Factors Influencing the Economic Behavior of the Food, Beverages and Tobacco Industry: A Case Study for Portuguese Enterprises

Kelly Murillo[1] and Eugénio Rocha [2]

Abstract

In today's world, it is increasingly important to conduct economic and financial analyzes of enterprises in all sectors to determine strengths, identify weaknesses and adopt strategies that allow them to be at the highest competitive level. In particular, the food sector plays an essential role in the economy of any country, representing a significant contribution of gross domestic product, total employment and personal disposable income of consumers.

In this work, we adopted a non-parametric approach that combines multidirectional efficiency analysis with other mathematical techniques such as cluster analysis, principal components and CN value, to examine the factors that influence the behavior of Portuguese enterprises engaged in the manufacture of food, beverages and tobacco products.

The results show, not only a characterization of the financial structure of the sector, and a diagnosis through indexes that identify the strategic positioning of the enterprises in terms of efficiency scores; but also an analysis of the variables that must be approached differently, to obtain better results, in terms of economic performance. In fact, although there is an increase in credit with the acquisition of long-term debts, there is no evidence that this implies the ability of enterprises to grow faster, which affects profitability.

Keywords: Multidirectional efficiency analysis, clustering analysis, NC-value, Food industry.

JEL Codes: C14, C38, D22

[1]**Corresponding Author** Department of Mathematics, University of Aveiro, Portugal. kellymurillo@ua.pt
[2]Department of Mathematics, University of Aveiro, Portugal. eugenio@ua.pt

1. Introduction

The foods products, beverages and tobacco sector (FBT sector) are of great importance in the economy growth of any country, since enterprises within the processing activities of this sector not only generate products for final consumption (many of which are essential daily products) but also intermediate products for other manufacturing activities (such as oils, fats and sugars).

The FBT sector is the largest manufacturing industry in the European Union and therefore a major contributor to Europe's economy. The socioeconomic status of the sector in any country is linked to the performance of enterprises involved in. In this sense, small and medium sized enterprises play a key role in this sector. More than 280000 of these enterprises which are representing 99% of the food and drink companies are responsible for almost 50% of the EU food and drink turnover.

Particularly, Portugal was one of the countries of the UE whose economy was affected by the crises in Europe in 2008 and subsequently in the country in 2009. Based on information compiled by the Central Balance-Sheet Database in Banco de Portugal (2011**)**, notwithstanding a decline in 2009, Manufacture of Food Products financing through trade credits grew by 9% in the 2006-2009 period. In 2009 the Manufacture of Food Products sector represented approximately 14% of the number of enterprises, 13% of the number of employees and 16% of turnover in Manufacturing. Compared to the overall results of international trade, Manufacture of Food Products was responsible for more than 6% of national exports of goods.

The aim of this paper is to evaluate the efficiency of Portuguese enterprises dedicated to Manufacture of FBT sector, between 2006 and 2013 inclusive, by developing a model based in Multidirectional Efficiency Analysis (MEA) and other mathematical techniques. This work presents an input-oriented model with important approaches. For the purposes of study, the model consider variable returns to scale (VRS), and combine the use of Principal component analysis (PCA) and Cluster Analysis (CA), which is used for identifying benchmarks for performing of Decision Making Units (DMUs). It is analyzed the relevance of enterprise size in this type of work. It is discussed how to choose the meaningful variables in the evaluation of efficiency and is boarded a fitting with a year relative importance order. The relative ranking obtained, inputs that are well/badly used are identified through an inefficiency index. On the other hand the index obtained by NC value, allows the comparisons between efficient and inefficient subgroups, identifying which variables contributed for distinguishing efficient from inefficient (i.e. identifying best practices).

Specifically in this approach we pretend to answer questions such as: in what years the enterprises of this sector presented a higher efficiency?, in what extent the European crisis affected the performance of these enterprises? and which are the variables that most influence the efficiency of enterprises?.

For measuring efficiency of Decision Making Units (DMUs) the Data envelopment analysis-DEA approach has been widely investigated and popularly applied to many fields and a variety of industries (software engineering, banking and insurance, sports to ecommerce etc); and in its use in a number of studies involving efficient frontier estimation in different sectors (governmental and non-profit, regulated and private). However, the analysis of efficiency, MEA model proposed in Bogetoft-Hougaard (1999) and further developed in Bogetoft-Hougaard (2004) and Asmild-Pastor (2010); it presents some advantages in contrast to DEA are known. It is not restricted to the assumption of weak or strong disposability; both the possible desirable output expansion and the possible input contraction are included in the efficiency evaluation and the efficiency improvement potential on each input or output variable can be separately measured.

In the analysis of efficiency, MEA have been used as alternative in the last years. In Gongbing et. al, (1995), is presented a nonradial DEA model with multidirectional efficiency involving undesirable outputs for the measurement of regional energy and environmental efficiency of Chinas transportation sector during the period 2006-2010. In Asmild et. al, (2009) is showed by MEA method that reform initiatives improve operating efficiency but potentially differently for cost drivers. In Wang et. al, (2013) is utilized MEA approach for evaluation of the environmental efficiency of industrial sectors of Chinese major cities. Murillo- Rocha (2018), study the effects of the Troika austerity measures on the Portuguese manufacturing firms in terms of efficiency scores, using a model based on MEA. However there are few studies in the literature about the efficiency of the Portuguese FBT sector. In the report of the Banco de Portugal (2011), the solvency of enterprises of the FBT sector, is evaluated by using two indicators. The first one relates Earnings before interest, taxes, depreciation, and amortization (EBITDA) with financial costs. The second one compares short-term financial debt to EBITDA. Data on activity in 2009 pointed to a significant decline in Manufacture of Food Products turnover (7%). However, EBITDA – grew by 5% and return on equity increased by 2.3 p.p., exceeding 5%. According to this report the EBITDA growth was due mostly to the fall in operating costs of enterprises in this sector, whereas developments in return on equity were mainly due to a strong contraction in financial costs (22%). In Machado (2011) it is presenting a general characterization of the Portuguese food and beverage sector, the midst of a political and economic crisis. In the report of EU-MERCI (2016), is presented an characterization of the Food and Beverage sector in the countries Portugal, Turkey, France, Czech Republic and Spain; during the period 2011-2016. In the Report of the Interreg of Central Europe, an overview of the state-of-the-art and economic trends of the European Food and Drink industry is given. It is based on the report of Food Drink Europe “Data and Trends 2014-2015”. Analysis show that the food sector together with other value chain related sectors represent one of the most important, potential fields to leverage improvement of socio-economic situation in remote areas.

Most of the studies on efficiency took advantage of the traditional DEA approach, in which the DMUs under measurement were restricted to the radial constraints on input and output variables. However, in this study, we use MEA model, since this method considers the improvement potential in each variable, it is more appropriated than the traditional radial DEA approach, to examine the impact of efficiency on enterprises.

The remainder of this paper is laid out as follows. In the next section, a general overview of the sector is given and the data sample is presented. Section 3, it is introducing the MEA approach and the more important aspects of the mathematical techniques are discussed. Section 4 the main results are discussed. Section 5 the concluding remarks are given. Finally, we present in the last Section, two appendixes. In the Appendix A, further details of the MEA model are given, and the Appendix B some supplementary figures are included.

2. Development of the Portuguese FBT sector and characterization of the data

The EU food and drink industry is the largest manufacturing industry in the EU and therefore a major contributor to Europe´s economy. Referring to economic trends of the European Food, with 15% share of turnover, 12,8% share of value added and 15% share of employment, the European food and drink sector is generating the highest turnover, value added and employment in the European Union manufacturing industry ahead of other manufacturing sectors, such as the automotive industry, see the Report of the Interreg Central Europe.

The FBT industry constitutes a sector of great economic impact in Portugal for two aspects: it is the manufacturing industry that contributes most to the Portuguese economy, both in terms of turnover and gross value added (GVA); and on the other hand, the sector has contributed to the trade balance, with a growth rate of exports higher than that of imports.

