

Appropriate Distribution of Cost Inefficiency Estimates as Predictor of Financial Instability

JUAN CÁNDIDO GÓMEZ-GALLEGO

Departamento Administración y Dirección de Empresas, UNIVERSIDAD CATÓLICA SAN ANTONIO, ESPAÑA. E-mail: jcandido@ucam.edu

JUAN GÓMEZ-GARCÍA

Departamento Métodos Cuantitativos para la Economía y la Empresa, UNIVERSIDAD DE MURCIA, ESPAÑA. E-mail: jgomezg@um.es

MARÍA CONCEPCIÓN PÉREZ-CÁRCELES

Departamento Administración y Dirección de Empresas, UNIVERSIDAD CATÓLICA SAN ANTONIO, ESPAÑA. E-mail: mc Perez@ucam.edu

ABSTRACT

The study develops a hypothesis validation process to test whether estimated cost inefficiency fits the theoretical distribution of the estimate of inefficiency, given the initial assumption of a half normal distribution for the inefficiency term in the stochastic frontier model. The application of this methodology to a sample of Spanish banks in the period 2002-2007 allows us to observe the evolution of the distributions of estimated cost inefficiency. The results confirm for the existence of a relationship between estimated cost inefficiency and the financial situation of these entities. In fact, an increase in the probability of higher levels of inefficiency is observed the period of study, coinciding with the beginning of financial instability in the Spanish banking system.

Keywords: Stochastic Frontier, Cost Inefficiency, Estimated Inefficiency.

La distribución de la ineficiencia estimada como predictor de inestabilidad financiera

RESUMEN

El estudio desarrolla un proceso de validación de hipótesis que pretende contrastar si la distribución de la ineficiencia en costes estimada se ajusta a la distribución teórica de la ineficiencia estimada, dado el supuesto inicial de half normalidad del término de ineficiencia en el modelo de frontera estocástica. La aplicación de esta metodología a una muestra de bancos españoles durante el periodo 2002-2007 permite observar la evolución de las distribuciones de la ineficiencia en costes estimada y la situación financiera de esas entidades. De hecho, las distribuciones muestran un incremento en la probabilidad de los mayores niveles de ineficiencia en los últimos años del estudio, coincidiendo con el comienzo de la inestabilidad financiera en el sistema bancario español.

Palabras clave: Frontera estocástica, ineficiencia en costes, ineficiencia estimada.

JEL Classification: C13, G21

Artículo recibido en mayo de 2012 y aceptado en octubre de 2012

Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. e-30311

1. INTRODUCTION

The present study aims to propose a solution for the emphasized, among others, by Berger and Humphrey (1997) about the inexistence of a priori justification for imposition of the distributional assumption on the inefficiency term in the stochastic frontier approach. One of the pioneering studies on this methodology, Aigner et al. (1977), establishes the assumption of an exponential or half normal distribution on the inefficiency term. However, there is no reason to justify the selection of a distribution that imposes a null mode. The specifications of more general functional forms, like truncated normal (Stevenson, 1980) and two parameters gamma (Greene, 1990) have partially resolved this problem; however, efficiency measures can be sensitive to distributional assumptions. Kumbhakar and Lovell (2000) provided estimates based on the estimated cost frontier in Greene (1990) and obtained correlations of inefficiency estimates that ranged from 0.747, between exponential and gamma, to 0.980, between half normal and truncated normal. Recently, it is concluded that distributional assumption is of considerable importance, due to differences in the estimated parameters with each assumption (Fried et al., 2008, chapter 2). One aspect to take into account is that the distribution of the estimate of inefficiency cannot be expected to coincide with the distribution assumed on cost inefficiency. In this sense, Wang and Schmidt (2009) and Bhandari (2011) have developed the distribution of the estimate of technical inefficiency with the assumptions of half normal and truncated normal distribution, respectively. The former proved that it is not appropriate to assume a half normal distribution for estimated inefficiency when inefficiency is distributed as a half normal, unless the random error variance is very small. Wang et al. (2011) propose tests in order to check the goodness of fit of the assumed distribution.

Given this discussion, the primary goal of the study is to establish a hypothesis validation process to test the validity of the half normal distributional assumption on the inefficiency term in the cost frontier. Furthermore, this paper aims to confirm whether the distribution of the cost inefficiency estimate can be considered a sufficiently significant indicator for it to be taken into account for the prediction of financial instability situations. A sample of entities of the Spanish banking system is selected in the period previous to the beginning of the current situation of instability in the financial markets (2002-2007).

