

Gavilan Ruiz, José Manuel; Ortega Irizo, Francisco Javier

## Article

# Productive efficiency analysis of quantitative economics journals through stochastic frontier analysis using panel data

Revista de Métodos Cuantitativos para la Economía y la Empresa

## Provided in Cooperation with:

Universidad Pablo de Olavide, Sevilla

*Suggested Citation:* Gavilan Ruiz, José Manuel; Ortega Irizo, Francisco Javier (2020) : Productive efficiency analysis of quantitative economics journals through stochastic frontier analysis using panel data, Revista de Métodos Cuantitativos para la Economía y la Empresa, ISSN 1886-516X, Universidad Pablo de Olavide, Sevilla, Vol. 30, pp. 297-311, <https://doi.org/10.46661/revmetodoscuanteconempresa.3496>

This Version is available at:

<https://hdl.handle.net/10419/286216>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-sa/4.0/>



## Productive efficiency analysis of quantitative economics journals through Stochastic Frontier Analysis using panel data

GAVILAN, JOSÉ M.

Faculty of Economics and Business Studies

University of Seville (Spain)

Correo electrónico: [gavi@us.es](mailto:gavi@us.es)

ORTEGA, FRANCISCO J.

Faculty of Economics and Business Studies

University of Seville (Spain)

Correo electrónico: [fjortega@us.es](mailto:fjortega@us.es)

### ABSTRACT

The main goal of a scientific journal is to diffuse new knowledge. The number of citations received by a journal can be considered as a measure of this objective and, in turn, as a measure of productivity in relation to the production process in which the journals are involved. In order to assess this production process, in this paper econometric models using data panel are employed to obtain measures of efficiency for those journals belonging simultaneously to the areas of “economics” and “social science, mathematical methods” in the Web of Science database. This efficiency is measured in terms of the distance between the actual production of the journals and their estimated maximum achievable number of citations based on their available resources.

**Keywords:** production, productivity, efficiency, scientific production, frontier production models, panel data models.

**JEL classification:** C23; C46; C51; D24.

**MSC2010:** 62J99; 62F10; 90B30.

## Análisis de eficiencia productiva de revistas de economía cuantitativa a través del análisis de frontera estocástica utilizando datos de panel

### RESUMEN

El principal objetivo de una revista científica es difundir nuevo conocimiento. El número de citas recibidas por una revista se puede considerar como una medida de este objetivo y, a su vez, como una medida de productividad en relación al proceso de producción en el que las revistas están involucradas. Para evaluar este proceso de producción, en este artículo se emplean modelos econométricos que usan datos de panel para obtener medidas de eficiencia para aquellas revistas que pertenecen simultáneamente a las áreas “economics” y “social science, mathematical methods” en la base de datos Web of Science. Esta eficiencia se mide en términos de la distancia entre la producción real de las revistas y su número de citas máximo alcanzable estimado en base a sus recursos disponibles.

**Palabras clave:** producción, productividad, eficiencia, producción científica, modelos de producción con frontera, modelos de datos de panel.

**Clasificación JEL:** C23; C46; C51; D24.

**MSC2010:** 62J99; 62F10; 90B30.



## 1. Introduction.

In general terms, any production process consists of fabricating a series of products (outputs) using certain available resources (inputs). Comparing the productivity of the different entities involved in the production process, which is the production obtained in relation to the resources used, constitutes the main goal in this framework. This comparison, or benchmarking, is usually established in terms of efficiency in the production process. The efficiency of an entity is set as the difference between its actual production and its frontier or maximum attainable production with its available resources. To this end, several mathematical and econometric models have been developed in the field of Economics.

In the setting of the process of publishing papers by the different scientific journals, it is clear that the objective is to diffuse new knowledge, whose extent of diffusion could be measured through the number of readers of such articles. Scientists reading the papers published in a journal, use the content of these papers to develop their own articles, and provide citations to the papers they have read. As a consequence, the number of citations received by the journal can be considered as a proxy of the diffusion of the knowledge (output of the production process). In order to apply the aforementioned models, a set of homogeneous journals (each of which reaching approximately the same number of potential readers) has to be selected, and several inputs have been chosen. Great care must be taken in the selection of these inputs since they determine the production frontier and, in consequence, the efficiency of each journal. Candidate inputs in the framework of this production process could include those related to the number of people involved in the process, the available budget, and the number of articles published.

Once the inputs and outputs have been selected, there are two alternative approaches to assess the efficiency of the selected group of journals: data envelopment analysis (DEA) and stochastic frontier analysis (SFA). In Coelli, Rao and Battese (1998) and Bogetoft and Otto (2011), detailed explanations of the two methodologies are shown. These two approaches establish the production frontier in terms of the selected inputs and outputs, and the efficiency is then measured. The DEA is a non-parametric non-stochastic approach based on linear programming. Therefore, in this approach, no source of randomness is taken into account. This drawback is solved in the SFA approach by the inclusion of random perturbations in econometric models.

