E C O N O M I C S B U L L E T I N

DEA investment strategy in the Brazilian stock market

Ana Lopes
Department of Management, Santa Catarina South University

Edgar Lanzer
Department of Management, Santa Catarina South University

Marcus Lima
Department of Management, Santa Catarina South University

Newton da Costa, Jr.
Department of Economics, Federal University of Santa Catarina

Abstract

This paper presents a multi-period investment strategy using Data Envelopment Analysis (DEA) in the Brazilian stock market. Results show that the returns based on the DEA strategy were superior to the returns of a Brazilian stock index in most of the 22 quarters analyzed, presenting a significant Jensen's alpha.

Citation: Lopes, Ana, Edgar Lanzer, Marcus Lima, and Newton da Costa, Jr., (2008) "DEA investment strategy in the Brazilian stock market." *Economics Bulletin,* Vol. 13, No. 2 pp. 1-10
Submitted: December 4, 2007. **Accepted:** March 26, 2008.
URL: http://economicsbulletin.vanderbilt.edu/2008/volume13/EB-07M00002A.pdf

1. Introduction

Since the seminal work of Markowitz (1952), several models of portfolio selection have been suggested. According to Cohen and Pogue (1967), the majority of these models sought to simplify the estimation process of the of variance-covariance matrix, with the intention of making the calculation faster in computational terms. Additionally, since the Markowitz model is static (for a single period), sophisticated multi-period or dynamic extensions have also been developed (e.g., Merton 1969, Samuelson 1969, Cover and Julian 2000). In order to estimate the expected returns, variances, and covariances, increasingly complex techniques for forecasting the behavior of stock prices and volatility also arose. See Poon and Granger (2003) for a review of these latest techniques.

On the other hand, Data Envelopment Analysis (DEA), has demonstrated a very satisfactory applicability in the evaluation (*ex-post*) of investment funds (e.g., Morey and Morey 1999, Gregorious 2003, Haslem and Scheraga 2003), although its initial applications had been predominantly to public organizations (e.g., Shen et al. 2005, Zhu 2003, Avkiran 2001, Calhoun 2003).

DEA is a technique taken from the operational research area developed by Charnes, Cooper and Rhodes (1978) and by Banker, Charnes and Cooper (1984). It is a powerful management tool, used for evaluating and comparing organizational units, noteworthy for the operational advantages offered in multi-attribute evaluations and also in the evaluation of the performance of multidimensional indicators in general.

The discussion of the preceding paragraphs immediately begs the following question: What is the potential use (*ex-ante*) of DEA in portfolio selection? This *ex-ante* potential does not appear to have been evaluated until now, constituting, therefore, the objective of the present work.

Next section presents the Data Envelopment Analysis technique. Section 3 presents the procedure for collecting data and the method employed. Section 4 describes the results and the final section concludes.

2. Performance Analysis using DEA

Performance measurement using DEA approach consists in determining the relative efficiency of a productive unit by considering its closeness to an *efficiency frontier*. DEA efficiency is not to be confused with mean-variance efficiency in the Markowitz model, where mean and variance are the only two parameters for optimization. In DEA approach, efficiency is the objective function value of a multi-criteria linear programming model. The objective of the DEA is to determine relative performance indicators among productive units, considering specific groups of inputs and outputs.

According to Cornuejols and Trick (2004), successively resolving the problem for all the productive units, a subgroup of those productive units considered to be efficient is obtained, which will serve as a basis for the determination of the efficiency frontier, and for the establishment of goals for the inefficient units. Therefore, each unit is compared only with similar units with the best performance, that is, those situated on the frontier of efficiency. Any productive unit included or excluded from the group under analysis modifies the production group and, as a result, the frontier itself.

Charnes, Cooper and Rhodes (1978) proposed a model that assumes constant returns to scale, called CRS or CCR. Subsequent works assumed different sets of suppositions, like the model developed by Banker, Charnes and Cooper (1984), which assumed variable returns to scale, called VRS or BCC. Both models can be classified as input oriented or output oriented, depending on the search for goals by the inefficient units.

