Portfolio optimization using DEA models in Brazilian stock market

May 29, 2017

Abstract: The main objective of this study is to monitor the performance of investment portfolios using the Data Envelopment Analysis (DEA) model as a classifier of possible assets to be chosen to make set the portfolios. Regarding to the specific objectives, the proposal was to comparatively assess the performance of the DEA portfolios integrated with the Markowitz optimization model, with measure of risk variation (variance and semi-variance). Besides comparing to the performance of structured portfolios with direct application of the Markowitz model on the sample, it involves varying the measure of risk as well. The data sample was based on assets of the Bovespa index which presented transactions on all trading days of the São Paulo Stock Exchange between July 1, 2013 and November 30, 2016. The results have shown that using the DEA technique for efficient asset classification, before to apply the investment portfolio technique, result in profitability superior to strategies that are based simply on the model of investment diversification proposed by Markowitz, with emphasis on the modeling that used semi-variance as a measure of risk. In addition, the DEA model has formed portfolios with smaller cardinalities than the portfolios that solely applied the Markowitz model, thus incurring lower transaction costs.

Keywords: Data envelopment analysis, Portfolio selection, Investment, Bovespa index.

1 Introduction

The investment diversification remains as a discussion point among academical community and market professionals. The decision of how to invest, which assets to choose and in which quantity is crucial to reach the investor's utility maximization objective. Nevertheless, there are many variables involved in this decision and, given the cognitive capacity of a single person, the task of processing whole information and decide for an optimal final solution becomes unfeasible.

In face of this need, many seeks in computational resources and mathematical models a path to improve their analysis ability and find better results. The repercussion of this pursuit can be observed in financial literature, which several empirical tests and simulations make use of technological advances to inquire the best strategy to invest, given the existence of a huge information volume available.

Currently, the portfolio optimization model proposed by Harry Markowitz in his seminal paper published in 1952, stands as the most popular approach in terms of portfolio optimization. Even after half century since its publication, the model still is issue of discussion and relevance, either in practical application or theoretical discussions. Over time, a lot of works made contributions, proposed new concepts and increment the studies about portfolio theory.

Regarding this seek, a pertinent approach to evaluate relative efficiency among a group elements is the data envelopment analysis (DEA), proposed by Charnes, Cooper, and Rhodes (1978), is frequently related in academical studies due to its versatility in several applications. In financial field, some papers used this methodology considering diverse specifications, inputs and outputs, being these discussed in detail ahead.

In line with those researches, our work tested the efficiency of portfolios composed using DEA models in Brazilian stock market. The optimization process was made in two steps: first of all, the efficient stocks were selected using DEA models. Second, the portfolios were constructed using pertinent approaches related in literature. As a additional contribution, we proposed an alteration in input data set for DEA model, alternative to risk measure used in present literature.

The present paper is organized as follow: In section 2 we provide a theoretical framework, in section 3 we present the methodological issues, where we discuss the data collect strategy and basic data description. In section 4 we discuss the results and in section 5 we present the final considerations.

2 Theoretical framework

2.1 Markowitz theory for portfolio selection

Markowitz (1952) seminal paper, modern financial theory forerunner, was presented a mathematical proposal to deal with trade off involving return maximization and risk minimization in an investment. According to Santos and Tessari (2012), the idea core is the investor to make a decision between risk and expected return to determinate the best resource allocation in his portfolio, picking lower risk among equal return portfolios, and analogously, picking a higher return among equal risk portfolios.

This way, the model can be expressed by a multiobjective formulation, which will seek a local optimum with intention of minimize risk and maximize expected return. This set of optimal solutions is called efficient frontier. The equations 1 to 4 present the model in a formal way:

$$\min_{w_i,\dots,w_n} \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \tag{1}$$

$$\max_{w_i,\dots,w_n} \sum_{i=1}^{n} w_i \mu_i \tag{2}$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1 \tag{3}$$

$$0 \leq w_i \leq 1, \qquad \forall i = 1, ..., n \tag{4}$$

Where:
w_i = investment weight for asset i;
w_j = investment weight for asset j;
σ_{ij} = covariance between assets i and j;
μ_i = expected return for asset i;

The same set of optimal solutions can be enhanced using a mono objective formulation. For this, it is introduced in model a variable to express the risk aversion of the investor, factor that will describe the behavior of the investor in face of investment risk options. The equations 5 to 7 present this formulation:

$$\min_{w_i,\dots,w_n} \lambda \left[\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}\right] - (1-\lambda)\left[\sum_{i=1}^{n} w_i \mu_i\right] \tag{5}$$

Subject to:

$$\sum_{i=1}^{n} w_i = 1 \tag{6}$$

$$0 \leq w_i \leq 1, \qquad \forall i = 1, ..., n \tag{7}$$

Where:
λ = risk aversion coefficient;
w_i = investment weight for asset i;
w_j = investment weight for asset j;
σ_{ij} = covariance between assets i and j;
μ_i = expected return for asset i;

Thus, if one hand the solution proposed by Markowitz provides the called efficient frontier, on the other hand, it does not indicate a specific point to be picked. Markowitz (1952) affirms that the values of mean, variance and covariance can be estimated combining statistical techniques and the judgment of the analyst. Henceforth, a set of mean-variance combinations can be derived and presented to the investor to associate the return-risk aimed. So, it is up to the investor and his characteristics as risk willing to decide for a specific point in set of feasible solutions.

A simple way to define this point is a choice for the best value of expected return per risk unit taken. This approach was proposed by Sharpe (1966), and it is called Sharpe Ratio (SR). Thus, the choice in this proposition will be done in a point of higher return per unit of risk, assuming the existence of a risk free investment. The model for this operation is presented in equation 8:

$$\max SR = \frac{R_P - R_f}{\sigma_P} \tag{8}$$

Where:
R_P = portfolio expected return;
R_f = risk free operation;
σ_P = portfolio standard deviation.