In the specific setting of scientific documentation, the DEA approach has been widely utilised. As a matter of fact, all the following papers use this approach. Abbott and Doucouliagos (2003) carry out an analysis of the efficiency of Australian universities. Similar studies include Bonaccorsi, Daraio and Simar (2006), where Italian universities are analysed, and Bonaccorsi and Daraio (2003), who consider institutes of the French INSERM and biomedical research institutes of the Italian CNR. Ruiz et al. (2010) examine the efficiency in the scientific production of a sample of Colombian research groups. Agasisti, Catalano, Landoni and Varganti (2012) analyse the production of 69 academic departments located in Italy. The DEA approach has also been applied to studies of efficiency in scientific production in a number of countries and regions. In Rousseau and Rousseau (1998), this approach is applied to a sample of 18 countries of the world, while in Guan and Chen (2010), 30 Chinese provinces are considered. In relation to the analysis of the efficiency applied to a group of journals, Lozano and Salmeron (2005) show the results of a DEA analysis applied to a group of journals of Operations Research/Management Sciences. Petridis et al. (2013) provide an evaluation of 54 forestry journals.

In this framework of scientific documentation, the SFA approach is used in Ortega and Gavilan (2013) to benchmark a group of quantitative economics journals. In that paper, the efficiency of the considered journals is estimated for the year 2011 using a cross-section database, which provides a snapshot of the situation of that year. However, this situation could be a fluke related to that particular year. Panel data provides evidence of a more reliable nature on the performance of the journals, and also enables this performance to be tracked over time. The SFA approach using panel data has been extensively applied to many fields, as, for example, in economics (Zhou, Li, & Li 2011; Kumbhakar & Zhang, 2013), healthcare (Greene, 2004; O'Donnell & Nguyen, 2013), universities (Abbott &

Doucouliafos, 2009; Sav, 2011), sports (Barros, Garcia-del-Barrio, & Leach, 2009; Park & Lee, 2012), fishing (Melfou, Theocharopoulos, & Papanagiotou, 2009), banking (Brissimis, Delis, & Tsionas, 2010; Parinduri & Riyanto, 2014), ports (Wang, Knox, & Lee, 2013), airlines (Kumbhakar, 1991; Merkert, Odeck, Brathen & Pagliari, 2012), agriculture and biological sciences (Maietta, 2002; Kumbhakar, Lien, & Hardaker, 2014), and energy (Filippini & Hunt, 2011; Stern, 2012). However, to the best of our knowledge, no application has yet been made in the field of the scientific documentation.

The main objective of this paper is to analyse the productive efficiency of a significant group of scientific journals that belong to the field of quantitative methods applied to Economics. To this end, a SFA model using panel data is utilised. The application of this methodology to the field of scientific documentation constitutes the main innovation of this paper; this approach is applied to generalize the cross-section model considered in Ortega and Gavilan (2013).

The remainder of this paper is arranged as follows. In Section 2, the set of journals analysed and the variables included in the model are selected. In Section 3, the econometric model employed to fit the panel data is presented. Section 4 is devoted to stating the main results provided by the model, and in Section 5 the conclusions reached in this paper are briefly shown.

## **2. Selection of journals and variables.**

As mentioned in the introduction, this paper generalizes the cross-sectional analysis carried out in Ortega and Gavilan (2013) to the case of panel data in order to attain better estimations and also to be able to follow the evolution over time of the efficiencies obtained. Therefore, for comparison purposes, the same set of journals and variables are to be considered.

With regard to the group of journals, a homogeneous group of which each has potentially the same quantity of readers is considered. Specifically, the 21 journals belonging simultaneously to the areas of “economics” and “social sciences, mathematical methods” of the JCR social sciences edition database for the analysed time span have been selected. Therefore, a group of journals relating to the field of quantitative economics has been chosen.

With respect to the output, a variable that measures the production (diffusion of the new knowledge) of the journals has to be selected. As argued in the introduction, the number of citations received can be considered as a proxy of that production (variable NC). It should be pointed out that, despite being the criteria most commonly used, this selection presents certain issues related to the fact that the number of citations received by a paper could be the consequence of reasons other than that of the value of its content (Callon, Courtial, & Penan, 1995; Ortega 2003).

In relation to the selection of the inputs, variables related to the quantity and quality of the articles published can be selected. Therefore, the total number of articles published by the journals can be considered. However, the citations received by a paper are retrospectively obtained since the citations refer to papers published before the citation is made (Basulto & Ortega, 2005; Gupta, 1997). This is the reason why the average number of articles published by the journal in the recent past (three previous years) has been selected (variable AP). Additionally, a variable related to the visibility of the journal is considered since the greater the number of people there are who read a paper, the more citations that paper will receive when these readers publish their own research. Moreover, it is clear that there is great interest in publishing in the journals with the highest levels of visibility. The impact factor of a journal can be taken as a measure of its visibility (Callon, Courtial, & Penan, 1995; Basulto & Ortega, 2005), since a better impact factor implies that the journal has been read by a greater quantity of researchers. In this sense, the act of selecting journals that cover the same subject area (quantitative economics) supports the choice of the impact factor as a visibility measure due to the similarity in the number of potential readers of these journals. As a consequence, although this selection is not standard practice, the average impact factor reached in the recent past (three previous years) by the journals is chosen

(variable IF). It could be considered that the impact factor is an output. However, it is precisely the fact of using data from the past that makes it an input for the future, since what a journal has achieved to date with respect to visibility is determinant of what it will be able to attain in the future.