Figure 1 presents an example of the estimation of the efficiency (or inefficiency), considering the return to scale. In this example, productive units subject to only one input and one output are considered, as shown in part (a) of the Figure 1. Part (b) shows the estimated efficiency frontier via CRS, and via VRS approach. In the case of the CRS model, with product/output orientation, the technical inefficiency of productive unit D can be estimated by the segment DD''. If we consider the VRS model, the technical inefficiency is DD'. The difference between these two measurements is called inefficiency of scale.

The CRS model that estimates the relative efficiency among a group of productive units is represented by the equation (1) that follows:

$$Max \quad \theta = \sum_{k=1}^{s} u_k Y_{kl}$$

subjetc to

$$\begin{aligned} & \sum_{i=1}^{m} v_i X_{il} = 1 \\ & \sum_{k=1}^{s} u_k Y_{kj} - \sum_{i=1}^{n} v_i X_{ij} \le 0, \quad j = 1,\ldots,n \\ & u_k, v_i \ge 0, \forall x, y \end{aligned} \tag{1}$$

In the above equation, θ indicates the level of efficiency of the unit under analysis. The vectors of inputs X and outputs Y of unit j are represented by X_{ij} and Y_{kj}. If $\theta = 1$, then unit 1 (unit under analysis) is on the frontier of efficiency and can serve as a reference for the rest. On the other hand, if $\theta < 1$, then unit 1 can still increase its outputs, maintaining all the inputs unaltered, and is, therefore, inefficient among the group of units analyzed. Each unit is compared with a virtual unit obtained by linear combination of all the units of the group. Each input and output of this virtual unit is the linear combination of the inputs and outputs of all the units of the group. The value of θ in this case is always less than or equal to one. The existence of a value of θ less than one indicates the possibility of constructing a virtual unit that can produce more, utilizing an equal (or lower) quantity of inputs than the unit analyzed.

In other words, it seeks to radially expand the vector of outputs, as much as possible, for the unit under analysis. The limit is the estimated efficiency frontier for the group of points observed (these points are determined by the other productive units). This problem must be resolved for each unit, generating its relative rate of efficiency. In relation to the orientation to inputs or to outputs, it is important to note that if a unit was efficient in the product/output oriented model, as presented here, the input oriented model will also be efficient. However, depending on the orientation utilized, the reference units and the indicators of retraction of inputs or expansion of outputs, for the inefficient units, can be different (Bhat et al. 2001).

3. Data and Method Utilized

The present work extends the one-period results of Powers and McMullen (2000) in the U.S. stock market to a multi-period investment strategy applied to the Brazilian stock market.

A total of seven attributes are taken into account, three of them are considered input to the DEA model: price to earnings ratio, beta, and return volatility for each stock, while the other four are considered outputs: earnings per share, and the last 12, 36, and 60 month return. The only difference to Powers and McMullen was that they used an additional output, 10 year or 120 month return.

The data utilized originated from the Economatica®. To be included in the sample the stock should belong to the IBrX-100 index (the Sao Paulo Stock Exchange value-weighted index) at the beginning of each of the 22 quarters along the period of Jan/2001 to Jun/2006. Along the study period companies comprising the IBxX-100 index accounted for more than 85% of the total capitalization in the Brazilian stock market.

The earnings per share and beta were the numbers in the immediately preceding quarter up to the day of portfolio formation, while the standard deviation of the returns was computed on a daily basis (closing price) over the previous 36 months.

The DEA technique helped to rebalance the portfolio at the beginning of each of the 22 quarters, from the first quarter of 2001 to the second quarter of 2006. In each quarter the portfolio return was estimated as an equally weighted average of the selected stock returns in that quarter.

