INSTITUTE
OF ECONOMIC STUDIES
Faculty of Social Sciences
Charles University

ROBUST PORTFOLIO OPTIMIZATION: A STOCHASTIC EVALUATION OF WORST-CASE SCENARIOS

Paulo Rotella Junior
Luiz Célio Souza Rocha
Rogério Santana Peruchi
Giancarlo Aquila
Karel Janda
Edson de Oliveira Pamplona

IES Working Paper 3/2022

Institute of Economic Studies,
Faculty of Social Sciences,
Charles University in Prague

[UK FSV – IES]

Opletalova 26
CZ-110 00, Prague
E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Institut ekonomických studií
Fakulta sociálních věd
Univerzita Karlova v Praze

Opletalova 26
110 00 Praha 1

E-mail : ies@fsv.cuni.cz
http://ies.fsv.cuni.cz

Bibliographic information:
Rotella Junior P., Rocha L. C. S., Peruchi R. S., Aquila G., Janda K., de Oliveira Pamplona E. (2022): " Robust Portfolio Optimization: A Stochastic Evaluation of Worst-Case Scenarios " IES Working Papers 3/2022. IES FSV. Charles University.

This paper can be downloaded at: http://ies.fsv.cuni.cz

Robust Portfolio Optimization: A Stochastic Evaluation of Worst-Case Scenarios

Paulo Rotella Junior[a,b]
Luiz Célio Souza Rocha[c]
Rogério Santana Peruchi[a]
Giancarlo Aquila[d]
Karel Janda[b,e]
Edson de Oliveira Pamplona[d]

[a]Deparment of Production Engineering, Federal University of Paraiba, João Pessoa, Brazil
[b]Institute of Economic Studies, Faculty of Social Sciences, Charles University, Prague, Czech Republic
[c]Department of Management, Instituto Federal do Norte de Minas Gerais, Almenara, Brazil
[d]Institute of Production and Management Engineering, Federal University of Itajubá, Brazil
[e]Faculty of Finance and Accounting, Prague University of Economics and Business, Prague, Czech Republic
Email (corresponding author): paulo.rotella@academico.ufpb.br

February 2022

Abstract:
This article presents a new approach for building robust portfolios based on stochastic efficiency analysis and periods of market downturn. The empirical analysis is done on assets traded on the Brazil Stock Exchange, B3 (Brasil, Bolsa, Balcão). We start with information on the assets from periods of market downturn (worst-case) and we group them using hierarchical clustering. Then we do stochastic efficiency analysis on these data using the Chance Constrained Data Envelopment Analysis (CCDEA) model. Finally, we use a classical model of capital allocation to obtain the optimal share of each asset. Our model is able to accommodate investors who exhibit different risk behaviors (from conservatives to risky investors) by

varying the level of probability in fulfilling the constraints (1-αi) of the CCDEA model. We show that the optimal portfolios constructed with the use of information from periods of market downturns perform better for the Sharpe ratio (SR) in the validation period. The combined use of these approaches, using also fundamentalist variables and information on market downturns, allows us to build robust portfolios, with higher cumulative returns in the validation period, and portfolios with lower beta values.

JEL: G11, G14, C38, C61
Keywords: Robust optimization, Stochastic evaluation, Chance Constrained DEA, Worst-case markets, Portfolios

1. INTRODUCTION

Diversification is a critical factor in reducing non-systematic risk in portfolio selection theory. There has been a growing amount of published research seeking to reconcile the benefits of diversification with investment practices (Kim et al., 2015). Portfolio selection involves allocating capital among a certain number of assets so that the investment provides a higher return while minimizing risks i.e., a risk-adjusted return that is satisfactory for investors, similar to what was proposed in Markowitz (1952), models (Leung et al., 2012).

Almost seventy years after the development of the Markowitz model (1952), this classic approach of mean-variance, which was a pioneering work, is still one of the most used models in asset allocation and management, and has given rise to many new approaches (Leung et al., 2012), that have been developed by various academics (Chen et al., 2020; Leung et al., 2012; Levy & Levy, 2014). Whether for researchers or for investors, the investment selection process remains a major challenge for financial management (Markowitz, 2014).

Among several tools for efficiency measurement e.g., conventional statistical methods, non-parametric methods and artificial intelligence methods, Data Envelopment Analysis (DEA) can effectively measure the relative efficiency of Decision Making Units (DMUs), which employ multiple inputs to produce multiple outputs (Emrouznejad & Tavana, 2014; Shi & Wang, 2020).

DEA is a non-parametric method that has been broadly used in different types of companies and organizations, helping managers from diverse areas, including the financial area (Azadi et al., 2015; Emrouznejad & Tavana, 2014; Kao, 2014). More recently, DEA continues to be used in efficiency evaluation and building portfolios (Amin & Hajjami, 2020; Choi & Min, 2017; Edirisinghe & Zhang, 2010; Lim et al., 2014; Rotela Junior et al., 2015).

Variations of classic models of Data Envelopment Analysis have since been presented. Among these models, some seek to include approximate information or uncertainty, using a DEA model with

Fuzzy coefficients (Azadi et al., 2015), or models like those proposed by Sengupta (1987), which combine Chance-Constrained Programming (CCP), from Charnes and Cooper (1963), with the Data Enveloping Analysis model (Jin et al., 2014).

The most widespread and traditional models of portfolio optimization theory, such as the reference models presented by Markowitz (1952), and Sharpe (1963), are recognized for being sensitive to small variations in inputs, and not considered robust (Kim et al., 2014; Kim et al., 2015). Consequently, researchers started to develop mathematical techniques on robust optimization. Several approaches have been used to increase the robustness of traditional models using mean-variance, and these approaches usually deal with solving max-min problems (Won & Kim, 2020; Xidonas et al., 2017). These techniques allow investors to incorporate risk into their portfolio optimization process considering estimate errors (Baltas & Yannacopoulos, 2019; Fabozzi et al., 2007, 2010; Sehgal & Mehra, 2020). Robust portfolio optimization has quickly become a widely applied approach among investors to incorporate uncertainty into their financial models (Kim et al., 2018).

According to Myers et al. (2009), the two main objectives for robustness process studies are: (*i*) ensure that the average response is as close as possible to an ideal value; (*ii*) ensure that the variance around this average is as small as possible.

Other relevant information for portfolio optimization theory was presented by Kim et al. (Kim et al., 2014, Kim et al. 2018; Kim et al., 2015). The authors state that the so-called robust models are achieved based on information coming from bear market periods. In other words, this information is more relevant than information coming from peak market periods, when seeking robustness. More recently, other researchers have analyzed market downturn conditions when proposing portfolio optimization models (Ashrafi & Thiele, 2021; Yu et al., 2019).

This study seeks to present a method for robust portfolio optimization based on the stochastic analysis of asset efficiency, using asset information taken from worst-case market scenarios. We used data from the Brazilian Stock Exchange (B3 - *Brasil, Bolsa, Balcão*).

In general terms, the presented model is the result of a combination of different mathematical techniques: Hierarchical Clustering, Chance-Constrained DEA (CCDEA), and the Sharpe approach. Hierarchical Clustering will be used to form different clusters, which contribute to diversifying the portfolio. The CCDEA model will allow us to stochastically identify the efficient assets in each group. Sharpe's approach will allow us to optimally allocate efficient assets in the portfolios. It is worth mentioning that information from worst-case market scenarios will be used, as this has been shown to be essential for building so-called robust portfolios. Finally, we will study the effect of adjusting risk criteria for meeting investor requirements with different risk profiles.

2. CHANCE CONSTRAINED DATA ENVELOPMENT ANALYSIS

Data Envelopment Analysis (DEA) has been gaining more popularity as a non-parametric efficiency technique for measuring the performance of financial assets, as can be seen in recent studies such as Adam & Branda (2021) and Choi & Min (2017). The most used classic models in literature are deterministic and do not consider random input and output variable errors. According to Azadi and Saem (2012), the generalized randomness in the evaluation processes comes from data collection errors.

One of the first attempts to fill this gap involved developing Chance Constrained Programming in mathematical models for Data Envelopment Analysis (Charnes & Cooper, 1959), to incorporate stochastic variations in the data.

Saen and Azadi (2011), define Chance Constrained Programming (CCP) as a type of approach for stochastic optimization, appropriate for solving optimization problems with random variables included in the constraints, and sometimes in the objective function, as was done by Charnes and Cooper (1959) . The major contribution can be found in the research carried out by Sengupta (1987). CCP can effectively reflect the reliability of satisfying, or even the risk of violating, a system with constraints under risk conditions. CCP does not require that all constraints be completely satisfied. The constraints are satisfied according to established probabilities (Azadi et al., 2012; Farzipoor Saen & Azadi, 2011).

The stochastic DEA model formulation is presented according to equations 1 through 4, where the *i*-th DMU, $\hat{x}_i = (\hat{x}_{i1}, \hat{x}_{i2}, ..., \hat{x}_{ia})^T$ and $\hat{y}_i = (\hat{y}_{i1}, \hat{y}_{i2}, ..., \hat{y}_{ib})^T$ respectively denote the stochastic variables for the input and output vectors, where $i=1,...,n$. The objective function of the stochastic model is formulated by Equation 1, where '*E*' represents an expected value from the sum of weighted $\hat{y}_{iq} (q = 1,...,b)$.

$$\max E\left(\sum_{q=1}^{b} u_q \hat{y}_{0q}\right) \quad (1)$$

Subject to:

$$E\left(\sum_{p=1}^{a} v_p \hat{x}_{0p}\right) = 1 \quad (2)$$

$$P\left(\frac{\sum_{q=1}^{b} u_q \hat{y}_{iq}}{\sum_{p=1}^{a} v_p \hat{x}_{ip}} \leq \beta_i\right) \geq 1 - \alpha_i \quad i = 1,2,...,n \quad (3)$$

$$u_q, v_p \geq 0 \quad (4)$$

u_q and v_p respectively indicate the weight of the multipliers associated with the *q*-th output and the *p*-th input. $u_1,\ldots,u_b$, $v_1,...,v_a$ are the weights that will be calculated by the optimization model. *P* is the probability, and superscript `^` means that $\hat{x}_{ip}$ and $\hat{y}_{iq}$ are random variables. Regarding the constraints, the model states that the proportion equal or inferior to β_i represents an efficiency level expected for the *i*-th DMU, the variation of which is in [0,1], and it is the desired level (Cooper et al., 1996; Jin et al., 2014; Rotela Junior et al., 2015). α_i is the risk criterion adopted by the decision maker. $1-\alpha_i$ is the probability of meeting the constraint requirements, which is a level of confidence (Jin et al., 2014; Rotela Junior et al., 2015), the variation of which is in [0,1].