Over time, several works[1] attempt to confirm or misrepresent Markowitz ideas, adding new concepts and evolving the portfolio theory discussion. It stands out the downside risk models, which enable by actual computational resources, constitute the post modern portfolio theory.

2.1.1 Alternative risk measure: semideviation

As discussed before, the mean-variance (MV) model proposed by Markovitz is based on the investor conflicting decision who desires maximize expected return and minimize the risk of an operation, this risk is determined by a measure of return dispersion. The most popular measure in literature is variance, since Markowitz (1952) and Markowitz (1959) indicated its use due to a more suitable computational cost.

According to Markowitz (1959), the definition of which risk measure adopted in portfolio analysis will depend on the distribution return format. Whether these present a symmetrical format, or all assets present the same deviation degree, the author suggests the use of variance as risk measure. However, whether the distribution return presents an asymmetrical format or the assets present different deviation degree among then, Markowitz (1959) proposes the use of a downside risk measure as alternative to deal with variance limitations, and indicates the semi deviation as an appropriate measure in those cases.

Nawrocki (1999) points out that this proposition comes to light on the work of Roy (1952), who affirmed that the investor will prefer safety of principal at first and will set some minimum acceptable return that will conserve the principal. The idea core is that the investor perceive that the risk is given in function of variation in losses domain, which should be avoided in an investment.

Whereas, Markowitz suggests that the calculation of the semi variance be done by computing only deviations bellow the mean of the return. According to Markowitz (1959), a variance based analysis focuses in both tails of distribution, preventing of big losses and big gains. Using a semi variance analysis, only the loss tail will considered, which improve the results. Markowitz also demonstrated that when the return distribution is normal, both measures, variance and semi variance, achieve good results. Yet, when the return distribution is not normal, the use of semi variance presents superior results.

The semi variance can be calculated using equations 9 and 10:

$$S^2 = E\left\{\min\left[\left(R_P - T\right), 0\right]^2\right\} \tag{9}$$

[1] Fabozzi, Gupta, and Markowitz, 2002; Tu and Zhou, 2011; DeMiguel, Garlappi, and Uppal, 2009; Kroll, Levy, and Markowitz, 1984; Duchin and Levy, 2009; Michaud, 1989; Statman, 1987; Uysal, Trainer Jr, and Reiss, 2001; Markowitz, 1976, 2010.

$$(R_P - T) = \begin{cases} (R_P - T) & \text{if} \quad (R_P - T) \leq 0 \\ 0 & \text{if} \quad (R_P - T) > 0 \end{cases} \tag{10}$$

Onde:
S^2 = semi variance;
R_P = portfolio standard expected return;
T = threshold point.

2.2 Data envelopment analysis

The Data Envelopment Analysis (DEA) was proposed for the first time by Charnes et al. (1978). It is a non parametric technique to evaluate relative performance, that works through individual optimization from observations from a data set, with the aim of finding an efficient frontier, and that determines the efficiency of each observation regarding this frontier.

In DEA models, each observation is called by the generic term Decision Making Unit (DMU). The model basic assumption is that DMUs have a similar set of inputs to produce a similar set of outputs, which makes the approach flexible, when compared to parametric models, which normally demands a comply of several assumptions (Cooper, Seiford, & Tone, 2005; Charnes, Cooper, Lewin, & Seiford, 1994).

The relative performance analysis output is a performance index which range between zero and one, indicating the degree of efficiency from each DMU regarding the others elements from data set. This result enables the comparison from various inefficiency sources for each input and output for each DMU. Thus, in DEA model conception, "the performance of a DMU is efficient if, and only if, it is not possible to improve any input or output without worsening any other input or output" (Cooper et al., 2005, p.45). Next topic will discuss the model BCC, specific model which will be applied in this work.

2.2.1 BCC model

The BCC model, also known as Variable Returns to Scale (VRS), was proposed by Banker, Charnes, and Cooper (1984) as an extension of CCR model (Charnes et al., 1978). The main difference between those two models is the inclusion of a variable free in signal, which represents the variable returns to scale. The BCC model aimed to measure efficiency of DMUs regarding an efficient frontier in which the returns to scale are considered as variables. According to Cooper et al. (2005), the model is presented in equations 11 to 14:

$$\max_{u,v,u_k} \quad uy_0 - u_k \tag{11}$$

Subject to:

$$vx_0 = 1 \tag{12}$$

$$-vX + uY - u_k \iota \leq 0 \tag{13}$$

$$v \geq 0, \quad u \geq 0, \quad u_k \quad free \quad in \quad signal \tag{14}$$

Where:
u = vector of outputs weights;
y_0 = output of DMU under analysis;
Y = matrix of outputs for all DMUs;
v = vector of inputs weights;
x_0 = input of DMU under analysis;
X = = matrix of inputs for all DMUs;
u_k = variable returns to scale;
ι = vector of ones;

2.3 Related work

Table 1 presents a resume of main works in literature which are related to our paper. Due to DEA methodology versatility, those works present diversified approaches in reference of sampling strategy, data periodicity and DEA model inputs. It is noted that three out of eight papers related used DEA models to composed portfolios without posterior optimization.

In general, all works present good results, regarding the specifications of each study. Considering Brazilian stock market, Rotela Junior, Pamplona, and Salomon (2014) and Lopes, Carneiro, Schneider, and Lima (2011) used DEA methodology for stock selection, and then optimized the portfolios using mean-variance model.

The two studies considered a 3 years sample and composed monthly portfolios, being first the sample of the study of Lopes et al. (2011) broader than of the Rotela Junior et al. (2014). Both present good results regarding the specifications of each study, which show the potential of the approach to the applications of the investments.

