

EVALUATION OF BUSINESS EFFICIENCY IN GENERATING INTANGIBLE ASSETS

Pedro Henrique de Melo ALBUQUERQUE¹
Rafael Dantas GUIMARÃES²

- **ABSTRACT:** *In this paper, we study the intangibility of organizations using data envelopment analysis as a tool for evaluating organizational efficiency. The companies from the Brazilian capital market are studied, focusing their efficiency in generating intangible assets. However, since there are only a few observations for the traditional analysis, we use the MCMC approach, associated with the data envelopment analysis with missing data, aiming at increasing the number of informative observations, thus allowing for a more complete analysis. Finally, we present how the types of organizations (according to their efficiency scores) are associated with the company's volatility, thus suggesting a relation between the efficiency in generating intangible assets and the investment risk of the organizations.*
- **KEYWORD:** *Data envelopment analysis; MCMC; intangible assets.*

1 Introduction

The intangibility of a company's net worth reflects the potential that this enterprise has to produce something more than simply physical goods: it reflects the degree of differentiation related to the goods and services it produces.

In the present paper, we are interested in evaluating the level of efficiency of publicly traded Brazilian companies in generating intangible assets, based on a set of inputs, namely, personnel expenses, publicity expenses and R&D expenses.

These inputs, according to IASC Foundation (2008) in Summaries of International Financial Reporting Standards, are the usual ones for the generation of intangible assets. Hereinafter, we define intangible assets as the difference between the market value and the book value of the companies listed on BMF & BOVESPA.

In Brazil, publicly traded companies are not required to provide statements with their expenses with personnel, publicity and R&D, thus making the efficiency evaluation of these companies considerably harder.

In order to overcome this issue, Section 3 of this paper presents the Markov Chain Monte Carlo method for imputation of missing values. However, differently from the traditional use of MCMC method, we have estimated a credibility interval for any missing values of inputs and outputs.

¹ Universidade de Brasília – Departamento de Administração e Instituto de Pesquisa Econômica e Aplicada – IPEA, Diretoria de Estudos e Políticas Regionais – DIRUR, Caixa Postal: 03784, CEP: 70076-900, Brasília, DF, Brasil. E-mail: pedroa@umb.br

² Instituto de Pesquisa Econômica e Aplicada – IPEA, Diretoria de Estudos e Políticas Regionais – DIRUR, Caixa Postal: 03784, CEP: 70076-900, Brasília, DF, Brasil. E-mail: rgdantas@gmail.com

Through this interval, we have used the suggestion of Despotis and Smirlis (2002) for estimating the efficiency of a decision-making unit (DMU) in the presence of uncertainty relative to input and output values.

Finally, we have estimated the efficiency levels of publicly traded Brazilian companies, relative to the production of intangible assets. The present paper can thus be synthesized in the following manner: Section 2 presents the importance of measuring a company's intangible assets and also presents a review of the main studies aimed at estimating companies' efficiency in producing intangible assets.

Section 3 presents the Markov Chain Monte Carlo method and how it can be used to generate credibility intervals for missing data. Section 4 presents the intervalar DEA method, used when only intervals for inputs and outputs are known. Section 5 presents the results for the Brazilian market and Section 6 concludes.

2 Intangible assets and their importance.

In a competitive world, it is crucial that enterprises be able to differentiate themselves in the market, whether by providing quality service, innovating or creating a widespread brand. These factors are created through time by organizations and constitute an intangible value, since they are abstract and not physical.

The part of the company that exceeds its book value relative to its market value reveals the company's created value. Kayo *et. al.* (2006) argues that the intangible assets are created based on the following inputs:

Human: Knowledge, talent, capacity, experience and employee experience and ability.

Innovation: Research and development, patents, technological know-how and exploration and operation rights.

Structural: Procedures, proprietary software, databases, IT systems, administrative systems and market intelligence.

Relationship: Brands, logos, trademarks, customer loyalty.

It is important to acknowledge that these inputs are not exclusive, but only an indication of the types of inputs that might contribute towards the creation of intangible assets.

