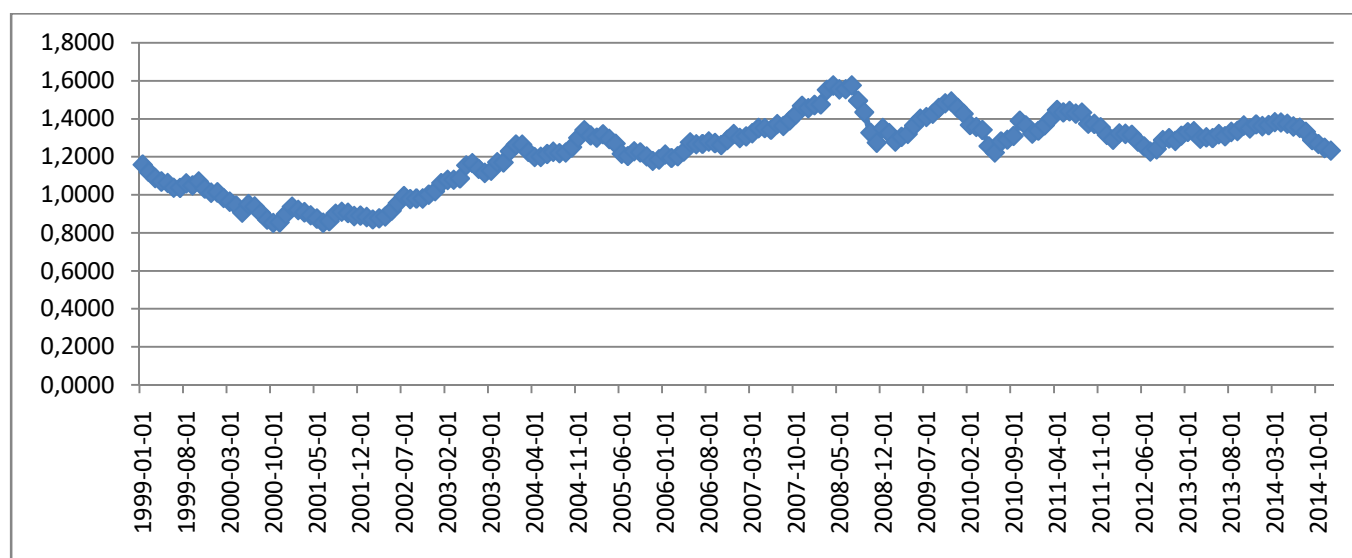
4. The evaluation of exchange rate forecast intervals

The exchange rate forecasts were built using VAR and Bayesian VAR models for monthly US/Euro exchange rate (ER) and consumer price index (CPI) over the period 1999:01-2014:12. The data are provided by Federal Reserve Bank of St. Louise from Federal Reserve Economic Database. The aggregation method for foreign exchange rate is based on the average. The consumer price index is the average monthly change in the goods and services prices, being paid by the consumers between two periods. The data series for CPI is seasonally adjusted and the reference base is the index from 1982-1984.

As we can see in Figure 1, there were some periods with values of exchange rate under 1: 2000:02-2002:10. This might be caused by the demand expansion by lowering the interest rates.

The stationary character of the data series is checked using ADF test. The series in level presents unit roots and the stationary character is ensured by the first differentiation. The variables corresponding to the new data series are denoted by Δcpi and Δer (see Table 1)

Figure 1. The evolution of monthly US/Euro foreign exchange rate (1999:01-2014:12)



Source: Own elaboration.

Table 1. The Augmented Dickey-Fuller (ADF) test for the transformed data series

Variable	Model	ADF statistic	1% critical value	5% critical value	10% critical value
Δcpi_t	Trend+intercept	-6.5162	-4.0101	-3.4348	-3.1411
	Intercept	-6.4737	-3.4667	-2.8771	-2.5750
	None	-4.2398	-2.5766	-1.9414	-1.6166
Δer_t	Trend+intercept	-5.7196	-4.0101	-3.4348	-3.1411
	Intercept	-5.7115	-3.4667	-2.8771	-2.5750
	None	-5.7126	-2.5766	-1.9414	-1.6166

Source: Author's computations.

Most of the lag length criteria indicated that the suitable lag equals 2 (see Table 2).