The formulation needs to be rewritten as proposed by Charnes and Cooper (1963), to provide a viable model from a computational point of view. Randomness is considered in this proposal, and the stochastic variable $\hat{x}_{ip}$ for each input can be represented as $\hat{x}_{ip} = \bar{x}_{ip} + a_{ip}\xi$, where p presents a variation in the interval [1,b], and i is in [1,n]; $\bar{x}_{ip}$ is the expected value of $\hat{x}_{ip}$ and a_{ip} is the standard deviation (Rotela Junior et al., 2015).

Similarly, the stochastic variable $\hat{y}_{iq}$ for each output can be represented as $\hat{y}_{iq} = \bar{y}_{iq} + b_{iq}\xi$, where q has a variation in [1,a], and i in [1,n]; $\bar{y}_{iq}$ is the expected value of $\hat{y}_{iq}$ and b_{iq} is the standard deviation. Thus, it is assumed that the random variable ξ follows a normal distribution, since part of the stochastic disorders suggest that the errors are the result of data collection.

In order to make the model solution simpler, it is convenient to present its equivalent deterministic formulation. The objective function (as presented in Equation 1), can be remodeled according to Equation 5:

$$E\left(\sum_{q=1}^{b} u_q \hat{y}_{0q}\right) = \sum_{q=1}^{b} u_q \bar{y}_{0q} \qquad (5)$$

The model constraints, according to Equations 2 and 3, when including the stochastic process, will be rewritten in Equations 6 and 7:

$$E\left(\sum_{p=1}^{a} v_p \hat{x}_{0p}\right) = \sum_{p=1}^{a} v_p \bar{x}_{0p} = 1 \qquad (6)$$

$$P_r\left(\frac{\sum_{q=1}^{b} u_q \hat{y}_{iq}}{\sum_{p=1}^{a} v_p \hat{x}_{ip}} \le \beta_i\right) = \tag{7}$$

$$P_r\left(\sum_{q=1}^{b} u_q \hat{y}_{iq} - \beta_i \sum_{p=1}^{a} v_p \hat{x}_{ip} \le 0\right) \ge 1-\alpha_i$$

$$i = 1, 2, \ldots, n$$

M_i and V_i are the average and variance of each random variable. These can be expressed as Equations 8 and 9:

$$M_i = \sum_{q=1}^{b} u_q \bar{y}_{iq} - \beta_i \sum_{p=1}^{a} v_p \bar{x}_{ip} \tag{8}$$

$$V_i = \left(\sum_{q=1}^{b} u_q \bar{y}_{iq} - \beta_i \sum_{p=1}^{a} v_p \bar{x}_{ip}\right)^2 \sigma^2 \tag{9}$$

In this way, the random variable $\left(\frac{\left(\sum_{q=1}^{b} u_q \hat{y}_{iq} - \beta_i \sum_{p=1}^{a} v_p \hat{x}_{ip}\right) - M_i}{\sqrt{V_i}}\right)$ follows a normal distribution of mean zero and variance one. Thus, Equation 7 can be expressed according to Equation 10, or even in its equivalent form as expressed by Equation 13.

$$\frac{-M_i}{\sqrt{V_i}} \ge \phi^{-1}\left(1-\alpha_i\right) \quad i = 1, 2, \ldots, n \tag{10}$$

Equation 10 can be written as 11:

$$\sum_{p=1}^{a} v_p \, \beta_i \left(\bar{x}_{ip} + \phi^{-1}\left(1-\alpha_i\right) a_{ip} \, \sigma\right) - \sum_{q=1}^{b} u_q \left(\bar{y}_{iq} + \phi^{-1}\left(1-\alpha_i\right) b_{iq} \sigma\right) \ge 0_i \tag{11}$$

$$i = 1, 2, \ldots, n$$

In this model, ϕ represents a function of standard normal distribution, and ϕ^{-1} is the inverse of the function. Finally, the CCDEA optimization model can be discussed. In this manner, the original proposal is presented as a linear model, according to Equations 12-15:

$$\max \sum_{q=1}^{b} u_q . \bar{y}_{oq} \tag{12}$$

Subjected to:

$$\sum_{p=1}^{a} v_p \bar{x}_{0p} = 1 \tag{13}$$

$$\sum_{p=1}^{a} v_p \beta_i \left(\bar{x}_{ip} + \Phi^{-1}(1-\alpha_i) a_{ip} \sigma \right) - \sum_{q=1}^{b} u_q \left(\bar{y}_{iq} + \Phi^{-1}(1-\alpha_i) b_{iq} \sigma \right) \geq 0_i \quad i = 1, 2, \ldots, n \tag{14}$$

$$u_q, v_p \geq 0 \tag{15}$$

The model of presented multipliers extends the applications of the Data Envelopment Analysis (DEA) to the financial area and helps in decision making. Besides the deterministic situation, efficiency can be measured considering random variables. The desired confidence levels of the model can be defined according to different situations of practical application, according to the particularities of each case.

3. PORTFOLIO SELECTION

Portfolio optimization and diversification concepts have been fundamental in understanding financial markets and developing financial decision making tools (Fabozzi et al., 2007). Diversification reduces portfolio risk, by investing in assets that are independent from one another i.e., portfolios comprise assets with low correlation (Kim et al., 2015).

Solutions for portfolio optimization are often influenced by poorly specified models or by errors in approximation, estimation, or even due to incomplete information. As demonstrated by Black and

Litterman (1992), a small variation in the expected return of assets can result in a large alteration in the allocation of investments in an optimized portfolio. In other words, classic models for portfolio optimization are not robust because they are susceptible to small variations in data input (Kim et al., 2015). As a matter of fact, Kim et al. (2014), affirmed that the main reason for questioning the Markowitz (1952) model is because of its high sensitivity towards small variations in input values.

Researchers have begun to incorporate risk by estimating errors directly in the portfolio optimization process using mathematical techniques for robust optimization. Different from traditional approaches where the inputs for the structure of portfolio allocation are deterministic, robust portfolio optimization incorporates the notion that these inputs have been estimated with errors (Fabozzi et al., 2007, Fabozzi et al. 2010). In this case, inputs like the expected return and asset covariance are not traditional predictions but rather sets of probabilities e.g., confidence intervals. The two best-known methodologies for dealing with risk are robust and stochastic optimization (Xidonas et al., 2020).

It was identified that the correlation between assets tends to increase during periods of low market (bear market) performance, so investors cannot benefit from diversification when it is most needed. Worse still, correlation within capital markets has been increasing in recent periods (Kim et al., 2015). Some solutions have been proposed to overcome this problem, like employing input variables that are less sensitive to historic data, or inserting risk sets on the input parameters of traditional models (Fabozzi et al., 2007).

Of the several approaches for increasing robustness in the mean-variance model, robust portfolio optimization applies robust optimization techniques for asset allocation by solving max-min problems (Won & Kim, 2020; Xidonas et al., 2017). According to Kim et al. (2015), even though worst-case optimization seems to be a natural extension of the mean-variance model for achieving robustness, more in-depth analysis on the importance of concentration in worst-case market scenarios has not been conducted.

The main contribution of the work conducted by Kim et al. (2015) was to demonstrate the importance of information on asset returns on the worst-performing days for achieving a robust portfolio. In other words, instead of selecting assets that always perform well in both bear markets and bull markets, assets that perform well in bear markets are usually considered. Therefore, these researchers believe that robustness can be achieved when worst-case market information is considered.

Furthermore, it is known that low beta assets perform better than high beta assets in crisis periods (market declines), since low beta assets reduce the overall risk and offer better returns (Kim et al., 2015).

4. MATERIALS AND METHOD

In order to establish a strategy for robust portfolio optimization using stochastic efficiency analysis, we used the following observations as starting points:

i. Preliminary results indicate that Data Envelopment Analysis (DEA) is well suited for determining portfolio composition. DEA allows assets to be evaluated using criteria that represent investor interests, supporting the classic models of mean and variance;

ii. DEA models that consider data randomness allow investors to reduce the search space for assets that perform well, by presenting good data discrimination;

iii. Risk variation in stochastic efficiency analysis can meet the needs of investors with different attitudes towards risk;

iv. One difficulty in applying CCDEA to portfolio optimization (depending on the data) is the contradictory behavior of the constraints, making it difficult to identify efficient assets;

v. Hierarchical Clustering allows individuals or assets to be grouped based on the similarity or dissimilarity of these initial groups;

vi. Robust portfolio models achieve robustness by concentrating on information gathered from crisis periods or market recessions i.e., poor performance days are fundamental for building portfolios that perform well under any market conditions.

4.1. Selecting the input and output variables

After defining the object of study, we set out to select a set of indicators that will be used as input and output variables in the efficiency analysis.

Efficiency considerations, which are essential for the definitions and assessments of interest in DEA, are presented through a definition of mathematical reason, which in a very simple way is represented by output divided by input. This simple ratio formulation can be extended to multiple outputs and inputs in order to be presented by more complex formulations (Cook & Zhu, 2014; Cooper et al., 2002). Then, following the premises presented in the extensive DEA literature, the variables in this study will be defined as outputs or inputs.