Therefore, we can think of these inputs as responsible for the creation of a company's intangible assets. It is natural to consider how these intangible assets are generated from the inputs. It is expected that greater input costs lead to greater levels of intangible assets. If this is not the case for a prolonged period, there is evidence that resources are not being used efficiently. Therefore, it is vital that companies measure the level of efficiency in generating intangible assets, either for strategic reasons or as a means to assess the quality of the organization's processes.

Acknowledging the efficiency in generating intangible assets is important both to the company's internal decision-making, such as inefficient resource reallocation, and to the reengineering of human, innovative and structural processes within the company aiming at increasing the efficiency in generating outputs, namely, the intangible assets.

Besides the internal importance, the acknowledgement and evaluation of other companies with relation to their own levels of intangible efficiency provides the market with essential strategic information concerning competition of the companies for clients.

In this sense, evaluating the efficiency level of companies in generating intangibles contributes to the improvement of organizational processes (if the company is considered inefficient) and stimulates competition among companies, thus allowing for better goods and services for the clients.

Only a few authors have evaluated business efficiency in generating intangible assets, but all with different approaches. Serrano-Cinca *et al.* (2005) evaluates the efficiency of internet companies, focusing mainly on some inputs (number of employees, expenses and assets) and on some outputs (number of website visits and income), but does not account neither for publicity expenses nor for R&D investments.

Wu *et al.* (2006) evaluate the efficiency of design companies in Taiwan, but again limit the analysis to only one group of inputs, the human component, since they focus mainly on intellectual capital. More recently, Hilmola *et al.* (2009) measure the efficiency of technology companies in Finland using intangible investment and average total assets as input, and revenue as output.

In the present paper, we wish to estimate the level of efficiency of Brazilian publicly traded companies in generating intangible assets. However, a considerable part of the data is incomplete, thus preventing us from using the traditional methodology.

In order to deal with missing values - due mainly to the fact that disclosure of some types of data is non-mandatory - we have used the Markov Chain Monte Carlo method, detailed in the next section.

3 Data envelopment analysis with missing data.

The presence of missing values in social, economic and financial data is a recurrent fact. A simple solution is to exclude observations that feature missing values, but this leads to loss of information.

Other suggestions are found in Little and Rubin (1987) and Shafer (1999), who provide statistical procedures for the imputation of missing values, aiming at reducing the loss of information and improving the estimation process.

In this paper, we have decided to use the Markov Chain Monte Carlo method to generate credibility intervals for missing values.

Since we might consider unobserved values as random variables, it is more interesting - methodologically - to work with a set of estimated values than to simply estimate a single value. Once we have obtained the credibility interval for the missing inputs and outputs, we use the Intervalar DEA method, which takes this uncertainty into consideration when estimating the efficiency of the decision-making units.

4 MCMC method for missing value imputation.

MCMC (Markov Chain Monte Carlo) methodology is a collection of techniques used to simulate pseudo-random data from probability distributions. Recently, MCMC has been subject of great interest among statisticians and thus has stimulated several applications and theoretical work. In broad lines, the goal of MCMC is to generate one or more values for a random variable Z , generally multidimensional. The method used in this paper was Gibbs Sampling with multiple chains.

Let $P(Z) = f(Z)$ be the density of Z , denominated target distribution. Instead of simulating directly from f , a sequence $\{Z^{(1)}, Z^{(2)}, \dots, Z^{(t)}\}$ is generated and each variable in the sequence depends, in some way, on the previous ones, and the stationary distribution (i.e. the limiting marginal distribution of $Z^{(t)}$ when $t \rightarrow \infty$) is the target f . For sufficiently large t , $Z^{(t)}$ is approximately a random realization of f . MCMC is useful when it is hard to draw directly from f and easy to obtain each variable of the sequence.

It is assumed that the data are arranged in a matrix with n rows and p columns, in which rows represent observation units and columns represent variables. Denoting the database by $\underline{Y} = (\underline{Y}_{obs}, \underline{Y}_{mis})$, where \underline{Y}_{obs} and \underline{Y}_{mis} are the observed and missing parts of the database, respectively. Let y_{ij} be an individual element of \underline{Y} , $i=1, 2, \dots, n$, $j=1, 2, \dots, p$. The i th row of \underline{Y} , expressed as a column vector is:

$$\underline{y}_i = (y_{i1}, y_{i2}, \dots, y_{ip})^T. \quad (1)$$

It is assumed that $\underline{y}_1, \underline{y}_2, \dots, \underline{y}_n$ are independent draws from a random vector, denoted symbolically by $(\underline{Y}_1, \underline{Y}_2, \dots, \underline{Y}_p)^T$, which has multivariate normal distribution with mean vector $\underline{\mu}$ and covariance matrix $\underline{\Sigma}$, i.e.,

$$\underline{y}_1, \underline{y}_2, \dots, \underline{y}_n \mid \Theta \sim iid \quad N(\underline{\mu}, \underline{\Sigma}), \quad (2)$$

where $\Theta = (\underline{\mu}, \underline{\Sigma})$ is the unknown parameter and $\underline{\Sigma}$ is positive, semi-definite. The density of an individual row is:

$$P(\underline{y}_i \mid \Theta) \propto |2\pi\underline{\Sigma}|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\underline{y}_i - \underline{\mu})^T \underline{\Sigma}^{-1}(\underline{y}_i - \underline{\mu})\right\}, \quad (3)$$

and discarding the proportionality constant, the global likelihood is:

$$L(\Theta \mid \underline{Y}) \propto |\underline{\Sigma}|^{-\frac{n}{2}} \exp\left\{-\frac{1}{2} \sum_{i=1}^n (\underline{y}_i - \underline{\mu})^T \underline{\Sigma}^{-1}(\underline{y}_i - \underline{\mu})\right\}. \quad (4)$$

Let $\underline{Y} = (\underline{Y}_{obs}, \underline{Y}_{mis})_{1 \times p}^T$ be a random vector of dimension $1 \times p$ with expectation equal to $\underline{\mu} = [\underline{\mu}_1^T, \underline{\mu}_2^T]_{1 \times p}^T$, with mean $\underline{\mu}_1^T$ and $\underline{\mu}_2^T$ for \underline{Y}_{obs} and \underline{Y}_{mis} , respectively. Let the variance and covariance matrix, $\underline{\Sigma}$, for the random vector \underline{Y} be given by:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}_{p \times p}, \quad (5)$$

where Σ_{11} is the variance and covariance matrix for $Y_{\sim obs}$, Σ_{22} the variance and covariance matrix relative to $Y_{\sim mis}$ and Σ_{12} the covariance matrix between $Y_{\sim obs}$ and $Y_{\sim mis}$, respectively. The goal of classic imputation is to initially search for the probability distribution:

$$P(\Theta | Y_{\sim obs}, Y_{\sim mis}), \quad (6)$$

where $\Theta = (\mu, \Sigma)$. However, in the present study we are interested in finding:

$$P(Y_{\sim mis} | Y_{\sim obs}, \Theta), \quad (7)$$

so that we can find single and interval estimates for the missing values. In order to find these probability distributions, we use the MCMC method, which can be summarized in two steps:

- 1) Step: Given an estimate for the mean vector and for the variance and covariance matrix, in this step the missing values are simulated independently for each observation, thus increasing the set of available operations. Mathematically, we can represent this step as:

Given a set of values $\Theta^{(t)}$ for the t th iteration, we generate values for $Y_{\sim mis}^{t+1}$ from the following conditional distribution:

$$P(Y_{\sim mis}^{t+1} | Y_{\sim obs}, \Theta^{(t)}). \quad (8)$$

- 2) P-Step: Given the complete sample set obtained from I-Step, in P-Step the posterior distribution for the mean vector and for the variance and covariance matrix is generated, formally:

Based on the complete vector, $\Theta^{(t+1)}$ is simulated from the conditional distribution:

$$P(\Theta^{(t+1)} | Y_{\sim obs}, Y_{\sim mis}^{(t+1)}). \quad (9)$$

The new missing values for the parameters are obtained and used again in I-Step. Proceeding with these steps for long enough, we obtain the following Markov chain:

$$(Y_{\sim mis}^{(1)}, \Theta^{(1)}), (Y_{\sim mis}^{(2)}, \Theta^{(2)}), \dots, \quad (10)$$

which converges in probability to $P(Y_{\sim mis}, \Theta | Y_{\sim obs})$.

