

## The predictive ability of relative efficiency for future earnings: an application using data envelopment analysis to Spanish SMEs

Reverte, Carmelo; Guzman, Isidoro

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

[www.peerproject.eu](http://www.peerproject.eu)

### Empfohlene Zitierung / Suggested Citation:

Reverte, C., & Guzman, I. (2008). The predictive ability of relative efficiency for future earnings: an application using data envelopment analysis to Spanish SMEs. *Applied Economics*, 42(21), 2751-2757. <https://doi.org/10.1080/00036840801964575>

### Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.



**THE PREDICTIVE ABILITY OF RELATIVE EFFICIENCY FOR FUTURE EARNINGS: AN APPLICATION USING DATA ENVELOPMENT ANALYSIS TO SPANISH SMEs**

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-07-0063.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	01-Feb-2008
Complete List of Authors:	REVERTE, CARMELO; Technical University of Cartagena, Department of Accounting and Finance GUZMAN, ISIDORO; Technical University of Cartagena, Department of Accounting and Finance
JEL Code:	C14 - Semiparametric and Nonparametric Methods < C1 - Econometric and Statistical Methods: General < C - Mathematical and Quantitative Methods, D24 - Production Capital and Total Factor Productivity Capacity < D2 - Production and Organizations < D - Microeconomics
Keywords:	relative efficiency, earnings prediction, DEA, SMEs



# ***THE PREDICTIVE ABILITY OF RELATIVE EFFICIENCY FOR FUTURE EARNINGS: AN APPLICATION USING DATA ENVELOPMENT ANALYSIS TO SPANISH SMEs***

## **1.INTRODUCTION**

Research on the time-series properties of earnings has documented that the transitory component of current earnings makes this number, when used by itself, a noisy predictor of future earnings (e.g., Brooks and Buckmaster, 1976; Beaver and Morse, 1978). This transitory component of earnings is due to temporary economic shocks that do not persist over time. The implication of this finding is that a measure that better captures the underlying persistent component of earnings may prove very useful in predicting future earnings when used along with current earnings. As pointed out by Lee (1999: 418), *“future empirical research needs to look beyond past earnings and book value, and seek to understand what other (preferably observable) financial and nonfinancial information might be useful in predicting future abnormal earnings (...) The aim is to understand what else, besides current earnings, might help us predict future earnings”*.

Our study, based on a large sample of 1939 Spanish small and medium enterprises (hereafter SMEs), is aimed at investigating whether an alternative measure of performance ignored in previous research on earnings forecasting -i.e., relative efficiency- is useful for predicting future earnings over and above current earnings and book value of equity. Relative efficiency captures the inherent ability of a firm -as compared to the other firms under analysis- to generate the maximum level of outputs given a certain level of inputs. In this paper we use a non-parametric technique based on linear programming referred to as Data Envelopment Analysis (DEA) in order to provide a numerical measure of the relative efficiency of business units in using available inputs to generate the maximum level of outputs. This technique has the advantage over other methodologies (such as the stochastic frontier analysis -SFA-) of avoiding the use of any *a priori* assumptions about the functional form of the production function.

Our findings highlight that our efficiency measure has an incremental predictive ability over and above current earnings and book value of equity for predicting future earnings. Moreover, we have further validated the models in a holdout sample and our results evidence the highest forecast accuracy of the model that includes our DEA-based efficiency measure as an additional predictor to current earnings and book value of equity.

The rest of the paper is structured as follows. In the next section we discuss the DEA methodology used in our empirical analysis to measure relative efficiency. In Section 3 we present the research design, including the methodology, variables, hypotheses and a description of the

sample. In section 4 we present the results from the empirical analysis. Finally, we summarise the main conclusions of our study.

## 2. EFFICIENCY MEASURE: DATA ENVELOPMENT ANALYSIS

Data Envelopment Analysis (DEA) is a non-parametric technique developed by Charnes *et al.* (1978) that using linear programming determines the relative efficiency of decision making units (DMUs) in the use of single or multiple inputs for the production of single or multiple outputs<sup>1</sup>. The advantage of this technique is that there is no need to make any *a priori* assumptions about the functional form of the production function. Instead, an empirical production function, in the form of a piece-wise linear frontier, is fitted to the observed data based on the amounts of outputs and inputs of the business units under analysis so that the relative efficiency is assessed against the benchmark of the so-called “best-practice frontier”. Those DMUs placed on the frontier are deemed to be efficient while those DMUs situated below the frontier are regarded as inefficient<sup>2</sup>.