Table 1: Related work

Reference	Sample	Opt	Results
Gardijan and Škrinjarić (2015)	It was considered 41 stocks listed in Zagreb stock exchange between Jan'17 and Jun'14.	N	Simulations using CCR and BCC models outperformed baselines. It stood out the BCC models among others.
Rotela Junior, Pamplona, and Salomon (2014)	It was considered 40 stocks from Ibovespa index between Apr'10 and Mar'13.	Y	The portfolio optimized using DEA + MV outperformed (31,48%p.y.) the optimization without DEA (22,20%p.y.) and Ibovespa index (-7,86%p.y.)
Lopes, Carneiro, Schneider, and Lima (2011)	The initial sample has 732 stocks from Ibovespa index from Jan'06 to Dez'08.	Y	The portfolio optimized using DEA + MV outperformed (0,58%p.m.) simple DEA portfolio (0,35%p.m.) and Ibovespa index (0,53%p.m.)
Edirisinghe and Zhang (2010)	It was considered 827 stocks from 9 industries from S&P500 index between 1997 and 2004.	Y	The study proposed the optimization of inputs and outputs according to company's industry. The DEA model selected 89 from 827 stocks and outperformed the baseline in all periods.
Edirisinghe and Zhang (2008)	It was considered 313 stocks from 6 industries from S&P500 index in 2002.	Y	It was made comparisons considering three investment horizons. The DEA models outperformed baselines in all simulations.
Chen (2008)	It was considered stocks from 8 major industries from Taiwan stock exchange between 2002 and 2004.	N	Simulations using CCR and BCC models outperformed baselines in all industries in terms of return and SR.
Lopes, Lanzer, Lima, and Costa Junior (2008)	It was considered stocks from IBr-X-100 index in the beginning of all quarters from Jan'01 to Jan'06.	N	The portfolio optimized using DEA outperformed (12,33%p.q.) IBrX-100 index (7,03%p.q.)
Edirisinghe and Zhang (2007)	It was considered 230 stocks from 6 industries from S&P500 index from Mar'96 to Dec'02.	Y	The DEA model selected 85 from 230 stocks and outperformed the baseline in all periods, considering 6 different levels of risk aversion.

3 Methodology

3.1 Data

To design this study, we used stocks from Ibovespa index in 31/01/2017. We selected only stocks which had daily trading between 01/07/2013 and 30/11/2016. After applying this sampling criterion, 47 companies, which represent 52 stocks negotiated in Bovespa remained in our sample. We present those companies in Table 2, with respective industries and Market Cap [2].

Table 2: Sample: company, industry and Market Cap

Company	Industry	Market Cap	Company	Industry	Market Cap
Ambev	Cons. Staples	270.3	Gerdau	Materials	19.7
Banco do Brasil	Financials	87.9	Hypermarcas	Health Care	17.8
Banco Bradesco	Financials	176.4	Itau Unibanco	Financials	69.1
Bradespar	Materials	6.5	JBS	Cons. Staples	35.1
Brasil Foods	Cons. Staples	35.6	Kroton	Cons. Discr.	21.7
Braskem	Materials	24.9	Lojas Americanas	Cons. Discr.	22.3
BR Malls	Real Estate	9.0	Lojas Renner	Cons. Discr.	15.6
BM&F Bovespa	Financials	34.0	Marfrig	Cons. Staples	4.3
CCR	Industrials	28.5	MRV	Cons. Discr.	6.0
Cielo	Info. Tech.	61.0	Multiplan	Real Estate	12.8
CEMIG	Utilities	12.1	Natura	Cons. Staples	11.2
CPFL	Utilities	25.8	Pão de Açúcar	Cons. Staples	20.7
COPEL	Utilities	7.6	Petrobras	Energy	205
COSAN	Energy	17.4	Raia Drogasil	Cons. Staples	20.7
CSN	Materials	15.7	Localiza	Industrials	7.8
CETIP	Financials	12.4	SABESP	Utilities	22.2
Cyrela	Cons. Discr.	5.5	Suzano	Materials	14.2
Ecorodovias	Industrials	4.9	Tim	Telec. Services	22.8
Engie	Utilities	23.7	Ultrapar	Energy	37.0
Embraer	Industrials	13.3	Usiminas	Materials	8.7
EDP	Utilities	8.5	Vale	Materials	156.7
Equatorial	Utilities	11.3	Vivo	Telec. Services	75.5
Estácio	Cons. Discr.	4.9	Weg	Industrials	26.3
Fibria	Materials	15.0	**TOTAL**		**1765.4**

The daily data used were adjusted for dividends, split and insplit. We collected data using Bloomberg terminal and made our analysis using: Microsoft Excel, R and SPSS.

3.2 Experiment design

Our propose in this work was: (1) use DEA models to pick efficient stocks in our sample, and then (2) optimize portfolios. Those steps are presented in Figure 1:

We use BCC methodology to classify the efficient stocks. The DEA model inputs were the semi variances for 12, 24 and 36 months and Price-to-Earnings index (PE). The DEA model outputs were the mean return for 12, 24 and 36 months and Earnings per Share index (EPS). We chose these variables based on previous works (Rotela Junior et al., 2014; Gardijan & Škrinjarić, 2015; Chen, 2008), and we contributed with field literature introducing the semi variance as measure for risk in DEA model. We calculated

[2] Market Cap in R$/Bi, values in 31/07/2017.