Conclusions from Siriopoulos and Tziogkidis (2010), were taken into consideration. They state that highly correlated inputs and outputs do not significantly affect the efficiency results. Therefore, we did not see a need to perform correlation tests for the possible input and output variables.

We chose to use input and output variables present in literature for DEA applications in the stock market (Powers & Mcmullen, 2002; Rotela Junior et al., 2014; Rotela Junior et al., 2015). It should be remembered that low beta assets have better returns while reducing the overall risk of the portfolio in worst-case market scenarios (Kim et al., 2015).

For this study, we chose to use the asset return, asset liquidity, and earnings per share (EPS) as output model variables. For input variables we chose to use the beta, the price-earnings (PE), and volatility (Kim et al., 2015; Powers & Mcmullen, 2002; Rotela Junior et al., 2014; Rotela Junior et al., 2015).

4.2. Selecting the sample and data collection

The sample comprised assets traded on the Brazilian Stock Exchange B3 (*Brasil, Bolsa, Balcão*). The B3 was founded after a merger between BM&FBOVESPA and Cetip, with participation in the Bovespa Index (Ibovespa). It is necessary that the assets present information from the long-term period.

Kim *et al.* (2015) used daily data for the return of the market index to identify worst-case market periods. The authors classified all the returns of the index in increasing order within a time interval. They then divided this period into *n* other periods. Within the longer period, the authors defined *n* as ten, and for defining the worst-case market period, the tenth period was selected, corresponding to the smallest values presented for the index. We used the same approach adopted by Kim et al. (2015), to identify crisis market periods.

We then proceeded to selecting the sample, however, we observed that only 61 companies in the Ibovespa index had all the necessary information for conducting the efficiency analysis. It is worth noting that Ibovespa index is a benchmark index of about 60 stocks traded on the Brazilian stock exchange. All information on the selected variables were collected by the software program Economática® for the stipulated period.

The information for this study corresponds to daily data from a 10-semester period. To validate the results, daily information from a 6-month period was used. It is important to highlight that the period used in this study was prior to the global covid-19 pandemic. It is worth noting that there is still no concrete evidence of the effects of the pandemic on stock markets.

The cumulative return was calculated for each proposed portfolio for the validation phase based on the results of the models adopted for optimization (identifying the ideal share).

5. OPTIMIZING ROBUST PORTFOLIOS

We started by collecting data in a database, and then proceeded using a proposal from Kim *et al.* (2015) by classifying the returns from Ibovespa in increasing order for the period adopted in this study. The remaining spreadsheet information followed such classification, and *n* was set equal to four for defining the worst-case market period, giving the model more than three hundred daily information pieces.

We calculated the average and variance for each of the variables adopted for the efficiency analysis using the data collected for each of the proposed scenarios i.e., worst-case (n = 4), and complete market information (n = 1),

We observed that the number of efficient assets was very reduced, even when varying the risk criterion. This led us to believe that the CCDEA model was composed of highly divergent constraints, making it more difficult to properly discriminate the analysis. Next, different forms of Hierarchical Clustering were tested, and the most viable option for each one of the considered scenarios was to group the DMUs by degree of similarity, considering the average and variance of the six variables adopted in this study.

This made it possible to group the DMUs by increasing the degree of similarity between the groups in which the efficiencies can be analyzed. Figure 1 shows the DMU grouping using Hierarchical Clustering for complete asset information (n = 1).

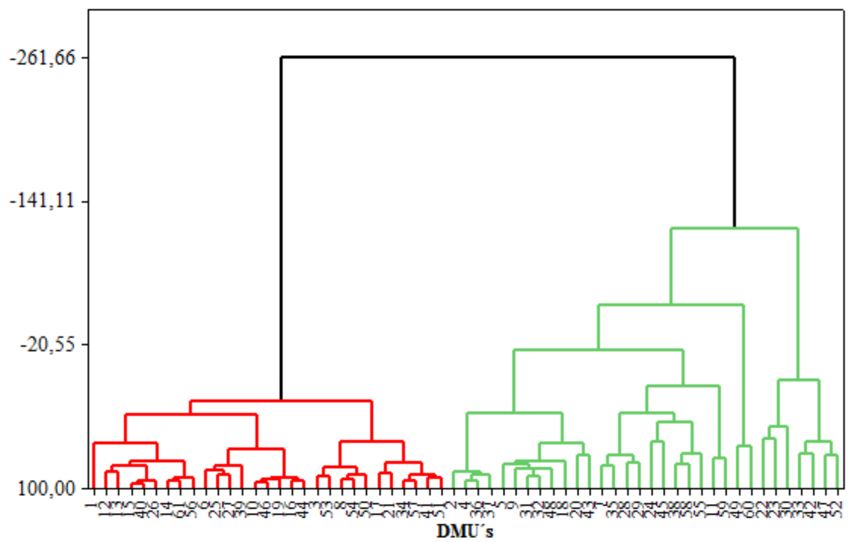

FIGURE 1. Dendrogram of the grouping considering complete market data information.

Figure 2 shows the DMU grouping for worst-case market periods, defined as n = 4. These Figures were obtained using the Minitab® software program.

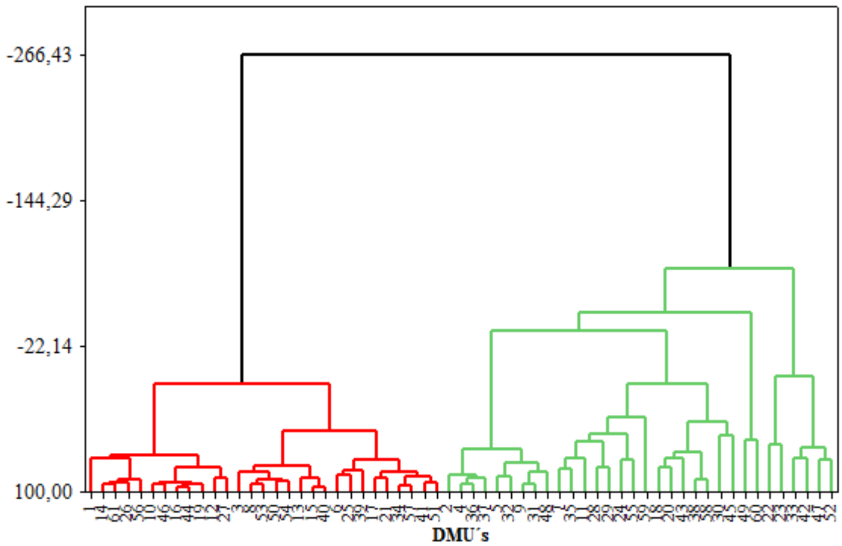

FIGURE 2. Dendrogram of the grouping considering information from worst-case market periods.

Two groupings were performed for each of the two proposed scenarios for the assets that make up the more similar groups.

Grouping can be done in a larger number of groups, however, given the number of variables, the model required a minimum number of DMUs for good data discrimination (Cooper et al., 2007), and so only two groups were formed. This solution can be adopted to facilitate meeting the constraints in the CCDEA model, since they respect the recommendations in the model application.

Table 1 shows the descriptive statistics of the DMU input and output variables that comprise group 1 and 2 considering information on the total market state.

Table 2 shows the descriptive statistics of the DMU input and output variables that comprise groups 1 and 2 considering information from the worst-case market periods.

TABLE 1 - DESCRIPTIVE STATISTICS OF THE INPUT AND OUTPUT VARIABLES OF THE MODEL, FOR GROUP 1 AND 2, CONSIDERING TOTAL MARKET STATE INFORMATION

	Group 1											
	Return		**Liquidity**		**EPS**		**Beta**		**PE**		**Volatility**	
	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	μ_4	σ_4^2	μ_5	σ_5^2	μ_6	σ_6^2
Mean	0.07	3.52	0.65	0.09	7.03	9.19	0.63	0.01	19.11	132.21	1.87	0.09

Median	0.07	3.39	0.59	0.03	6.31	1.89	0.61	0.01	19.49	18.88	1.77	0.08
Standard Deviation	0.03	0.78	0.26	0.22	3.13	18.73	0.18	0.01	9.55	327.64	0.25	0.06
Minimum	0.01	2.17	0.29	0.00	2.55	0.18	0.30	0.00	1.00	1.98	1.45	0.01
Maximum	0.13	4.79	1.30	1.19	14.42	92.35	0.99	0.02	42.94	1706.57	2.31	0.27
Group 2												
Mean	-0.02	6.02	1.62	0.27	3.36	286.57	1.05	0.01	12.27	3365.30	2.39	0.19
Median	-0.02	5.41	1.04	0.10	6.32	31.91	1.06	0.01	10.24	1393.75	2.33	0.14
Standard Deviation	0.06	2.29	1.63	0.59	10.38	650.30	0.22	0.01	15.98	4451.30	0.45	0.15
Minimum	-0.17	3.10	0.35	0.01	-26.73	0.54	0.59	0.00	-30.46	2.07	1.70	0.03
Maximum	0.08	10.70	7.05	2.59	17.39	3173.20	1.50	0.05	47.87	14978.71	3.29	0.60

TABLE 2 - DESCRIPTIVE STATISTICS OF THE INPUT AND OUTPUT VARIABLES OF THE MODEL, FOR GROUP 1 AND 2, CONSIDERING INFORMATION FROM THE WORST-CASE MARKET PERIODS

Group 1												
	Return		**Liquidity**		**EPS**		**Beta**		**PE**		**Volatility**	
	μ_1	σ_1^2	μ_2	σ_2^2	μ_3	σ_3^2	μ_4	σ_4^2	μ_5	σ_5^2	μ_6	σ_6^2
Mean	-0.98	3.15	0.65	0.09	7.12	9.46	0.63	0.01	19.39	123.06	1.86	0.09
Median	-0.97	3.12	0.60	0.03	6.32	2.12	0.60	0.01	19.27	18.35	1.77	0.07
Standard Deviation	0.29	0.70	0.26	0.22	3.13	20.24	0.18	0.01	8.48	318.48	0.24	0.06
Minimum	-1.52	1.95	0.29	0.01	2.62	0.20	0.29	0.00	7.77	2.14	1.45	0.01
Maximum	-0.46	4.53	1.31	1.19	14.61	102.62	0.98	0.02	40.20	1686.56	2.29	0.23
Group 2												
Mean	-1.86	4.42	1.62	0.31	3.25	308.54	1.05	0.01	12.70	3152.21	2.37	0.20
Median	-1.78	4.12	1.04	0.10	6.44	35.82	1.06	0.01	10.61	1127.46	2.31	0.14
Standard Deviation	0.39	1.93	1.63	0.69	10.92	692.02	0.22	0.01	18.73	4112.77	0.44	0.15
Minimum	-2.79	2.01	0.36	0.01	-29.27	0.58	0.60	0.00	-30.61	1.89	1.70	0.03
Maximum	-1.14	9.85	7.10	3.17	18.03	3312.95	1.51	0.05	65.25	13506.95	3.25	0.64

Similar to other studies, negative data were transformed by adding a value that makes the most negative value in the series positive for the same variable, without changing the efficiency analysis (Cook & Zhu, 2014). It is worth mentioning that the input and output variables were independent in this study.