Assuming that $Y \sim N(\mu, \Sigma)$, with μ and Σ as previously described and that for the parameters we have $\mu | \Sigma \sim W^{-1}(\mu_0, \frac{1}{\tau} \Sigma)$, where μ_0 and $\tau > 0$ are fixed, known values and $\Sigma \sim W^{-1}(m, \psi)$, where m is the number of degrees of freedom and ψ is a precision matrix (Anderson, 1984), the posterior distribution (Anderson, 1984; Shafer, 1999) is given by:

$$\Sigma | Y \sim W^{-1} \left(n + m, (n-1)S + \psi + \frac{n\tau}{n+\tau} (\bar{y} - \mu_0)(\bar{y} - \mu_0)' \right) \quad (11)$$

$$\mu | (\Sigma, Y) \sim N \left(\frac{1}{n+\tau} (n\bar{y} + \tau\mu_0), \left(\frac{1}{n+\tau} \Sigma \right) \right),$$

where n is the number of variables. Since there is no available information, we use Jeffrey's prior, where:

$$\Sigma^{(t+1)} | Y \sim W^{-1} (n-1, (n-1)S)$$

$$\mu^{(t+1)} | (\Sigma^{(t+1)}, Y) \sim N \left(\bar{y}, \frac{1}{n} \Sigma^{(t+1)} \right). \quad (12)$$

5 Data Envelopment Analysis

Data Envelopment Analysis (DEA) is used as a non-parametric method for measuring efficiency of decision-making units (DMUs) such as companies, individuals or any other populational element of interest.

This method was first introduced in operations research literature by Charnes *et al.* (1978). The original CCR model had as target only technologies featuring constant returns to scale, i.e., technologies represented by functions $f(x)$ such that $f(\lambda x) = \lambda f(x)$, $\forall \lambda$. Banker *et al.* (1984) have extended the CCR model in order to accommodate processes that feature variable returns to scale, which was a considerable advance in the literature of efficient frontiers.

In the following years, the methodological contributions have increased exponentially, due to a great number of researchers producing many works on CCR and BCC models. The fast diffusion of DEA as an acceptable method for analyzing efficiency can be seen by the following figure: Seiford (1994), in his doctorate thesis on DEA, presents as reference as much as 472 published articles on the topic.

In a more recent compilation, Taveres (2002) includes 3183 references from 2152 different authors. In the present moment, a search on the web for the term "Data envelopment analysis" returns as much as 246000 hits. In parallel, computational development of software and hardware allows for easy access to data manipulation, thus permitting fast solutions for the linear programming problems involved in data envelopment analysis.

Among the many available models, we have looked for the ones that are capable of incorporating uncertainty from missing values in the dataset. We have adopted as model the intervalar DEA suggestion of Despotis and Smirlis (2002), shown next.

6 Intervalar DEA formulation

It is assumed that there are n units, each with m inputs, producing s outputs. For each unit (DMU) j ($j = 1, \dots, n$), let y_{rj} , ($r = 1, \dots, s$), be the level of its r th output and x_{ij} , ($i = 1, \dots, m$), the level of its i th input.

Differently from the traditional DEA, the intervalar DEA assumes that some of the inputs x_{ij} and outputs y_{rj} are not precisely known. We only know that $x_{ij} \in [x_{ij}^L, x_{ij}^U]$ e $y_{rj} \in [y_{rj}^L, y_{rj}^U]$, where the intervals must be constant and strictly positive.

In this sense, the units can take any values within this interval in order to maximize its efficiency. Therefore, using intervals instead of single values, the resulting DEA-CCR is non-linear and the weights for the outputs and inputs are given by $u_1, \dots, u_r, \dots, u_s$ e $v_1, \dots, v_i, \dots, v_m$ respectively.