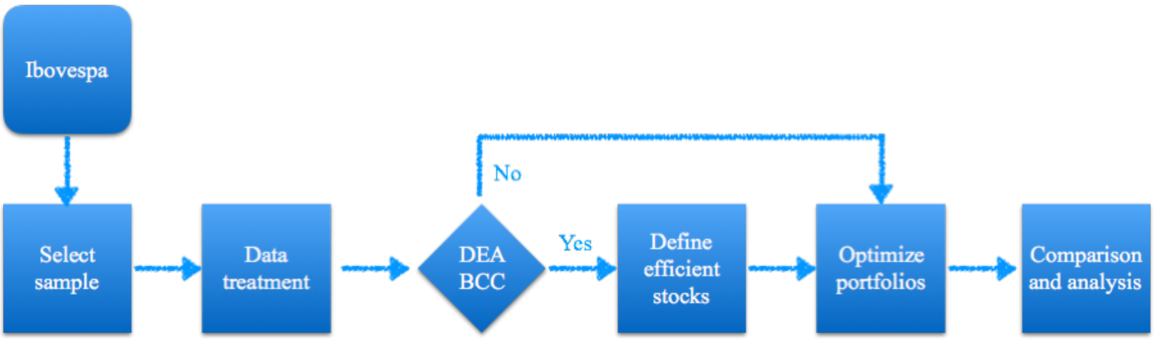

Figure 1: Research flow

the metrics using returns daily series and made the optimization using software R. Our experiment considered as efficient only stocks which achieved an index equal to 1.

Then, we calculated the efficient frontier using the assets classified as efficient using DEA model. Once drawn the frontier, we chose the maximum SR point to evaluate the performance of the experiment. Moreover, the choice of risk measure to optimize portfolios was made according to risk measure used as input in DEA model. For calculate SR index we considered interbank deposit certificate (CDI) as risk free rate. Was stipulated constraints for full investment and short sales were not allowed.

In order to evaluate and performance compare, we built portfolios without previous DEA classification, portfolios without optimization, using DEA efficiency as buy signals using naive weights (1/N), and a portfolios base on Ibovespa index (IBOV). For better understanding, we present in Table 3 a resume of our strategies, its criterion and codes to be used henceforward.

Table 3: Strategies

Code	DEA		MV Model	
	Inputs	Outputs	Target	Risk
MaxSR-DEA-SV	SV (12/24/36) PE	Ret. (12/24/36) EPS	MaxSR	SV
MaxSR-SV	NA NA	NA NA	MaxSR	SV
N-DEA-SV	SV (12/24/36) PE	Ret. (12/24/36) EPS	NA	NA
MaxSR-DEA-Var	Variance (12/24/36) PE	Ret. (12/24/36) EPS	MaxSR	Variance
MaxSR-Var	NA NA	NA NA	MaxSR	Variance
N-DEA-Var	Variance (12/24/36) PE	Ret. (12/24/36) EPS	NA	NA
IBOV	NA	NA	NA	NA

For all strategies we built 36 portfolios, and operated considering buy at opening price of current of first business day of the month and sell at opening price of the following first business day of the month. We unconsidered taxes and transactions costs in all scenarios.

4 Discussion

4.1 Historical data

Using basic data analysis we can verified a higher degree of heterogeneity in our sample, regarding information of closing prices and returns. Besides, we perceived more volatility in return considering data second half, starting in a drop followed by a growth resumption. Some tickers, as Cetip, Equatorial and Raia Drogasil, presented regular growth and return levels during whole period. It is noted that Bradespar, Gerdau, Usiminas and Vale exhibited downward trend during our analysis, with a higher volatility in second half. Figure 3 in appendix displays data behavior in a graphical presentation.

Still on this subject, concerning descriptive statistics, the tickers of Bradespar, CCR, Gerdau, Petrobras, Usiminas and Vale exhibited high standard deviation and range scores, clue to confirm our conclusions above. Further, we found range scores higher than mean scores, associated with a downward behavior observed in graphical presentation, which indicated a devaluation of those stocks in our analysis period. Table 8 in appendix presents the complete data description.

4.2 DEA optimization and portfolio's cardinality

Concerning the DEA model application, we considered as efficient only assets which reached a index equal to 1. The Table 4 presents the cardinality for each month in each strategy. We found a median cardinality equal to 12.5 (N-DEA-SV strategy), achieving the minimum in 14^{th} and 30^{th} months, with 6 assets. The month with highest cardinality was the 3^{rd}, with 19 assets.

We observed a stock reduction when we compared strategies which used simple optimization and strategies which DEA models preceded optimization process, reducing the median from 8 stocks to 3 stocks. This reduction is a highlight point in our analysis, whereas our experiment was design free of tax and costs. The fewer stocks were hold in a portfolio, lower will be the operation costs. Thus, keeping the levels of return and risk, a portfolio with fewer stocks will be preferred over a portfolio which has multiple elements.

Table 4: **Portfolio's cardinality**

Month	MaxSR-DEA-SV	MaxSR-SV	N-DEA-Var	MaxSR-DEA-Var	MaxSR-Var	N-DEA-Var
1	4	10	15	5	9	14
2	2	7	7	2	7	8
3	4	7	19	4	7	18
4	2	6	8	2	6	8
5	2	5	17	3	5	17
6	3	6	9	3	6	9
7	3	6	18	4	6	13
8	4	7	9	3	8	13
9	7	10	16	8	10	16
10	4	10	8	5	10	7
11	2	7	18	2	7	18
12	2	12	11	3	11	11
13	4	7	13	4	7	14
14	1	5	6	1	6	6
15	2	7	15	2	7	15
16	3	8	10	3	8	13
17	1	8	12	1	8	12
18	2	9	7	2	10	7
19	3	9	15	3	9	14
20	6	12	12	6	12	12
21	3	8	16	2	8	16
22	4	9	9	3	9	9
23	3	7	16	3	7	16
24	2	9	11	3	9	12
25	2	6	17	3	7	16
26	3	8	12	3	8	12
27	2	9	18	2	9	16
28	4	6	13	4	7	13
29	5	7	14	5	7	15
30	2	7	6	2	7	5
31	4	5	17	4	5	15
32	4	8	10	3	8	11
33	6	16	11	6	14	12
34	2	9	8	2	10	8
35	6	9	17	6	7	17
36	5	8	13	5	10	13
Median	**3.0**	**8.0**	**12.5**	**3.0**	**8.0**	**13.0**

Table 5 shows the efficiency by stock. The median analysis indicated that the assets were classified as efficient in 9 of 36 available periods by DEA model. The stocks BRAP4 and ESTC3 were the more frequent tickers, 14 times, and the stocks GOAU4 and SUZB5 were the less frequent ticker, classified as efficient only 4 times.