The paper is organized as follows: Section 2 explains the methodology, Section 3 presents the data used, Section 4 shows the results and Section 5 highlights the main conclusions.

2. METHODOLOGY

The present study applies the stochastic production frontier model proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977). Through

modification of the specification of the error term, the stochastic cost frontier model can be obtained:

$$\ln C_i = \alpha_0 + \sum_p^r \beta_p \ln y_{ip} + \sum_m^n \delta_m \ln w_{im} + \varepsilon_i \quad i = 1, 2, \dots, I \quad (1)$$

where $\ln C_i$ is the cost logarithm, $\ln y_i$ is the outputs logarithm, $\ln w_i$ is the price of the inputs logarithm and ε_i is the perturbation composed of the sum of u_i , interpreted as cost inefficiency effects, and v_i , as random effects of the decision-making unit i .

Following the assumption of a half normal distribution proposed by Aigner et al. (1977) on the inefficiency term, Jondrow et al. (1982) consider the distribution of u conditional on ε to make estimates about technical inefficiency, u . Specifically, a point estimate of u , \hat{u} , can be made through the mean of the conditional distribution. In the case of a costs frontier, the expression is:

$$\hat{u} = E(u/\varepsilon) = \sigma_* \left[\frac{f(-\varepsilon\lambda/\sigma)}{1 - F(-\varepsilon\lambda/\sigma)} + \left(\frac{\varepsilon\lambda}{\sigma} \right) \right] \quad (2)$$

where f and F represent the standard normal density and distribution functions, respectively, and $\sigma_*^2 = \sigma_u^2 \sigma_v^2 / \sigma^2$, $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\lambda = \sigma_u / \sigma_v$.

As Wang and Schmidt (2009) did for technical inefficiency, from Equation 2 it is possible to obtain the theoretical distribution of the cost inefficiency estimate, \hat{u} , when u is assumed to be distributed as a half normal.

Since function $\hat{u} = E(u/\varepsilon) = h(\varepsilon)$ is monotonic (strictly decreasing), it can be inverted $\varepsilon = h^{-1}(\hat{u}) = g(\hat{u})$ and the density of the estimate can be calculated through the following expression:

$$f_{\hat{u}}(\hat{u}) = f_{\varepsilon}(g(\hat{u})) \cdot |g'(\hat{u})| \quad (3)$$

where f_{ε} and $f_{\hat{u}}$ represent the densities of ε and \hat{u} .

The density of ε (Aigner et al., 1977) is the following:

$$f_{\varepsilon}(\varepsilon) = (2/\sigma) \cdot f(\varepsilon/\sigma) \cdot F(-\varepsilon\lambda/\sigma) \quad (4)$$

The Jacobian term $|g'(\hat{u})|$ is obtained from the equation:

$$g'(\hat{u}) = \frac{\sigma^2}{\sigma_u^2 \cdot [-1 + c'(g(\hat{u})\lambda/\sigma)]} \quad (5)$$

where $c'(s) = -s \frac{f(s)}{1-F(s)} + \left(\frac{f(s)}{1-F(s)} \right)^2$.

Substituting Equations 4 and 5 in Equation 3, the theoretical distribution is

$$f_{\hat{u}}(\hat{u}) = \frac{2\sigma \cdot f\left(\frac{g(\hat{u})}{\sigma}\right) \cdot F\left(\frac{-g(\hat{u}) \cdot \lambda}{\sigma}\right)}{\sigma_{\hat{u}}^2 | -1 + c'(g(\hat{u}) \cdot \lambda/\sigma) |} \quad (6)$$

Finally, it would be of great interest to ascertain whether point estimates obtained in Equation 2 from a real sample of data emerge from the theoretical distribution of inefficiency estimate, given the initial hypothesis (Equation 6). In order to test if the discrepancy between these distributions is significant, the Kolmogorov Smirnov statistics of goodness of fit is applied, following Wang et al. (2011).

Validation of the model requires testing the existence of inefficiency effects (Dios-Palomares, 2002), the functional form of the frontier and the temporal variability of the inefficiency effects.