In contrast to the national manufacturing industry, the FBT sector has maintained a higher performance than the average national economy since 2010. De facto, despite the crisis the FBT sector, grew between 2013 and 2016 in terms of number of companies and value additional gross

The FBT sector was responsible for more than 108,000 jobs in 2016, becoming one of the sectors that generate the most employment in Portugal, EU-MERCI (2016). However, not only at the national level, the performance of the FBT sector has been outstanding. As for external competitiveness, exports of the food and beverage industry in 2016 increased by 2.0% compared to 2015, reaching 4,550 million (8.95% of Portugal's total exports of goods).
Among the main export destinations are Spain, Angola, France and the United Kingdom, representing a total of 58.15%. The main suppliers are Spain, France, the Netherlands and Germany, accounting for 64% overall. In 2016, 1.4% of the EU-28 food and beverage industry exports corresponded to Portugal, EU-MERCI (2016).

Among the sectors of the food and beverage industry, the largest volume of business is the beverage industry (20%), followed by the preparation and preservation of meat and meat products (15%), see Figure 1. Although the sector performed better than the total manufacturing industry, in the period from 2012 to 2016, there were two years of negative growth, (2012 and 2013). For this reason we are interested in studying the performance of companies during the years 2006 to 2013, and discovering the factors that led to this negative growth.

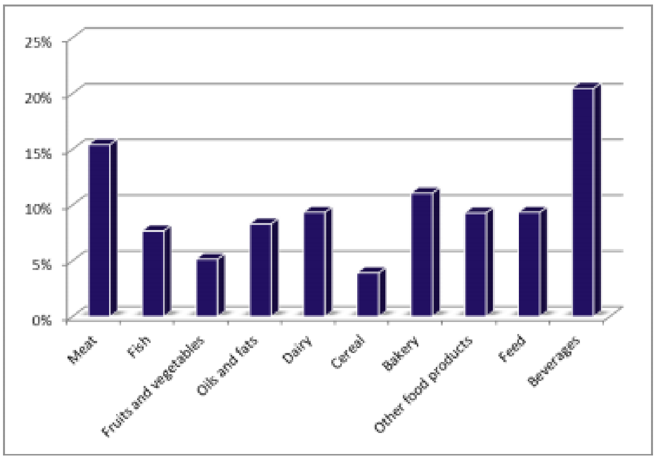

Figure 1: Turnover by sector in Portugal

Source: Report of EU-MERCI.

The data set in this study comprises financial information and other characteristics of non-financial Portuguese enterprises of section C (Manufacturing), of NACE Rev. 2 - Statistical classification of economic activities in the European Community. Specifically we consider 2092 enterprises of subgroup: Manufacture of food products, beverages and tobacco products, between 2006 and 2013 inclusive. The data set used was collected from Amadeus (Bureau van Dijk) database, which contains comparable financial information for public and private enterprises from all over Europe. The set contents only the common enterprises in all years between 2006 and 2012 for each NACE considered.

We extracted from Amadeus database the following variables: (1) number of employees (**NE**), defined by the total number of employees included in the enterprise's payroll; (2) cash and cash equivalent (**CASH**), detail of the other current assets, which is only the amount of cash at bank and in hand of the enterprise; (3) issued share capital (**CAPITAL**) issued share capital (authorized capital); (4) total assets (**TASSETS**) the sum of fixed assets and current assets; (5) long term debt (**LTDEBT**) defined as the total enterprise's debt due for repayment beyond one year; (6) profit margin (**PROFITM**) calculated as net income divided by revenues, or net profits divided by sales; (7) liquidity ratio (**LIQR**) defined by the sum of cash and marketable securities, divided by current assets; (8) solvency ratio (**SOLVR**}) calculated by adding the enterprise's post-tax net profit and depreciation, and dividing the sum by the quantity of long-term and short-term liabilities; (9) sales (**SALES**) the net sales; (10) EBIT margin (**EBITM**) the difference between all operating revenues and all operating expenses (gross profit-other operating); (11) EBITDA margin (**EBITDAM**) the sum of operating profit and depreciation; (12) current liabilities (**CLIAB**) the sum of loans, creditors and other current liabilities; and (13) cash flow (**CASHFLOW**) the sum of profit for period and depreciation.

During the 2006-2013 period, the FBT sector in Portugal maintained lower efficiency in relation to some manufacturing sectors: CHEM (manufacturing of coke, refined petroleum products and chemical products); MED (Manufacture of pharmaceutical products, chemical and botanical chemicals) and TRANSP (Manufacture of transport equipment). However, in the same period, the FBT sector has a level efficiency more similar in relation to other sectors: TEXT (Manufacture of textiles, clothing, leather and related products); MATER (Manufacture of wood, rubber, plastics and paper products); and EQUIP (Manufacture of computer, electronic and optical products), see Murillo et. al (2018). Note that the previous period involves the period in which Portugal was directly affected by the economic crisis in the world at the end of 2009. In fact, the period involved three significant periods in Portugal: pre-crisis (2006-2008), pre-troika (2009-2011) and troika (2012-2013). It is hoped that in some way the consequences of the crisis over the FBT sector may have been reflected in the efficiency of their enterprises.

2. Methodology

In the next sections, we establish some important details on the model and the main techniques used in this work.

3.1 Multidirectional Efficiency Analysis (MEA)

For measuring the efficiency of the Portuguese FBT sector, in this study is used the non-parametric deterministic MEA model introduced, in Bogetoft-Hougaard (1999). In what follows, we give a general description of the model and fix notation. Further details are presented in the Appendix A.

Denote by E the set of enterprises and T the set of years. Let $n = (e, t) \epsilon N$ be a tuple identifying the enterprise $e \epsilon E$ and the year $t \epsilon T$. We consider that any given tuple $n \epsilon N$ produces J outputs $y_j(n)$, $j \epsilon [J]$, using I inputs $x_i(n)$, $i \epsilon [I]$. The first $1 < D \leq I$ inputs are the so-called discretionary inputs (variables that into the optimization process), which will be represented by the indices d such that 1< d < I, with $i \epsilon [D]$ referring to the discretionary inputs, while $i \epsilon [I]$\-d} referring to the non-discretionary inputs.

The dataset Z =-$z(n)\}_{n\in N}$ is the set of values $z(n) = (x(n), y(n))$ for all $n \in N$, where $x(n) \in R^I$ is the vector of all the inputs $x_i(n)$, and $y(n) \in R^J$ is the vector of all the outputs $y_j(n)$, for a given tuple $n \epsilon N$.

The technical efficiency of each enterprise will be measured by calculating the MEA score.

Definition 1 (MEA score): For a given data set Z=-$z(n)\}_N$ with $z(n) = (x(n), y(n))$, the MEA score of each $n \in N$ is then defined as

$$\mathrm{MEA}_Z(n) = \frac{\dfrac{1}{\gamma^*(n)} - \dfrac{1}{D}\sum_{i=1}^{D} \dfrac{x_i(n) - \alpha_i{}^*(n)}{x_i(n)}}{\dfrac{1}{\gamma^*(n)} + \dfrac{1}{J}\sum_{j=1}^{J} \dfrac{\beta_j{}^*(n) - y_j(n)}{y_j(n)}} \tag{5}$$

where $\lambda \in \Lambda^N$ and $\alpha_i{}^*(n)$, $\beta_j{}^*(n)$ and $\gamma^*(n)$ represent the corresponding optimal solutions to the linear optimization problems $P_i^{\alpha}(z, n)$, $P_j^{\beta}(z, n)$ and $P^{\gamma}(z, n, \alpha*, \beta*)$, defined in the Appendix A.

The value $\mathrm{MEA}_Z(\mathrm{n})$ varies between 0 and 1, with fully efficient enterprises having efficiency scores equal to 1, and null efficient enterprises having scores equal to 0.

Before continuing, we emphasize an important aspect of the proposed model, for its correct application and interpretation. The performance model used in this study uses an input-oriented model to test whether the enterprises under evaluation can reduce its inputs while maintaining the outputs at their current levels.

Using MEA model the inefficiency of the three inputs variables used in this study can be analysed individually. In fact, one of the great advantages of MEA is that it allows estimating the level of influence of each variable individually on the model. Since, an input-oriented model is used; we introduce the following definition to compute the number of times each input was used inefficiently.