All the attribute data used were standardized. This procedure makes the numeric instances more balanced, reducing the risk of imprecision in the computation. The standardization was performed according to equation (2):

$$Z_{ij} = \frac{(X_{ij} - \overline{X}_j)}{\hat{\sigma}_j}, \qquad (2)$$

where Z_{ij} is the standardized result for indicator j of stock i; X_{ij} is the value of the indicator j of stock i; $\overline{X}_j$ is the average of indicator j for all the stocks; and $\hat{\sigma}_j$ is the standard deviation of indicator j for all the stocks.

After the indicators standardization, a re-scaling is necessary, since the DEA model does not accept negative values. For this, the minimum value of each indicator column was determined according to the following formula:

$$RZ_{ij} = \text{Abs}\ (\text{Min}\ Z_j) + Z_i\ , \qquad (3)$$

where Z_{ij} is the re-scaling for each j attribute.

Finally, all the attributes in all the quarters were divided by the respective maximum, as shown by equation (4):

$$MRZ_{ij} = RZ_{ij}\ /\text{Column maximum}\ , \qquad (4)$$

where MRZ_{ij} is the normalization of stock i in attribute j.

The DEA model used to select the stocks with the best performance was the CCR type "oriented to product/output" (see Charnes et al. 1978). This approach is presented in equation (5):

$$\boldsymbol{Max} \quad \boldsymbol{h} = \sum_{j=1}^{4} \boldsymbol{u}_{jl} \boldsymbol{O}_{jl} + \mu_i$$

Subject to

$$\sum_{i=1}^{3} \boldsymbol{v}_{il} \boldsymbol{I}_{il} + \rho_i = 1 \tag{5}$$

$$\sum_{k=1}^{4} \boldsymbol{u}_{kj} \boldsymbol{O}_{kj} - \sum_{i=1}^{3} \boldsymbol{v}_{ij} \boldsymbol{I}_{ij} + \mu_i \leq 0, \ \boldsymbol{for} \quad \boldsymbol{j} = 1, \ldots \boldsymbol{n} \ ,$$

where O_{kj} and I_{ij} are, respectively, the value of the kth *output* attribute of stock j and the value of the ith *input* attribute of the same stock. The values of μ_i and ρ_i are slack variables that quantify the inefficiency in stock j. All the weights are non-negative and are subordinated to the restrictions described below. The analysis is carried out for each of the 100 stocks.

The objective function (5) is to provide the maximum efficiency value for stock j**,** determined by means of the weighting factors of each output and input. The DEA method selects efficient alternatives with values of 100% in accordance with the model presented.

To prevent stocks being classified as efficient despite having undesirable levels of various attributes, an additional restriction imposed on the weighting of the output indicators was placed. This action is known in the DEA literature as setting "multiplier bounds". Here the multiplier bounds were selected in order that dominance of one output over another was limited to a factor of five. This is possible through the following restriction:

$$\frac{1}{5} \leq \frac{\boldsymbol{u}_1}{\boldsymbol{u}_2}, \frac{\boldsymbol{u}_1}{\boldsymbol{u}_3}, \frac{\boldsymbol{u}_1}{\boldsymbol{u}_4}, \frac{\boldsymbol{u}_1}{\boldsymbol{u}_5}, \frac{\boldsymbol{u}_2}{\boldsymbol{u}_3}, \frac{\boldsymbol{u}_2}{\boldsymbol{u}_4}, \frac{\boldsymbol{u}_2}{\boldsymbol{u}_5}, \frac{\boldsymbol{u}_3}{\boldsymbol{u}_4}, \frac{\boldsymbol{u}_3}{\boldsymbol{u}_5}, \frac{\boldsymbol{u}_4}{\boldsymbol{u}_5} \leq 5 \tag{6}$$

Similarly, the restriction imposed on the weighting of the *input* indicators were:

$$\frac{1}{5} \leq \frac{\boldsymbol{v}_1}{\boldsymbol{v}_2}, \frac{\boldsymbol{v}_1}{\boldsymbol{v}_3}, \frac{\boldsymbol{v}_2}{\boldsymbol{v}_3} \leq 5 \tag{7}$$

In this model, a stock is considered efficient if none of its output attributes (earnings per share, returns for 1, 3 and 5 years) can be increased, nor any of its input indicators (price to earnings, beta and return volatility) reduced, without reducing some output attribute or increasing some input attribute.