The panel of data containing the values of the three aforementioned variables, NC, AP and IF, for each of the selected 21 journals and for the years 2010, 2011 and 2012, have been obtained from the JCR Social Science edition database of the Web of Science and are shown in Table 1. The journal QME-QUANT MARK ECON did not belong to the JRC database in the year 2007, and therefore its averages AP and IF for the year 2010 have been calculated using the information of only two (instead of three) previous years.

Other related inputs, for example those linked to the labour force or capital factor, could be considered. However, the variable representing the total number members of the editorial boards (as a variable representing the human factor) was found not to be of significance in Ortega and Gavilan (2013) and therefore has not been considered in the proposed model. With regard to the capital factor, it would be desirable to include reference to this in the model as, for example, the budget available for each journal. However, to date, this information has not been made available.

**Table 1. Journals selected and panel of data obtained.**

Abbreviated title of journal	NC <sub>2010</sub>	IF <sub>2010</sub>	AP <sub>2010</sub>	NC <sub>2011</sub>	IF <sub>2011</sub>	AP <sub>2011</sub>	NC <sub>2012</sub>	IF <sub>2012</sub>	AP <sub>2012</sub>
ASTIN BULL	482	0.603	28.00	408	0.779	33.00	434	0.641	31.67
ECONOMET J	516	0.654	36.00	510	0.725	32.00	573	0.765	30.00
ECONOMET REV	666	1.225	26.67	650	1.351	25.67	701	1.204	26.00
ECONOMET THEOR	1440	0.753	61.67	1507	0.842	67.67	1739	0.871	60.33
ECONOMETRICA	19858	3.612	53.00	19659	3.683	57.67	21481	3.387	58.00
EMPIR ECON	716	0.427	58.00	783	0.571	67.67	879	0.630	75.00
INSUR MATH ECON	1572	1.004	109.00	1567	1.205	123.00	1682	1.142	101.00
INT J GAME THEORY	944	0.442	42.00	780	0.527	43.67	910	0.473	41.33
J APPL ECONOMET	2135	1.334	48.00	2315	1.417	45.00	2600	1.578	48.67
J BUS ECON STAT	2616	1.513	32.67	2919	1.701	34.33	3117	1.678	41.67
J ECONOMETRICS	8246	1.894	145.33	8523	1.836	133.00	9614	1.689	126.00
J MATH ECON	938	0.400	69.33	893	0.466	88.33	855	0.432	87.00
J PROD ANAL	1045	0.596	34.67	1101	0.643	33.33	1033	0.753	35.00
JAHRB NATL STAT	118	0.145	23.00	113	0.239	30.67	152	0.289	37.00
MATH FINANC	1184	1.145	28.33	1098	1.168	28.67	1165	1.171	30.33
OXFORD B ECON STAT	1449	0.843	38.00	1501	0.993	37.67	1677	1.091	37.67
QME-QUANT MARK ECON	197	1.167	15.00	240	1.078	15.00	294	1.200	14.67
QUANT FINANC	701	0.779	64.67	712	0.701	77.00	938	0.649	100.00
REV ECON STAT	7462	2.171	59.33	7639	2.557	66.00	8266	2.701	79.33
SOC CHOICE WELFARE	779	0.576	72.00	725	0.615	71.00	820	0.578	68.00
THEOR DECIS	569	0.563	32.33	511	0.650	43.00	570	0.588	51.67

Source: Own elaboration.

### 3. Model.

The origins of the SFA econometric models are set in the (almost simultaneously published) papers of Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977), who share the composed

error structure for inefficiencies and statistical noise. The econometric model selected in this paper is a Cobb-Douglas production model where all the inputs and the output are taken in logarithmic terms.

Since one of the main objectives is to analyse the possible stability of the efficiencies over time and the changes of position of the journals in the yearly ranking, in this paper the efficiencies are left completely free and unrestricted in time instead of considering a panel data model based on the assumption of time-invariant efficiency as in the early models, or imposing any functional restriction to the time evolution of the efficiencies as in Kumbhakar (1990) or Battese and Coelli (1992). Moreover, a variable  $t$  is added to take into account the possible technological changes over time. This variable has been centred by taking the values -1 (for 2010), 0 (for 2011) and 1 (for 2012). Therefore, by considering the inputs and the output selected to assess the efficiency of the selected group of journals for the period 2010-2012, the following model is proposed:

$$\log(\text{NC}_{it}) = \beta_0 + \beta_1 \cdot \log(\text{IF}_{it}) + \beta_2 \cdot \log(\text{NA}_{it}) + \beta_3 \cdot t + v_{it} + \varepsilon_{it}, \quad i = 1, \dots, 21; \quad t = -1, 0, 1.$$

This is a Coob-Douglas type model. The more general translog model has been considered; however, this approach is not suitable since it presents the wrong skewness problem.