Definition 2 (Inefficiency index): The inefficiency index for each given input is given by

$$\mathrm{R_i(n)} = \frac{\sum_{\mathrm{n=1}}^{\mathrm{N}} \gamma(\mathrm{n})\big(\mathrm{x_i(n)} - \alpha_{\mathrm{i}}{}^*(\mathrm{n})\big)}{\sum_{\mathrm{n=1}}^{\mathrm{N}} \mathrm{x_i(n)}}. \qquad (6)$$

for $i \in [I]$ and tuple $n \in N$.

The definition 2 is based on the ideas in Bogetoft -Otto (2011).

3.2 Other mathematics techniques

To continue, four aspects about the model will be discussed.

3.2.1 What clusters should be used?

In efficiency studies, the Cluster Analysis (CA) can show the degree of sensitivity of the efficiency score for a particular enterprise to the presence of the other enterprises in the sample that make up the reference technology to that enterprise. CA can also compare the efficiency score of each enterprise to the other scores, Hirschberg-Lye (2001). In the enterprises efficiency, it is common to include the size of the enterprises as a cluster. In this sense many studies include the enterprises' size as a quantitative variable (e.g., sales) or as a nominal variable (e.g., micro, small, medium and large enterprises). This is because it is often assumed that the effects of size, as measured by sales, on the capital structure of enterprises may vary depending on whether the enterprise is in fact micro, small, medium or large Ramalho-Silva, (2009). The division corresponding to the size enterprises is based on definitions adopted by the European Commission (recommendation 2003/361/EC).

In this study, in contrast with the conventional division by size, a spectral grouping algorithm is used to obtain the clusters. The method uses Partition Around Medoid (PAM) based on the distance GDM2 (GDM distance measure for ordinal data), (Walesiak (1993), Karun-Isaac 2013)). The main advantage using this method is that, PAM determines the optimal clustering procedure for the data set. The cluster analysis is performed by embedding the data into the subspace of the eigenvectors of an affinity matrix. The result of the processing the files contains: the number of clusters found for each enterprise and the number of changes for each enterprise (mean and deviation standard), with a maximal number of five possible variants.

3.2.2 How to choose the meaningful variables?

Techniques to obtain a better understanding of the relations and mathematical significances of the variables of the section 3.2 will be applied. We will select the better input-output variables to the study using the Principal Component Analysis (PCA), with the test of dimensionality *test-dim*. The PCA analysis proposed in Pearson (1901), transforms a number of correlated variables into a number of uncorrelated variables, Abdi-Williams (2010). One time applying the PCA, the *test-dim* is performed. It allows testing the number of axes in multivariate analysis. The procedure is based on the computation of the RV coefficient, Dray (2007).

3.2.3 *How to compare groups with differentt levels of efficiency?*

In order, to compare the behavior of input and output variables between two groups G_0 and G_1 with different levels of efficiency, we use the NC-value. It indicator is defined to measure the overlapping of Gaussian distribution functions of the groups, Inman-Bradley (1989). This procedure requires that first we define which the enterprises in each group are; then the indicator will make a comparison of the behavior of each group against each variable. The mean and the standard deviation of the variable in each group, generate a normal distribution. Then, we compute the NC-value between the normal distributions of the groups. Such is done for each sub-dataset, variable and year of interest. The smaller the NC-value, the less common the behavior of the two groups with respect to the selected variables. A higher NC-value means greater intersection between the behavior of the groups G0 and G1.

3.2.4 *How can we tell if a model fits the data?*

The macroeconomic variables defined every year in a country determine the socio-economic level of this, affecting positive or negatively the performance of their enterprises. Therefore it is important to study the behavior of enterprises over the years with respect to its variables. We will apply model fitting by least squares in our data.

The model fitting, first defines a function that takes a set of parameters and returns a predicted data set; then it obtains an 'error function' that provides a number representing the difference between the original data and the model's prediction for any given set of model parameters; and finally it finds the parameters that minimize this difference.
For comparing the behavior of the variables on the study period, will be applied the conventional method of fitting (least squares), which finds the line minimizing the sum of distances between observed points and the fitted line.

4. Analysis and Results

Before applying the MEA analysis, some specific characteristics of the model used, are presented.

The algorithm results of CA (Section 3.2.1), indicates that the best clusters to use in the study corresponding to the variables NE (cluster 1) and SALES (cluster 2). The results of applying PCA to dataset (Section 3.2.2) showed the independent variables, and, therefore for the estimation of efficiency of the enterprises, were considering the inputs: NE, TASSETS, LTDEBT, CLIAB and the outputs: PROFITM, LIQR, SOLVR, EBITM, and EBITDAM. An example of the descriptive statistics of the variables, are exhibits in Table 1.

Table 1: Descriptive statistics for inputs and outputs (2006, cluster 1)

	Sum	Mean	Standard deviation
NE	25038,0	91,7	127,3
TASSETS	4814257,0	17634,6	52586,6
LTDEBT	708951,0	2596,8	4910,9
CLIAB	2269499,0	8313,1	32475,6
SFUNDS	1823040,0	6677,8	19883,8
CASH	127801,0	468,1	2564,9
CAPITAL	740652,0	2713,0	6610,7
PROFITM	407,7	1,4	8,0
LIQR	274,8	1,0	0,7
PROFITM	407,7	1,4	8,0
LIQR	274,8	1,0	0,7
SOLVR	9444,6	34,5	16,4
EBITM	923,9	3,3	7,3
EBITDAM	2447,7	8,9	7,7

Source: Authors Calculations from Eurostat (2006-2013).

<u>Efficiency ratios</u>

We calculated the MEA score of each enterprise considering the inputs and outputs defined by the PCA. A summary of the resulting efficiencies of the enterprises are provided in the Table 2 (cluster 1) and Table 3 (cluster 2). These values give an idea of the overall performance of the different enterprises over the years.

Consider EFF as subset of unities such as 0, 6 $\leq$ MEA (e, t) $\leq$ 1,0, (see Equation 5). The tables Table 2-Table 3 present information about the percentages of two quantities: Total

Efficiency (EFFT), calculated as the coefficient between the number of efficient enterprises and the number total of enterprises; and Full efficient (FULLEFF), the percentages of enterprises with MEA efficiency score equal to 1. The quantities into the parentheses represent the proportion of enterprises by clusters in each year (considering 273 enterprises for the cluster 1; and 280 enterprises, for the cluster 2).

Table 2: Efficiency ratios by cluster 1

	EFFT	FULLEF
2006	67,0% (183)	7,3% (20)
2007	71,4% (195)	7,6% (21)
2008	68,9% (188)	8,7% (24)
2009	74,4% (203)	8,1% (22)
2010	84,6% (231)	8,4% (23)
2011	77,7% (212)	9,5% (26)
2012	74,0% (202)	6,9% (19)
2013	75,5% (206)	9,9% (27)

Table 3: Efficiency ratios by cluster 2

	EFFT	FULLEF
2006	63,2% (177)	7,8% (22)
2007	63,6% (178)	11,4% (32)
2008	63,6% (178)	11,8% (33)
2009	66,8% (187)	7,14% (20)
2010	76,8% (215)	10,7% (30)
2011	70,7% (198)	8,93% (25)
2012	64,6% (181)	6,07% (17)
2013	63,6% (178)	8,93% (25)

As we can see in the Table 2-Table 3, the 2010 was a representative year. In this year the total EFFT have a higher percentage in both clusters (by above of 60%). In particular 2013 (cluster 1) also was characterized by greater percentage in FULLEFF; contrary to cluster 2 (2007 and 2008).

Inefficiency index

Detailed results for MEA directional with respect to cluster 1 and cluster 2 are presented in Figure 2. The tables show the contribution of variables in the calculation of the model. Percentages represent the number of times each input was used inefficiently (excess inputs). The NE was the variable less used inefficiently on the two clusters. LTDEBT was the variable more used inefficiently in the cluster 1. In the case of cluster 2, was CLIAB. Interestingly, most of the variables in cluster 2 have a lower percentage to be used inefficiently than the cluster 1.