Table 5: Efficient stocks

Asset	N-DEA-SV	N-DEA-Var	Asset	N-DEA-SV	N-DEA-Var
ABEV3	7	7	GOAU4	6	4
BBAS3	9	11	HYPE3	5	6
BBDC3	9	9	ITSA4	8	8
BBDC4	5	6	ITUB4	4	5
BRAP4	12	14	JBSS3	9	9
BRFS3	9	9	KROT3	10	10
BRKM5	8	9	LAME4	9	9
BRML3	10	11	LREN3	11	12
BVMF3	10	10	MRFG3	8	8
CCRO3	7	6	MRVE3	8	7
CIEL3	9	9	MULT3	8	8
CMIG4	9	8	NATU3	7	6
CPFE3	8	6	PCAR4	5	6
CPLE6	12	11	PETR3	9	11
CSAN3	12	12	PETR4	9	11
CSNA3	13	13	RADL3	11	10
CTIP3	8	7	RENT3	10	11
CYRE3	11	9	SBSP3	10	10
ECOR3	9	6	SUZB5	4	4
EGIE3	7	7	TIMP3	8	9
EMBR3	10	8	UGPA3	6	6
ENBR3	11	10	USIM5	5	6
EQTL3	11	9	VALE3	6	7
ESTC3	14	14	VALE5	9	9
FIBR3	12	12	VIVT4	8	9
GGBR4	11	10	WEGE3	7	7
			Median	**9**	**9**

4.3 Models performance

Henceforth, we follow with performance models analysis, starting in Table 6, regarding the basic descriptions of the returns. MaxSR-DEA-SV achieved highest mean return and lowest variation coefficient among strategies, outperforming nearly four times IBOV, which presented the worst scores in that comparison. We noted an improvement in terms of mean and dispersion when we compared strategies with and without DEA classification, a trace of efficiency in process of selecting best stocks before dealing with portfolio optimization.

Table 6: Descriptive statistics

Portfolio	Mean	SD	CV	Median	Asym.	Kurt.
MaxSR-DEA-SV	2,87	8,1	2,82	2,27	-0,05	0,24
MaxSR-SV	0,74	6,65	9,03	2,24	-0,31	-1,24
N-DEA-SV	1,13	7,3	6,47	0,9	0,38	-0,89
MaxSR-DEA-Var	2,55	7,37	2,89	2,43	-0,2	-0,16
MaxSR-Var	0,94	6,54	6,95	1,97	-0,33	-1,05
N-DEA-Var	1,05	7,69	7,33	0,76	0,19	-1,14
IBOV	0,67	6,95	10,44	0,39	0,3	-0,54

Table 7 presents specific performance indicators. According to strategies accuracy, the Ratio indicator represents the proportion between the average result in gain moments and the average result in loss moments. That is, a Ratio of 0.50 indicates that one loss operation consumes in average 50% of one gain operation profits. Thus, we look for scores lower than 1 and as closer as possible to 0. In our analysis, N-DEA-Var attained the best Ratio, and MaxSR-Var the worst. We noted that strategies which used DEA models and optimization achieved better results in terms of Ratio, though they did not outperformed N-DEA-SV, N-DEA-Var and IBOV strategies.

Table 7: General report

Portfolio	Card.	Acur.	Ratio	Return %				
				2014	2015	2016	Total	Mean
MaxSR-DEA-SV	3,3	72,22	1,05	11,77	44,22	47,43	103,42	26,71
MaxSR-SV	8,0	58,33	1,09	13,86	24,11	-11,45	26,51	8,15
N-DEA-SV	12,6	52,78	0,77	0,97	-7,32	46,99	40,65	12,04
MaxSR-DEA-Var	3,4	69,44	0,96	16,32	41,86	33,52	91,69	24,22
MaxSR-Var	8,1	61,11	1,13	14,78	25,21	-6,14	33,85	10,21
N-DEA-Var	12,5	52,78	0,81	7,65	-7,62	37,78	37,81	11,28
IBOV		52,78	0,88	-0,63	-12,34	36,94	23,97	7,42

Regarding cumulative returns per year, in 2014 neither strategy achieved a highlight performance, hinted by closer graph curves represented in Figure 2. The best strategy in this year was MaxSR-DEA-Var, which achieved 16.32%, followed by MaxSR-Var and MaxSR-SV, 14.78% and 13.86%, respectively.

In the beginning of 2015 we observed a split, which naive strategies (N-DEA-SV, N-DEA-Var and IBOV) achieved a bottom performance, strategies without DEA models remained in a middle position and strategies using DEA models and optimization (MaxSR-DEA-SV and MaxSR-DEA-Var) outperformed the others, scenario which remained during 2015 and 2016. In this period, MaxSR-DEA-SV reached the best result, achieving 44.22% and 47.43%, respectively. That strategy attained a overall return of 103.42%, equivalent to an average return of 26.71% per year.