The efficiency level (β_i) was equal to 1. We observed in the studied data that a good range of discrimination was obtained in the analysis units when the risk criterion (α_i) was varied between 0.5 and 0.6. This range will change according to the data submitted to the CCDEA model.

The variation within the range stipulated in the previous step may be related to investor's aversion to risk. In this study, a variation of 0.01 was chosen within the relevant range defined for probability variation to fulfill constraints (1-α_i), generating eleven portfolios for each market state.

The assets pre-selected by the CCDEA model were submitted to the Sharpe approach (Sharpe, 1963) for optimal allocation within the portfolio. A similar strategy of pre-specification for assets was adopted by Chakrabarti (2021).

We use the Capital Asset Pricing Model (CAPM), presented by Sharpe (1964), to analyze the results and to identify any abnormal returns (Rotela Junior et al., 2014; Rotela Junior et al., 2015).
The Sharpe ratio is the most commonly used metric for measuring and comparing portfolio performance (Homm & Pigorsch, 2012; Kourtis, 2016).

As previously mentioned, daily information was used for the validation period. The accumulated returns were calculated in the validation period for each portfolio according to the participations defined by the optimization models.

5.1. Results and analysis

Efficiency assessments were carried out for the proposed groups, and α_i was considered. Table 3 shows the efficiency results with different compliance probability levels (defined by 1- α_i) of the constraints for groups 1 and 2, respectively, when supplied with worst-case market scenario information.

We evaluated the efficiency of the proposed groups considering the risk criteria adopted. Table 3 shows the efficiency results for groups 1 and 2, respectively, when submitted to the CCDEA model under different probability levels for fulfilling the model constraints, supplied with complete market information taken from the stipulated period.

Likewise, Table 4 presents the efficiency results for groups 1 and 2, respectively. However, the CCDEA model was supplied with information taken from market downturns. It is important to highlight that the likelihood of complying with the constraints of the optimization model increases by reducing the risk criterion (α_i), making the model more critical. Therefore, fewer assets will be efficient.

TABLE 3 - DESCRIPTIVE STATISTICS FOR THE EFFICIENCIES OF GROUP 1 AND 2, CONSIDERING FULL MARKET INFORMATION

					Group 1						
(1-α_i)	**40%**	**41%**	**42%**	**43%**	**44%**	**45%**	**46%**	**47%**	**48%**	**49%**	**50%**
Mean	1.25	1.19	1.13	1.08	1.04	1.00	0.96	0.93	0.90	0.88	0.86
Median	1.24	1.17	1.11	1.06	1.02	0.99	0.96	0.95	0.92	0.89	0.87
Standard Deviation	0.17	0.16	0.15	0.15	0.14	0.14	0.13	0.13	0.12	0.12	0.11
Minimum	0.97	0.92	0.88	0.83	0.80	0.76	0.73	0.70	0.68	0.67	0.66
Maximum	1.56	1.49	1.40	1.34	1.27	1.22	1.17	1.12	1.08	1.04	1.00
					Group 2						
(1-α_i)	**40%**	**41%**	**42%**	**43%**	**44%**	**45%**	**46%**	**47%**	**48%**	**49%**	**50%**
Mean	2.25	2.07	1.93	1.81	1.69	1.50	1.34	1.22	1.10	1.00	0.82
Median	1.80	1.68	1.58	1.47	1.37	1.25	1.11	1.02	0.96	0.90	0.82
Standard Deviation	2.12	2.02	1.84	1.77	1.71	1.37	1.16	1.04	0.81	0.63	0.15
Minimum	1.10	1.02	0.96	0.90	0.83	0.74	0.66	0.60	0.55	0.51	0.47
Maximum	13.53	12.77	11.70	11.24	10.81	8.77	7.52	6.77	5.39	4.34	1.00

TABLE 4 - DESCRIPTIVE STATISTICS FOR THE EFFICIENCIES OF GROUP 1 AND 2, CONSIDERING WORST-CASE SCENARIO MARKET INFORMATION

					Group 1						
(1-α_i)	**40%**	**41%**	**42%**	**43%**	**44%**	**45%**	**46%**	**47%**	**48%**	**49%**	**50%**
Mean	0.96	0.95	0.94	0.92	0.91	0.90	0.89	0.88	0.87	0.86	0.85
Median	0.97	0.95	0.93	0.92	0.90	0.89	0.87	0.87	0.87	0.86	0.85
Standard Deviation	0.15	0.15	0.14	0.14	0.13	0.13	0.13	0.12	0.12	0.12	0.12
Minimum	0.70	0.68	0.69	0.68	0.68	0.67	0.66	0.66	0.66	0.65	0.65
Maximum	1.22	1.20	1.17	1.15	1.13	1.10	1.08	1.06	1.04	1.02	1.00
					Group 2						
(1-α_i)	**40%**	**41%**	**42%**	**43%**	**44%**	**45%**	**46%**	**47%**	**48%**	**49%**	**50%**
Mean	1.25	1.14	1.09	1.03	0.99	0.95	0.89	0.85	0.82	0.79	0.75
Median	1.08	1.03	0.97	0.93	0.91	0.88	0.85	0.83	0.81	0.79	0.76
Standard Deviation	1.09	0.79	0.70	0.62	0.57	0.54	0.38	0.31	0.27	0.25	0.21
Minimum	0.39	0.39	0.38	0.38	0.37	0.36	0.32	0.33	0.35	0.31	0.31
Maximum	6.79	5.05	4.47	4.01	3.68	3.48	2.37	1.83	1.49	1.34	1.00

Clustering analysis was conducted to form a cluster of assets with a certain degree of similarity, seeking better DMU discrimination, since the divergence between the CCDEA model constraints are reduced when these DMUs are grouped. However, efficient assets in the groups for each market state will be gathered and optimized as proposed by Sharpe.

Since traditional DEA models ignore diversification between investment opportunities (Adam & Branda, 2021), Sharpe's model could be necessary to build the portfolio. Efficient assets from Table 3 were submitted to Sharpe's proposal taking information from the total market state into account for each risk criterion (α_i) adopted. Then, the assets were allocated in eleven portfolios by varying the risk criterion (α_i) when the model is supplied with complete market information, or $n = 1$. It is interesting to observe that not all efficient assets considered will be used in the allocation when submitted to the Sharpe model.

Eleven portfolios were proposed from varying the risk criterion for the total market state (TS) information. These were identified as TS-1 to TS-11, to simplify discussion.

Another eleven portfolios were proposed according to the risk criterion (α_i) for the worst-case (WS) market information. It is interesting to notice that there were fewer efficient assets when the model was supplied with WS market information. When submitted to Sharpe's proposal, only some were selected for the portfolios. These portfolios were identified as WS-1 to WS-11, to simplify discussion.

Table 5 presents the adopted risk criterion (α_i), portfolio beta (β), return results (R_E and R), standard deviation (SD), Sharpe ratio (S_R), and number of assets (N) for each TS optimized portfolio. And Table 6 presents the same content, but for WS optimized portfolios.

TABLE 5 - RESULTS BY RISK CRITERIA OF THE OPTIMIZED PORTFOLIOS, BASED ON COMPLETE MARKET INFORMATION

	TS-1	TS-2	TS-3	TS-4	TS-5	TS-6	TS-7	TS-8	TS-9	TS-10	TS-11
α_i	60%	59%	58%	57%	56%	55%	54%	53%	52%	51%	50%
β	0.702	0.700	0.691	0.691	0.763	0.713	0.680	0.689	0.668	0.674	0.616
R_E	1.07%	1.07%	1.06%	1.07%	1.09%	1.07%	1.06%	1.06%	1.05%	1.06%	1.03%
SD	8.80%	8.51%	8.71%	8.90%	8.69%	8.87%	8.97%	8.95%	8.70%	8.61%	9.65%
R	-3.05%	-2.68%	-3.33%	-2.10%	-1.00%	-0.22%	0.35%	0.39%	2.10%	1.54%	2.96%
SR	-0.467	-0.440	-0.504	-0.356	-0.240	-0.145	-0.079	-0.074	0.120	0.055	0.200
N	57	54	52	48	45	41	34	26	19	17	12
AAR	-1.59%	-1.12%	-1.84%	0.05%	1.48%	2.57%	3.90%	4.70%	7.44%	6.53%	9.31%

TABLE 6 - Results by risk criterion of optimized portfolios, based on information from periods of market downturns

	WS-1	WS-2	WS-3	WS-4	WS-5	WS-6	WS-7	WS-8	WS-9	WS-10	WS-11
α_i	60%	59%	58%	57%	56%	55%	54%	53%	52%	51%	50%
β	0.458	0.458	0.446	0.444	0.444	0.438	0.432	0.429	0.429	0.429	0.429
R_E	0.96%	0.96%	0.96%	0.96%	0.96%	0.96%	0.95%	0.95%	0.95%	0.95%	0.95%
SD	7.67%	7.67%	8.16%	8.05%	8.05%	8.38%	8.36%	8.32%	8.32%	8.32%	8.32%
R	2.52%	2.52%	2.65%	3.85%	4.03%	5.34%	5.73%	5.73%	5.73%	5.73%	5.73%
SR	0.202	0.203	0.206	0.358	0.358	0.366	0.525	0.575	0.575	0.575	0.575
N	11	11	11	10	10	10	9	8	8	8	8
AAR	5.58%	5.58%	5.42%	6.62%	6.62%	7.24%	8.91%	9.16%	9.16%	9.16%	9.16%

The main discussion shows the importance of information taken from periods of market crisis and recession and how this information contributes to robust portfolio optimization. Tables 5 and 6 analyze and compare the TS optimized portfolios (portfolios TS-1 to TS-11) and WS optimized portfolios (portfolios WS- 1 to WS-11). The portfolios are compared in pairs according to the α_i value adopted.