2.1. Testing the existence of inefficiency effects

As argued in Coelli (1995), if we consider a frontier model with the assumption of a half normal distribution for inefficiency, it will be possible to test the existence of technical inefficiency by contrasting the null hypothesis $\sigma_u^2 = 0$ with the alternative hypothesis $\sigma_u^2 > 0$. This hypothesis can be tested by using several statistical models. One of these is the Wald test, $W = \widehat{\rho}_{MV} / S_{\widehat{\rho}_{MV}}$, which is frequently used in applications of the stochastic frontier model. This statistic is asymptotically distributed as a standard normal and contrasts $\rho = 0$ against $\rho > 0$, where $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ and the critical region is defined by $\left| \frac{\widehat{\rho}_{MV}}{S_{\widehat{\rho}_{MV}}} \right| >$

$Z_{\alpha/2}$.

2.2. Testing the model of stochastic frontier

Below, we specify a series of hypothesis tests using the one-sided generalized likelihood ratio (LR) test to check several aspects in the specification of the stochastic frontier model (Zajc, 2006). This statistic is asymptotically distributed as a chi-square random variable with a number of degrees of freedom equal to the number of restrictions and is calculated as:

$$LR = -2\{\ln[L(H_0)] - \ln[L(H_1)]\} \quad (7)$$

where $L(H_0)$ and $L(H_1)$ are the value of the log likelihood function under the null and alternative hypothesis, respectively. The critical region of the test with size α is defined by $LR_{\text{exp}} > \chi_{\text{restrictions}; \alpha}^2$.

a) Testing the temporal variability of inefficiency effects. According to Lee and Schmidt (2003), $u_{it} = \beta(t) \cdot u_i$, where $\beta(t) = \exp\{-\eta(t - T)\}$; so u_{it} is distributed as an exponential function of time, where u_i represents random i.i.d.

variables as a truncated normal and η is a unknown parameter to be estimated. In this way, the null hypothesis of inefficiency effects with non significant variation through time, $H_0: \eta = 0$, is contrasted with the alternative hypothesis, $H_1: \eta \neq 0$, where the critical region is defined, from Equation 7, by $LR_{exp} > \chi_{1,\alpha}^2$.

b) Testing the functional form of the frontier. To ascertain whether the most appropriate functional form to impose on the cost frontier in the present study is the Cobb-Douglas or the translog, the null hypothesis H_0 : Cobb – Douglas is contrasted with the alternative hypothesis H_1 : translog by means of a whole significance contrast of the parameters of the translog that do not appear in the Cobb-Douglas, with a critical region defined as $LR_{exp} > \chi_{10,\alpha}^2$.

3. DATA

Following earlier studies, such as Altunbas and Chakravarty (2001), Maudos and Pastor (2003), Weill (2003), Pasiouras (2008), Delis et al. (2009), the intermediation approach, suggested by Sealey and Lindley (1977), has been adopted to select inputs and outputs of banking activity (see Table 1).

The data set necessary in this study has been selected from balances and profit and loss accounts of Spanish banking entities existing during the period 2002-2007 and is available from the Spanish Bank Association (AEB). A representative sample of 26 DMUs of the Spanish banking system is studied. The sample includes only Spanish entities. They are mainly commercial banks with a more or less marked regional character.

Table 1
Definition of variables

Variable	Name	Definition
y_1	Credit to clients	Sum of outstanding balance of credits provided to clients (millions of euros)
y_2	Securities portfolio	Fixed and variable interest investments, property of the entity (millions of euros)
x_1	Loanable Funds	Sum of the balance of deposits of clients in the entity (millions of euros)
x_2	Physical Capital	Value of tangible fixed assets of the entity (millions of euros)
x_3	Employees	Number of employees of the entity (thousand)
w_1	Price Loanable Funds	Interests / Total liability (euros)
w_2	Price Physical Capital	Recovery and maintenance / Physical Capital (euros)
w_3	Price of Labour	Personal expenses / Employees (thousand euros)
C	Total Cost	$\sum w_i x_i$ (millions of euros)

Source: Own elaboration.

Table 2 shows that total costs have increased more than one hundred percentage points. In the outputs, the growth rates of credits and securities portfolio are positive, more than two hundred percentage points in the last case. However, in the inputs estimated, only the price of physical capital shows a negative trend due to the increase of employees.

