

# DEA Portfolio Modeling

## The Case of Socially Responsible Investing

Stephan M. Gasser\*  
Margarethe Rammerstorfer<sup>†</sup> Karl Weinmayer<sup>‡</sup>

January 15, 2019

### Abstract

In this paper, we expand the literature on multi-criteria portfolio modeling for social responsible investments using multi-directional efficiency analysis (MEA). We apply a positive screening according to MEA efficiency scores, but also exploit the information contained in the efficiency score directly in order to compute portfolio weights. We conduct a broad empirical analysis testing multiple portfolio strategies based on public equity market data of social responsible investments from the USA going back to 2005. We show that the explicit consideration of a social responsibility variable in the MEA portfolio models has a positive effect on the financial and social performance of all asset allocations strategies. Furthermore we find that portfolios constructed using multi-directional efficiency analysis outperform the mean-variance portfolios and that these strategies are also strictly preferred to a naive portfolio strategy since they offer the same financial performance, while surpassing it in terms of social performance. These results clearly outline the benefits of MEA portfolio modeling, not only for socially responsible investors, but also traditional investors.

**Keywords:** Socially Responsible Investments, Data Envelopment Analysis, DEA, Portfolio Optimization, International Financial Markets, Multi-directional Efficiency Analysis, Sustainable Finance

**JEL classification:** G11, G15, A13

---

\*WU (Vienna University of Economics and Business), Department of Finance, Accounting and Statistics, Welthandelsplatz 1, 1020 Vienna, Austria; *email:* stephan.gasser@wu.ac.at

<sup>†</sup>WU (Vienna University of Economics and Business), Department of Finance, Accounting and Statistics, Welthandelsplatz 1, 1020 Vienna, Austria; *email:* margarethe.rammerstorfer@wu.ac.at

<sup>‡</sup>MU (Modul University Vienna), Department of International Management, Am Kahlenberg 1, 1190 Vienna, Austria; *email:* karl.weinmayer@modul.ac.at

# 1 Introduction

In recent years, an increasing number of both institutional and individual investors have started to change their views with regards to their investment targets. They no longer solely care about financial returns and risk but also about the social responsibility of their investment activities, appealing to the already popular sentiment of “doing well while doing good” (see e.g. Hamilton et al., 1993). As a result, roughly \$8.72 trillion were invested in US-domiciled socially responsible assets at the beginning of 2016, an increase of 33% in comparison to 2014 (US SIF, 2016).

Consequently, socially responsible investing has moved into the focus of the academic sector, where a diverse range of subtopics is now being covered. One of the largest associated areas of interest is the subject of constructing socially responsible investment portfolios. In this field numerous papers have been published developing models that allow for the incorporation of multiple preferences regarding return, risk and social responsibility in the portfolio decision making process. This paper expands the literature on socially responsible investments and multi-criteria portfolio modeling by relying on multi-directional efficiency analysis (MEA), an extension of data envelopment analysis (DEA).

In the context of asset pricing and management, DEA has proven to be a valuable stock selection tool. Usually, DEA portfolio management follows a two-stage process. First, DEA is applied in order to identify highly efficient stocks, this is usually referred to as stock screening. Second, a portfolio optimization, such as a mean-variance optimization (see Markowitz, 1952) is applied to compute portfolio weights. This second step typically aims to maximize financial performance only, which implies that this optimization is potentially in conflict with step one if a multi-objective orientation - also including non-financial variables - is considered during the DEA. With the analysis proposed in this article, we

overcome this problem by preserving the overall preferences defined via the MEA variables and incorporating them directly into the asset allocation of the final portfolio.

This paper explores how the proposed MEA methodology performs in a variety of asset allocation strategies in an out-of-sample analysis compared to selected, well established benchmarks. The empirical analysis is based on a data set of 428 US public equities, with the observation period encompassing the years 2005 to 2014. We use quarterly stock price data (Thomson Reuters Datastream) and Thomson Reuters ASSET4 ESG (environment, social, governance) scores as social responsibility measure. Furthermore, all stocks are grouped into sectors according to the Global Industry Classification Standard (GICS). We compute quarterly out-of-sample portfolio returns and volatility, Sharpe Ratios, ESG scores and Delta Ratios (a ratio relating social responsibility and risk similar to the Sharpe Ratio, see Gasser et al., 2017).