The stochastic frontier is  $\beta_0 + \beta_1 \cdot \log(\text{IF}_{it}) + \beta_2 \cdot \log(\text{NA}_{it}) + \beta_3 \cdot t + v_{it}$  where the stochastic part, representing the statistical noise, is described by the random variable  $v_{it}$  which, as usual, is considered to have a normal distribution with zero mean and variance  $\sigma_v^2$ , that is  $v_{it} \sim N(0, \sigma_v^2)$ . The other part of the composite error  $\varepsilon_{it}$  is the random variable  $u_{it}$ , which represents the inefficiency of the journals (distance between the production frontier or maximum quantity of citations attainable given the levels of the inputs and the citations actually received). This variable, as a distance, has to be selected positive and, as it is commonly set, it is taken following a half-normal distribution, that is  $u_{it} \sim \text{HN}(0, \sigma_u^2)$ . This is a normal distribution truncated to its positive values, in other words, it is the absolute value of a normal variable,  $u_{it} \sim \left| N(0, \sigma_u^2) \right|$ . All the considered random variables  $v_{it}$  and  $u_{it}$  are assumed to be independent. Other possibilities considered for modelling the inefficiency are the exponential (especially in the Bayesian methodology) and gamma distributions. In this paper, the half-normal model has been selected because it is the most commonly used. Furthermore, the main results remain unchanged when considering these other options.

It is important to point out that the selected model is neutral at the stochastic frontier. The possibility of utilising a non-neutral model has been discarded since it has problems of convergence, and of validity and reliability of the statistics tests.

Once the model is estimated, the quantity  $u_{it}$  provides the levels of the inefficiency of the journals in the year  $t$  (Jondrow, Lovell, Materov, & Schmidt, 1982), that is to say, it is considered that the inefficiencies are time varying. Furthermore, the likelihood ratio test has been carried out to test whether the inefficiencies can be accepted to be constant over time, which is clearly rejected since a virtually zero p-value is obtained. When the variables are expressed in logarithmic terms, as it is the case, it is more common to consider the efficiency measures  $\text{Eff}_{it} = \exp(-u_{it})$  because they take values between 0 (minimum level of efficiency) and 1 (maximum level of efficiency) and are easier to interpret than the inefficiency  $u_{it}$ .

The model considered is estimated through the maximum likelihood method by using the version 1.1-0 of the package *frontier* in the software R (Coelli & Henningsen, 2013). In Coelli (1996), a

description of the method for various kinds of models is presented. In this method, a reparameterisation of the model considered is set. Specifically, the parameters  $\sigma_{\varepsilon}^2 = \sigma_v^2 + \sigma_u^2$  (total variance of the composite error) and  $\gamma^2 = \sigma_u^2 / (\sigma_v^2 + \sigma_u^2)$  are utilised. This last parameter is related to the proportion of the total variance of the composite error, which is due to inefficiency. This parameter is crucial in the determination of the levels of inefficiency of the journals, and clearly ranges from 0 (where the whole composite error stems from random effects and there exist no differences in efficiency in the journals) to 1 (where the whole composite error is due to inefficiency and no statistical noise exists). Therefore, its statistical significance determines the presence of different levels of efficiency in the journals considered.

#### 4. Results.

As prior step to the estimation of the considered model, it has been checked that the wrong skewness problem is not present, since the skewness coefficient of the residuals is negative (specifically, its value is -0.2190).

The results of the maximum likelihood estimation of the parameters of the model, along with their standard errors, z-values and p-values are shown in Table 2, where the p-values have been marked according to their significance at 5% (\*), 1% (\*\*) or 0.1 (\*\*\*)

**Table 2. Results of the estimation of the proposed model through maximum likelihood.**

Parameters	Estimations	Standard errors	Z-values	p-values
$\beta_0$	4.1944	0.5740	7.3071	0.0000***
$\beta_1$	1.4022	0.0875	16.0303	0.0000***
$\beta_2$	0.8984	0.1295	6.9352	0.0000***
$\beta_3$	-0.0927	0.0711	-1.3029	0.1927***
$\sigma_{\varepsilon}^2$	0.3895	0.1579	2.4668	0.0136***
$\gamma$	0.7061	0.2749	2.5691	0.0102***

Source: Own elaboration.

The first significant conclusion that can be drawn from these results is that the inputs chosen in the econometric model (IF and NA) are highly significant in the establishment of the frontier (maximum attainable number of citations) and, as expected, both are positive, that is, an increase in the available resources of a journal leads to an increase in its level of citations. However, the variable introduced in order to account for technological change over time remains insignificant at the usual levels. Therefore, there is no evidence of that change.



With regard to the major parameter  $\gamma$ , it is significant at the usual 5% level. The estimated value 0.7061 establishes that 46.61% of the composite error is due to the variations in the inefficiencies of the analysed journals and the rest is statistical noise (Coelli, 1995; Ortega & Gavilan, 2014).

Once the model has been estimated, the specific efficiencies for each considered journal and year, whose determination forms the main goal of this type of model, are obtained and presented in Table 3 together with the average efficiency of the three years studied and the corresponding ranks to these averages. In order to graphically illustrate the situation, the obtained efficiencies are also shown in Figure 1. The efficiencies estimated for the journals show great stability over time, which is the reason because their averages of the three years and the corresponding ranks have been added in Table 3, thereby representing a summary of the three years. As a consequence of this stability, the three average efficiencies for the years considered are also very stable ranging between 0.66 and 0.71.