Therefore, for the j_0 th DMU we have:

$$\begin{aligned} \text{Minimize: } Z_{j_0} &= \sum_{r=1}^s u_r y_{rj_0} \\ \text{Subject to:} \\ \sum_{i=1}^m v_i x_{ij_0} &= 1 \\ \sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} &\geq 0 \\ \text{with } j &= 1, \dots, n, \text{ for } u_r, v_i \geq \epsilon \quad \forall r, i. \end{aligned} \tag{13}$$

Note that the model (13) is non-linear since not only the weights but also the input and output levels are variables to be solved. Despotis and Smirlis (2002) suggest the following strategy for linearizing the problem: expressing the values x_{ij} and y_{rj} in terms of new variables s_{ij} and t_{rj} that allocate the levels of input and output inside the known intervals $[x_{ij}^L, x_{ij}^U]$ and $[y_{rj}^L, y_{rj}^U]$ in the following manner:

$$\begin{cases} x_{ij} = x_{ij}^L + s_{ij}(x_{ij}^U - x_{ij}^L) & \text{for } i = 1, \dots, m; j = 1, \dots, n; 0 \leq s_{ij} \leq 1. \\ y_{rj} = y_{rj}^L + t_{rj}(y_{rj}^U - y_{rj}^L) & \text{for } r = 1, \dots, s; j = 1, \dots, n; 0 \leq t_{rj} \leq 1. \end{cases} \tag{14}$$

Using these expressions, the terms $v_i x_{ij}$ and $u_r y_{rj}$ take the following forms:

$$\begin{cases} v_i x_{ij}^L + v_i s_{ij} (x_{ij}^U - x_{ij}^L) \\ u_r y_{rj}^L + u_r t_{rj} (y_{rj}^U - y_{rj}^L), \end{cases} \quad (15)$$

respectively. In these expressions, the new terms $v_i s_{ij}$ (inputs) and $u_r t_{rj}$ (outputs) are switched by new variables $q_{ij} = v_i s_{ij}$ and $p_{rj} = u_r t_{rj}$ that meet the following conditions: $0 \leq q_{ij} \leq v_i, 0 \leq p_{rj} \leq u_r, \forall i, j, r$. Thus, we obtain:

$$s_{ij} = \frac{q_{ij}}{v_i} \quad \text{and} \quad t_{rj} = \frac{p_{rj}}{u_r} \quad \text{with } v_i, u_r \geq \epsilon, \quad (16)$$

such that $0 \leq s_{ij}, t_{rj} \leq 1$ for all i, j and r . Applying the described transformations on the model (13), we obtain the following linear model:

$$\text{Minimize: } Z_{j_0} = \sum_{r=1}^s u_r y_{rj_0}^L + p_{rj_0} (y_{rj_0}^U - y_{rj_0}^L)$$

Subject to

$$\sum_{i=1}^m v_i x_{ij_0}^L + q_{ij_0} (x_{ij_0}^U - x_{ij_0}^L) = 1 \quad (17)$$

$$\sum_{r=1}^s u_r y_{rj}^L + p_{rj} (y_{rj}^U - y_{rj}^L) - \sum_{i=1}^m v_i x_{ij}^L + q_{ij} (x_{ij}^U - x_{ij}^L) \leq 0 \quad \text{with } j = 1, \dots, n.$$

$$\text{for } p_{rj} - u_r \leq 0; r = 1, \dots, s; j = 1, \dots, n \quad q_{ij} - v_i \leq 0; i = 1, \dots, m; j = 1, \dots, n, u_r$$

$$p_{rj} \geq 0, q_{ij} \geq 0, \forall r, i, j.$$

The goal of the model is, then, to estimate the weights u_r, v_i and the levels of input and output. It is interesting to note that the DEA-CCR model is a particular case of (17) when the size of the intervals is zero.

Therefore, in the mathematical programming problem (17), we want to find the efficiency score for the j_0 th DMU. Since some input and output values are unknown, it is necessary to initially estimate the lower limit and then the upper limit for the DMU's efficiency score. When unit j_0 is evaluated by model (17), the levels of input and output, together with their respective weights, are adjusted in order to favor unit j_0 .

In this manner, the efficiency score obtained by unit j_0 in model (17) is not worse than any other efficiency score this unit might obtain. We can, however, build intervals for the efficiency scores. More specifically, the lower limit for the efficiency score is given by:

$$\begin{aligned}
 &\text{Maximize: } L_{j_0} = \sum_{r=1}^s u_r y_{rj_0}^L \\
 &\text{Subject to} \\
 &\sum_{i=1}^m v_i x_{ij_0}^U = 1 \\
 &\sum_{r=1}^s u_r y_{rj_0}^L - \sum_{i=1}^m v_i x_{ij_0}^U \leq 0 \\
 &\sum_{r=1}^s u_r y_{rj}^U - \sum_{i=1}^m v_i x_{ij}^L \leq 0 \text{ for } j = 1, \dots, n \text{ e } j \neq j_0 \\
 &u_r, v_i \geq \epsilon, \forall r, i.
 \end{aligned} \tag{18}$$