Stocks considered to be efficient were selected to make up a portfolio at the beginning of a quarter, based on historical data of the attributes of the various stocks of the sample. In each of the 22 quarters DEA-portfolio was composed by an investment of the same proportion for each efficient stock, that is, the portfolio was equally weighted. The acquisition of the stocks on the first day of a quarter and the sale on the last day of the same quarter was simulated. For the calculation of the return for each stock, the closing price on the first and last day of the quarter was used. The same procedure was adopted for calculating the IBrX100 index returns.

Comparisons with the performance of the DEA- portfolio were also made with the CDI (Interbank deposit certificate) rate for the same quarters. This last series is generally used as a proxy for the risk free rate in the Brazilian capital market.

4. Results

The application of DEA to stock selection resulted in a different number of stocks for each quarter along the 22 quarters analyzed. The average number of stocks in each quarter was six. However, this number varied from one, in two consecutive quarters to twelve, in one quarter. This number could go up to twelve as long as there are four outputs and three inputs, thus forming twelve partial output/input ratios that could be dominant over the other stocks.

Also, it must be remembered that the procedure adopted was that each DEA-efficient stock would make up an equal fraction of the portfolio in one quarter and that it could be a candidate equally qualified to make up the portfolio in the following quarter.

Figure 2 presents the 22 quarterly returns for each of the three series and its cumulative (geometric) returns. One can observe the high variability of the DEA and IBrX100 series when compared to the CDI one.

Table 1 reports the 22 quarterly returns for each of the three series. Jarque-Bera statistics show that one cannot reject the null hypothesis of normality for any of the series. DEA-portfolio presented higher quarterly returns than the IBrX100 index in 17 out of the 22 quarters analyzed. When the comparison is with CDI quarterly rates, DEA outperforms CDI in 13 out of 22 quarters.

When performing t-tests on the equality between two series, DEA average returns were found to be marginally higher than IBrX-100 returns (p-value of .15), and significantly higher than CDI (p-value of .03), both results were obtained using one-tail t-tests. Interestingly, the null hypothesis "IBrX-100 average returns are higher than CDI average returns" could not be rejected (p-value of .20).

Also, Table 1 reports a statistically significant Jensen's alpha for the DEA-portfolio (p-value of .08). It is the coefficient from a regression of the DEA-portfolio's excess returns on the benchmark's (IBrX-100) excess returns, where excess return is a portfolio's return in excess of the riskless return (proxied by the CDI rate). Jensen's alpha is a well known risk-adjusted measure of portfolio performance used both in the mutual fund industry and in the academy.

5. Conclusions

This paper presented a multi-period investment strategy based on Data Envelopment Analysis (DEA) technique to select DEA-efficient stocks traded in the Sao Paulo Stock Exchange, along the period of Jan/2001 to Jun/2006.

Along the 22 quarters analyzed, the technique was capable of generating superior performance when compared with both market average, proxied by the IBrX-100 index, and CDI (Brazilian interbank deposit certificate) quarterly rates. When the comparison was made between total returns DEA-portfolio series was only marginally superior to the IBrX-100 series. However, when the comparison was made with excess returns DEA strategy could achieve a significant Jensen's alpha.