Assuming that there are  $n$  DMUs that use a vector of  $m$  inputs to produce a vector of  $s$  outputs, DEA allows to determine the efficiency ( $h_z$ ) of each DMU ( $DMU_z$ ) by solving the following linear programming problem:

$$\text{Max } h_z = \frac{\sum_r u_r y_{rz}}{\sum_i v_i x_{iz}} \quad [1]$$

subject to:

$$h_z = \frac{\sum_r u_r y_{rf}}{\sum_i v_i x_{if}} \leq 1 \quad f = 1, 2, \dots, n \quad [2]$$

$$u_r \geq 0, \quad r = 1, \dots, s$$

$$v_i \geq 0, \quad i = 1, \dots, m$$

where  $y_r$  and  $x_i$  represent the amount of output  $r$  and input  $i$ , respectively while  $u_r$  and  $v_i$  represent their corresponding weights.

The previous formulation corresponds to the original model developed by Charnes *et al.* (1978), known as the CCR model, which assumes that all DMUs operate under constant returns to scale. This assumption was relaxed in the model proposed by Banker *et al.* (1984) –known as the

<sup>1</sup> The DEA model can have two different orientations: a) an output orientation, in which the maximum amount of outputs for a given level of inputs is pursued; b) an input orientation, in which it is pursued to consume the minimum amount of inputs to achieve a given level of outputs.

<sup>2</sup> For a review of the DEA approach and its methodological problems, see for instance Färe and Zelenyuk (2002), Sengupta (2002), Färe *et al.* (2004) and Barnum and Gleason (2005, 2006).

BCC model- that allows for variable returns to scale. Banker (1993) proved that in large samples the BCC estimator is consistent because the distance between the DEA estimate and the true efficiency score goes towards zero.

The output-oriented BCC model (Banker *et al.*, 1984) to estimate the efficiency of each DMU is given by solving the following linear programming:

$$\text{Max } \theta_z \quad [3]$$

s.a.:

$$\sum_{f=1}^n \lambda_f x_{if} + S^i = x_{iz} \quad i = 1, \dots, m \quad [4]$$

$$\theta_z y_{rz} - \sum_{f=1}^n \lambda_f y_{rf} + S^o = 0 \quad r = 1, \dots, s \quad [5]$$

$$\sum_{f=1}^n \lambda_f = 1 \quad [6]$$

$$\theta_z \geq 0; \lambda_f \geq 0 \quad f = 1, \dots, n \quad [7]$$

where  $\theta_z$  is the efficiency score for the DMU  $z$  being evaluated,  $x_{if}$  is the amount of input  $i$  consumed by DMU  $f$ ,  $y_{rf}$  is the amount of output  $r$  produced by DMU  $f$ ,  $x_{iz}$  is the amount of input  $i$  consumed by the DMU  $z$  being evaluated,  $y_{rz}$  is the amount of output  $r$  produced by the DMU  $z$  being evaluated,  $\lambda_f$  is the weight on the inputs/outputs of DMU  $f$  and  $S^o$  and  $S^i$  are the slacks of the linear programming model.

The optimal value ( $\theta_z$ ) of the linear program ranges between 1 and  $\infty$ , such that it is equal to 1 for efficient units and higher than 1 for inefficient units. The estimated relative efficiency measure ( $\varphi_z = 1/\theta_z$ ) varies between 0 and 1. Therefore, a DMU with a relative efficiency measure ( $\varphi_z$ ) equal to 1 is regarded as efficient whereas if it takes a value of, for example, 0.85 this would imply that the DMU being evaluated will have to increase its output an 15% of the estimated value in the efficient frontier.

### 3. RESEARCH DESIGN

#### 3.1. Methodology, variables and hypothesis

As efficiency is sector-sensitive, we have splitted our sample of SMEs in several sectors and computed the efficiency of each firm in comparison with the rest of firms in the same sector. After analysing the data, firms were grouped in the following three sectors: a) Manufacturing; b) Service sector and c) Construction.

The methodology used in our empirical study is based on three stages. In the first one, we obtain for each year and firm the DEA-based efficiency scores. As in Abad *et al.* (2006), the proposed DEA model takes the inputs and outputs that correspond with those of a typical production function (see Table 1).