It is important to highlight that the WS optimized portfolios had better results when reading the Sharpe ratio (S_R) considering different α_i values.

Tables 5 and 6 present the expected return values for the portfolios that were calculated as presented before. We needed to calculate the beta values (β) for each of the portfolios (also shown in the Tables). For the TS optimized portfolios (TS-1 to TS-11), the expected returns (R_E) vary from 1.03% to 1.09% a.m. For the WS optimized portfolios (WS-1 to WS-11), the expected returns (R_E) were concentrated between 0.95% and 0.96% a.m.

The obtained mean profitability (R) were -3.05%, -2.68%, -3.33%, -2.10%, -1.00%, -0.22%, 0.35%, 0.39%, 2.10%, 1.54%, and 2.96% for TS-1 to TS-11, respectively. The obtained mean profitability (R) were 2.52%, 2.52%, 2.65%, 3.85%, 4.03%, 5.34%, 5.73%, 5.73%, 5.73%, 5.73%, and 5.73% for WS-1 to WS-11, respectively.

The accumulated abnormal return (AAR) of the portfolios was obtained from information collected during the validation period. The portfolios were compared in pairs, and one was optimized from the TS (n = 1) period, and the other from the WS (n = 4) period. Figures 3 and 4 show the accumulated return in pairs established according to the probability level (1-α_i) of fulfilling constraints from the CCDEA model.

Figure 3 shows the AAR of the portfolio pairs when a risk range (α_i) criterion of 60% to 55% is adopted. Figure 4 shows the AAR of the portfolio pairs when a risk range (α_i) criterion of 54% to 50% is adopted.

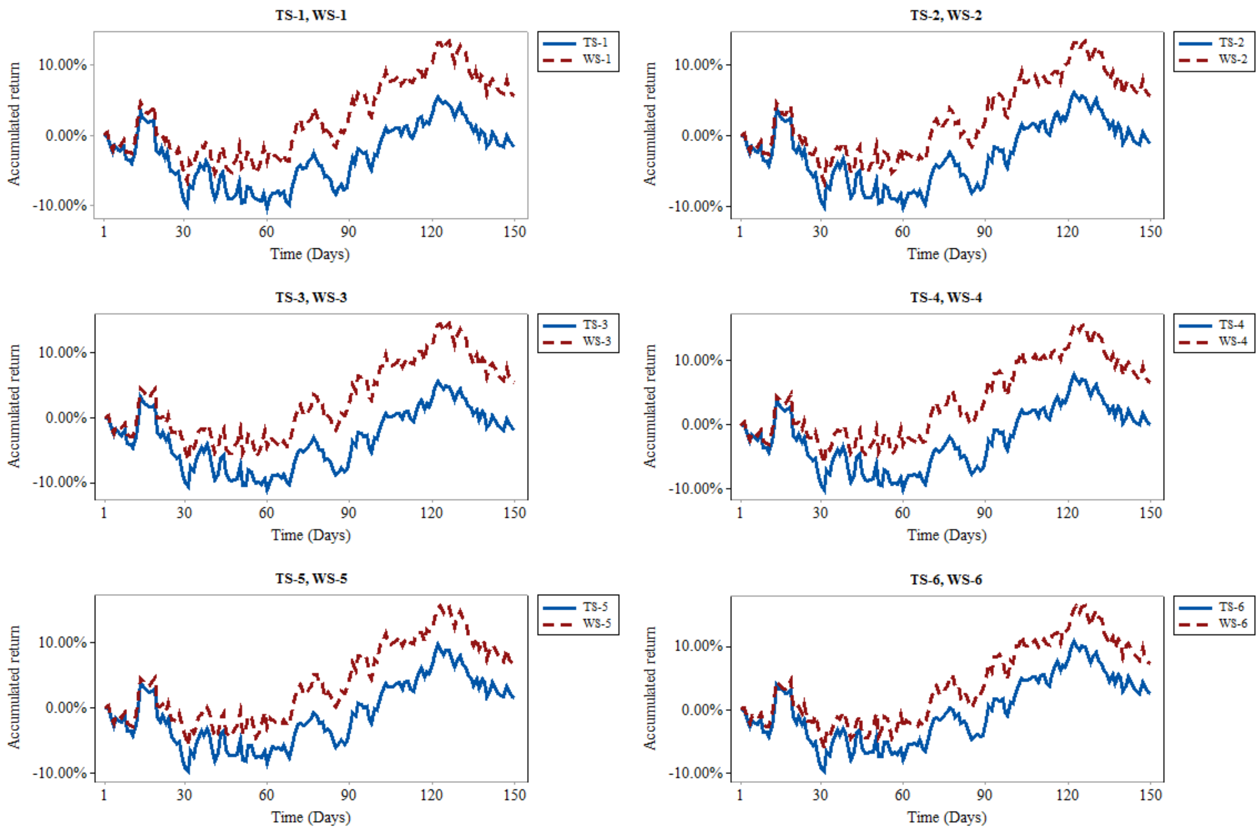

FIGURE 3. Accumulated return of the pairs of portfolios considering risk criterion variation (60% - 55%).

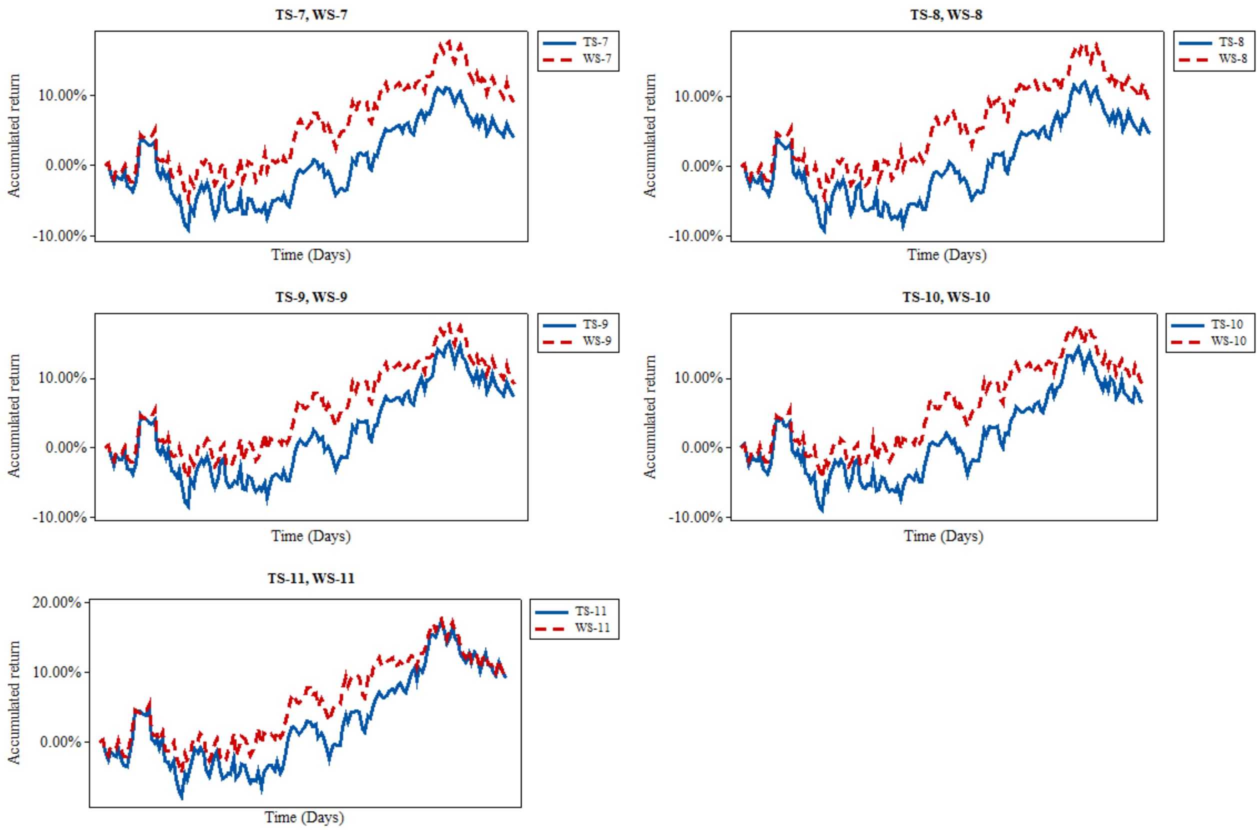

FIGURE 4. Accumulated return of the pairs of portfolios considering risk criterion variation (54% - 50%).