We find that the incorporation of a social responsibility variable in the MEA improves the financial and social performance of the resulting portfolios, such that all our ESG-based models are able to outperform the mean-variance portfolio and are also strictly preferred to the naive portfolio benchmark. These results clearly outline the benefits of MEA portfolio modeling and also provide validation to argue that not only socially responsible investors but also conventional investors, who focus solely on risk and return, should be interested in social responsibility measures when constructing their portfolios.

The structure of his article is organized as follows: in Section 2, we describe the existing literature both on building portfolios with a view to social responsibility considerations as well as the related applications of DEA and MEA. In Section 3, the general MEA framework is described, while Section 4 details the data set and provides a descriptive data analysis. Section 5 describes the results of our empirical analysis before Section 6 concludes.

## 2 Literature Review

The study conducted in this paper is connected to two major strands of literature. The first one relates to the fairly large number of papers that have been published on the subject of socially responsible investing over the last 20 years, aiming to investigate the empirical relationship between financial performance and social responsibility. More recently it has been observed that investors, institutions and foundations wish to actively incorporate this dimension into their investment decision making process. In light of this development, an increasing number of papers have been published focusing on formulating and introducing theoretical models that allow for an asset allocation that incorporates measures for return, risk and social responsibility at the same time.

Bilbao-Terol et al. (2012) for example, introduce a goal programming model for SRI portfolio selection that aims to enable investors to match ethical and financial preferences. On the basis of a UK mutual funds data set they demonstrate that investor's risk attitudes tend to influence the loss of return as a result of choosing SRIs. Ballesterro et al. (2012) and Bilbao-Terol et al. (2013) also focus on SRIs and propose different models to incorporate investor preferences into the portfolio optimization process. Using their "Financial-Ethical Bi-Criteria model", Ballesterro et al. find that ethical investments are accompanied by risk exposure increases, while the results of Bilbao-Terol et al. on the basis of their "Hedonic Price Method" and a data set of 160 investment funds suggest that the financial penalties associated with SRIs are relatively minor for highly risk-averse investors. Even more recently, Hirschberger et al. (2013) develop a multiparametric algorithm for the computation of the non-dominated set of portfolios in a tri-criterion optimization, and Gasser et al. (2017) propose a Markowitz model modification to set up a three-dimensional capital allocation plane illustrating the complete set of feasible optimal portfolios on the basis of a return/risk/social responsibility optimization.

The second strand of literature relates to Data Envelopment Analysis (DEA), which has already been implemented in the context of performance measures and as a tool for asset selection in general, but has not yet been employed as a portfolio management and asset allocation model for socially responsible investing. The comparison of companies and industries according to their DEA efficiencies have a long history in several different research areas. An excellent overview on using stochastic frontier analysis can be found in Lovell (1993). Data envelopment analysis, a special form of frontier analysis, traces back to Charnes et al. (1978) and combines operations research, econometrics and management science aspects within one single efficiency measure. Even today, the number of areas where DEA is applied is still increasing. In the financial sector the fields of application of DEA are numerous, from measuring the performance of specific assets like funds or stocks to portfolio selection on the basis of DEA's efficiency scores. In their article on DEA for the performance assessment of mutual funds, Basso and Funari (2016) provide an overview of most of the recent applications of DEA for asset pricing purposes.

Powers and McMullen (2000) use DEA for stock selection based on multiple financial criteria, where the DEA efficiency scores of individual assets are determined via eight variables: 1-, 3-, 5-, and 10-year returns as well as earnings per share (outputs) price-to-earnings ratio, beta, and sigma (inputs). Most interestingly, they find that highly efficient stocks can be classified as quite robust to unfavorable changes of the considered variables. For the group of all inefficient firms the authors also analyze how much certain variables need to change in order for an inefficient company to become efficient. They conclude that DEA is helpful to distinguish strong performers and others and thus is a good companion to common selection models. In 2008, Chen uses DEA for stock selection at the Taiwan stock market and obtains similar results, such that he confirms the superior performance of firms that were selected using DEA.