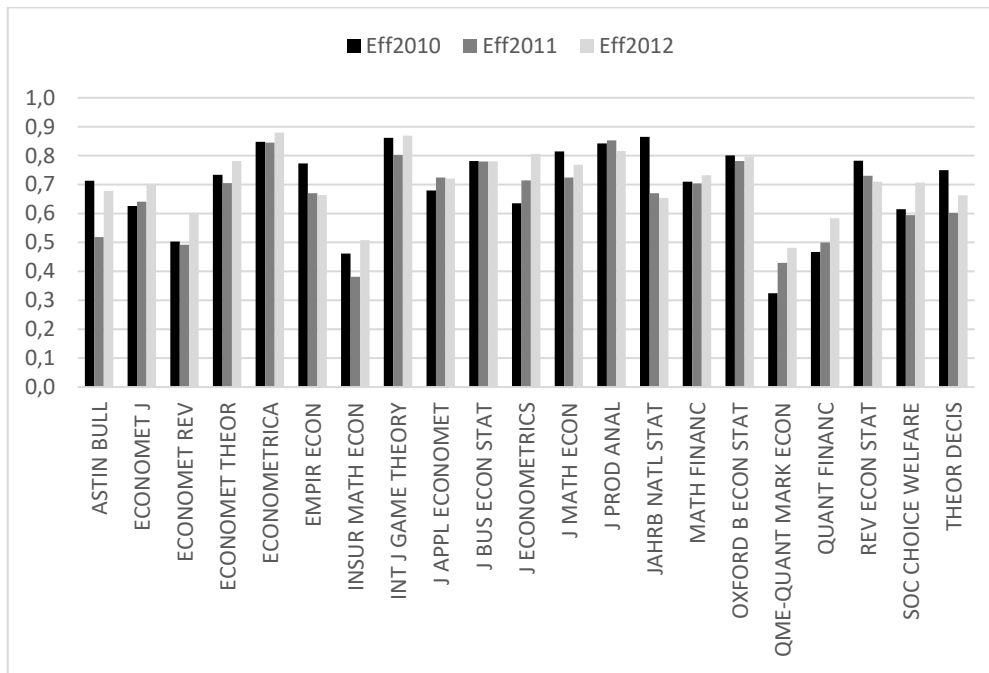
The stability of the efficiencies and their corresponding ranks can be quantified through Pearson's coefficients of correlation and Spearman's rank correlation, which are shown in Table 4. All Pearson's correlations are above 0.79 and all Spearman's rank correlations are above 0.62. As expected, the correlations between two consecutive years are greater than the correlations between 2010 and 2012.

**Table 3. Estimated efficiencies, their means by journal and by year and the ranks of the average efficiencies by year.**

	Eff <sub>2010</sub>	Eff <sub>2011</sub>	Eff <sub>2012</sub>	Averages	Rank of the Averages
ASTIN BULL	0.713	0.519	0.678	0.637	17
ECONOMET J	0.626	0.641	0.702	0.656	15
ECONOMET REV	0.503	0.492	0.601	0.532	18
ECONOMET THEOR	0.734	0.705	0.781	0.740	8
ECONOMETRICA	0.848	0.845	0.880	0.857	1
EMPIR ECON	0.773	0.670	0.663	0.702	13
INSUR MATH ECON	0.462	0.381	0.508	0.450	20
INT J GAME THEORY	0.862	0.802	0.869	0.844	2
J APPL ECONOMET	0.680	0.724	0.720	0.708	12
J BUS ECON STAT	0.781	0.780	0.780	0.780	5
J ECONOMETRICS	0.635	0.714	0.806	0.719	10
J MATH ECON	0.815	0.725	0.769	0.769	6
J PROD ANAL	0.842	0.853	0.816	0.837	3
JAHRB NATL STAT	0.865	0.670	0.654	0.730	9
MATH FINANC	0.710	0.704	0.732	0.715	11
OXFORD B ECON STAT	0.801	0.781	0.796	0.793	4
QME-QUANT MARK ECON	0.324	0.430	0.481	0.412	21
QUANT FINANC	0.467	0.500	0.583	0.517	19
REV ECON STAT	0.782	0.730	0.710	0.741	7
SOC CHOICE WELFARE	0.615	0.594	0.707	0.639	16
THEOR DECIS	0.750	0.602	0.663	0.672	14
Averages	0.695	0.660	0.710	0.688	

Source: Own elaboration.

**Figure 1. Estimated efficiencies of each journal for each one of the three years considered.**



Source: Own elaboration.

**Table 4. Correlations between the efficiencies of different years.**

	2010	2011	2012		2010	2011	2012
2010	1	0.8485	0.7967	2010	1	0.8681	0.8482
2011	0.7935	1	0.9250	2011	0.8677	1	0.9257
2012	0.6260	0.9078	1	2012	0.7308	0.8902	1

Source: Own elaboration.

Pearson's correlations are above the diagonal and Spearman's rank correlations are below the diagonal. Correlations involving all 21 journals considered are shown in the left-hand-side table and the correlations excluding the journal JAHRB NATL STAT are presented in the right-hand-side table.