Table 2. Lag length criteria for VAR model

Lag	LogL	LR	FPE	AIC	SC	HQ
0	189.8130	NA	0.000381	-2.196644	-2.159899	-2.181734
1	219.9155	59.14865	0.000281	-2.501935	-2.391701*	-2.457207*
2	225.6149	11.06554*	0.000275*	-2.521811*	-2.338088	-2.447264
3	227.0109	2.677775	0.000284	-2.491356	-2.234143	-2.386990
4	228.1586	2.174527	0.000294	-2.457995	-2.127294	-2.323811
5	232.1550	7.478671	0.000294	-2.457953	-2.053763	-2.293950
6	233.5947	2.660437	0.000303	-2.428008	-1.950328	-2.234186
7	238.7297	9.369111	0.000299	-2.441283	-1.890113	-2.217642
8	240.7302	3.603297	0.000306	-2.417897	-1.793239	-2.164437

Source: Author's computations.

The corresponding equations of the VAR model are represented below:

$$\Delta er = 0.3190269538 * \Delta er(-1) - 0.06897735018 * \Delta er(-2) - 0.003635677141 * \Delta cpi(-1) - 0.002265028041 * \Delta cpi(-2) + 0.00357020649;$$

$$\Delta cpi = 3.427668687 * \Delta er (-1) + 1.876892616 * \Delta er (-2) + 0.4382151242 * \Delta cpi (-1) - 0.2301331038 * \Delta cpi (-2) + 0.3058719426.$$

Table 3. VAR Residual Portmanteau test errors' autocorrelation

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	0.073084	NA*	0.073499	NA*	NA*
2	0.967128	NA*	0.977761	NA*	NA*
3	3.356869	0.5000	3.408704	0.4919	4
4	5.282142	0.7270	5.378492	0.7165	8
5	12.64331	0.3955	12.95365	0.3724	12
6	14.12884	0.5891	14.49130	0.5622	16
7	24.67919	0.2140	25.47607	0.1838	20
8	25.84405	0.3611	26.69608	0.3188	24
9	28.17443	0.4552	29.15130	0.4049	28
10	29.28258	0.6048	30.32581	0.5514	32
11	35.70315	0.4826	37.17184	0.4148	36
12	42.93435	0.3466	44.92894	0.2730	40

Source: Author's computations.

In the first period (see Table 4), the variation in exchange rate is not due to consumer price index changes. In the second period, 0.41% of the variance in exchange rate is due to CPI changes. Starting from the third period, more than 1.2% of the variation in US/EUR foreign exchange rate is explained by CPI.