For the upper limit, we have the following model:

$$\begin{aligned}
 &\text{Maximize: } U_{j_0} = \sum_{r=1}^s u_r y_{rj_0}^U \\
 &\text{Subject to} \\
 &\sum_{i=1}^m v_i x_{ij_0}^L = 1 \\
 &\sum_{r=1}^s u_r y_{rj_0}^U - \sum_{i=1}^m v_i x_{ij_0}^L \leq 0 \\
 &\sum_{r=1}^s u_r y_{rj}^L - \sum_{i=1}^m v_i x_{ij}^U \leq 0 \text{ for } j = 1, \dots, n \text{ e } j \neq j_0 \\
 &u_r, v_i \geq \epsilon, \forall r, i.
 \end{aligned} \tag{19}$$

The idea behind the construction of these limits consists in searching for the worst and best possible scenarios for each DMU, considering all other DMUs. In this sense, the upper limit of DMU i , for example, can be found by using the upper estimate of its output, the lower estimate of its inputs, the lower estimates of the other DMUs' outputs and the upper estimates their inputs. Therefore, we estimate the best possible location of DMU i on an efficiency frontier that is built in the most inefficient manner, thus making the DMU i 's efficiency index the greatest possible.

Based on the efficiency scores, we can classify the DMUs as:

- 1) $E^{++} = \{j \in J \mid h_j^L = 1\}$: The units that are efficient in any case, i.e., any combination of inputs and outputs.

- 2) $E^+ = \{j \in J \mid h_j^L < 1 \text{ e } h_j^* = 1\}$: The units that are efficient in a maximal sense but there are still some combinations of inputs and outputs that do not reach the optimal.
- 3) $E^- = \{j \in J \mid h_j^* < 1\}$: Units are definitely inefficient.

7 Results

Using data from the Bloomberg system, we have obtained 449 companies listed in the system by July 2010. For these organizations, we have kept in the database those that featured at least one valid observation, thus resulting in 403 organizations.

In the suggested model, we are interested in the efficiency in generating intangible assets, considering the companies from BM&F. Therefore, we define as output only the first variable (Stated intangibles - last filing) in Table 1.

Table 1 - Variables used in the model

Stated intangibles (Last Filing).	Personnel expenditures in 2005
Sales, marketing and advertisement expenditures in 2005	Personnel expenditures in 2006
Sales, marketing and advertisement expenditures in 2006	Personnel expenditures in 2007
Sales, marketing and advertisement expenditures in 2007	Personnel expenditures in 2008
Sales, marketing and advertisement expenditures in 2008	Personnel expenditures in 2009
Sales, marketing and advertisement expenditures in 2009	Gross profit in 2005
Research and development expenditures in 2005	Gross profit in 2006
Research and development expenditures in 2006	Gross profit in 2007
Research and development expenditures in 2007	Gross profit in 2008
Research and development expenditures in 2008	Gross profit in 2009
Research and development expenditures in 2009	

In the suggested model, we are interested in the efficiency in generating intangible assets, considering the companies from BM&F. Therefore, we define as output only the first variable (Stated intangibles - last filing).

All other variables were considered inputs for the generation of this output. Provided that the greater part of the observations from Bloomberg had missing values (due to the fact that providing some statements is not demanded by Brazilian regulation), we cannot consider those as zeros.

Therefore, using the Markov Chain Monte Carlo method described in section 2 we have built two databases featuring, respectively, the first quartile and the third quartile from the estimated distribution of the missing values, according to MCMC method.

Once having obtained these intervals, we proceed with the model suggested by Despotis and Smirlis (2002). For this dataset, we have obtained the following scale of efficiency in producing intangible assets, presented in Figure 1.