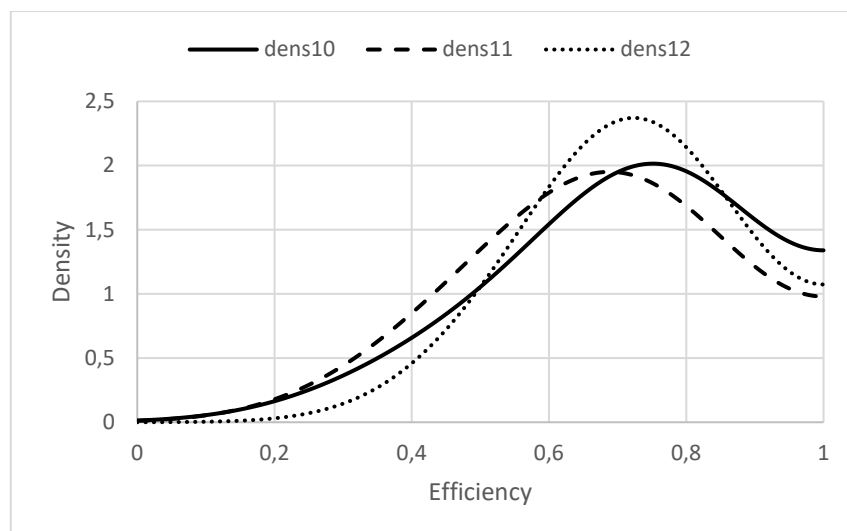
The journals ECONOMETRICA, INT J GAME THEORY, and J PROD ANAL are always ranked in the first positions. ECONOMETRICA always presents by far the greatest production (number of citations) of the group of analysed journals. However, the number of citations of INT J GAME THEORY and J PROD ANAL are not especially outstanding, although they are reached with moderate resources (visibility, IF, and number of articles published, AP), both in the recent past.

On the other hand, the journals QUANT FINANC, INSUR MATH ECON, and QME-QUANT MARK ECON always occupy the last positions. In the case of INSUR MATH ECON, although it presents a high number of citations in relation to the group of journals analysed and a visibility (IF) in the recent past near to the average of the group, it publishes a very high number of articles in relation to the group (about twice the average of the group). The journal QUANT FINANC has much fewer citations than INSUR MATH ECON and, although its IF is below the average of the group, it publishes a number of articles far above the average of the group. As for the journal QME-QUANT MARK, it presents one of the lowest number of citations (the second minimum) and, although it has the lowest number of articles published it presents an IF near to the average of the group.

The greatest exception to the aforementioned stability in the efficiencies reached for the analysed journals over time is the journal JAHRB NATL STAT, which presents the highest efficiency in 2010 and falls to the positions 13 in 2011 and 17 in 2012. This journal always has the lowest number of citations and IF. However, in 2010 it publishes the lowest number of articles of the group, while reaching the first position in the rank of that year. Nevertheless, as time goes on, this journal maintains the level of citations in 2011 and increases it by 29% in 2012 but doubles its IF and publishes 61% more articles than in 2010, which is why it does not maintain the level of efficiency reached in 2010 and loses numerous positions in the rankings of 2011 and 2012. These are precisely the kind of conclusions that can be drawn using a panel data approach, and cannot be observed through a cross-sectional analysis for a particular year. Specifically, if a cross-sectional analysis for 2010 is carried out, a misleading conclusion for the journal JAHRB NATL STAT could be reached, since its high efficiency is a fluke related to that particular year, which is not maintained over time. If this journal is removed from the sample, then the stability of the efficiencies increases significantly since all Pearson's correlations remain above 0.84, and all Spearman's rank correlations are above 0.73, and the correlations involving the year 2010 are now greater (Table 4).

In order to have a general overview of the efficiencies reached for the journals as a group, the densities fitted to all 21 journals for each of the three years considered is shown in Figure 2. These densities have been fitted while bearing in mind that the efficiencies range from 0 (minimum level efficiency) to 1 (maximum level of efficiency) (Silverman, 1986). In 2010, the average efficiency of the group is 0.695 and the mode 0.751 with 62% of the journals (13 out of 21) having a level of efficiency above 0.70. With regard to 2011, the average efficiency is 0.660 and the mode 0.689 with 52% (11 out of 21) of the journals whose efficiency is above 0.70, thereby indicating a slight decrease in relation to 2010. As for 2012, the average efficiency is 0.710 and the mode 0.724 with 62% of the journals (13 out of 21) having a level of efficiency above 0.70, which constitute very similar figures to those of 2010. Therefore, a great stability in time is observed and the group of journals operates at a high level of efficiency during the period of time analysed, whereby its global efficiency for the three years considered stands at 0.688.

**Figure 2. Densities fitted to the estimated efficiencies of the journals for each of the three years considered.**



Source: Own elaboration.

On the other hand, Table 2 shows that the two inputs considered (IF and AP) in the process of producing citations (diffusion of knowledge, NC) are highly significant, that is to say, the journals having, in the recent past, a greater number of articles published and higher levels of visibility measured through the impact factor are, potentially, able to reach a greater number of citations (production frontier). This is reinforced by the fact that the coefficient of linear correlation between each input (IF and AP) and the estimated frontier is high and significantly different from 0 (Table 5). However, unlike the production frontier, the level of efficiency is not determined by the two inputs considered, as the low coefficients of linear correlation and non-significantly different from zero in Table 5 show. There is a minor direct relationship between the levels of efficiency of the journals and their actual production in the form of NC, the output considered, since their linear correlation takes the value 0.3551, with a *p*-value for the significance test equal to 0.0043.