It is important to observe that the WS optimized portfolios (WS-1 to WS-11) had better S_R values. Regardless of the risk criterion adopted, the beta values of the proposed portfolios were lower than portfolios TS-1 to TS-11. Kim et al. (2015), conclude that robust portfolios optimized using stochastic models, achieve expected robustness since they focus mainly on crisis period information. The authors believe that these robust portfolios tend to comprise assets with low beta values. Most importantly, these assets tend to perform better when compared to high beta value assets in any other period classification.

After developing the accumulated return graphs, we conducted a statistical test to compare the obtained series of abnormal accumulated returns for each pair of portfolios associated by the risk criterion. We decided to use the Mann-Whitney non-parametric test. The P-value results obtained in these tests (when portfolios are analyzed in pairs) were less than 0.05. This result allows us to affirm that the abnormal accumulated return of WS optimized portfolios is statistically higher than the accumulated returns of TS optimized portfolios for the entire period. Figure 5 presents a boxplot diagram for the pairs of portfolios associated by the risk criterion.

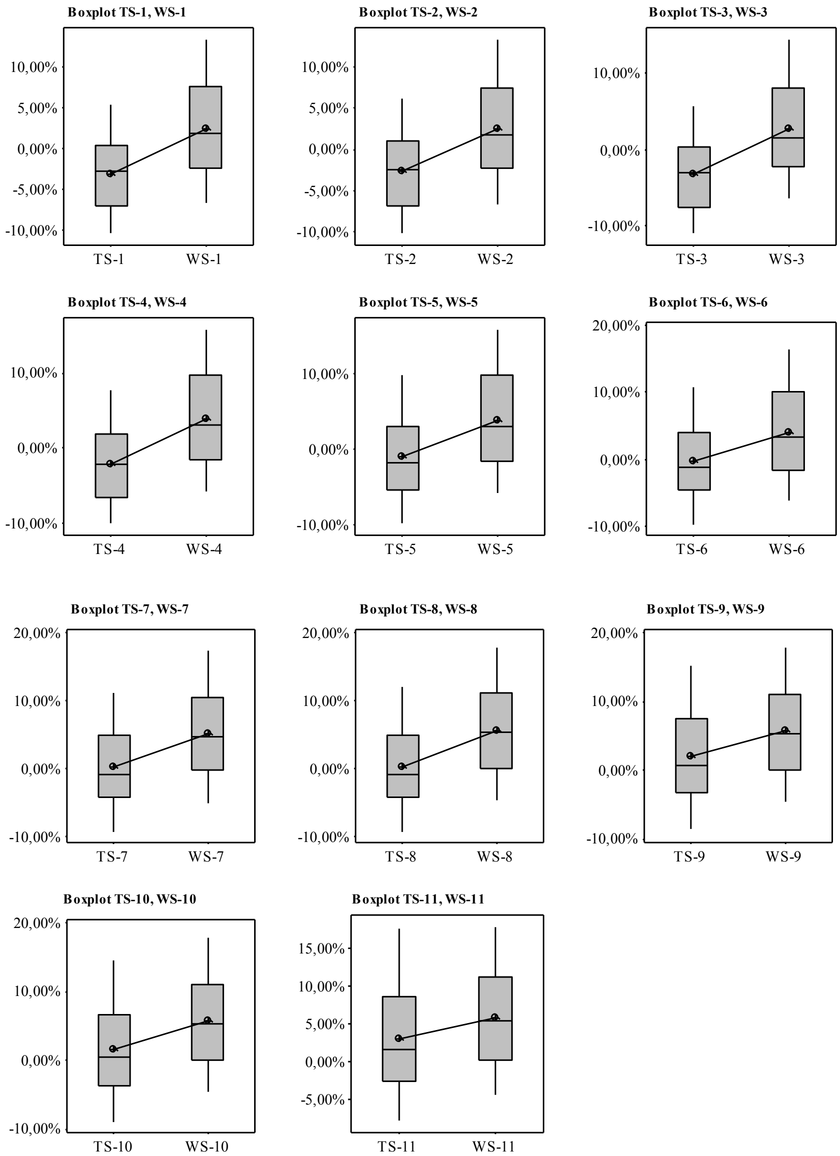

FIGURE 5. Boxplot of the accumulated returns from pairs of portfolios, by risk criterion (60% - 50%).

The results have shown that the proposed method supplied with WS information has better results than TS information for the total period.

However, to confirm the applicability of this method, we needed to compare these portfolios with other portfolios build using classic models for portfolio optimization. Models presented by Markowitz (1952), and Sharpe (1963), were used for this comparison.

TS (n=1) information assets were ideally allocated using the Markowitz (1952) model and by maximizing the Sharpe ratio of the assets. This portfolio was named the Comparative Markowitz (MC) portfolio. Using the same set of information, assets were ideally allocated using the Sharpe model (1963). This portfolio was named the Comparative Sharpe (SC) portfolio. Table 7 shows the share of assets in the comparative portfolios. Additionally, the shares are again presented for four of the eleven proposed portfolios, with two of them (TS-1 and TS-11) supplied with the same set of information as the Comparative portfolios, and two other portfolios (WS-1 and WS-11) from WS (n=4) market periods. For this comparison, we chose to consider only portfolios with risk criteria 60% and 50%, which are limit values for the adopted range.

TABLE 7- Assets allocation in the comparative and proposed portfolios

(α_i)	**-**	**-**	**60%**	**60%**	**50%**	**50%**
Portfolios	**MC**	**SC**	**TS-1**	**WS-1**	**TS-11**	**WS-11**
DMU1	0.171	0.046	0.046	0.176	0.139	0.235
DMU2	0.000	0.014	0.015	0.000	0.061	0.000
DMU3	0.000	0.025	0.025	-	-	-
DMU4	0.000	0.024	0.024	0.000	0.091	0.000
DMU5	0.000	0.007	0.008	0.000	-	-
DMU6	0.090	0.038	0.038	0.093	-	-
DMU7	0.000	0.017	0.017	0.000	-	-
DMU8	0.000	0.017	0.017	-	-	-
DMU9	0.000	0.012	0.013	-	-	-
DMU10	0.000	0.027	0.028	-	-	-
DMU11	0.000	0.017	0.017	0.027	0.063	0.067
DMU12	0.120	0.033	0.033	0.095	0.096	0.127
DMU13	0.000	0.020	0.021	0.056	0.074	0.101
DMU14	0.000	0.028	0.029	-	-	-
DMU15	0.000	0.019	0.019	0.041	-	-
DMU16	0.014	0.029	0.030	-	-	-
DMU17	0.000	0.022	0.023	-	-	-
DMU18	0.000	0.003	0.003	-	-	-
DMU19	0.103	0.032	0.032	-	-	-
DMU20	0.000	0.006	0.007	-	-	-

DMU21	0.000	0.013	-	-	-	-
DMU22	0.000	0.004	0.004	0.000	-	-
DMU23	0.000	0.006	0.007	0.000	-	-
DMU24	0.000	0.004	0.004	0.000	0.046	0.019
DMU25	0.083	0.027	0.027	-	-	-
DMU26	0.000	0.020	0.020	0.097	-	-
DMU27	0.106	0.028	0.028	-	-	-
DMU28	0.000	0.011	0.011	-	-	-
DMU29	0.000	0.012	0.012	0.000	-	-
DMU30	0.000	0.001	0.001	-	-	-
DMU31	0.000	0.006	0.006	-	-	-
DMU32	0.000	0.006	0.006	-	-	-
DMU33	0.000	0.005	0.005	-	-	-
DMU34	0.032	0.023	0.023	-	-	-
DMU35	0.000	0.012	0.012	0.000	-	-
DMU36	0.000	0.021	0.021	0.000	-	-
DMU37	0.000	0.021	0.021	0.000	-	0.000
DMU38	0.000	0.010	0.010	0.000	-	-
DMU39	0.000	0.021	0.022	-	-	-
DMU40	0.000	0.021	0.021	0.045	-	-
DMU41	0.000	0.021	0.021	-	-	-
DMU42	0.000	0.003	0.003	-	-	-
DMU43	0.000	0.006	0.006	-	-	-
DMU44	0.000	0.020	0.021	-	-	-
DMU45	0.000	0.000	0.000	0.000	-	-
DMU46	0.017	0.029	0.029	-	-	-
DMU47	0.000	0.000	0.000	-	-	-
DMU48	0.000	0.003	0.003	-	-	-
DMU49	0.000	0.005	0.005	0.000	0.046	0.000
DMU50	0.042	0.023	0.024	-	-	-
DMU51	0.000	0.022	0.022	-	-	-
DMU52	0.000	0.000	0.000	-	-	-
DMU53	0.000	0.013	0.013	-	-	-
DMU54	0.000	0.023	0.023	0.046	-	-
DMU55	0.000	0.011	0.011	0.000	0.050	0.015
DMU56	0.043	0.038	0.038	0.166	0.134	0.214
DMU57	0.000	0.020	0.020	-	-	-
DMU58	0.169	0.001	0.001	0.000	-	-
DMU59	0.009	0.011	0.011	0.000	-	-
DMU60	0.000	0.011	0.012	0.000	0.077	0.000
DMU61	0.000	0.033	0.034	0.158	0.123	0.222

The same validation period employed in the previous comparison was used to validate the results obtained in the portfolios.

Table 8 shows some parameters for the proposed portfolios (TS-1, TS-11, WS-1, and WS-11) and the comparative portfolios (MC and SC), like the adopted risk (α_i), the portfolio beta (β), the return results (R_E and R), the standard deviation (SD), the Sharpe ratio (S_R), and the number of assets (N) that comprise the portfolio.