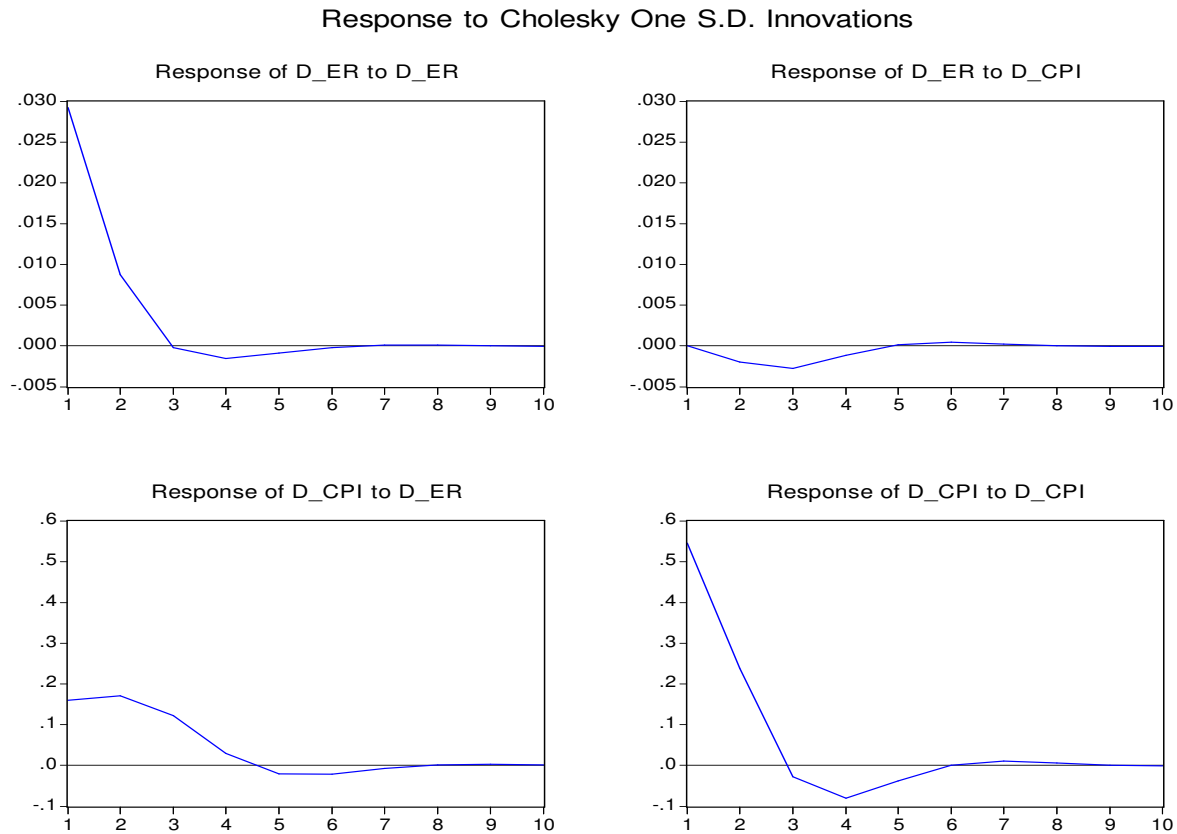
Table 4. Variance decomposition of Δer

Period	S.E.	Δer	Δcpi
1	0.029250	100.0000	0.000000
2	0.030594	99.58078	0.419218
3	0.030717	98.79185	1.208151
4	0.030776	98.65042	1.349579
5	0.030788	98.64855	1.351445
6	0.030792	98.62717	1.372829
7	0.030793	98.62223	1.377769
8	0.030793	98.62225	1.377746
9	0.030793	98.62185	1.378146
10	0.030793	98.62172	1.378283

Source: Author's computations.

In the second, the third and the fourth period (see Figure 2), a shock in the consumer price index determined a negative response of exchange rate.

Figure 2. Impulse-response functions



Source: Own elaboration.

A Bayesian VAR model of order 2 is also built. The means of posterior coefficients are used as coefficient to make forecasts for the months of 2014. The VAR model is also used to make forecasts for 2014.

The BVAR model has the following form:

$$Y_i = X_i * \Phi_i + u_i, \quad \text{where } u_i \sim N(0, s^2);$$

$$\Phi_i | \tau_i \sim \tau_i * N(0, V1) + (1 - \tau_i) * N(0, V2), \quad V1 > V2$$

$\tau_i = 1$ suggests that a variable was chosen;
 $\tau_i = 0$ shows that Φ_i can be excluded.

The estimation algorithm is a Gibbs sampler with hierarchical priors:

First stage: $s^2 \sim \text{IG}(a,b)$, $\text{Beta}_i | \tau_i \sim \tau_i * \text{N}(0,V1) + (1-\tau_i) * \text{N}(0, V2)$.
 Second stage: $\tau_i | \pi_i \sim \text{Bernoulli}(\pi_i)$.
 Third stage: $\pi_i \sim \text{Beta}(a', b')$.

Table 5. Estimation of Bayesian VAR model

Constant	Posterior Phi1 coefficients		Posterior Phi2 coefficients		Posterior covariance matrix of the VAR system	
0.6894	1.35	4.83	-0.35	-4.35	0.32	0
0.0104	0	1.3	0	-0.33	0	0.01

Source: author's computations

Starting from the point forecasts based on VAR and BVAR models, some forecast intervals were built by considering the forecast error from previous period.

The prediction intervals are built as follows:

$$(er_t(k) - z_{\alpha/2} \cdot e(k-1), er_t(k) + z_{\alpha/2} \cdot e(k-1)), \quad k = 1, \dots, K.$$

where:

$er_t(k)$ is the point forecast of exchange rate for period $(t+k)$; and

$z_{\alpha/2}$ is the quantile $\alpha/2$ of standard normal distribution.

Table 6. The point forecasts and prediction intervals for US/Euro foreign exchange rate