---

**Table 1. DEA model (output orientation)**

---

**Output:**

- Revenues

**Inputs:**

- Cost of Materials consumed	—→	Materials
- Personnel expenses	—→	Labor
- Depreciation expense	—→	Capital
- Other operating expense	—→	Overhead

---

Regarding the DEA methodology, we have computed the efficiency scores based on Banker *et al.* (1984) that consider variable returns to scale (BCC model). Moreover, we have chosen an output orientation as we are interested in relating our efficiency measure to future earnings. Therefore, our model will be oriented towards achieving the maximum level of revenues for a given level of inputs.

In the second stage, we estimate the following regression models (see Table 2) in order to test the incremental predictive ability of our relative efficiency measure over and above current earnings and book value of equity.

---

**Table 2. Earnings prediction models**

---

**Model 1 :**  $NI_{i,t+1} = \alpha_0 + \alpha_1 NI_{i,t} + \alpha_2 BV_{i,t} + e_{i,t}$

**Model 2 :**  $NI_{i,t+1} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 BV_{i,t} + \beta_3 EFFIC_{i,t} + e_{i,t}$

---

Variables:  $NI_{i,t+1}$ : Net income of firm  $i$  in year  $t+1$ ;  $NI_{i,t}$ : Net income of firm  $i$  in year  $t$ ;  $BV$ : Book value of equity for firm  $i$  at the end of year  $t$ ;  $EFFIC_t$ : Efficiency score of firm  $i$  in year  $t$ .

The first model includes current earnings and book value of equity as regressors. Regarding the former, there is a huge number of studies that have shown the predictive ability of current earnings for predicting future earnings (e.g., Ball and Watts, 1972; Albretch *et al.*, 1977; Watts and Leftwich, 1977). Concerning book value of equity, the Ohlson valuation model (1995) shows its usefulness for predictive purposes, being that higher the lower the persistence of current earnings. In this respect, Shroff (1999) –for the US- and Reverte (2003) –for Spain- have corroborated the predictive ability of book value of equity for the case of listed firms. The second model includes our relative efficiency measure as an additional regressor in order to test the following hypothesis: *The*

DEA efficiency measure in year  $t$  is positively associated with next year's ( $t+1$ ) earnings, after controlling for the effect of current earnings and book value of equity. To the extent our efficiency measure captures the permanent component of earnings, it should be able to predict future earnings over and above current earnings and book value. Therefore, we expect  $\beta_3$  to be positive and statistically significant.

Finally, in the last stage of our study, we validate our models in a holdout sample different to that used in their estimation by computing the differences in earnings forecasted by models (1) and (2) and actual earnings corresponding to our sample of firms for year 2004. In order to assess the forecast accuracy of both models, we compute the mean absolute percentage error, whose formulation is shown on Table 3. This measure has two advantages. First, as forecast errors are taken in absolute value, it avoids negative and positive errors to compensate each other. Second, it is not sensitive to scale as errors are divided by the actual value of the forecasted variable.

**Table 3. Forecast accuracy measure**

$$\text{Mean absolute percentage error} = \frac{1}{N} \sum \left| \frac{e_i}{y_i} \right|$$

Notes:

$e_i$ : forecast error (difference between the actual and forecasted value).

$y_i$ : actual value of forecasted variable.

$N$ : number of observations in the forecast period.

Once we have computed the mean absolute percentage error, we examine whether the differences in the forecast errors derived from models (1) and (2) are statistically significant by applying the non-parametric Wilcoxon signed-rank test, which is appropriate when the distribution of the errors departs from normality.

### 3.2. Sample

Data used in our study are obtained from the database SABI. We have chosen small and medium-sized firms (SMEs) with the legal form of public or limited liability companies<sup>3</sup>. The sample period ranges from 1999 to 2003, including additionally the year 2004 in order to validate our models. To ensure reliability of our data, we include only SMEs with a favourable audit report. As a result of the previous filters, our sample comprises 1939 SMEs and a total of 7756 firm/year observations for our four-year estimation period. The composition of the sample by sector is as follows: a) Manufacturing sector (802 firms/3208 observations), b) service sector (957 firms /3828

<sup>3</sup> In order to define a firm as SME, we have followed the *Recommendation 2003/361/EC of the European Commission on SME definition* issued on 6<sup>th</sup> May 2003, selecting the maximum levels proposed in this Recommendation, namely: total assets, up to 43 million euros; sales revenues, up to 50 million euros, and average number of employees, up to 250.

observations) and c) construction sector (180 firms/720 observations). Table 4 reports the descriptive statistics of our regressors (i.e., earnings, book value of equity and efficiency) for each of the three sectors and also for the aggregate sample. As it can be observed, the mean efficiency scores are quite similar across sectors, ranging from 81.7% for the manufacturing sector to 85.4% for the service sector. The mean efficiency value for the pooled sample is 83.6%.