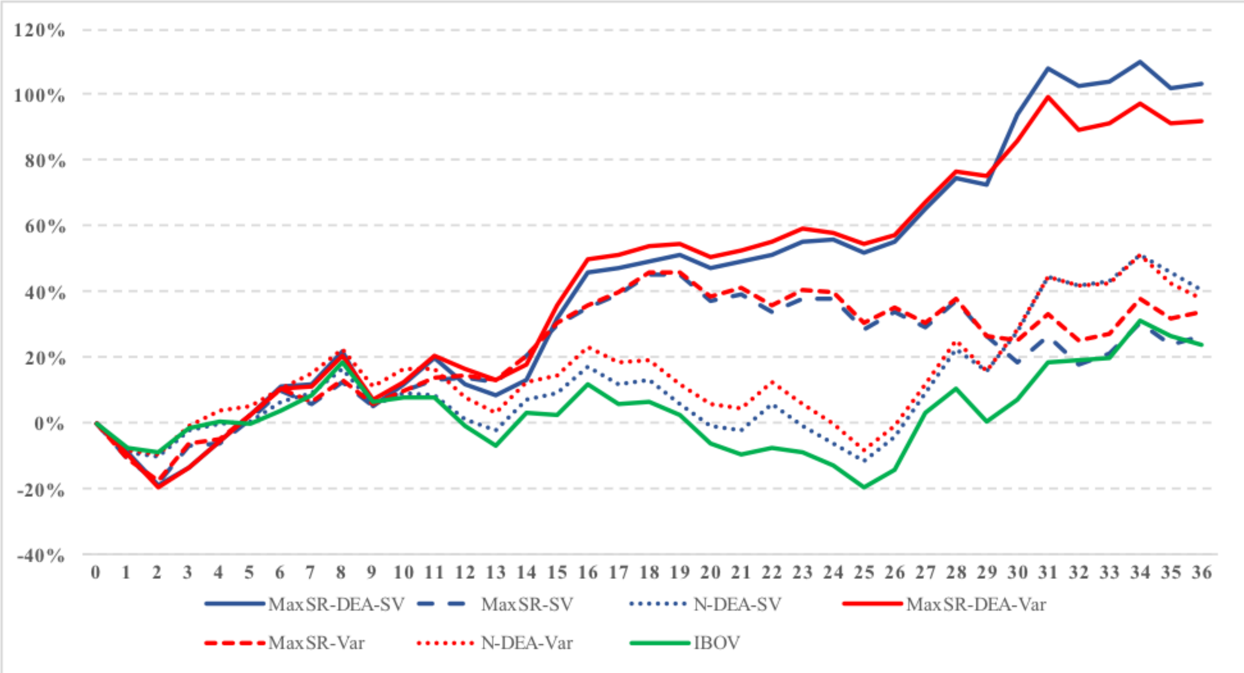

Figure 2: Cumulative return by strategy

5 Final considerations

In our work we tested the efficiency of portfolios composed using DEA models in Brazilian stock market. We conducted simulations which combined DEA models and Markowitz model to optimize investments. The data comprehended a 3 and a half years period in a sample composed by 52 stocks from Ibovespa index. They were composed 36 monthly portfolios for each strategy stipulated. The following strategies were applied and compared: (1) DEA + Mean-semi variance (maximum Sharpe Ratio); (2) Mean-semi variance (maximum Sharpe Ratio); (3) DEA semi variance (1/N allocation); (4) DEA + Mean-variance (maximum Sharpe Ratio); (5) Mean-variance (maximum Sharpe Ratio); (6) DEA variance (1/N allocation); (7) buy-and-hold Ibovespa.

Our findings corroborated previous works results (Edirisinghe & Zhang, 2007; Chen, 2008; Lopes, Lanzer, Lima, & Costa Junior, 2008; Edirisinghe & Zhang, 2008; Edirisinghe & Zhang, 2010; Lopes et al., 2011; Rotela Junior et al., 2014; Gardijan & Škrinjarić, 2015), indicating that the use of DEA models to define efficient stocks as a preliminary stage of investment optimization process outperform a simple optimization process, with no previous classification criterion.

Furthermore, our work contributes with the discussions of the financial studies using semi variance as an input for DEA models and as a risk measure for portfolio optimization. Our implementation using this configuration outperformed portfolios using mean-variance model in terms of return and cardinality.

As a proposal for future studies, we suggest the use of others downside risk measures. We also suggest the testing of other sets of inputs and outputs for DEA models and the use of specific indicators for companies on common industries in Brazilian stock market.

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, *30*(9), 1078–1092.

Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (1994). *Introduction. in data envelopment analysis: Theory, methodology, and applications*. New York: Springer.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, *2*(6), 429–444.

Chen, H.-H. (2008). Stock selection using data envelopment analysis. *Industrial Management & Data Systems*, *108*(9), 1255–1268.

Cooper, W. W., Seiford, L. M., & Tone, K. (2005). *Introduction to data envelopment analysis and its uses* (2nd ed.). Harlow: Pearson Education.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953.

Duchin, R. & Levy, H. (2009). Markowitz versus the talmudic portfolio diversification strategies. *The Journal of Portfolio Management*, *35*(2), 71–74.

Edirisinghe, N. C. P. & Zhang, X. (2007). Generalized dea model of fundamental analysis and its application to portfolio optimization. *Journal of Banking & Finance*, *31*(11), 3311–3335.

Edirisinghe, N. C. P. & Zhang, X. (2008). Portfolio selection under dea-based relative financial strength indicators: Case of us industries. *Journal of the Operational Research Society*, *59*(6), 842–856.

Edirisinghe, N. C. P. & Zhang, X. (2010). Input/output selection in dea under expert information, with application to financial markets. *European journal of operational research*, *207*(3), 1669–1678.

Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The Journal of Investing*, *11*(3), 7–22.

Gardijan, M. & Škrinjarić, T. (2015). Equity portfolio optimization: A dea based methodology applied to the zagreb stock exchange. *Croatian Operational Research Review*, *6*(2), 405–417.

Kroll, Y., Levy, H., & Markowitz, H. M. (1984). Mean-variance versus direct utility maximization. *The Journal of Finance*, *39*(1), 47–61.