Month	Point forecasts based on VAR model	Point forecasts based on BVAR model	Forecast intervals based on VAR model (intervals' limits)		Forecast intervals based on BVAR model (intervals' limits)		Actual values
2014:01	1.3798	1.3702	1.3446	1.4149	1.3350	1.4054	1.3618
2014:02	1.3815	1.3778	1.3521	1.4108	1.3485	1.4071	1.3665
2014:03	1.3820	1.3806	1.3803	1.3836	1.3790	1.3822	1.3828
2014:04	1.3830	1.389	1.3791	1.3868	1.3692	1.3851	1.3810
2014:05	1.3845	1.39	1.3637	1.4053	1.3379	1.3692	1.3739
2014:06	1.3862	1.3903	1.3338	1.4386	1.3228	1.3379	1.3595
2014:07	1.3880	1.3907	1.3200	1.4559	1.2763	1.3228	1.3533
2014:08	1.3896	1.3902	1.2757	1.5035	1.1899	1.2763	1.3315
2014:09	1.3912	1.3905	1.1907	1.5918	1.1494	1.1899	1.2889
2014:10	1.3929	1.3947	1.1475	1.6382	1.1097	1.1494	1.2677
2014:11	1.3945	1.3982	1.1060	1.6830	1.0826	1.1097	1.2473
2014:12	1.3961	1.4025	1.0762	1.7161	0.0000	1.0826	1.2329

Source: author's computations.

All the forecast intervals based on VAR models includes the actual values, but only 5 out of 12 intervals based on BVAR model contain the registered values of exchange rate (see Table 6). Starting from the half of 2014, the exchange rate underwent a fast decrease that was not anticipated by the models. Moreover, the predictions based on BVAR model anticipated a higher increase than VAR model forecasts, fact that explains the differences between forecast intervals.

The forecast intervals based on the point predictions of VAR and Bayesian VAR models are assessed using likelihood ratio test and chi-square test. Considering an *ex-ante* probability of 0.8, we assess the hypothesis that the empirical probability is 0.8.

Table 7. The evaluation of prediction intervals for US/Euro foreign exchange rate

<i>LR_{CC}</i> statistic		Chi-square statistic	
VAR model predictions	BVAR model predictions	VAR model predictions	BVAR model predictions
4.4629	3.7523	3	25.5208

Source: author's computations.

For a level of significance of 5%, both tests suggest that the probability that forecast intervals include the actual values is higher than 0.8 for the forecasts based on VAR models, but this probability is lower than 0.8 for intervals based on BVAR predictions.

5. Conclusions

In this study, point forecasts are built for US/EURO foreign exchange rate for months in 2014. The forecasting method is represented by econometric models like: VAR model and Bayesian VAR model. The results indicated that all the prediction intervals based on VAR model included all the actual values unlike the intervals based on BVAR model.

In a future research, other types of forecast intervals should be developed and other econometric models should be chosen.

References

Aye, G.C.; Balcilar, M.; Bosch, A.; Gupta, R. and Stofberg, F. (2013): *The out-of-sample forecasting of non-linear models of real exchange rate behaviour: the case of the South African rand*. Department of Economics Working Papers, University of Pretoria, 2013-04.

- Boero, G. and Marrocu, E. (2004): "The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts". *International Journal of Forecasting*, 20(2), pp. 305-320.
- Cai, C.X. and Zhang, Q. (2016): "High-Frequency Exchange Rate Forecasting. *European Financial Management*. 22(1), pp. 120-141.
- Ca'Zorzi, M.; Kocięcki, A. & Rubaszek, M. (2015): "Bayesian forecasting of real exchange rates with a Dornbusch prior". *Economic Modelling*, 46, pp. 53-60.
- Christoffersen, P.F. (1998): "Evaluating Interval Forecasts". *International Economic Review*, 39(4), pp. 841-62.
- Fattouh, B.; Mouratidis, K. and Harris, L. (2008): *South Africa's real exchange rate and commodities boom: A Markov regime switching approach*, CSAE Conference, Economic Development in Africa.
- Giacomini, R. and White, H. (2006): "Tests of conditional predictive ability". *Econometrica*, 74, pp. 1545-1578.
- Kilian, L. and Taylor, M. (2003): "Why is it so difficult to beat the random walk forecast of exchange rate". *Journal of International Economics*, 60, pp. 85-107.
- Lam, L.; Fung, L. and Yu, I. (2008): *Comparing forecast performance of exchange rate models*. Working Paper Hong Kong Monetary Authority, 08/2008.
- Lee, Y.S. and Scholtes, S. (2014): "Empirical prediction intervals revisited". *International Journal of Forecasting*, 30(2), pp. 217-234.
- Meese, R. and Rogoff, K. (1983): "Empirical Exchange Rate Models of the Seventies". *Journal of International Economics*, 14, pp. 3-24.
- Molodtsova, T. and Papell, D.H. (2009): "Out-of-sample exchange rate predictability with Taylor rule fundamentals". *Journal of International Economics*, 77(2), pp. 167-180.
- Mtonga, E. (2006): "The real exchange rate of the rand and competitiveness of South Africa's trade". MPRA paper 1192.
- Olave, P. and Miguel, J.A. (2012): "Bootstrap method in exchange rate forecasting". University of Zaragoza, Spain.