<b>Panel A: Manufacturing sector (N=3208 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
Mean	611.22	5018.69	0.817
Median	374.00	3680.00	0.810
Standard deviation	838.69	4475.07	0.099
Minimum	-4650.00	-2654.00	0.082
Maximum	10440.00	33853.00	1.000
<b>Panel B: Service Sector (N=3828 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
Mean	453.69	3493.08	0.854
Median	277.00	2360.05	0.856
Standard deviation	680.56	3603.72	0.088
Minimum	-4201.00	-1459.00	0.171
Maximum	7379.00	31869.00	1.000
<b>Panel C: Construction sector (N=720 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
Mean	502.80	3019.87	0.828
Median	328.00	2246.00	0.826
Standard deviation	696.75	2835.88	0.121
Minimum	-4048.00	-733.00	0.088
Maximum	5857.00	19916.00	1.000
<b>Panel D: Aggregate sample (N=7756 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
Mean	523.42	4080.29	0.836
Median	316.00	2785.00	0.835
Standard deviation	755.05	4009.44	0.097
Minimum	-4650.00	-2654.00	0.082
Maximum	10440.00	33853.00	1.000

**Table 5. Correlations among explanatory variables**

<b>Panel A: Manufacturing sector (N=3208 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
<b>NI<sub>t</sub></b>		0.624	0.425
<b>BV<sub>t</sub></b>			0.209
<b>Panel B: Service Sector (N=3828 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
<b>NI<sub>t</sub></b>		0.567	0.413
<b>BV<sub>t</sub></b>			0.216
<b>Panel C: Construction Sector (N=720 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
<b>NI<sub>t</sub></b>		0.638	0.383
<b>BV<sub>t</sub></b>			0.207
<b>Panel D: Aggregate sample (N=7756 obs.)</b>			
	<b>NI<sub>t</sub></b>	<b>BV<sub>t</sub></b>	<b>EFFIC<sub>t</sub></b>
<b>NI<sub>t</sub></b>		0.604	0.386
<b>BV<sub>t</sub></b>			0.170

Notes: NI<sub>t</sub>: Net income of firm *i* in year *t* (thousands of euros); BV: Book value of equity for firm *i* at the end of year *t* (thousands of euros); EFFIC<sub>t</sub>: Efficiency score of firm *i* in year *t*.



Table 5 shows the correlation matrix among the regressors. It can be observed that our efficiency measure is not highly correlated with both current earnings and book value of equity in the three sectors. Specifically, for the pooled sample, that correlations are 0.39 and 0.17, respectively. Hence, we can assert that the information captured by our efficiency measure is not already contained in both current earnings and book value of equity, avoiding any collinearity problem.

## 4.RESULTS

### 4.1. Cluster analysis

In order to get a first approximation of descriptive nature on the relationship between our efficiency measure and future earnings for our sample of Spanish SMEs, we have performed a  $k$ -means cluster analysis. Specifically, for each of the three sectors and for the pooled sample, we have identified three clusters based on the current value of the efficiency measure ( $EFFIC_t$ ). Then, we compute for each cluster the mean value of next-year's earnings ( $NI_{t+1}$ ) and, finally, we perform the Kruskal-Wallis test in order to examine whether the differences in the mean value of next-year's earnings across the three clusters are statistically significant. Results from this cluster analysis are reported on Table 6.