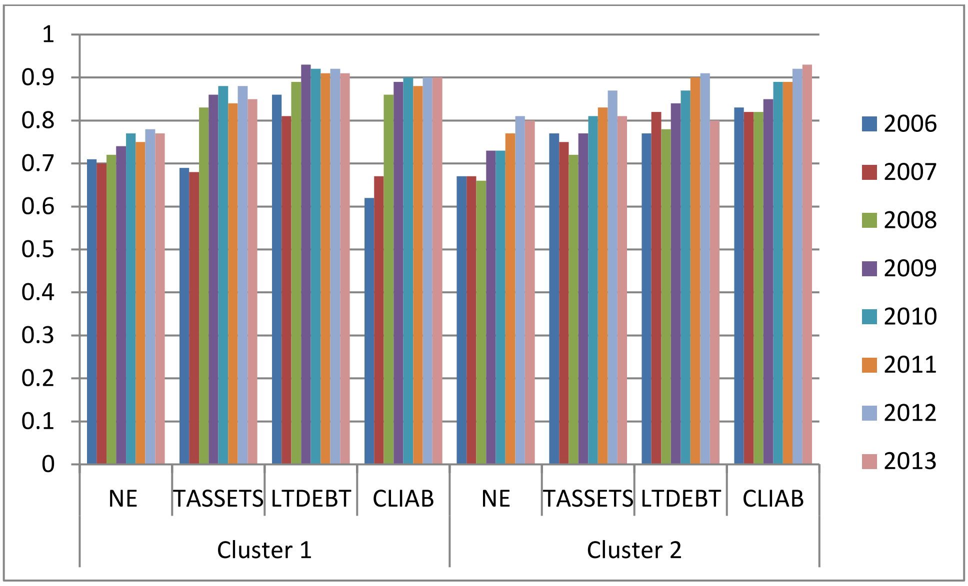

Figure 2: Inefficiency index by clusters.

<u>Groups with different levels of efficiency</u>

We calculate the NC-value for each variable (see Section 3.2.3), considering two groups with different level of efficiency. The group G_1 composed by the more efficient enterprises, corresponding to the units such that $0,6 \leq MEA(e, t) \leq 1,0$; and G_0 composed by the less efficient enterprises, corresponding to the units such that 0; $0, 0 \leq MEA(e, t) \leq 0, 4$. In Figure 3 and Figure 4, are represented the behavior of the two groups for LTDEBT and LIQR, in the two clusters. For the first variable, the NC-value is equal to 71, 2 (cluster 1) and the NC-value is 322, 5 (cluster 2). For the second variable, the NC-value is equal to 330, 9 (cluster 1) and the NC-value is 220, 5 (cluster 2).

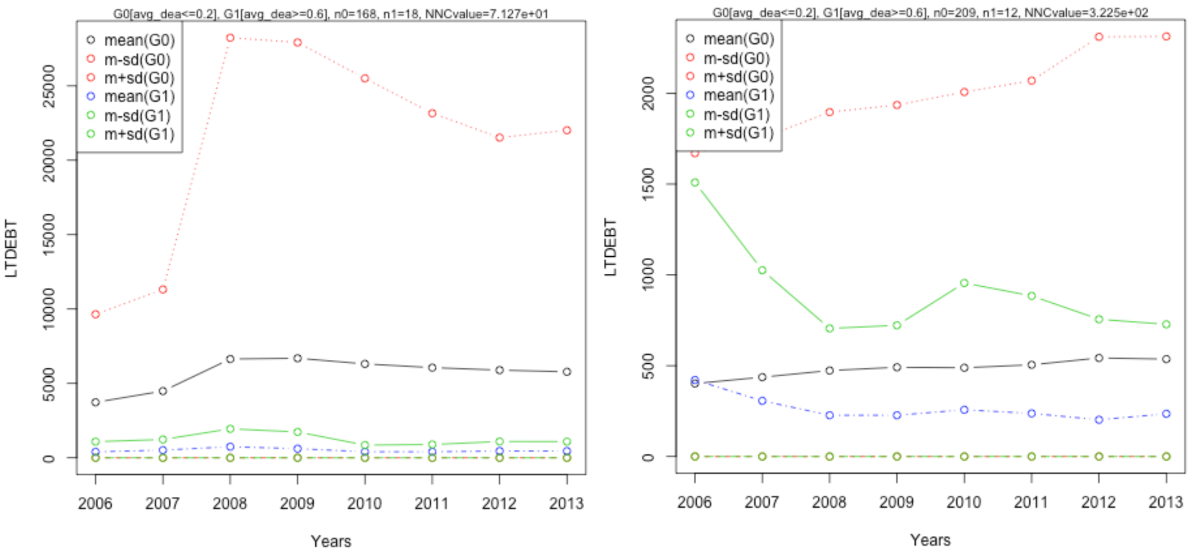

Figure 3: NC-value in the LTDEBT, for (a) cluster 1 (left side); (b) cluster 2 (right side).

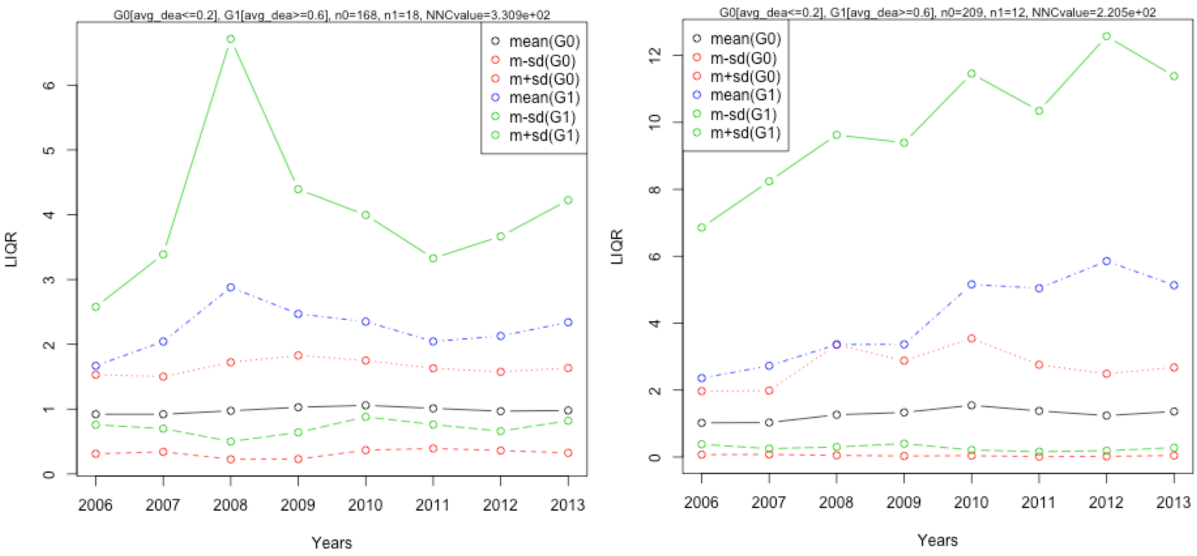

Figure 4: NC-value in the LIQR, for (a) cluster 1 (left side); (b) cluster 2 (right side).

The blue line represents the mean of the efficient group and the black line the less efficient one. We can see that, the less efficient group; obtain more quantitative of debt, during all study period (Figure 3), independent to the clusters. For this reason its liquidity is notably less than the efficient group, in special to the 2008 in the cluster 1; and from to the 2010 in the cluster 2 (Figure 4).

In Table 4, we present the NC-value for each variable, corresponding to each cluster. Note that the difference between the efficiency and inefficiency groups in the PROFITM results is very small with respect to the other cases in special in the cluster 2. The bigger difference between the two groups is presented in EBITDAM for the cluster 1 (379,5) and in CAPITAL for the cluster 2 (520,1).