Lopes, A. L. M., Carneiro, M. L., Schneider, A. B., & Lima, M. V. A. (2011). Markowitz na otimização de carteiras selecionadas por data envelopment analysis–dea. *Gestão e Sociedade*, *4*(9), 640–656.

Lopes, A. L. M., Lanzer, E., Lima, M. V. A., & Costa Junior, N. d. (2008). Dea investment strategy in the brazilian stock market. *economics bulletin*, *13*(2), 1–10.

Markowitz, H. M. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77–91.

Markowitz, H. M. (1959). *Portfolio selection, efficent diversification of investments*. New York: J. Wiley.

Markowitz, H. M. (1976). Markowitz revisited. *Financial Analysts Journal*, 47–52.

Markowitz, H. M. (2010). Portfolio theory: As i still see it. *Annual Review of Financial Economics*, *2*(1), 1–23.

Michaud, R. O. (1989). The markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, *45*(1), 31.

Nawrocki, D. N. (1999). A brief history of downside risk measures. *The Journal of Investing*, *8*(3), 9–25.

Rotela Junior, P., Pamplona, E. d. O., & Salomon, F. L. R. (2014). Otimização de portfólios: Análise de eficiência. *Revista de Administração de Empresas*, *54*(4), 405.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica*, *20*(3), 431–449.

Santos, A. A. P. & Tessari, C. (2012). Técnicas quantitativas de otimização de carteiras aplicadas ao mercado de ações brasileiro (quantitative portfolio optimization techniques applied to the brazilian stock market). *Revista Brasileira de Finanças*, *10*(3), 369.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, *39*(1), 119–138.

Statman, M. (1987). How many stocks make a diversified portfolio? *Journal of Financial and Quantitative Analysis*, *22*(03), 353–363.

Tu, J. & Zhou, G. (2011). Markowitz meets talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, *99*(1), 204–215.

Uysal, E., Trainer Jr, F. H., & Reiss, J. A. (2001). Revisiting mean-variance optimization. *The Journal of Portfolio Management*, *27*(4), 71–81.