**Table 5. Linear correlation between the explanatory variables and: (i) the estimated frontier, (ii) the efficiency.**

	Frontier		Efficiency	
	Correlation	p-value	Correlation	p-value
IF	0.8870	0.0000	0.1607	0.2082
AP	0.5365	0.0000	-0.0927	0.4699

Source: Own elaboration.

## 5. Conclusions.

In this paper, the process of diffusing new knowledge by means of scientific journals through the publication of papers is set as an input-output process in order to assess the efficiency in that process of a homogeneous set of journals. These 21 journals belong simultaneously to the areas of “economics” and “social sciences, mathematical methods” for the years 2010, 2011 and 2012. The output and the inputs of that process have been carefully selected. The output, a proxy of the diffusion of new knowledge, consists of the number of citations received by the journal in a specific year (NC). Two inputs have been included in the model: one to take into account the number of articles published by the journal in the recent past (NA), and the other measures the visibility reached by the journal through the average of its impact factor in the recent past (IF). Other candidate inputs include the total number of members of the editorial boards of the journals as a representation of the labour force, which has not been included in this model since it yielded insignificant results (Ortega & Gavilan, 2013), and an input related to the capital factor as the budget available for the journal, which has also been left out of the model due to the unavailability of that information.

Once the inputs and outputs have been selected, the SFA approach is carried out to benchmark the selected group of journals. To this end, a Cobb-Douglas econometric model is applied to a panel of data, and a variable is also added to take into consideration the possible technological changes over time. This panel data approach, in which the inefficiencies are left completely free and unrestricted over time, generalizes the cross-sectional approach established in Ortega and Gavilan (2013), thereby providing better estimations and enabling the efficiency of the journals to be tracked over time.

The maximum likelihood estimation of the proposed model shows that the two selected inputs are highly significant in the determination of the production frontier (the maximum attainable number of citations by the journals). However, there is no evidence of technological change over time (Table 2).

The efficiencies obtained for the journals and for each of the years analysed, show great stability across time, and present high correlations between the years considered (Pearson's correlations above 0.79 and Spearman's rank correlations above 0.62). The journals *ECONOMETRICA*, *INT J GAME THEORY* and *J PROD ANAL* are always ranked in the first positions. On the other hand, the journals *QUANT FINANC*, *INSUR MATH ECON* and *QME-QUANT MARK ECON* are always in the last positions. The greatest exception to the aforementioned stability in the efficiencies over time is the journal *JAHRB NATL STAT*, which presents the highest efficiency in 2010 and falls to position 13 in 2011, and 17 in 2012. It should be borne in mind that this conclusion can be drawn using a panel data approach but cannot be observed through a cross-sectional analysis for a particular year.

The densities, fitted to all 21 journals for each of the years considered (Figure 2) in order to attain a general overview of the journals as a group, again show great stability over time and also show that the group of journals operates at a high level of efficiency during the period of time analysed: the global efficiency for the three years under consideration is of 0.688.

As aforementioned, the two inputs (IF and AP) considered in the process of diffusing new knowledge are highly significant. However, unlike the production frontier, the level of efficiency is not determined by the two inputs considered, as the low coefficients of linear correlation and non-significantly different from zero show (Table 5). A slight direct relation between the levels of efficiency of the journals and their number of citations is observed.

The study carried out in this paper could be extended to benchmark several sets of journals across time, and could also be extended through the inclusion of further relevant inputs in the econometric model.

## References

- Abbott, M., & Doucouliagos, C. (2003). The efficiency of Australian universities: A data envelopment analysis approach. *Economics of Education Review*, 22, 89-97.
- Abbott, M., & Doucouliagos, C. (2009). Competition and efficiency: overseas students and technical efficiency in Australian and New Zealand universities. *Education Economics*, 17, 31-57.
- Aigner, D.J., Lovell, C.A.K., & Schmidt, P. (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21-37.
- Agasisti, T., Catalano, G., Landoni, P., & Varganti, R. (2012). Evaluating the performance of academic departments: an analysis of research-related output efficiency. *Research Evaluation*, 21, 2-14.
- Barros, C.P., Garcia-del-Barrio, P., & Leach, S. (2009). Analysing the technical efficiency of the Spanish Football League First Division with a random frontier model. *Applied Economics*, 41, 3239-3247.
- Basulto, J., & Ortega, F.J. (2005). Modelling citation age data with right censoring. *Scientometrics*, 62, 329-342.