Table 4: NC value for all variables by the clusters

	Cluster 1	Cluster 2
PROFITM	37,6	0,2
CLIAB	67,4	264,6
LTDEBT	71,2	322,5
SALES	97,2	340,5
TASSETS	97,8	500,1
EBITM	125,4	16,0
NE	136,1	500,1
CASHFLOW	169,4	432,1
CAPITAL	182,5	520,1
CASH	235,1	183,9
SOLVR	266,6	184,8
LIQR	330,8	220,5
EBITDAM	379,5	190,6

<u>Characterization of the FBT sector using model fitting</u>

Many problems in analyzing data involve describing how variables are related. A standard tool to measure the relation between a pair of variables is the correlation. Before applying model fitting, in order to get an idea of the relation between the variables of this study, are presented, the correlation matrices for inputs and outputs, using Pearson (Table 5) and Spearman (Table 6) correlations. The p-values measure how the variables are compatible. Here all the correlations shown have $p<0.05$, indicating strong evidence. When such is not the case, we replaced the correlation value by "NA".

Table 5: Correlation matrix Pearson

- Correlation matrix (pearson) follows with P-value<=0.05:

	NE	TASSETS	CAPITAL	LTDEBT	SFUNDS	CLIAB	CASH	PROFITM	LIQR	SOLVR	EBITM	SALES	CASHFLOW	EBITDAM
NE	1.0000	0.7411	0.7575	0.5470	0.7667	0.6052	0.3963	0.0684	NA	0.0732	0.0627	0.7397	0.6587	NA
TASSETS	0.7411	1.0000	0.7732	0.4138	0.9211	0.9481	0.4070	0.0763	NA	NA	0.0757	0.7629	0.8592	0.0605
CAPITAL	0.7575	0.7732	1.0000	0.6099	0.7618	0.6513	0.2647	NA	NA	0.0586	NA	0.6357	0.5848	NA
LTDEBT	0.5470	0.4138	0.6099	1.0000	0.3939	0.2481	0.0668	NA	NA	NA	NA	0.3622	NA	NA
SFUNDS	0.7667	0.9211	0.7618	0.3939	1.0000	0.7639	0.6551	NA	NA	NA	NA	0.7446	0.8466	0.0721
CLIAB	0.6052	0.9481	0.6513	0.2481	0.7639	1.0000	0.2041	0.0540	NA	NA	0.0612	0.6883	0.8069	0.0458
CASH	0.3963	0.4070	0.2647	0.0668	0.6551	0.2041	1.0000	0.0815	NA	NA	0.0569	0.3980	0.5239	NA
PROFITM	0.0684	0.0763	NA	NA	NA	0.0540	0.0815	1.0000	NA	0.3169	0.9539	0.0516	NA	0.7481
LIQR	NA	NA	NA	NA	NA	NA	NA	NA	1.0000	0.2774	NA	NA	NA	0.0483
SOLVR	0.0732	NA	0.0586	NA	NA	NA	NA	0.3169	0.2774	1.0000	0.2674	NA	NA	NA
EBITM	0.0627	0.0757	NA	NA	NA	0.0612	0.0569	0.9539	NA	0.2674	1.0000	0.0501	NA	0.8298
SALES	0.7397	0.7629	0.6357	0.3622	0.7446	0.6883	0.3980	0.0516	NA	NA	0.0501	1.0000	0.7162	NA
CASHFLOW	0.6587	0.8592	0.5848	NA	0.8466	0.8069	0.5239	NA	NA	NA	NA	0.7162	1.0000	0.0736
EBITDAM	NA	0.0605	NA	NA	0.0721	0.0458	NA	0.7481	0.0483	NA	0.8298	NA	0.0736	1.0000

Table 6: Correlation matrix Spearman

- Correlation matrix (spearman) follows with P-value<=0.05:

	NE	TASSETS	CAPITAL	LTDEBT	SFUNDS	CLIAB	CASH	PROFITM	LIQR	SOLVR	EBITM	SALES	CASHFLOW	EBITDAM
NE	1.0000	0.6751	0.5711	0.3262	0.633	0.631	0.4235	-0.0545	NA	NA	NA	0.8249	0.550	NA
TASSETS	0.6751	1.0000	0.7307	0.5051	0.817	0.901	0.4853	-0.0687	NA	NA	0.0444	0.8646	0.699	0.0724
CAPITAL	0.5711	0.7307	1.0000	0.3873	0.722	0.645	0.3883	NA	0.0718	NA	NA	0.6942	0.506	-0.0656
LTDEBT	0.3262	0.5051	0.3873	1.0000	0.372	0.361	NA	NA	0.0704	-0.204	NA	0.4092	0.341	NA
SFUNDS	0.6334	0.8173	0.7216	0.3719	1.000	0.637	0.5613	NA	0.2614	0.482	NA	0.7827	0.689	NA
CLIAB	0.6305	0.9011	0.6449	0.3606	0.637	1.000	0.3793	NA	-0.2844	-0.235	NA	0.8077	0.636	NA
CASH	0.4235	0.4853	0.3883	NA	0.561	0.379	1.0000	NA	0.3851	0.299	NA	0.5138	0.445	0.0838
PROFITM	-0.0545	-0.0687	NA	NA	NA	NA	NA	1.0000	0.2667	0.336	0.9142	NA	0.368	0.6792
LIQR	NA	NA	0.0718	0.0704	0.261	-0.284	0.3851	0.2667	1.0000	0.581	0.2017	0.0442	NA	NA
SOLVR	NA	NA	NA	-0.2037	0.482	-0.235	0.2993	0.3362	0.5812	1.000	0.2377	NA	NA	NA
EBITM	NA	0.0444	NA	NA	NA	NA	NA	0.9142	0.2017	0.238	1.0000	0.0619	0.434	0.7950
SALES	0.8249	0.8646	0.6942	0.4092	0.783	0.808	0.5138	NA	0.0442	NA	0.0619	1.0000	0.717	-0.0433
CASHFLOW	0.5505	0.6992	0.5059	0.3407	0.689	0.636	0.4451	0.3683	NA	NA	0.4339	0.7172	1.000	0.4786
EBITDAM	NA	0.0724	-0.0656	NA	NA	NA	0.0838	0.6792	NA	NA	0.7950	-0.0433	0.479	1.0000

For appliying model fitting, first we consider the data set without dividing by clusters, and we selected the variables more relatively stable during all study period: CAPITAL (for the inputs) and the ratio EBITM (for the outputs). Then the equation of the line between CAPITAL (resp. EBITM) and each one of the other inputs (resp. outputs) is generated. To continue, the points of intersection between each two straight correlated with CAPITAL (resp. EBITM) are calculated, generanding a relative order (see Murillo et al, 2018). Using this information is obtained a relative order graph, which represents the general behavior of the variables in groups over the years: Figures 5 for CAPITAL (NE and SALES) and Figures 6 for EBITM (PROFITM and SOLVR). The relative order graphs for all variables inputs/CAPITAL and outputs/EBITM are shown in the Appendix B, Figure 7-Figure 8.

Note that, in each graph are represented the minimum and maximum values for each variable and intersections are considered only within this range. We use three divisions for CAPITAL variable and four divisions for EBITM. The p-value is represented by $\log_{10}p$ and the relative order (columns) represents the number of times in which a enterprise changes of cluster.

For better understeand the behaviour of the enterprises and facilitate the description on the graphic, for the CAPITAL case, we will call large enterprises, to the one represented by a red and continuous line; the medium enterprises are represented by a blue line and small enterprises with a black line color. Note that, this division does not correspond to the division European Commission (recommendation 2003/361/EC) (see Section 3.2.1).