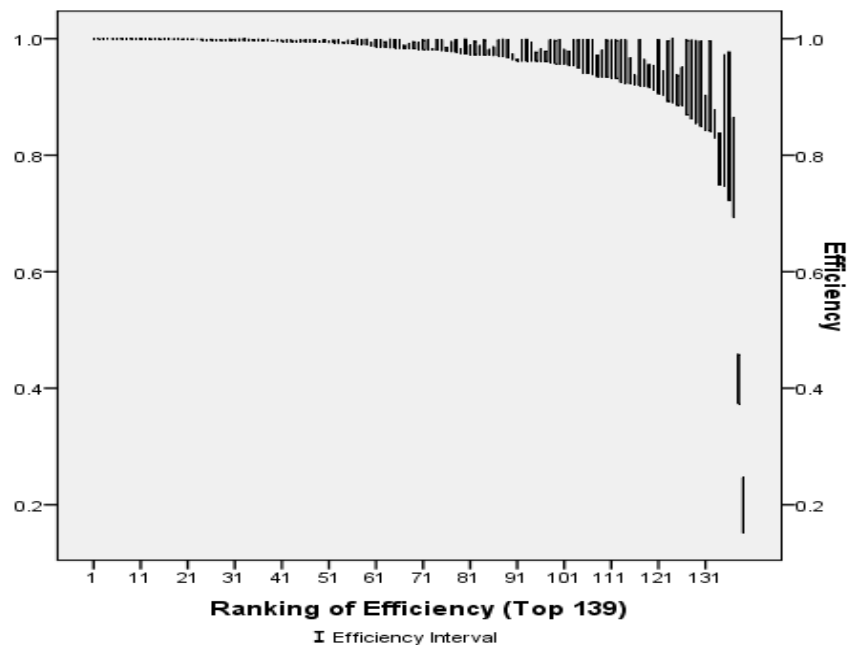


Figure 1 - Ranking of efficiency.

Using the data envelopment analysis with missing data, only 139 DMUs were informative in the sense of having an efficiency interval smaller than 1. In the other cases, the decision making units had either zero efficiency or its interval was big enough to make it not informative.

Based on the efficiency scores, we can rank all 403 DMUs. It is clear from the Table 2 that only a small part of the publicly-traded companies are efficient in a broad sense, i.e., belong to classes E^{++} and E^+ .

Table 2 - Frequency distribution by type of efficiency score

Type of score	Frequency	Percentage
E^{++}	4	0.99%
E^+	17	4.22%
E^-	382	94.79%

A question that naturally arises is how these companies differ in terms of their stock volatility in the capital markets. In order to check for the existence of relation between the efficiency score and the total volatility in 2009, we have used the exact test of Freeman (1951), where the classes of return were the quartiles of the total volatilities of all 404 companies.

For the test suggested by Freeman (1951), we have obtained a critical value of 0.0462 from the data presented in Table 3, thus rejecting the null hypothesis of independence between quartile classes and efficiency scores. Therefore, the results suggest a dependence relation between efficiency in generating intangible assets and the company's volatility.

Table 3 - Contingency table between semi-variance and efficiency scores

	Q1[0, 1.89154)	Q2 [1.9154, 2.45951)	Q3 [2.45951, 3.74924)	Q4 [3.74924, ∞)
E^{++}	0.50%	1.99%	1.00%	0.75%
E^+	0.75%	0.00%	0.00%	0.25%
E^-	35.82%	18.91%	19.90%	20.15%

For the test suggested by Freeman (1951), we have obtained a critical value of 0.0462, thus rejecting the null hypothesis of independence between quartile classes and efficiency scores. Therefore, the results suggest a dependence relation between efficiency in generating intangible assets and the company's volatility.

Conclusion

In this paper, the main goal was to study how the publicly-traded companies are distributed according to their efficiency in generating intangible assets. For this, we have used the Markov Chain Monte Carlo method for the imputation of the quartiles of the missing observations.

The companies' efficiency was then evaluated according to the model of Despotis and Smirlis (2002), which uses intervals instead of single values, thus increasing the number of valid observations for the analysis.

We show that approximately 1% of Brazilian companies can be considered efficient in generating intangible assets, which suggests some vulnerability in the capital markets, since the greater the presence of intangible assets the smaller the probabilities of default Thorburn (2000).

Finally, we show how the types of efficiency score are related to the volatility levels of the BM&F companies, corroborating the hypothesis that the intangible asset is one of the tools organizations have that help stabilize their market volatility and decrease their probability of default.