- Battese, G.E., & Coelli, T.J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis*, 3, 153-169.
- Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. Springer-Verlag.
- Bonaccorsi, A., & Daraio, C. (2003). A robust nonparametric approach to the analysis of scientific productivity. *Research Evaluation*, 12, 47-69.
- Bonaccorsi, A., Daraio, C., & Simar, L. (2006). Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics*, 66, 389-410.
- Brissimis, S.N., Delis, M.D., & Tsionas, E.G. (2010). Technical and allocative efficiency in European banking. *European Journal of Operational Research*, 204, 153-163.
- Callon, M., Courtial, J.P., & Penan, H. (1995). *Cienciometría. La medición de la actividad científica: de la Bibliometría a la vigilancia tecnológica*. Gijón: Ediciones Trea (in Spanish).
- Coelli, T.J. (1995). Estimators and hypothesis tests for a stochastic frontier function: A Monte Carlo analysis. *Journal of Productivity Analysis*, 6, 247-268.
- Coelli, T.J. (1996). *A guide to FRONTIER version 4.1: a computer program for frontier production function estimation*. CEPA Working Paper 96/07, Department of Econometrics, University of New England, Armidale, Australia.
- Coelli, T.J., & Henningsen, A. (2013) *Frontier: Stochastic Frontier Analysis. R package version 1.1-0*. <http://CRAN.R-Project.org/package=frontier>.
- Coelli, T.J., Rao, D.S.P., & Battese, G.E. (1998). *An introduction to efficiency and productivity analysis*. Boston: Kluwer Academic Publishers.
- Filippini, M., & Hunt, L.C. (2011). Energy Demand and Energy Efficiency in the OECD Countries: A Stochastic Demand Frontier Approach. *The Energy Journal*, 32, 59-80.
- Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, 10, 959-980.
- Guan, J., & Chen, K. (2010). Modeling macro-R&D production frontier performance: an application to Chinese province-level R&D. *Scientometrics*, 82, 165-173.
- Gupta, B.M. (1997). Analysis of distribution of the age of citations in theoretical population genetics. *Scientometrics*, 40, 139-162.
- Jondrow, J., Lovell, C.A.K, Materov, I.S., & 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. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of Econometrics*, 46, 201-211.
- Kumbhakar, S.C. (1991). Estimation of technical inefficiency in panel data models with firm- and time-specific effects. *Economics Letters*, 36, 43-48.
- Kumbhakar, S.C., Lien, G., & Hardaker, J.B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis*, 41, 321-337.
- Kumbhakar, S.C., & Zhang, R. (2013). Labor-use efficiency and employment elasticity in Chinese manufacturing. *Economia e Politica Industriale*, 40, 5-24.
- Lozano, S., & Salmeron, J.L. (2005). Data Envelopment Analysis of OR/MS journals. *Scientometrics*, 64, 133-150.
- Maietta, O.W. (2002). The decomposition of cost inefficiency into technical and allocative components with panel data of Italian dairy farms. *European Review of Agricultural Economics*, 27, 473-495.
- Meeusen, W., & van den Broeck, J. (1977). Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18, 435-444.
- Melfou, K., Theocharopoulos, A., & Papanagiotou, E. (2009). Assessing productivity change with SFA in the sheep sector of Greece. *Operational Research*, 9, 281-292.
- Merkert, R., Odeck, J., Brathen, S., & Pagliari, R. (2012). A Review of Different Benchmarking Methods in the Context of Regional Airports. *Transport Reviews*, 32, 379-395.
- O'Donnell, C.J., & Nguyen, K. (2013). An econometric approach to estimating support prices and measures of productivity change in public hospitals. *Journal of Productivity Analysis*, 40, 323-335.
- Ortega, F.J. (2003). Propuesta de mejora del índice agregado de impacto. Aplicación a la valoración de sexenios de investigación. *Revista Española de Documentación Científica*, 26, 403-417 (in Spanish).
- Ortega, F.J., & Gavilan, J.M. (2013). The measurement of production efficiency in scientific journals through stochastic frontier analysis models: application to quantitative economics journals. *Journal of Informetrics*, 7, 959-965.
- Ortega, F.J., & Gavilan, J.M. (2014). A comparison between maximum likelihood and Bayesian estimation of stochastic frontier production models. *Communications in Statistics - Simulation and Computation*, 43, 1714-1725.
- Parinduri, R.A., & Riyanto, Y.E. (2014). Bank Ownership and Efficiency in the Aftermath of Financial Crises: Evidence from Indonesia. *Review of Development Economics*, 8, 93-106.

- Park, I., & Lee, Y.H. (2012). Efficiency Comparison of International Golfers in the LPGA. *Journal of Sports Economics*, 13, 378-392.
- Petridis, K., Malesios, C., Arabatzis, G., & Thanassoulis, E. (2013). Efficiency analysis of forestry journals: Suggestion for improving journals' quality. *Journal of Informetrics*, 7, 505-521.
- Rousseau, S., & Rousseau, R. (1998). The scientific wealth of European nations: Taking effectiveness into account. *Scientometrics*, 42, 75-87.
- Ruiz, C. F., Bonilla, R., Chavarro, D., Orozco, L. A., Zarama, R., & Polanco X. (2010). Efficiency measurement of research groups using Data Envelopment Analysis and Bayesian networks. *Scientometrics*, 83, 711-721.
- Sav, G.T. (2011). Cost Efficiencies and Rankings of Flagship Universities. *American Journal of Economics and Business Administration*, 3, 596-603.
- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall: London.
- Stern, D.I. (2012). Modeling international trends in energy efficiency. *Energy Economics*, 34, 2200-2208.
- Wang, G.W., Knox, K.J., & Lee, P.T.-W. (2013). A study of relative efficiency between privatised and publicly operated US ports. *Maritime Policy and Management*, 40, 351-366.
- Zhou, X., Li, K.-W., & Li, Q. (2011). An analysis on technical efficiency in post-reform China. *China Economic Review*, 22, 357-372.