Table 2
Averages of the variables in the cost function

Variable	Mean	Std. Dev	2002	2003	2004	2005	2006	2007
y1	17370.86	35225.42	10957.88	12350.17	14433.60	18043.33	22465.88	25974.31
y2	3399.43	9665.49	1465.77	2189.94	2401.13	5612.46	4120.17	4607.12
x1	14843.13	30506.07	11215.37	11645.28	11559.88	15347.69	18421.77	20868.77
x2	226.62	462.90	220.79	213.42	209.12	240.27	251.10	225.00
x3	3.59	6.79	3.52	3.54	3.57	3.58	3.62	3.74
w1	0.04	0.04	0.04	0.03	0.05	0.03	0.04	0.06
w2	0.23	0.32	0.19	0.25	0.32	0.25	0.17	0.18
w3	61.50	47.09	56.11	53.30	58.17	57.32	61.93	82.19
C	936.58	2130.15	739.51	666.06	682.01	867.83	1091.88	1572.17

Note: 26 banks, period 2002-2007, 156 observations

Source: Own elaboration.

4. RESULTS

Firstly, the presence of cost inefficiency is confirmed in the stochastic frontier model, as indicated by from the significance of the parameter ρ , $\widehat{\rho}_{MV} = 0.754$ and $s_{\widehat{\rho}_{MV}} = 0.127$, explained in Section II.

Regarding the assumptions in the stochastic frontier model, the results of the test are presented in Table 3.

Table 3
Hypothesis Contrasts

Test	Ln(H ₀)	Ln(H ₁)	LR	f.d.	$\chi^2_{\alpha,1}$	Result	p-value
a	-79.707	-65.325	28.765	1	10.827	RH ₀	0.001
b	-115.780	-79.707	72.146	10	29.588	RH ₀	0.001

Notes: RH₀: the null hypothesis is rejected; a) temporal variability of inefficiency effects; b) functional form of the frontier

Source: Own elaboration.

From the results presented in Table 3, the following conclusions are obtained:

- a) Inefficiency effects vary over time, and so it is assumed that the inefficiency error component is distributed as an exponential function of time.
- b) The functional form of the cost frontier that best fits the data used in the study is the translogarithmic one.

The point estimation proposed by Jondrow et al. (1982) has been carried out by using software Frontier 4.1. Table 4 shows the mean, the mode and the percentiles 2.5 and 97.5 of estimated cost inefficiency for each unit.

Table 4
Inefficiency estimates for each unit in the period 2002-2007

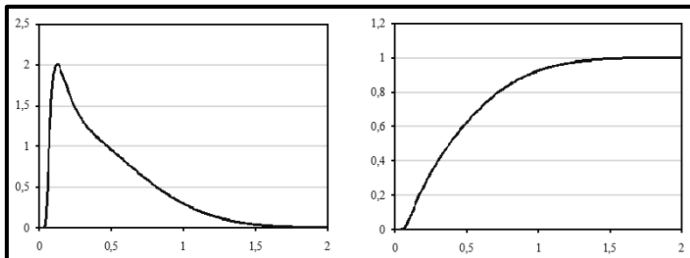
i	$E(u_i/\varepsilon_i)$	$Mo(u_i/\varepsilon_i)$	$P_{12.5}$	$P_{97.5}$	i	$E(u_i/\varepsilon_i)$	$Mo(u_i/\varepsilon_i)$	$P_{12.5}$	$P_{97.5}$
1	0.56	0.48	0.01	1.00	14	0.09	0.00	0.00	0.41
2	0.69	0.65	0.05	1.00	15	0.02	0.00	0.00	0.14
3	0.52	0.47	0.07	1.00	16	0.03	0.00	0.00	0.15
4	0.08	0.00	0.00	0.34	17	0.14	0.00	0.00	0.48
5	0.10	0.00	0.00	0.43	18	0.07	0.00	0.00	0.26
6	0.65	0.52	0.00	1.00	19	0.92	0.82	0.00	1.00
7	0.12	0.00	0.00	0.46	20	0.05	0.00	0.00	0.19
8	0.07	0.00	0.00	0.31	21	0.43	0.34	0.02	1.00
9	0.13	0.02	0.00	0.52	22	0.33	0.20	0.01	1.00
10	0.06	0.00	0.00	0.27	23	0.41	0.36	0.00	1.00
11	0.25	0.11	0.00	0.93	24	0.37	0.25	0.02	0.96
12	0.04	0.00	0.00	0.17	25	0.30	0.20	0.00	1.00
13	0.25	0.17	0.00	1.00	26	0.47	0.37	0.00	0.53

Source: Own elaboration.