The results show that, for the publicly-traded companies in Brazil, a minority is capable of efficiently generating intangible assets. In this sense, there are only a few organizations capable of generating differentiability in relation to their products and services - since intangibility, such as the power of the brand, patents and personnel, is the most important feature of a company (Blackston, 2003)

ALBUQUERQUE, P. H. M.; GUIMARÃES, R. D. Avaliação da eficiência organizacional na geração de ativos intangíveis. *Rev. Bras. Biom.*, São Paulo, v.29, n.4, p.643-656, 2011.

- RESUMO: Neste trabalho, estudamos a intangibilidade das organizações utilizando análise envoltória de dados como uma ferramenta para avaliar a eficiência organizacional. As empresas do mercado de capitais brasileiro são estudadas, com foco na eficiência na geração de ativos intangíveis. No entanto, uma vez que existem apenas algumas observações para a análise tradicional, usamos a abordagem MCMC, associada com a análise envoltória de dados para dados faltantes, visando aumentar o número de observações informativas, permitindo assim uma análise mais completa. Por fim, apresentamos como os tipos de organizações (de acordo com sua pontuação de eficiência) estão associadas com a volatilidade da empresa, sugerindo uma relação entre a eficiência na geração de ativos intangíveis e os riscos de investimento das organizações.
- PALAVRAS-CHAVE: Análise envoltória de dados; MCMC; ativos intangíveis.

References

ANDERSON, T. W. *An introduction to multivariate statistical analysis*. 2.ed. New York: John Wiley & Sons, 1984. 688p.

BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage. Sci.*, Hanover, v.30, n.9, p.1078-1092, 1984.

BLACKSTON, M. Observations: building brand equity by managing the brand's relationships. *J. Advert. Res.*, New York, v.40, n.6, p.101-105, 2003.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. *Eur. J. Oper. Res.*, Amstermam, v.2, p.429-444, 1978.

DESPOTIS, D. K.; SMIRLIS, Y. G. Data envelopment with imprecise data. *Eur. J. Oper. Res.*, Amsterdam, v.140, p.24-36, 2002.

FREEMAN, G. H.; HALTON, J. H. Note on an exact treatment of contingency, goodness of fit and other problems of significance. *Biometrika*, Oxford, v.38, n.1-2, p.141-149, 1951.

HILMOLA, O. P.; TORKKELI, M.; SAVOLAINEN, V. V. Intangibles and business performance a technical efficiency approach. *Int. J. Entrep.sh. Innov. Manage.*, London, v.9, n.4, p.364-379, 2009.

KAYO, E.; KIMURA, H.; LEIVA, D. M.; NAKAMURA, W. Ativos intangíveis, ciclo de vida e criação de valor. *RAC*, London, v.10, n.3, p.73-90, 2006.

LITTLE, R. J. A.; RUBIN, D. B. *Statistical analysis with missing data*. New York: Wiley Series in Probability and Statistics, 1987. 382p.

SEIFORD, L. R. *A DEA Bibliography 1978-1992. Data envelopment analysis: theory, methodology and application*. Boston: Kluwer Academic Publishers, 1994. p.437-470.

SERRANO-CINCA, C.; FUERTES-CALL, Y.; MAR-MOLINERO, C. Measuring DEA efficiency in Internet companies. *Decis. Support Syst.*, Amsterdam, v.38, n.4, p.557-573, 2005.

SHAFER, J. L. Analysis of incomplete multivariate data. *Comput. Stat. Data Anal.*, Amsterdam, v.30, n.4, p.478-479, 1999.

TAVERES, G. *A bibliography of data envelopment analysis (1978-2001)*. RUTCOR Research Report RRR 01-02. Piscataway: Rutgers University, 2002.

THORBURN, K. S. Bankruptcy auctions: costs, debt recovery, and firm survival. *J. Financ. Econ.*, Lausanne, v.58, n.3, p.337-368, 2000.

WU, W.-Y.; TSAI, H.-J.; CHENG, K.-Y.; LAI, M. Assessment of intellectual capital management in Taiwanese IC design companies: using DEA and the Malmquist productivity index. *R&D Manage.*, v.36, n.5, p.531-545, 2006.

Received in 08.11.2011

Approved after revised in 05.03.2012