Edirisinghe and Zhang (2008) use financial statement data and develop a relative financial strength (RFS) indicator on the basis of a DEA approach that visualizes firms' fundamental strength. They show further that the RFS indicator can be used for stock selection purposes within a mean-variance optimization.

Dia (2009) proposes a four-step methodology for the portfolio selection of assets, in which portfolios are formed using a model that optimizes the weighted sum of all the assets' efficiency scores with respect to investor preferences.

Pätäri et al. (2010) compose portfolios using DEA scale efficiency for Finish non-financial stocks. The performance of those portfolios is evaluated on basis of average return and risk-performance metrics. They find that DEA efficiency scores have a positive effect on the value of the resulting portfolio and that this effect is most evident for shorter investment horizons. They further conclude that the use of DEA is useful if several variables need to be taken into consideration or if the number of stocks in the sample is large. In 2012, Pätäri et al. again examine the applicability of data envelopment analysis as a selection tool for equity portfolios. Here, they combine value investing and momentum investing for constructing portfolios for the sample period of 1994 until 2010. The performance is measured on the basis of returns and several risk-adjusted performance metrics. They show that the DEA approach improves portfolio performance in a statistically significant way.

For the Malaysian stock market, Ismail et al. (2012) investigate the effectiveness of a DEA model on portfolio selection for investors over long horizons and use efficiency scores to select only efficient firms. Their "Technical Efficiency Portfolio" seems to produce significantly higher cumulative abnormal returns over a 36-month holding period than their comparatives. Similar to the earlier studies, they conclude that DEA is an appropriate method for asset selection as it leads to superior portfolio performance.

Bahrani and Khedri (2013) create a portfolio of efficient companies by using Data Envelopment Analysis on a data set from Teheran stock exchange. They find that it is not possible to generate a return beyond the average return of the market by using the constant returns to scale of Charnes et al. (1978). However, using the variable returns to scale model of Banker et al. (1984) improves the performance of the resulting portfolio.

Lim et al. (2014) use DEA cross-efficiency evaluation (i.e. cross efficiencies between stocks) for portfolio selection. With a data set from the Korean stock market, the resulting portfolios yield higher risk-adjusted returns than various benchmark portfolios.

All in all, the above mentioned articles show that Data Envelopment Analysis provides a useful tool for portfolio management. Its flexibility in the selection of relevant parameters makes this method a perfect fit for application to socially responsible investing. Previous studies use DEA in combination with financial data only and focus solely on the optimization of financial performance. Therefore, it is a novel approach to apply DEA to socially responsible investing by explicitly incorporating a social responsibility measure as a non-financial variable for determining overall firm efficiency with respect to both dimensions. We focus here on multi-directional efficiency analysis, a recent extension of DEA, which has not been applied to portfolio management before and which features some useful properties compared to other DEA models with respect to the weighting of variables in determining efficiency scores, the non-negativity of variables and adjustment of inputs and outputs via different scaling factors.

We also investigate the usefulness of efficiency weighting - the direct transformation of efficiency scores into portfolio weights - and explore the relevance of including less efficient firms in the portfolio decision process, since previous studies use DEA efficiency scores purely for asset screening based on its highest efficiency.

### 3 Methodology

Data Envelopment Analysis is a nonparametric approach to measure the relative efficiency of different decision making units or firms. Based on the seminal work of Koopmans (1951) and Debreu (1951) on linear programming, Farrell (1957) developed a procedure to analyse the efficiency of different operating units with respect to several input and output variables. He distinguished the efficiency of a certain operating unit into an efficiency coming from the technical part (technical efficiency) and a second part - the so called allocative efficiency - which describes the efficient allocation of resources on a certain output value (or maximization of output given a certain amount of input). This DEA methodology implies a radial scaling for the input factors (input orientation) or output factors (output orientation) or both (input- and output orientation) and efficiency is measured based on an weighted ratio of outputs over inputs.

For the sake of this article we intend to disentangle improvement potential for inputs and outputs separately without assuming a specific relationship between the factors included or making any assumptions about the tradeoff between the improvement (potential) of factors. We rely on a fairly recent extension of DEA called multi-directional efficiency analysis (MEA), which was first introduced in Bogetoft and Hougaard (1999) and further developed in Tone (2001) and Asmild et al. (2003). In contrast to DEA, multi-directional efficiency allows the analysis of the improvement potential for each of the factors included in the analysis separately. Since we are interested in the improvement potential of both inputs and outputs, we therefore specify a MEA model with mixed orientation in the following.