<b>Table 6. Cluster Analysis</b>			
<b>Panel A: Manufacturing Sector (N=3208 obs.)</b>			
<b>Cluster</b>	<b>N</b>	<b>EFFIC<sub>t</sub></b>	<b>NI<sub>t+1</sub></b>
1	1045	0.711	266.12
2	1448	0.824	549.21
3	715	0.955	1218.17
Kruskal-Wallis test for $NI_{t+1}$ : $\chi^2 = 460.03$ ( $p$ -value: 0.0001)			
<b>Panel B: Service Sector (N=3828 obs.)</b>			
<b>Cluster</b>	<b>N</b>	<b>EFFIC<sub>t</sub></b>	<b>NI<sub>t+1</sub></b>
1	845	0.737	208.19
2	1819	0.845	370.76
3	1164	0.952	810.86
Kruskal-Wallis test for $NI_{t+1}$ : $\chi^2 = 452.56$ ( $p$ -value: 0.0001)			
<b>Panel C: Construction Sector (N=720 obs.)</b>			
<b>Cluster</b>	<b>N</b>	<b>EFFIC<sub>t</sub></b>	<b>NI<sub>t+1</sub></b>
1	229	0.694	333.26
2	278	0.826	550.34
3	213	0.975	938.12
Kruskal-Wallis test for $NI_{t+1}$ : $\chi^2 = 49.12$ ( $p$ -value: 0.0001)			
<b>Panel D: Aggregate sample (N=7756 obs.)</b>			
<b>Cluster</b>	<b>N</b>	<b>EFFIC<sub>t</sub></b>	<b>NI<sub>t+1</sub></b>
1	2096	0.718	261.17
2	3559	0.835	454.68
3	2101	0.956	949.51
Kruskal-Wallis test for $NI_{t+1}$ : $\chi^2 = 798.55$ ( $p$ -value: 0.0001)			

Notes: N: Number of observations in each cluster;  $NI_{t+1}$  = Net income in t+1 (thousands of euros);  $EFFIC_t$ : Efficiency score in year t.

As can be shown on Table 6, in the three sectors the higher the value of efficiency in the current period the higher the value of next-year's earnings. For cluster 1, comprised by those observations with the lowest values of the efficiency measure (0.718 for the aggregate sample), the mean value of next-year's earnings is 261.17 millions euros. On the contrary, in cluster 3, comprised by those observations with the highest values of the efficiency measure (0.956 for the aggregate sample), the mean value of next-year's earnings is 949.51 millions euros. The Kruskal-Wallis test evidences that the differences in the mean value of next-year's earnings across the three clusters are statistically significant at the 1% level for the three sectors and also for the aggregate sample.

## 4.2. Estimation results

Table 7 reports the results for the pooled sample period (1999-2003) regarding the estimation of the two earnings prediction models (1) and (2) for the three sectors and for the aggregate sample.

Table 7. Estimation results of models (1) and (2)

<b>Model 1 : <math>NI_{i,t+1} = \alpha_0 + \alpha_1 NI_{i,t} + \alpha_2 BV_{i,t} + e_{i,t}</math></b>						
<b>Model 2 : <math>NI_{i,t+1} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 BV_{i,t} + \beta_3 EFFIC_{i,t} + e_{i,t}</math></b>						
<b>Panel A: Manufacturing sector (N=3208 obs.)</b>						
<b>Model</b>	<b>Intercept</b>	<b><math>NI_t</math></b>	<b><math>BV_t</math></b>	<b><math>EFFIC_t</math></b>	<b>F-Test (<math>\beta_3 = 0</math>)</b>	<b>Adj. R<sup>2</sup></b>
Model 1	56.93* (3.45)	0.82* (17.87)	0.01 (1.60)			0.640
Model 2	-387.84* (-3.28)	0.79* (15.89)	0.01*** (1.80)	561.52* (3.65)	13.32 (p-value=0.000)	0.644
<b>Panel B: Service Sector (N=3828 obs.)</b>						
	<b>Intercept</b>	<b><math>NI_t</math></b>	<b><math>BV_t</math></b>	<b><math>EFFIC_t</math></b>	<b>F-Test (<math>\beta_3 = 0</math>)</b>	<b>Adj. R<sup>2</sup></b>
Model 1	53.70* (4.31)	0.76* (15.50)	0.02* (3.32)			0.603
Model 2	-284.91* (-2.20)	0.74* (13.56)	0.01* (3.37)	407.28* (2.57)	6.58 (p-value=0.010)	0.605
<b>Panel C: Construction sector (N=720 obs.)</b>						
	<b>Intercept</b>	<b><math>NI_t</math></b>	<b><math>BV_t</math></b>	<b><math>EFFIC_t</math></b>	<b>F-Test (<math>\beta_3 = 0</math>)</b>	<b>Adj. R<sup>2</sup></b>
Model 1	67.33 (1.21)	0.47* (4.27)	0.10* (3.89)			0.424
Model 2	-338.94*** (-1.65)	0.43* (3.75)	0.10* (3.97)	508.68*** (1.89)	3.57 (p-value=0.059)	0.428
<b>Panel D: Aggregate sample (N=7756 obs.)</b>						
	<b>Intercept</b>	<b><math>NI_t</math></b>	<b><math>BV_t</math></b>	<b><math>EFFIC_t</math></b>	<b>F-Test (<math>\beta_3 = 0</math>)</b>	<b>Adj. R<sup>2</sup></b>
Model 1	61.77* (6.23)	0.78* (23.93)	0.02* (4.20)			0.600
Model 2	-307.12* (-3.92)	0.75* (21.37)	0.02* (4.42)	452.16* (4.60)	21.16 (p-value=0.000)	0.603