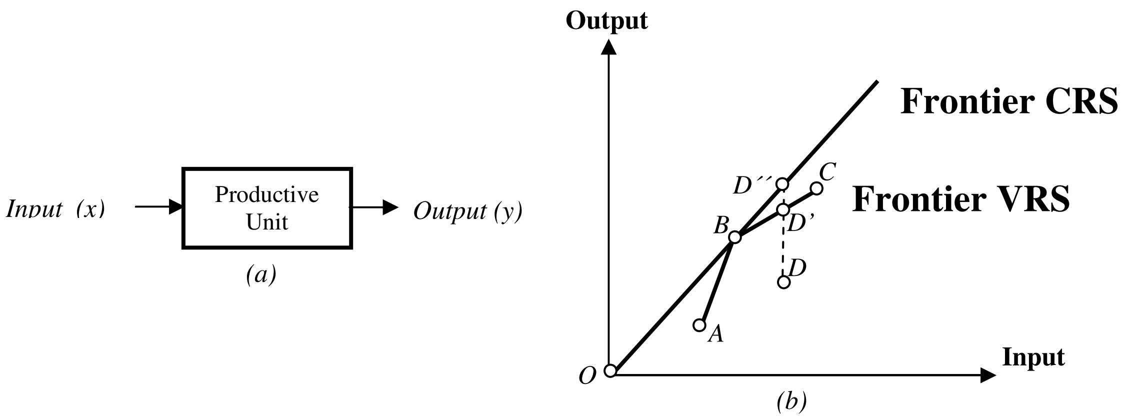

Figure 1. (a) Productive unit considered with one input and one output; (b) Graphic demonstration of the calculation of efficiency (or inefficiency) of scale.

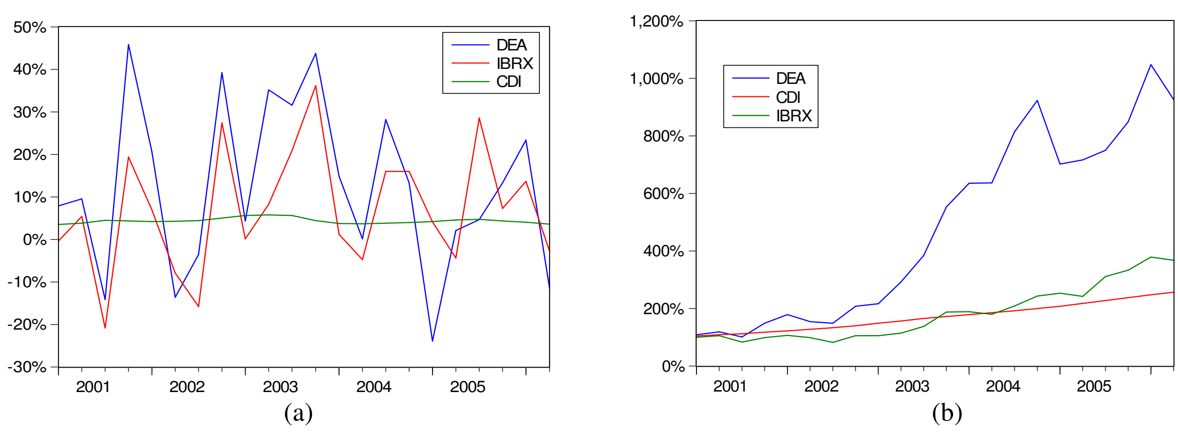

Figure 2. (a) Quarterly Returns, and (b) Cumulative Returns for the DEA strategy, IBrX100 index, and CDI series

Table 1
Descriptive statistics for the quarterly returns

Quarter/Year	DEA (%)	IBrX100 (%)	CDI (%)
Q1/2001	7.90	-.44	3.56
Q2/2001	9.50	5.35	3.84
Q3/2001	-14.19	-20.87	4.49
Q4/2001	45.91	19.40	4.38
Q1/2002	20.69	6.98	4.20
Q2/2002	-13.64	-7.93	4.26
Q3/2002	-3.58	-15.76	4.43
Q4/2002	39.28	27.42	4.99
Q1/2003	4.32	.12	5.67
Q2/2003	35.22	8.31	5.78
Q3/2003	31.60	20.86	5.61
Q4/2003	43.80	36.18	4.40
Q1/2004	14.90	1.22	3.76
Q2/2004	.14	-4.71	3.67
Q3/2004	28.18	16.02	3.86
Q4/2004	13.30	16.03	3.99
Q1/2005	-23.99	4.12	4.18
Q2/2005	2.06	-4.34	4.56
Q3/2005	4.63	28.54	4.74
Q4/2005	13.30	7.26	4.31
Q1/2006	23.36	13.68	4.04
Q2/2006	-11.53	-2.80	3.58
Mean (1)	**12.33**	**7.03**	**4.38**
Sdt. Dev.	**19.66**	**14.33**	**0.65**
Jarque-Bera(2) (p-value)	**.69 (.71)**	**.24 (.89)**	**2.99 (.22)**
Jensen's Alpha(3) (p-value)	**5.26 (.08)**	**-**	**-**
Cumulat. Geometric Returns (%)	**927.08**	**368.42**	**256.50**