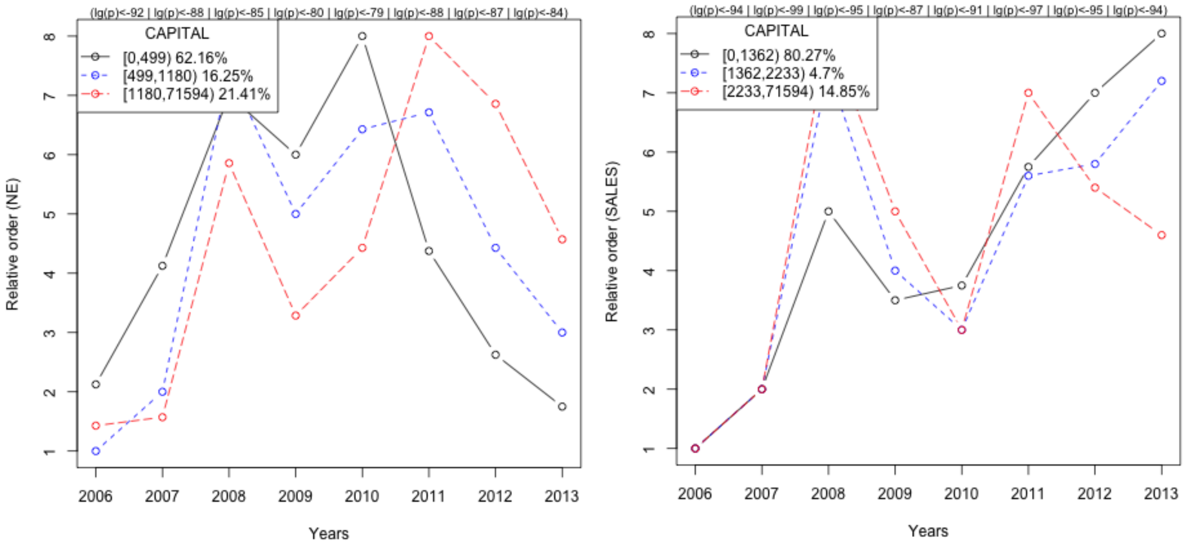

Figure 5: Relative order CAPITAL: (a) NE (left side); (b) SALES (right side).

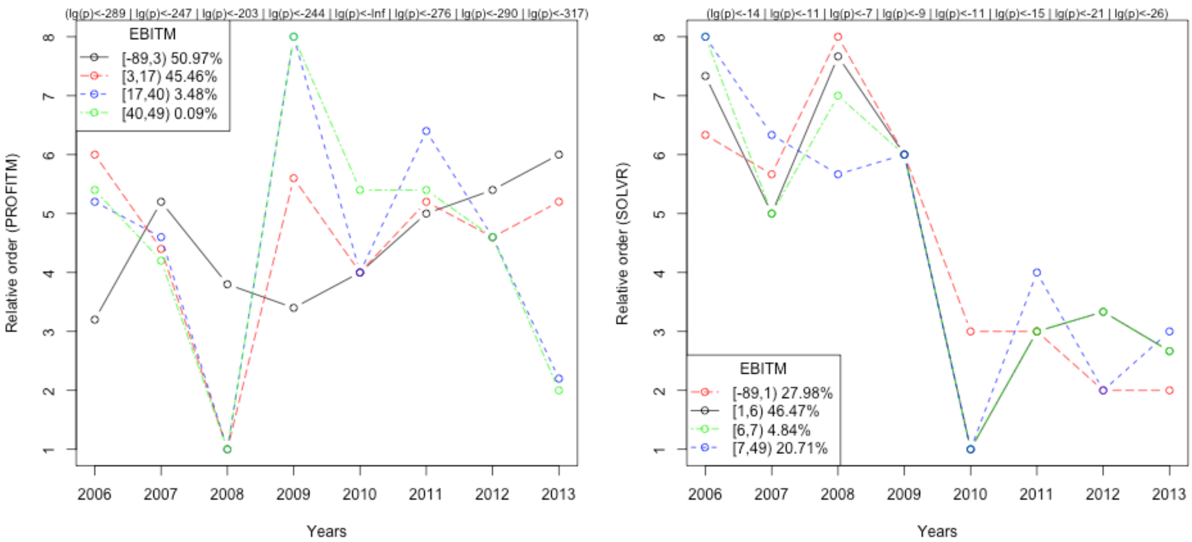

Figure 6: Relative order EBITM: (a) PROFITM (left side); (b) SOLVR (right side).

Interesting observations can be made from the relative order graphs (see Appendix B). The most representative aspects of these variables over the years are summarized to follow. Decrease in NE for all companies for the years 2008-2009 and 2011-2013. However, small enterprises reflected an increase in their number of companies for the 2006-2010 over the other companies. For all companies, SALES increased in 2007-2008 and decreased in 2009-2010. In the whole period, large companies had less CLIAB. Large companies acquired more LTDEBT in 2008 and 2011 than the other companies, while medium-sized companies had an increase between 2007 and 2009. Between the years 2008-2010 small companies had higher TASSETS than other companies. On the other hand, since EBITM indicates the ability of an enterprise to be profitable, and ultimately to generate profits we want to make a special look at what was the behavior of the variables related to EBITM during and after the crisis period? The enterprises increased the acquisition of debts during the period 2007-2010. However CASH decreased in the 2010 for all enterprises. Solvency decreased at the end of the 2010-2013 study periods, and liquidity increased between 2010-2011 and 2012-2013.

Concluding Remarks

This work uses a nonparametric model that combine the Multidirectional Efficiency Analysis, with other mathematical techniques to examine the performance of the Portuguese foods products, beverages and tobacco sector (FBT sector) in the period 2006-2013. The study, allowed establishing differences between the efficiency patterns of the sector, considering different indicators: efficiency ratios, inefficiency index and the NC value between groups with different efficiency levels. The study also allowed comparisons between variables, using model fitting.

With the analysis presented in this document, interesting results were obtained. Firstly, we showed that there are significant differences for all stages of study, in the enterprises when are organized by the clusters NE (cluster 1) and SALES (cluster 2). Particularly, overcoming the difficulties generated by the crisis, in the year 2010 was a good year for the sector, reaching a percentage of total efficiency above 60% in both clusters. All the input variables showed an inefficiency of use of the enterprises between 0.62 and 0.93, for all years and clusters. The NE was the variable less used inefficiently on the two clusters.

At another stage of the analysis, two groups with different levels of efficiency were selected, according to their MEA score. The bigger difference between the two groups is presented on the variables EBITDAM for the cluster 1 (379,5) and CAPITAL for the cluster 2 (520,1). On the other hand, the less efficient group; increased the acquisition of debts (LTDEBT and CLIAB), during all study period and its liquidity was notably less than the efficient group, in special to the 2008 in the cluster 1; and from to the 2010 in the cluster 2. Note that although there is an increase in credit with the acquisition of long-term debts, there is no evidence that this implies the ability of enterprises to grow faster, which affects profitability. With respect to efficiency levels, it should be noted that in most cases large enterprises in Portugal, were most affected by the crisis in the different variables. In fact, large enterprises suffered more removals in periods 2008-2010. Small enterprises showed greater stability capacity different variables.

Important indicators of financial and economic performance of the enterprises, as profitability, liquidity, solvency, interest coverage and efficiency; are includes in the study.

These indicators are measured by the variables EBITM, LIQR, SOLV, PROFITM and EBITDAM, allowing obtaining a real characterization of the efficiency of the sector in economic terms. On the other hand, the study period involves the period in which Portugal was directly affected by the world financial crisis at the end of 2009. Therefore, the period can be divided in three significant periods in Portugal: pre-crisis (2006-2008), pre-troika (2009-2011) and troika (2012-2013). In this sense, it is hoped that in some way the consequences of the crisis over the FBT sector reflected in the efficiency of the enterprises and in special on the indicators of financial and economic performance. Although the companies increased the acquisition of debts during the period 2007-2010, to face the difficulties of the crisis; the results of the ratios are varied. Solvency decreased at the end of the 2010-2013 study periods, and liquidity increased between 2010-2011 and 2012-2013.

Given the impact that the results may have on this type of study, using the most appropriate analysis model is vital. The performance model used in this study uses an input-oriented model to test whether the enterprises under evaluation can reduce its inputs while maintaining the outputs at their current levels. The results of this type of study are important in determining the failures that avoid improving the quality of companies and establishing strategies to strengthen sector performance. Since the FBT sector is one of the sectors that generate more employment in Portugal, an interesting future work would be to use benchmarking to study the efficiency of the sector in countries of the European Union, establishing a comparative ranking and considering a study more extensive period.

Acknowledgement

The authors were supported by FCT - Fundação para a Ciência e a Tecnologia, within project UID/MAT/04106/2019 (CIDMA). Murillo was also supported by national funds (OE), through FCT, I.P., in the scope of the framework contract foreseen in the numbers 4, 5 and 6 of the article 23, of the Decree-Law 57/2016, of August 29, changed by Law 57/2017, of July 19.).