Stuart, A.; Ord, J.K. and Arnold, S. (1999): *Kendall's Advanced Theory of Statistics*, 6th ed., vol. 2A. London: Edward Arnold.

Wallis, K. (2003): "Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts". *International Journal of Forecasting*, 19, pp. 165-175.

Wright, J.H. (2008): "Bayesian model averaging and exchange rate forecasts". *Journal of Econometrics*, 146(2), pp.329-341.

Volumen 23 (junio 2017)

Stochastic Frontier Models with Dependent Errors based on Normal and Exponential Margins

Modelos de frontera estocástica con errores dependientes basados en márgenes normal y exponencial

Gómez-Déniz, Emilio; Pérez-Rodríguez, Jorge V.

Páginas 3–23

Hacia una visión holística de la distancia en los negocios internacionales: el caso colombiano

Towards Holistic Version of Distance in International Business: The Colombian Case

Caicedo Marulanda, Carolina; Pla Barber, José; León Darder, Fidel; Mora Rodríguez, Jhon James.

Páginas 24–47

Extreme Value Theory: An Application to the Peruvian Stock Market Returns

Teoría de valores extremos: una aplicación a los retornos bursátiles peruanos

Rodríguez, Gabriel

Páginas 48–74

El contagio en el fracaso empresarial como consecuencia de la proximidad geográfica: un análisis con los estadísticos join-count aplicado al sector servicios

The Contagion on Business Failure by the Geographical Proximity: An Analysis with the Join-Count Tests in the Service Sector

Rodríguez Fuentes, Christian Camilo; Maté Sánchez-Val, María Luz; López Hernández, Fernando A.

Páginas 75–97

Pronosticando la inflación mensual en Colombia un paso hacia delante: una aproximación "de abajo hacia arriba"

Forecasting the Colombian Monthly Inflation One Step Ahead: A "Bottom to Top" Approach

Alonso, Julio César; Rivera, Andrés Felipe

Páginas 98–118

Estimating the Price Elasticity of Demand for Cannabis: A Geographical and Crowdsourced Approach

La estimación de la elasticidad en los precios de la demanda de cannabis: un enfoque geográfico y basado en la colaboración abierta distribuida

Halcoussis, Dennis; Lowenberg, Anton D.; Roof, Zach

Páginas 119–136

Proceso de reestructuración bancaria: análisis de las entidades participantes, fusionadas y nacionalizadas

Bank Restructuring: Analysis of Participating, Merged and Nationalized Entities

Jimber del Río, Juan Antonio; Caridad y Ocerin, José María; Ceular Villamandos, Nuria

Páginas 137–163

La técnica de escalamiento lineal por intervalos: una propuesta de estandarización aplicada a la medición de niveles de bienestar social

Interval Linear Scaling Technique: Proposal for Standardization Applied to the Measurement of Social Well-Being Levels

Actis di Pasquale, Eugenio; Balsa, Javier

Páginas 164–193

Links between the Eurozone Stock Markets: A New Perspective, Considering the Capitalization Level

Relación entre los índices bursátiles europeos: una nueva perspectiva a partir de los niveles de capitalización

Gabriel, Vítor; Saraiva, Helena

Páginas 194–209

La tasa de variación del PIB en un modelo simple de determinación de la renta

The Variation of GDP in a Simple Model of Income Determination

Escartin, Eduardo; Velasco, Francisco; González-Abril, Luis

Páginas 210–222

Qualitative Analysis of the Goodwin Model of the Growth Cycle

Análisis cualitativo del modelo de Goodwin de ciclos de crecimiento

Serebriakov, Vladimir; Dohnal, Mirko

Páginas 223–233

Elección contable para la valoración de las inversiones inmobiliarias. Contribución de las técnicas de minería de datos para determinar patrones de decisión

Accounting Choice for Measuring Investment Properties. Data Mining Techniques Contribution to Determine Decision Patterns

De Vicente Lama, Marta; Molina Sánchez, Horacio; Ramírez Sobrino, Jesús N.; Torres Jiménez, Mercedes

Páginas 234–256

Forecast Intervals for US/EURO Foreign Exchange Rate

Intervalos de pronóstico para los tipos de cambio US/EURO

Simionescu, Mihaela

Páginas 257–271