Let  $N$  be the number of DMUs analysed within a certain sector or country in each period  $t = 1, \dots, T$ .<sup>1</sup> Let DMU $_j$  with  $j \in \mathbb{N}$  at time  $t$  produce outputs  $y_{r,j}^t$ , with  $r = 1, \dots, m$  by using inputs  $x_{i,j}^t$  with  $i = 1, \dots, n$ . A certain DMU $_j$  under analysis is designated

---

<sup>1</sup>In this analysis we refer to quarterly observations as explained in section 4.



as  $DMU_o$  with  $o = 1, \dots, N$  and the production plan  $(x_o^t, y_o^t)$ . In order to analyze the improvement potential for  $DMU_o$ , an ideal reference point  $(x_o^{t*}, y_o^{t*})$  is detected for each point  $t$  by solving a system of linear programs for each variable included. The ideal reference point for each input variable  $x_{i,j}^t$  is given by

$$\begin{aligned}
& \text{minimize} && d_{i,o}^t, \\
& \text{subject to} && \sum_{j=1}^N \lambda_j x_{i,j}^t \leq d_{i,o}^t, \\
& && \sum_{j=1}^N \lambda_j x_{-i,j}^t \leq x_{-i,o}^t, \quad -i = 1, \dots, i-1, i+1, \dots, n, \\
& && \sum_{j=1}^N \lambda_j y_{r,j}^t \geq y_{r,o}^t, \quad r = 1, \dots, m, \\
& && \lambda_j \geq 0, \quad j = 1, \dots, N,
\end{aligned} \tag{1}$$

and for each output variable  $y_{r,j}^t$  by

$$\begin{aligned}
& \text{maximize} && \delta_{r,o}^t, \\
& \text{subject to} && \sum_{j=1}^N \lambda_j x_{i,j}^t \leq x_{i,o}^t, \quad i = 1, \dots, n, \\
& && \sum_{j=1}^N \lambda_j y_{r,j}^t \geq \delta_{r,o}^t, \\
& && \sum_{j=1}^N \lambda_j y_{-r,j}^t \geq y_{-r,o}^t, \quad -r = 1, \dots, r-1, r+1, \dots, m, \\
& && \lambda_j \geq 0, \quad j = 1, \dots, N.
\end{aligned} \tag{2}$$

The ideal reference point corresponds to the maximum improvement potential inherent in each input and output variable separately.<sup>2</sup> The solution to this system of linear programs is generally outside the production set  $P$  and is given by  $(\mathbf{d}_o^{t*}, \boldsymbol{\delta}_o^{t*}) \notin P$ . This implies that it is not possible to implement the ideal production plan due to technological

---

<sup>2</sup>The higher the distance to the optimal reference point, the larger is the improvement potential.

boundaries, which are given by the set of DMUs. The technological constraints are represented by the efficient frontier, which serves as a benchmark for measuring the relative efficiency of all DMUs. However, movement in the direction of this ideal is still possible. The distance between the current production plan and the efficient frontier represents the potential improvement direction  $\beta_o^t$ , which is given by

$$\begin{aligned}
& \text{maximize} && \beta_o^t, \\
& \text{subject to} && \sum_{j=1}^N \lambda_j x_{i,j}^t \leq x_{i,o}^t - \beta_o^t (x_{i,o}^t - d_{i,o}^{t*}), \quad i = 1, \dots, n, \\
& && \sum_{j=1}^N \lambda_j y_{r,j}^t \geq y_{r,o}^t + \beta_o^t (\delta_{r,o}^{t*} - y_{r,o}^t), \quad r = 1, \dots, m, \\
& && \lambda_j \geq 0, \quad j = 1, \dots, N.
\end{aligned} \tag{3}$$

The solution  $(\lambda^*, \beta^*)$  gives the realizable improvement potential compared to the benchmarks spanning the efficient frontier with  $\beta_o^{t*} = [0, 1]$ . A value of 0 implies that no further improvement is possible and this DMU is situated on the efficient frontier and even helps defining it. In order to have a straightforward interpretation of our results, we transform the MEA scores into Farrell efficiency scores  $\eta_j^t$  with 1 representing the most efficient firms.