References

Abdi, H. & Williams L. J. (2010). Principal component analysis. Wiley Interdisciplinary Reviews: Computational Statistics, 2, 433-459.

Asmild, M., Holvad, T., Hougaard, J. L., & Kronborg, D. (2009). Railway Reforms: Do They Inuence Operating Efficiency? Transportation, 36 (5), 617-638. doi:10.1007/s11116-009-9216-x.

Asmild, M., & Pastor, J. T. (2010). Slack Free MEA and RDM with Comprehensive Efficiency Measures. Omega, 38 (6), 475-483. doi:10.1016/j.omega.2009.12.004

Banco de Portugal, (2011) S.ectoral Analysis of Manufacture of Food Products, Central Balance- Sheet Studies, report.

Banker, R. D., Charnes, A., & Cooper, W. (1984). Some Models for Estimating Technical and

Scale Inefficiencies in Data Envelopment Analysis. Management science, 30 (9), 1078-1092.doi:10.1287/mnsc.30.9.1078

Bogetoft, P., & Hougaard, J. L. (1999). Efficiency Evaluations Based on Potential (Non-proportional) Improvements. Journal of Productivity Analysis, 12 (3), 233-247. doi:10.1023/A:1007848222681

Bogetoft, P., & Hougaard, J. L. (2004). Super EfficiencyEvaluations Based on Potential Slack. European Journal of Operational Research, 152 (1), 14-21. doi:10.1016/S0377-2217(02)00642-2

Bogetoft, P. & Otto, L. (2011). Benchmarking with DEA, SFA, and R. International Series in Operations Research and Management Science.Springer Science & Business Media, (157).

Chung, Y. H., Fre, R., & Grosskopf S., (1997). Productivity and Undesirable Outputs: A Directional Distance Function Approach. Journal of Environmental Management, 51, 229-240,.

Dray, S. (2007). On the Number of Principal Components: A Test of Dimensionality Based on Measurements of Similarity Between Matrices. Computational Statistics & Data Analysis, 52 (4), 2228-2237. doi:10.1016/j.csda.2007.07.015

EU-MERCI (2016). Analysis of Food and Beverage sector in different Countries. Horizon 2020 Project Nr. 693845, available at: http://www.eumerci portal.eu/documents/20182/38527/10+-+Other+countries.pdf.

Gongbing, B., Pingchun W., Feng Y. & Liang L. (1995). Energy and Environmental Efficiency of Chinas Transportation Sector: A Multidirectional Analysis Approach, Mathematical Problems in Engineering.

Inman, H. F. & Bradley Jr E. L. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities, Communications in Statistics - Theory and Methods 18:10 (1989), 3851{3874.

Interreg Central Europe Report. Food sector related knowledge integration. Food sector global market trend analysis, I-CON project.

Hirschberg, J. G. & Lye J. N. (2001), Clustering in a Data Envelopment Analysis using Bootstrapped E_ciency Scores, Economics, University of Melbourne, Melbourne, 3010, Australia.

Karun, K. & Isaac, E. (2013), Cogitative analysis on k-means clustering algorithmand its variants, Int J Adv Res Comp Communi Eng., 4 (2), 1875-1880.

Machado D. M (2011). Portugal Food Processing sector, Gain report PT1/02, Global Agricultural information Network, 1-24.

Murillo, K. P & Rocha, E. M. (2018). The Portuguese manufacturing sector during 2013-2016 after the troika austerity measures.World Journal of Applied Economics, 4(1), 21-38. doi:10.22440/wjae.4.1.2

Murillo, K. P., Rocha, E. M. & Ramalho, J.S. (2018). About the efficiency behavior of the Portuguese manufacturing firms during the financial crisis. Libertas Mathematica (new series) 38(1), 1-27.

Noor, A. R., Susila, M. & Behrouz, A. (2013). Scale directional distance function and its application to the measurement of eco-efficiency in the manufacturing sector. Annals of Operations Research, 211(1), 381-398.

Pearson, K. (1901). LIII. On Lines and Planes of Closest Fit to Systems of Points in Space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2 (11), 559-572. doi:10.1080/14786440109462720.

Ramalho J. S & Da Silva J. V (2009). A two-part fractional regression model for the financial leverage decisions of micro, small, medium and large firms, Quantitative Finance 9(5), 621-636.

Walesiak, M. (1993). Multivariate statistical analysis in marketing research. Wroclaw University of Economics, Research Papers (654).

Wang, K., Wei, Y.-M., & Zhang, X. (2013). Energy and Emissions Efficiency Patterns of Chinese Regions: A Multi-directional Efficiency Analysis. Applied Energy, (104) 105-116. doi:10.1016/j.apenergy.2012.11.039.

APPENDIX A

In this section, some complementary details of the MEA model (Section 3.1.) are presented.

Suppose that any given tuple $n \epsilon N$ produces J outputs $y_j(n), j \epsilon [J]$, using I inputs $x_i(n), i \epsilon [I]$ and the first $1 < D \leq I$ inputs are discretionary.

The dataset Z =-$z(n)\}_{n \in N}$ is the set of values $z(n) = (x(n), y(n))$ for all $n \in N$, where $x(n) \in R^I$ is the vector of all the inputs $x_i(n)$, and $y(n) \in R^J$ is the vector of all the outputs $y_j(n)$, for a given tuple $n \epsilon N$.

Considering the variable returns to scale (VRS), a model for the efficiency measurement of decision-making units (Banker, et. al. 1984), we define the set

$$\Lambda^N = \left\{\lambda \epsilon R^N : \sum_{n=1}^{N} \lambda_n = 1\right\}. \qquad (1)$$

The MEA score for a specific observation $z(\bar{n})=(x(\bar{n}).y(\bar{n}))$ is found by solving the following linear optimization programs:

***$P_m^\alpha(z, \bar{n})$*:**
min $\alpha_m(\bar{n})$ such that
$\sum_n \lambda_n x_m(n) \leq \alpha_m(\bar{n})$
$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}), i \epsilon [I]\backslash\text{-}m\}$
$\sum_n \lambda_n y_l(n) \geq y_l(\bar{n}), l \epsilon [J]$

***$P_j^\beta(z, \bar{n})$*:**
max $\beta_j(\bar{n})$ such that
$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}), i \epsilon [I]$ (2)
$\sum_n \lambda_n y_j(n) \geq \beta_j(\bar{n}), j \epsilon [J]$
$\sum_n \lambda_n y_l(n) \geq y_l(\bar{n}), l \epsilon [J]\backslash\text{-}j\}$

***$P^\gamma(\alpha,\beta,z,\bar{n})$*:**
max $\gamma(\bar{n})$ such that
$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n})\text{-}\gamma(\bar{n})(x_i(\bar{n})\text{-}\alpha_i{*}(\bar{n})), i \epsilon [M]$
$\sum_n \lambda_n x_i(n) \leq x_i(\bar{n}), i \epsilon [I]\backslash\text{-}m\}$
$\sum_n \lambda_n y_l(n) \geq y_l(\bar{n})+\gamma(\bar{n})(\beta_l{*}(\bar{n})\text{- } y_l(\bar{n})), l \epsilon [J],$

where $\lambda \epsilon \Lambda^N$. $\alpha_i{}^*(\bar{n})$ and $\beta_j{}^*(\bar{n})$ represent the corresponding optimal solutions to the linear optimization problems $P_i^\alpha(z.\ n)$ and $P_j^\beta(z.\ n)$.

By the directional contribution of each input and output variable, the definition of the MEA score (Equation 1) is obtained. In fact, for the input $x_i(n), i \epsilon [I]$ the contribution in Z=-$z(n)\}_N$ is given by

$$\text{meff}_i(n) = \frac{x_i(n) - \gamma^*(n)(x_i(n) - \alpha_i{}^*(n))}{x_i(n)} \chi_{[D]}(i), \qquad (3\ )$$

where $\chi_{[D]}$ is the characteristics function of the set $[D]$. That means $\chi_{[D]}(i) = 1$ if $i \epsilon [D]$; and $\chi_{[D]}(i) = 0$ if $i \notin [D]$.