TABLE 8 - Results for comparative portfolios and proposed portfolios

	MC	SC	TS-1	WS-1	TS-11	WS-11
α_i	-	-	60%	60%	50%	50%
β	0.475	0.702	0.702	0.458	0.616	0.429
R_E	0.96%	0.96%	0.96%	0.96%	0.95%	0.95%
SD	7.20%	8.60 %	8.80%	7.67%	9.65%	8.32%
R	-1.00%	-3.00%	-3.00%	2.52%	2.96%	5.73%
SR	-0.272	-0.460	-0.467	0.202	0.200	0.575
N	13	58	57	11	12	8
AAR	1.35%	-1.57%	-1.59%	5.58%	9.31%	9.16%

The previous analysis already showed the advantage of using WS information when optimizing robust portfolios. However, here the optimized portfolios are compared using efficiency stochastic analysis using CCDEA associated to the hierarchical grouping and Sharpe's proposal for optimized Comparative portfolios using classic and deterministic models for portfolio optimization.

It is noteworthy that the Comparative Markowitz portfolios (MC) and Comparative Sharpe portfolios (SC) had only 13 and 58 assets after the optimization, and S_{Rs} equal to -0.272 and -0.460, respectively. Again, we see that WS optimized portfolios perform better.

Portfolios WS-1 and WS-11 had beta values equal to 0.458 and 0.429, while Comparative portfolios MC and SC had beta values equal to 0.475 and 0.702.

The cumulative return for the portfolios (AAR) was obtained based on the information from the period considered in the validation. Again, the results shown in Table 8 show abnormal returns. Figure 6 shows the accumulated returns for the analyzed portfolios.

Figure 6 (a) shows the accumulated returns of the proposed portfolios (when a risk criterion (α_i)=60% is adopted) and the comparative portfolios: MC, SC, and the Ibovespa Brazil Sao Paulo Stock

Exchange Index (BVSP). Figure 6 (b) shows the accumulated return for proposed portfolios (when a risk criterion (α_i)=50% is adopted) and the same comparative portfolios.

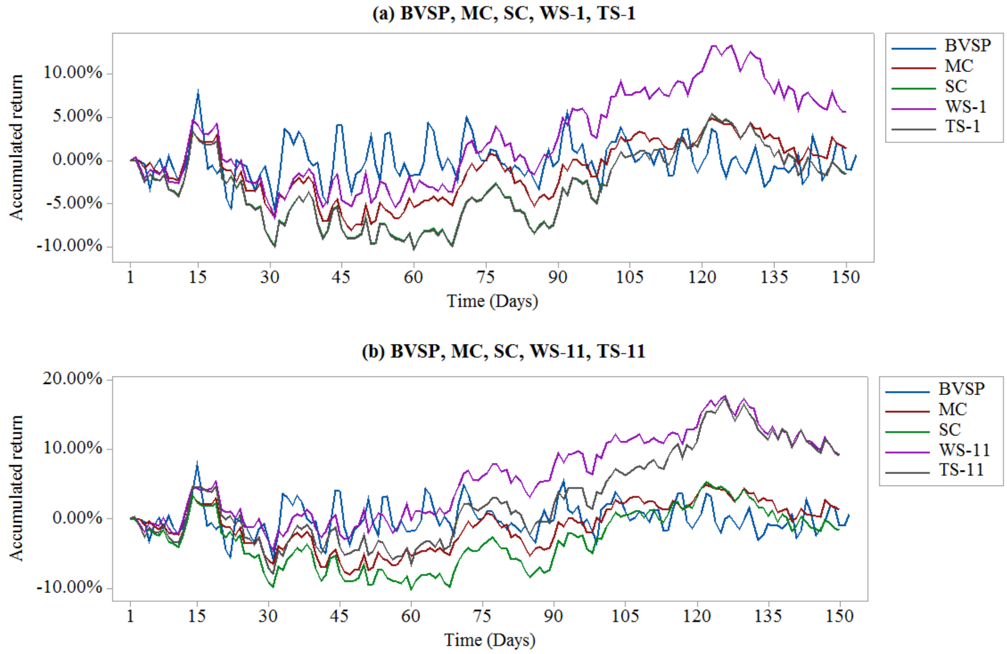

FIGURE 6. Accumulated return of proposed and comparative portfolios.

After presenting the accumulated return graphs, we performed a statistical test for the comparison between the series of accumulated returns obtained for the analyzed portfolios. We used the Kruskal-Wallis test (one-way analysis of variance), which allowed us to determine whether the medians of two or more groups differed. The *P-value* results obtained were less than 0.05. We see that the accumulated returns of WS optimized portfolios (WS-1 and WS-11) are statistically higher than the accumulated returns of TS optimized portfolios (TS-1 and TS-11), when using the same risk criterion. They are also statistically higher than the accumulated returns of the BVSP, MC and SC comparative portfolios.

Figure 7 shows the boxplot diagram of the analyzed portfolios.

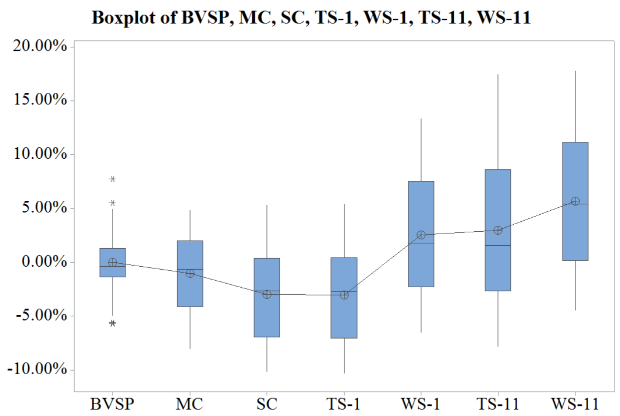

FIGURE 7. Boxplot diagram for proposed and comparative portfolios.

6. CONCLUSIONS

We were able to reduce the search space when identifying efficient assets. This study used stochastic information for the different adopted variables. Subsequently, the efficient assets were submitted to approaches that promoted ideal asset allocation within the portfolios. It is interesting to note that both commonly used and fundamentalist variables were considered in the asset allocation.

We identified that it is possible to represent more flexible models by considering the different risk profiles of investors (conservative or risky) by varying the probability of meeting the constraints ($1-\alpha i$) in the CCDEA model. The more rigorous one is when fulfilling these constraints, the smaller the resulting values in the efficiency analysis, resulting in fewer efficient assets.

Varying the risk criterion (α_i) between the 0.5 and 0.6, which varied in units of 0.01, allowed us to better discriminate the analysis units. We were able to build eleven portfolios for each adopted scenario, TS-1 to TS-11 for Total Market State (TS) information, and WS-1 to WS-11 for Worst-Case Scenario (WS) information. Additionally, we used Comparative Markowitz (MC) and Comparative Sharpe portfolios (SC), optimized from classic portfolio optimization models, as proposed by Markowitz (1952), and Sharpe (1963). These portfolios were supplied with information without any regard to differentiation in market states.

When portfolios TS-1 to TS-11 were compared to portfolios WS-1 to WS-11, WS optimized portfolios performed better according to the Sharpe ratio (S_R) for the validation period. These results are in alignment with the notion that robust optimization for building robust portfolios should specifically concentrate on information obtained in Worst-Case scenarios (Kim et al., 2015).

In general, portfolios formed using our proposed method (WS-1 to WS-11) performed better as measured by the Sharpe ratio (S_R), and according to the accumulation of abnormal returns in the validation period. The averages of the series of abnormal returns were statistically higher than for the comparative portfolios.

Another important fact worth emphasizing is that the WS optimized portfolios had lower beta values compared to TS optimized portfolios and the MC and SC Comparative portfolios. Robust optimization portfolios tend to comprise assets with lower beta values that perform well under any market state (bull or bear market).

The Data Enveloping Analysis (DEA) is already being used for portfolio optimization, and we can see that the stochastic approach of Chance Constrained Data Envelopment Analysis (CCDEA) helps reduce the search space for efficient assets considering different variables. Hierarchical Clustering allowed us to better discriminate the data submitted to the CCDEA model, even with a reduction to the risk criterion.

Finally, we suggested that this proposed method be applied to different stock markets, to more mature stock markets, to more assets, with data coming from longer historical series, and to validate data from different periods. We also suggest that future research compare this proposed method with other methods of robust portfolio optimization.

Acknowledgments

This paper is part of a project GEOCEP that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 870245. The authors would like to thank the Brazilian National Council for Scientific and Technological Development - CNPq Brazil (Processes

406769/2018-4, 308021/2019-3, 302751/2020-3, and 303909/2020-0), the Minas Gerais Research Funding Foundation - FAPEMIG Brazil (Process APQ-00378-21), and the Coordination for the Improvement of Higher Education Personnel - Capes Brazil. The authors also thank the Czech Science Foundation - Czech Republic (grant number 22-19617S) for the financial support and research incentive. The views expressed here are those of the authors and not necessarily those of our institutions. All remaining errors are solely our responsibility.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

Adam, L., & Branda, M. (2021). Risk-aversion in data envelopment analysis models with diversification. *Omega (United Kingdom)*, *102, 102338*. https://doi.org/10.1016/j.omega.2020.102338

Amin, G. R., & Hajjami, M. (2020). Improving DEA cross-efficiency optimization in portfolio selection. *Expert Systems with Applications*, 114280. https://doi.org/10.1016/j.eswa.2020.114280

Ashrafi, H., & Thiele, C. (2021). *A study of robust portfolio optimization with European options using polyhedral uncertainty sets*. *8*. https://doi.org/10.1016/j.orp.2021.100178

Azadi, M., Jafarian, M., Saen, R. F., & Mirhedayatian, S. M. (2015). A new fuzzy DEA model for evaluation of efficiency and effectiveness of suppliers in sustainable supply chain management context. *Computers and Operations Research*, *54*, 274–285. https://doi.org/10.1016/j.cor.2014.03.002

Azadi, M., & Saen, R. F. (2012). Developing a new chance-constrained DEA model for suppliers selection in the presence of undesirable outputs. *International Journal of Operational Research*, *13*(1), 44–66. https://doi.org/10.1504/IJOR.2012.044027

Azadi, M., Saen, R. F., & Tavana, M. (2012). Supplier selection using chance-constrained data envelopment analysis with non-discretionary factors and stochastic data. *International Journal of Industrial and Systems Engineering*, *10*(2), 167–196. https://doi.org/10.1504/IJISE.2012.045179

Baltas, I., & Yannacopoulos, A. N. (2019). Portfolio management in a stochastic factor model under the existence of private information. *IMA Journal of Management Mathematics*, *30*(1), 77–103. https://doi.org/10.1093/imaman/dpx012

Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28

Chakrabarti, D. (2021). Parameter-free robust optimization for the maximum-Sharpe portfolio problem. *European Journal of Operational Research*. https://doi.org/10.1016/j.ejor.2020.11.052

Charnes, A., & Cooper, W. W. (1959). *Chance-Constrained Programming*. *August 2015*.