Notes: White-adjusted  $t$ -statistics are reported in parentheses. \*, \*\*, \*\*\* indicate significance at 1%, 5% and 10%, respectively. The  $F$ -value tests the hypothesis that the coefficient on  $EFFIC$  is zero in model (2).

It can be observed that, as hypothesised, our efficiency measure is positively and significantly associated with next-year's earnings. The  $F$ -test indicates that efficiency has an incremental explanatory power over and above current earnings and book value of equity for predicting next-year's earnings in the three sectors (manufacturing, services and construction). For the pooled sample, the  $F$ -test is statistically significant at the 1% level ( $p$ -value=0.00). This finding corroborates our main hypothesis, i.e., the more efficient firms in the use of available resources –as compared to the rest of firms in the sample- have higher earnings in the following period than less efficient firms.

### 4.3. Validation results

In the predictive ability studies the relevant test is the validation of the models in a holdout sample different to that used in their estimation. In this respect, our hypothesis is as follows: *The model including our efficiency measure (model 2) predicts more accurately next-year's earnings than the model that only includes current earnings and book value of equity (model 1).*

We first compute the differences in earnings forecasted by models (1) and (2) and actual earnings corresponding to our sample of firms for year 2004. In order to assess the forecast accuracy of both models, we compute the mean absolute percentage error.

<b>Table 8. Validation of the models</b>		
<b>Panel A: Manufacturing Sector</b>		
	<b>Model 1</b>	<b>Model 2</b>
<i>Mean absolute percentage error</i>	2.87	2.63
Wilcoxon test: $Z=-3.02$ ( $p$ -value: 0.0025)		
<b>Panel B: Service Sector</b>		
	<b>Model 1</b>	<b>Model 2</b>
<i>Mean absolute percentage error</i>	1.41	1.31
Wilcoxon test: $Z=-3.62$ ( $p$ -value: 0.0003)		
<b>Panel C: Construction Sector</b>		
	<b>Model 1</b>	<b>Model 2</b>
<i>Mean absolute percentage error</i>	1.53	1.38
Wilcoxon test: $Z=-1.71$ ( $p$ -value: 0.0873)		
<b>Panel D: Aggregate Sample</b>		
	<b>Model 1</b>	<b>Model 2</b>
<i>Mean absolute percentage error</i>	2.18	2.10
Wilcoxon test: $Z=-3.87$ ( $p$ -value: 0.0001)		

Notes: Mean absolute percentage error (defined in Table 3)

Model 1 :  $NI_{i,t+1} = \alpha_0 + \alpha_1 NI_{i,t} + \alpha_2 BV_{i,t} + e_{i,t}$

Model 2 :  $NI_{i,t+1} = \beta_0 + \beta_1 NI_{i,t} + \beta_2 BV_{i,t} + \beta_3 EFFIC_{i,t} + e'_{i,t}$

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 8 reports, for the three sectors and the pooled sample, the validation results as well as the Wilcoxon signed-rank test in order to examine whether the differences in the forecast errors derived from models (1) and (2) are statistically significant. Our results document that the model including the efficiency measure (model 2) has the highest forecast accuracy for the three sectors since the forecast errors are lower than model 1. Moreover, the Z-statistic of the Wilcoxon test indicates that the differences between the forecast errors of both models are statistically significant ( $p$ -value= 0.0001 for the pooled sample).