A Appendix

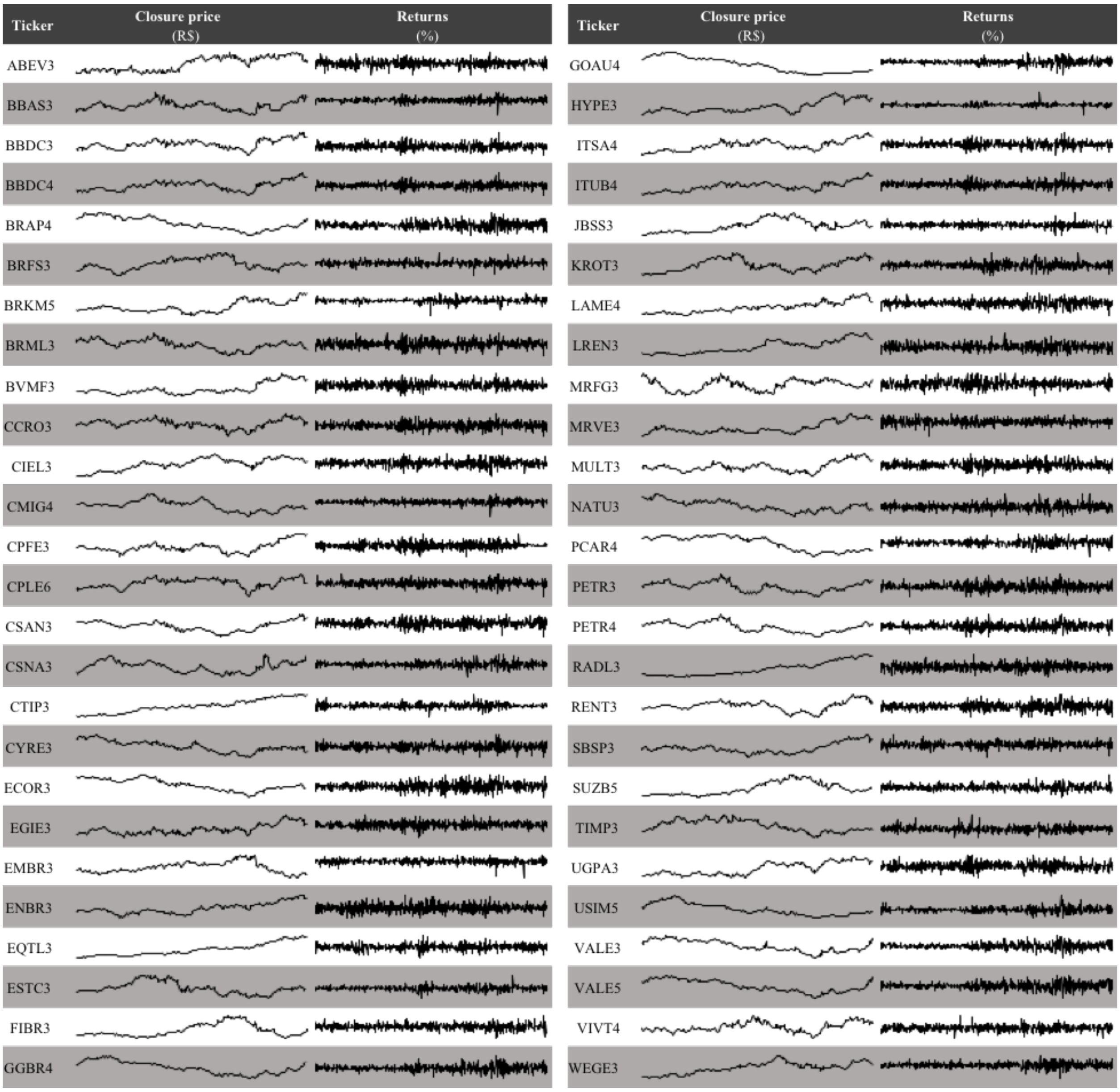

Figure 3: Historical prices behavior

Table 8: Historical data - descriptive statistics

Ticker	Mean	Range	Min	Max	SD	Asym.	Kurt.
ABEV3	0.0002	0.1132	-0.0566	0.0566	0.0138	-0.061	1.2403
BBAS3	0.0006	0.3722	-0.2379	0.1343	0.031	-0.1929	5.0266
BBDC3	0.0004	0.1864	-0.0724	0.1141	0.0214	0.3726	1.5637
BBDC4	0.0005	0.2159	-0.0934	0.1225	0.0223	0.2547	1.9972
BRAP4	-0.0001	0.2384	-0.1089	0.1295	0.0315	0.1101	0.9758
BRFS3	0.0002	0.1865	-0.1005	0.0861	0.0167	-0.0461	3.5495
BRKM5	0.0008	0.3378	-0.2204	0.1174	0.0262	-0.6753	8.2724
BRML3	-0.0004	0.1655	-0.0855	0.08	0.0236	0.0731	0.8947
BVMF3	0.0005	0.1823	-0.0859	0.0964	0.0228	0.1896	1.1218
CCRO3	0.0000	0.1815	-0.0858	0.0958	0.0226	-0.1157	0.8837
CIEL3	0.0006	0.1444	-0.0823	0.0621	0.0178	-0.1636	1.4013
CMIG4	-0.0004	0.3738	-0.2364	0.1374	0.0294	-0.4204	5.9096
CPFE3	0.0004	0.1653	-0.0794	0.0859	0.0204	0.0751	1.1648
CPLE6	0.0003	0.225	-0.1319	0.0931	0.0249	-0.1416	1.7738
CSAN3	0.0002	0.1628	-0.0953	0.0674	0.0214	-0.0542	0.8754
CSNA3	0.0011	0.417	-0.2295	0.1875	0.0423	0.2277	2.8272
CTIP3	0.001	0.1593	-0.0791	0.0803	0.0136	0.0922	3.8608
CYRE3	-0.0005	0.1809	-0.0843	0.0966	0.0209	0.0558	1.8843
ECOR3	-0.0006	0.1711	-0.0831	0.088	0.0245	0.1181	1.1154
EGIE3	0.0002	0.1437	-0.0799	0.0638	0.0159	-0.0575	1.7433
EMBR3	-0.0002	0.2341	-0.1678	0.0663	0.0205	-1.1974	8.6402
ENBR3	0.0005	0.1491	-0.066	0.083	0.021	0.0694	0.8022
EQTL3	0.0013	0.1358	-0.0644	0.0714	0.0152	-0.1727	1.507
ESTC3	0.0002	0.3774	-0.1644	0.213	0.0297	0.0915	5.2716
FIBR3	0.0004	0.2182	-0.1128	0.1054	0.0239	-0.0706	1.8055
GGBR4	0.0002	0.273	-0.1238	0.1492	0.0314	0.234	1.7945
GOAU4	-0.0011	0.372	-0.2094	0.1626	0.0367	-0.0222	3.4001
HYPE3	0.0007	0.3456	-0.1538	0.1918	0.0184	0.629	19.6095
ITSA4	0.0006	0.1839	-0.086	0.0979	0.0198	0.1844	1.8032
ITUB4	0.0008	0.1949	-0.0912	0.1037	0.0208	0.2586	2.0037
JBSS3	0.0006	0.3498	-0.1591	0.1907	0.0304	0.1914	4.5067
KROT3	0.0008	0.2459	-0.1107	0.1352	0.0286	0.1033	2.1055
LAME4	0.0006	0.162	-0.0884	0.0736	0.0205	-0.1562	1.0701
LREN3	0.0007	0.1528	-0.0593	0.0934	0.0195	0.2489	0.829
MRFG3	-0.0002	0.2009	-0.1058	0.0951	0.0282	0.1229	1.1831
MRVE3	0.0008	0.2155	-0.1336	0.0819	0.0243	-0.032	1.2547
MULT3	0.0002	0.1398	-0.0694	0.0705	0.0183	0.1437	0.7665
NATU3	-0.0005	0.184	-0.0796	0.1044	0.023	0.3786	1.9011
PCAR4	-0.0007	0.1812	-0.1099	0.0713	0.02	-0.1429	2.0192
PETR3	0.0003	0.27	-0.1203	0.1497	0.0358	0.2369	0.988
PETR4	0.0001	0.2825	-0.1316	0.1509	0.036	0.1346	1.2533
RADL3	0.0013	0.1513	-0.0664	0.085	0.019	0.1169	0.5926
RENT3	0.0002	0.1522	-0.0719	0.0803	0.0211	-0.0782	1.0079
SBSP3	0.0004	0.2248	-0.1238	0.101	0.0232	-0.2533	2.0058
SUZB5	0.0006	0.2302	-0.1029	0.1273	0.0231	0.2303	2.0256
TIMP3	0.0001	0.2316	-0.089	0.1426	0.0236	0.4645	3.1216
UGPA3	0.0004	0.1127	-0.057	0.0557	0.0151	0.0802	0.8934
USIM5	-0.0007	0.4718	-0.1704	0.3014	0.043	0.6473	4.9035
VALE3	0.0001	0.2945	-0.1567	0.1378	0.0324	0.0886	1.8621
VALE5	0.0001	0.236	-0.1284	0.1075	0.03	0.0545	1.3121
VIVT4	0.0001	0.1585	-0.0717	0.0869	0.0172	0.0677	1.7133
WEGE3	0.0005	0.1641	-0.0955	0.0686	0.0162	-0.1848	2.2577