The analysis of Table 4 provides consistent results, in trend and mean value, with those obtained by previous studies of efficiency in the Spanish banking system. The average cost efficiency estimate in the period was 0.72, that is, an average entity representative of the Spanish banking sector registered costs 28% higher than the potential minimum costs of the efficient bank with the same conditions. The study carried out by Maudos and Pastor (2003) about cost efficiency in Spanish banking obtained a result of 0.91 in the period 1985-1996, whereas Maudos and Fernández de Guevara (2007) presented an average of 0.88 for Spanish banking and of 0.86 for UE-15 banking in the period 1993-2000. More recently, Casu and Girardone (2009) confirmed a decreasing trend in Spanish banking during the period 2000-2005, with an average cost efficiency of 0.75.

The theoretical distribution of \hat{u} , when u is assumed to be? half normal, can be obtained by substituting the significant estimates $\hat{\sigma}^2 = 0.449$ and $\hat{\lambda} = 1.750$ in Equation 6. Figure 1 shows the shape of this distribution.

Figure 1
Density and distribution functions of average \hat{u} in the Banking System,
period 2002-2007

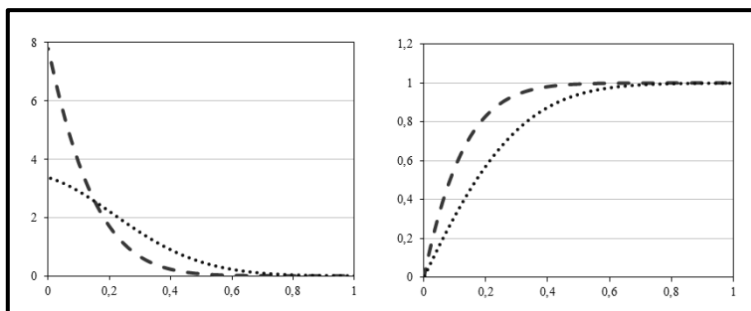


Source: Own elaboration.

The application of the Kolmogorov Smirnov test to contrast the goodness of fit of estimated inefficiency (Table 4) to the theoretical distribution of \hat{u} assuming u half normal, represented in Figure 1, returns a value of the statistic of 0.125. Therefore, the hypothesis of data from the assumed theoretical distribution is not rejected and, consequently, the initial assumption of half normal distribution of u is accepted with a p-value 0.410.

With regard to the second objective, Figure 2 represents the distribution of cost inefficiency estimate of the Spanish banking system in the first and last years of the period 2002-2007 to ascertain whether the behaviour previous to the financial crisis provides any indication of the coming crisis.

Figure 2
Density and distribution functions of average \hat{u} in the Banking System, 2002 and 2007



Notes: 2002 with lines and 2007 with dots.

Source: Own elaboration.

In Figure 2 left, density in 2002 shows a smaller dispersion than in 2007, with it being more probable to reach small values of inefficiency in the first year

of the period. In fact, distribution functions in 2002 and 2007, represented in Figure 2 right, reach the value 0.80 for the inefficiency value of 0.20 and 0.40, respectively. That is to say, cost inefficiency in 2007 achieves values up to 0.40 with a probability of 80% while in 2002, with the same probability, the value of cost inefficiency of 0.20 is not exceeded. This change in the shape of the distribution and in the behavior of cost inefficiency coincides with the beginning of the financial instability in the Spanish banking sector.

5. CONCLUSIONS

The assumption of the half normal distribution on the inefficiency term in a stochastic frontier model applied to the Spanish banking system is accepted as an appropriate assumption through the test of goodness of fit between the theoretical density of cost inefficiency estimate and the distribution of the estimated cost inefficiency of the units selected for the sample.

As regards the inefficiency estimate, the average entity representative of the Spanish banking system in the period 2002-2007 presented costs 28% higher than the potential minimum costs of an efficient entity with the same conditions. This result is consistent with point estimates published in previous studies on efficiency in the Spanish banking sector.

The most important conclusion of the study is obtained from density and distribution functions of cost inefficiency estimate, that provide more details in the estimate. The density of the cost inefficiency estimate in 2007 shows a significantly larger dispersion than the corresponding density in 2002, and hence the probability of obtaining higher levels of inefficiency grew from 2002 to 2007, when the financial crisis in Spanish banking sector began. An important management and political implication can be derived from these results because it is demonstrated that the analysis of the behavior of the distribution of cost inefficiency estimate in Spanish banking system is an indicator to take into account in unstable financial situations.