## 5. CONCLUDING REMARKS

Research on earnings prediction has documented that the transitory component of current earnings makes that figure less useful for predicting future earnings (e.g., Brooks and Buckmaster, 1976; Beaver and Morse, 1978). The implication of this finding is that a measure that better captures the underlying persistent component of earnings might prove very useful in predicting future earnings when used along with current earnings. Our study, based on a large sample of 1939 Spanish SMEs, investigates whether an alternative measure of performance ignored in previous research on earnings forecasting -i.e, relative efficiency- has predictive ability over and above current earnings and book value for predicting future earnings. Relative efficiency captures the inherent ability of a firm to make the most productive use of available resources as compared to the rest of firms in the sample. In order to measure the relative efficiency of our sample firms we use the Data Envelopment Analysis (DEA) non-parametric technique.

Our findings highlight that our efficiency measure has an incremental predictive ability over and above current earnings and book value of equity for predicting future earnings in the three sectors considered (manufacturing, services and construction). Moreover, we have further validated the models in a holdout sample and our results evidence again in the three sectors the highest forecast accuracy of the model that includes our efficiency measure as an additional predictor to current earnings and book value of equity. Therefore, our research suggest the usefulness of relative efficiency to capture the persistent component of earnings and, as a result, support its inclusion as an additional predictor in earnings forecasting models.

## REFERENCES

- Abad, C., Banker, R.D. and Mashruwala, R. (2006) Eficiencia relativa como indicador del rendimiento futuro, *Revista Iberoamericana de Contabilidad de Gestión (Iberoamerican Journal of Management Accounting)*, **8**, 177-202.
- Albretch, W.S., L. Lookabill and J. McKeown (1977) The time series properties of annual earnings, *Journal of Accounting Research*, **15**, 226-44.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

- Ball, R. and R. Watts (1972) Some time series properties of accounting income, *Journal of Finance*, **27**, 663-681.
- Banker, R. D., Charnes A. and Cooper W.W. (1984) Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis, *Management Science*, **30**, 1078-92.
- Banker, R. D. (1993) Maximum likelihood, consistency and data envelopment analysis: A statistical foundation, *Management Science*, **39**(10), 1265-73.
- Barnum D. T. and Gleason J. M. (2005) Technical efficiency bias caused by intra-input aggregation in data envelopment analysis, *Applied Economics Letters*, **12**, 785-88.
- Barnum D. T. and Gleason, J. M. (2006) Measuring efficiency in allocating inputs among outputs with DEA, *Applied Economics Letters*, **13**, 333-36.
- Beaver, W. and D. Morse (1978) What determines price-earnings ratios?, *Financial Analysts Journal*, July-August, **34**, 65-76.
- Brooks, L. and D.A. Buckmaster (1976) Further Evidence of the Time Series Properties of Accounting Income, *The Journal of Finance*, **31** (5), 1359-73.
- Charnes, A., Cooper, W.W. and Rhodes, E. (1978) Measuring the Efficiency of Decision Making Units, *European Journal of Operational Research*, **2**, 429-44.
- Färe, R. and Zelenyuk, V. (2002) Input aggregation and technical efficiency, *Applied Economics Letters*, **9**, 635-36.
- Färe, R., Grosskopf, S. and Zelenyuk, V. (2004) Aggregation bias and its bounds in measuring technical efficiency, *Applied Economics Letters*, **11**, 657-60.
- Lee, C.M. (1999): Accounting-Based Valuation: Impact on Business Practices and Research, *Accounting Horizons*, **13** (4), 413-25.
- Ohlson, J. (1995) Earnings, book values and dividends in security valuation, *Contemporary Accounting Research*, **11**(2), 661-87.
- Reverte Maya, C. (2003) Evidencia empírica sobre la utilidad de la información financiera para la predicción de los resultados futuros, *Revista Española de Financiación y Contabilidad (Spanish Journal of Accounting and Finance)*, Vol. **31** (111), 189-224.
- Sengupta, J.K. (2002) Economics of efficiency measurement by the DEA approach, *Applied Economics*, **34**, 1133-39.
- Shroff, P. K. (1999) The variability of earnings and non-earnings information and earnings prediction”, *Journal of Business Finance & Accounting*, **26**(7)&(8), 863-82.
- Watts, R. L. and R. Leftwich (1977) The time series of annual accounting earnings”, *Journal of Accounting Research*, **15**, 253-71.
- White, H. (1980): “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”, *Econometrica*, **48**, 817-38.