Notes

(1) DEA Portfolio mean returns are higher than IBrX-100 index returns with a p-value of .15 (one tail t-test) and higher than CDI rate with a p-value of .03 (one tail t-test). (2)Jarque-Bera tests the null hypothesis that the distribution function of the returns is normal. (3)Jensen's Alpha of the DEA portfolio is the intercept from a regression of the portfolio's excess returns on the benchmark's excess returns. Jensen's Alpha is significant (p-value of .08)

References

Avkiran, N. K. (2001) "Investigating technical and scale efficiencies of Australian universities through data envelopment analysis" *Socio-Economic Planning Sciences* 35, 57-80.

Banker, R. D. A., A. Charnes, and W. W. Cooper (1984) "Some models for estimating technical and scale inefficiencies in data envelopment analysis" *Management Science* 30, 1078-1092.

Bhat, R., B. B. Verma, and E. Reuben (2001) "A note on data envelopment analysis [DEA]" *Journal of Health Management* 3, 309-328.

Calhoun, J. (2003) "Data envelopment analysis of relative efficiencies of institutions of higher learning" Association for the Study of Higher Education, Portland, OR.

Charnes, A., W. W. Cooper, and R. Rhodes (1978) "Measuring the efficiency of decision making units" *European Journal of Operational Research* 2, 429-444.

Cornuejols, G., and M. Trick (2004) "Quantitative methods for the management sciences" Course notes, Graduate School of Industrial Administration, Carnegie Mellon University, 45-760**.** In: <http://mat.gsia.cmu.edu/QUANT/NOTES/>.

Cohen, K. J., and J. A. Pogue (1967) "An empirical evaluation of alternative portfolio-selection models" *Journal of Business* 40, 166-193.

Cover, T., and D. Julian (2000) "Performance of universal portfolios in the stock market" *Information Theory Proceedings*. IEEE International Symposium on Information Theory – Sorrento, Italy

Gregoriou, G. N. (2003) **"**Performance appraisal of hedge funds using data envelopment analysis" *Journal of Wealth Management*, 88-95.

Haslem, J. A., and C. A. Scheraga (2003) "Data envelopment analysis of Morningstar's large-cap mutual funds" *Journal of Investing* Winter, 41-48.

Markowitz, H. (1952) "Portfolio selection" *Journal of Finance* 7, 77-91.

Merton, R. (1969) "Lifetime portfolio selection under uncertainty: The continuous-time case" *Review of Economics and Statistics* 51, 247-257.

Morey, M. R., and R. C. Morey (1999) "Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking" *Omega* 27, 241-258.

Poon, S. H., and C. W. J. Granger (2003) "Forecasting volatility in financial markets: a review" *Journal of Economic Literature* 41, 478-539.

Powers, J., and P. R. McMullen (2000) “Using data envelopment analysis to select efficient large market cap securities” *Journal of Business and Management* 7, 31- 42.

Samuelson, P. A. (1969) “Lifetime portfolio selection by dynamic stochastic programming” *Review of Economics and Statistics* 51, 239-246.

Shen, Y., K. Eggleston, L. Lau, and C. Schmid (2005) “Hospital ownership and financial performance: a quantitative research review” NBER working paper 11662.

Zhu, J. (2003) *Quantitative models for performance evaluation and benchmarking*, Massachusetts: Kluwer.