For the outputs $j \epsilon [J]$ the contribution is given by

$$\text{meff}_j(n) = \frac{y_j(n)}{y_j(n) + \gamma^*(n)(\beta_j{}^*(n) - y_j(n))}. \qquad (4)$$

APPENDIX B

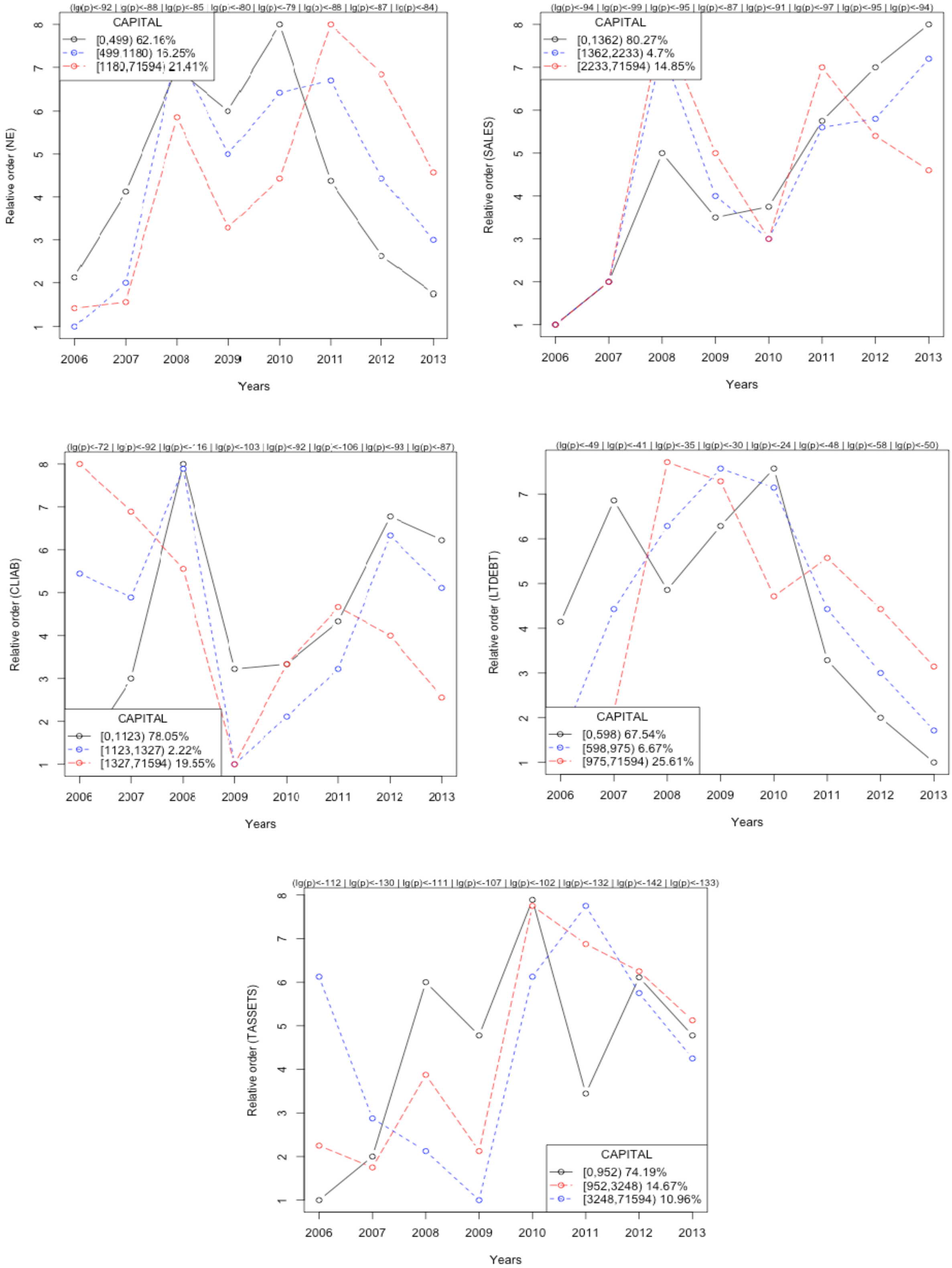

Figure 7: Relative order inputs/CAPITAL.

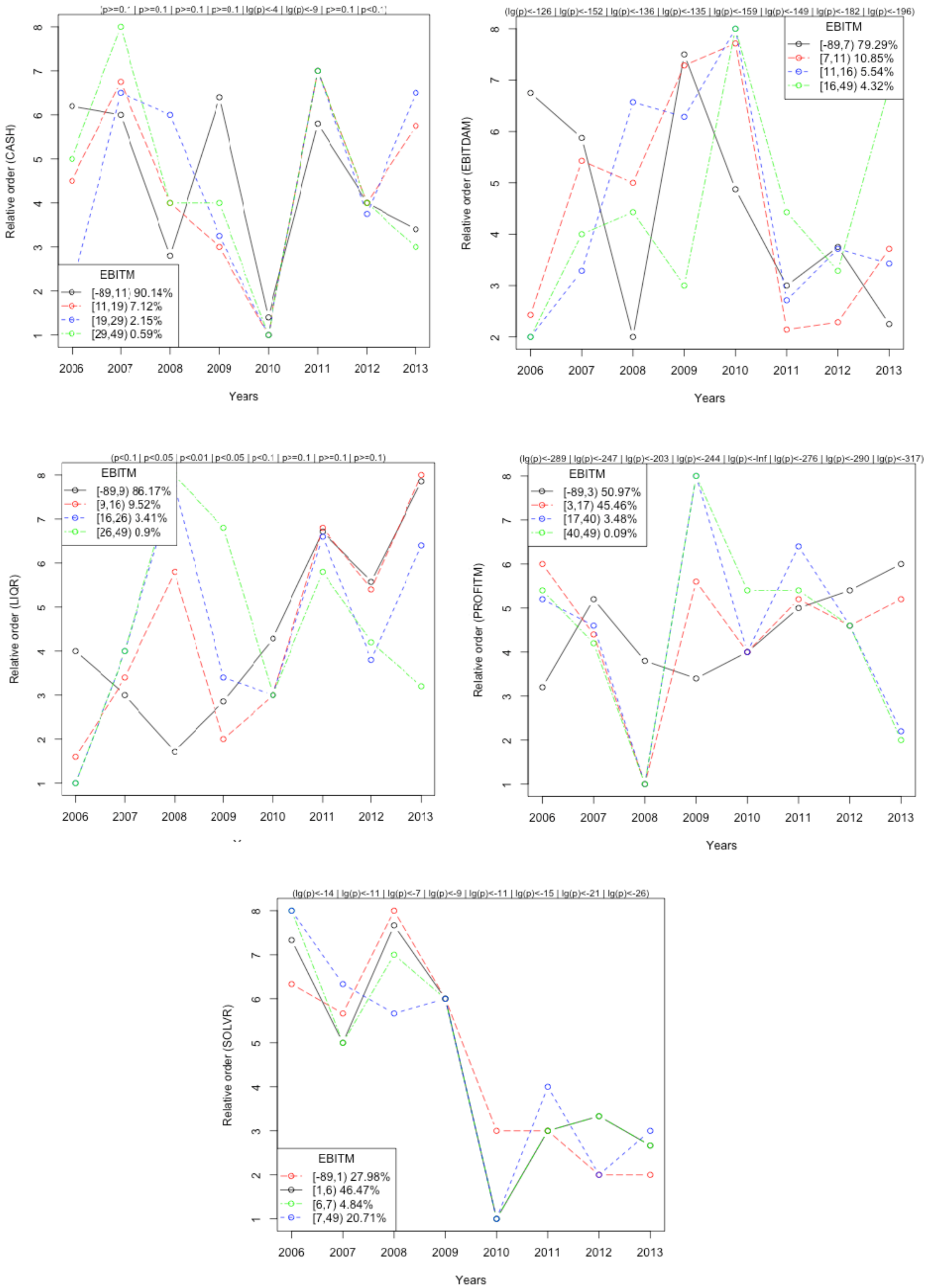

Figure 8: Relative order outputs/EBITM.