Charnes, A., & Cooper, W. W. (1963). Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints. *Operations Research*, *11*(1), 18–39. https://doi.org/10.1287/opre.11.1.18

Chen, C., Liu, D., Xian, L., Pan, L., Wang, L., Yang, M., & Quan, L. (2020). Best-case scenario robust portfolio for energy stock market. *Energy*, *213*, 118664. https://doi.org/10.1016/j.energy.2020.118664

Choi, H. S., & Min, D. (2017). Efficiency of well-diversified portfolios: Evidence from data envelopment analysis. *Omega (United Kingdom)*, *73*, 104–113. https://doi.org/10.1016/j.omega.2016.12.008

Cook, W., & Zhu, J. (2014). *Data Envelopment Analysis – A Handbook on the modeling of internal structures and networks*. Springer International Publishing.

Cooper, W. W., Huang, Z., & Li, S. X. (1996). Satisficing DEA models under chance constraints. *Annals of Operations Research*, *66*, 279–295. https://doi.org/10.1007/BF02187302

Cooper, W. W., Seiford, L. M., & Tone, K. (2002). Data Envelopment Analysis: A comprehensive text with models, applications, references and DEA-Solver Software. In K. A. Publishers (Ed.), *Journal of Chemical Information and Modeling* (1st ed., Vol. 53, Issue 9).

Cooper, W. W., Seiford, L., & Tone, K. (2007). *Data envelopment analysis: a comprehensive text with models, application, references and DEA-Solver Software*. Springer International Publishing.

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. https://doi.org/10.1016/j.ejor.2010.06.027

Emrouznejad, A., & Tavana, M. (2014). *Peformance Measurement with Fuzzy Data Envelopment Analysis*. Springer International Publishing.

Fabozzi, F. J., Huang, D., & Zhou, G. (2010). Robust portfolios: Contributions from operations research and finance. *Annals of Operations Research*, *176*(1), 191–220. https://doi.org/10.1007/s10479-009-0515-6

Fabozzi, F. J., Kolm, P. N., Pachamanova, D. A., & Focardi, S. M. (2007). Robust Portfolio Optimization. *The Journal of Portfolio Management*, *33*(3), 40–48. https://doi.org/10.3905/jpm.2007.684751

Farzipoor Saen, R., & Azadi, M. (2011). A chance-constrained data envelopment analysis approach for strategy selection. *Journal of Modelling in Management*, *6*(2), 200–214. https://doi.org/10.1108/17465661111149584

Homm, U., & Pigorsch, C. (2012). Beyond the Sharpe ratio: An application of the Aumann-Serrano index to performance measurement. *Journal of Banking and Finance*, *36*(8), 2274–2284. https://doi.org/10.1016/j.jbankfin.2012.04.005

Jin, J., Zhou, D., & Zhou, P. (2014). Measuring environmental performance with stochastic environmental DEA: The case of APEC economies. *Economic Modelling*, *38*, 80–86. https://doi.org/10.1016/j.econmod.2013.12.017

Kao, C. (2014). Efficiency decomposition for general multi-stage systems in data envelopment analysis. *European Journal of Operational Research*, *232*(1), 117–124. https://doi.org/10.1016/j.ejor.2013.07.012

Kim, J. H., Kim, W. C., & Fabozzi, F. J. (2014). Recent Developments in Robust Portfolios with a Worst-Case Approach. *Journal of Optimization Theory and Applications*, *161*(1), 103–121. https://doi.org/10.1007/s10957-013-0329-1

Kim, J. H., Kim, W. C., & Fabozzi, F. J. (2018). Recent advancements in robust optimization for investment management. *Annals of Operations Research*, *266*(1–2), 183–198. https://doi.org/10.1007/s10479-017-2573-5

Kim, W. C., Kim, J. H., Mulvey, J. M., & Fabozzi, F. J. (2015). Focusing on the worst state for robust investing. *International Review of Financial Analysis*, *39*, 19–31. https://doi.org/10.1016/j.irfa.2015.02.001

Kourtis, A. (2016). The Sharpe ratio of estimated efficient portfolios. *Finance Research Letters*, *17*, 72–78. https://doi.org/10.1016/j.frl.2016.01.009

Leung, P.-L., Ng, H.-Y., & Wong, W.-K. (2012). An improved estimation to make Markowitz's portfolio optimization theory users friendly and estimation accurate with application on the US stock market investment. *European Journal of Operational Research*, *222*(1), 85–95. https://doi.org/10.1016/j.ejor.2012.04.003

Levy, H., & Levy, M. (2014). The benefits of differential variance-based constraints in portfolio optimization. *European Journal of Operational Research*, *234*(2), 372–381. https://doi.org/10.1016/j.ejor.2013.04.019

Lim, S., Oh, K. W., & Zhu, J. (2014). Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. *European Journal of Operational Research*, *236*(1), 361–368. https://doi.org/10.1016/j.ejor.2013.12.002

Markowitz, H. (1952). PORTFOLIO SELECTION*. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research*, *234*(2), 346–355. https://doi.org/10.1016/j.ejor.2012.08.023

Myers, R., Montgomery, D. ., & Anderson-Cook, C. (2009). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments* (John Wiley & Sons (ed.); 3rd ed.). Hoboken.

Powers, J., & Mcmullen, P. (2002). No Title. *Journal of Business and Management*, *7*(2), 31–42.

Rotela Junior, P, Pamplona, E. O., & Salomon, F. R. (2014). Otimização de Portfólios: Análise de Eficiência. *Revista de Administração de Empresas*, *54*(4), 405–413. https://doi.org/10.1590/S0034-759020140406

Rotela Junior, Paulo, Pamplona, E. de O., Rocha, L. C. S., Valerio, V. E. de M., & Paiva, A. P. (2015). Stochastic portfolio optimization using efficiency evaluation. *Management Decision*, *53*(8), 1698–1713. https://doi.org/10.1108/MD-11-2014-0644

Sehgal, R., & Mehra, A. (2020). Robust portfolio optimization with second order stochastic dominance constraints. *Computers and Industrial Engineering*, *144*(January), 106396. https://doi.org/10.1016/j.cie.2020.106396

Sengupta, J. K. (1987). Data envelopment analysis for efficiency measurement in the stochastic case. *Computers and Operations Research*, *14*(2), 117–129. https://doi.org/10.1016/0305-0548(87)90004-9

Sharpe, W. F. (1963). A Simplified Model for Portfolio Analysis. *Management Science*, *9*(2), 277–293. https://doi.org/10.1287/mnsc.9.2.277

Sharpe, W. F. (1964). Capital Asset Prices: a Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Shi, H. L., & Wang, Y. M. (2020). A Merger and Acquisition Matching Method That Considers Irrational Behavior from a Performance Perspective. *IEEE Access*, *8*, 45726–45737. https://doi.org/10.1109/ACCESS.2020.2976608

Siriopoulos, C., & Tziogkidis, P. (2010). How do Greek banking institutions react after significant events?-A DEA approach. *Omega*, *38*(5), 294–308. https://doi.org/10.1016/j.omega.2009.06.001

Won, J.-H., & Kim, S.-J. (2020). Robust trade-off portfolio selection. *Optimization and Engineering*, *21*(3), 867–904. https://doi.org/10.1007/s11081-020-09485-z

Xidonas, P., Mavrotas, G., Hassapis, C., & Zopounidis, C. (2017). Robust multiobjective portfolio optimization: A minimax regret approach. *European Journal of Operational Research*, *262*(1), 299–305.

https://doi.org/10.1016/j.ejor.2017.03.041

Xidonas, P., Steuer, R., & Hassapis, C. (2020). Robust portfolio optimization: a categorized bibliographic review. *Annals of Operations Research*, *292*(1), 533–552. https://doi.org/10.1007/s10479-020-03630-8

Yu, J.-R., Chiou, W.-J. P., Lee, W.-Y., & Chuang, T.-Y. (2019). Realized performance of robust portfolios: Worst-case Omega vs. CVaR-related models. *Computers & Operations Research*, *104*, 239–255. https://doi.org/10.1016/j.cor.2018.12.004

IES Working Paper Series

2022

1. Klara Kantova: *Parental Involvement and Education Outcomes of Their Children*
2. Gabriel Nasser, Doile de Doyle, Paulo Rotella Junior, Luiz Célio Souza Rocha, Priscila França Gonzaga Carneiro, Rogério Santana Peruchi, Karel Janda, Giancarlo Aquila: *Impact of Regulatory Changes on Economic Feasibility of Distributed Generation Solar Units*
3. Paulo Rotella Junior, Luiz Célio Souza Rocha, Rogério Santana Peruchi, Giancarlo Aquila, Karel Janda, Edson de Oliveira Pamplona: *Robust Portfolio Optimization: A Stochastic Evaluation of Worst-Case Scenarios*

All papers can be downloaded at: http://ies.fsv.cuni.cz

Univerzita Karlova v Praze, Fakulta sociálních věd
Institut ekonomických studií [UK FSV – IES] Praha 1, Opletalova 26
E-mail : ies@fsv.cuni.cz http://ies.fsv.cuni.cz