REFERENCES

- AIGNER, D.; LOVELL, C.A. and SCHMIDT, P. (1977). "Formulation and estimation of stochastic frontier production models". *Journal of Econometrics*, 6, pp. 21-37.
- ALTUNBAS, Y. and CHAKRAVARTY, S.P. (2001). "Frontier cost functions and bank efficiency". *Economics Letters*, 72, pp. 233-240.

- BERGER, A.N. and HUMPHREY, D.B. (1997). "Efficiency of financial institutions: international survey and directions for future research". *European Journal of Operational Research*, 98, pp. 175-212.
- BHANDARI, A.K. (2011). "On the distribution of estimated technical efficiency in stochastic frontier models: revisited". *International Journal of Business and Economics*, 10, pp. 69-80.
- CASU, B. and GIRARDONE, C. (2009). "Testing the relationship between competition and efficiency in banking: a panel data analysis". *Economic Letters*, 105, pp. 134-137.
- COELLI, T.J. (1995). "Estimators and hypothesis tests for a stochastic frontier function: a monte carlo analysis". *Journal of Productivity Analysis*, 6(4), pp. 247-268.
- DELIS, M.; KOUTSOMANOLI-FILLIPAKI, A.; STAIKOURAS, C. and KATERINA, G. (2009). "Evaluating cost and profit efficiency: a comparison of parametric and nonparametric methodologies". *Applied Financial Economics*, 19(3), pp. 191-202.
- DIOS-PALOMARES, R. (2002). "Análisis de interpretación de los parámetros de relación de varianzas en el modelo de frontera estocástica". *Estudios de Economía Aplicada*, 20(2), pp. 365-379.
- FRIED, H.O.; KNOX LOVELL, C.A. and SCHMIDT, S.S. (2008). *The Measurement of Productive Efficiency and Productivity Growth*. Oxford University Press.
- GREENE, W.H. (1990). "A gamma-distributed stochastic frontier model". *Journal of Econometrics*, 46, pp. 141-164.
- JONDROW, J.; LOVELL, C.A.K.; MATEROV, Y.S. and SCHMIDT, P. (1982). "On the estimation of technical inefficiency in the stochastic frontier production function model". *Journal of Econometrics*, 19, pp. 233-238.
- KUMBHAKAR, C. and LOVELL, C.A.K. (2000). *Stochastic Frontier Analysis*. Cambridge University Press.
- LEE, Y.H. and SCHMIDT, P. (2003). *A Production Frontier Model with Flexible Temporal Variation in Technical Inefficiency*. Fried H.O., Lovell C.A.K., Schmidt S.S. eds. Oxford University Press.
- MAUDOS, J. and FERNÁNDEZ DE GUEVARA, J. (2007). "The cost of market power in banking: social welfare loss vs. cost efficiency". *Journal of Banking and Finance*, 31(7), pp. 2103-2125.
- MAUDOS, J. and PASTOR, J.M. (2003). "Cost and profit efficiency in the Spanish banking sector (1985-96): a non-parametric approach". *Applied Financial Economics*, 13(1), pp. 1-12.
- MEEUSEN, W. and VAN DEN BROECK, J. (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error". *International Economic Review*, 18(2), pp. 435-444.
- PASIOURAS, F. (2008). "Estimating the technical and scale efficiency of Greek commercial banks: the impact of credit risk, off-balance sheet activities, and

- international operations". *Research in International Business and Finance*, 22, pp. 301–318.
- SEALEY, C.W.JR. and LINDLEY, J.T. (1977). "Inputs, outputs, and a theory of production and cost at depository financial institutions". *The Journal of Finance*, 32(4), pp. 1251-1266.
- STEVENSON, R.E. (1980). "Likelihood functions for generalized stochastic frontier estimation". *Journal of Econometrics*, 13, pp. 57-66.
- WANG, W.S.; AMSLER, C. AND SCHMIDT, P. (2011). "Goodness of fit test in stochastic frontier models". *Journal of Productivity Analysis*, 35(2), pp. 95-118.
- WANG, W.S. and SCHMIDT, P. (2009). "On the distribution of estimated technical efficiency in stochastic frontier models". *Journal of Econometrics*, 148, pp. 36-45.
- WEILL, L. (2003). "Banking efficiency in transition economies: the role of foreign ownership". *Economies of Transition*, 11, pp. 569-592.
- ZAJC, P. (2006). "A comparative study of bank efficiency in central and eastern Europe: the role of foreign ownership."