In order to use MEA for portfolio management, we select common input and output variables, that have been introduced in the existing literature on portfolio selection and performance evaluation (see e.g. Branda and Kopa, 2014; Chen and Lin, 2006; Gregoriou, 2006; Lin, 2009; Pendaraki, 2012; Zhao and Shi, 2010). Therefore, we consider various risk measures as input factors and expected return as well a social responsibility measure on the output side. The factors incorporated in the different MEA model specifications are summarized in Table 1.

MEA model	Inputs	Outputs
Model 1	Volatility ( $\sigma$ )	Expected return ( $\bar{\mu}$ )
Model 2	Volatility ( $\sigma$ )	Expected return ( $\bar{\mu}$ ), ESG Score ( $\theta$ )
Model 3	Volatility ( $\sigma$ ), Value at Risk (VaR, 95%), Expected Shortfall (ES, 95%)	Expected return ( $\bar{\mu}$ ), ESG Score ( $\theta$ )

**Table 1:** Overview of input and output variables

Model 1 represents the base scenario and only considers volatility and expected return. Model 2 is extended by considering the ESG Score as social responsibility measure (see section 4), whereas for Model 3, further risk measures are included.

For each model we implement three asset allocation strategies, which differ from each other in how the information gathered from the MEA models is used. In Strategy 1 efficiency-weighted portfolios are constructed by directly transforming efficiency scores  $\eta_j^t$  into portfolio weights  $w_j^t$ :

$$w_j^t = \frac{\eta_j^t}{\sum_{j=1}^N \eta_j^t} \quad (4)$$

For Strategy 2 we use the efficiency scores as basis for a screening, in which only the fully efficient firms are taken into consideration (efficiency-screening) and an efficiency-weighted portfolio is implemented.<sup>3</sup>

In Strategy 3 an efficiency-screening is again applied and followed by a mean-variance optimization in the second step, which is the standard approach proposed in the literature on DEA portfolio management (see Section 1).

4

<sup>3</sup>A screening process with respect to certain efficiency levels would also be possible but since the decision where to cut off the sample would be arbitrary, we only focus on fully efficient firms (i.e. firms with an efficiency score of 1 in our MEA model, which results in an equally-weighted portfolio with  $w_j = \frac{1}{N}$ , if efficiency-weighting is applied).

<sup>4</sup>Since the Strategies 2 and 3 both rely on screening techniques, the application of these strategies might be more restricted compared to Strategy 1, since the reduction in the asset universe could lead to adverse diversification effects.

## 4 Data

For the empirical analysis in this paper, we rely on a comprehensive data set of US public equities. Since an unbiased and independent social responsibility measure is a prerequisite, we use the constituent list of the Thomson Reuters ASSET4 database<sup>5</sup> to build the initial basis for our data set. This database provides us with ESG scores (see Section 1) on hundreds of the largest publicly traded stocks in the US. The ESG score is an aggregate score indicating a company's total social responsibility on a scale between 0 (lowest possible ESG rating) and 100 (best-rated company) and makes companies comparable across markets.

Since ESG scores are updated on a quarterly basis, we also obtain quarterly stock prices from Thomson Reuters Datastream. All stocks are then grouped into sectors according to their GICS (Global Industry Classification Standard) industry codes. Our observation period covers 10 years (40 quarters), ranging from 2005 to 2014 and only sectors for which more than 10 firms are available are included in the final data set.<sup>6</sup> Furthermore, we use the US 3 Month Treasury Bill Rate as a risk-free rate. Table 2 provides descriptive statistics for the data set.

## 5 Results

Since we test three different MEA model specifications as indicated in Table 1, we will first focus on the results of each model specification individually, which allows us to observe the performance and usefulness of the different asset allocation strategies. Within each MEA model, we therefore compare the three portfolio allocation strategies among each other

---

<sup>5</sup>The ASSET4 database covers public equities from major equity indices worldwide. It consists of 250 key performance indicators from four category pillars (Economic, Environmental, Social and Corporate Governance Performance).(Thomson Reuters, 2012)

<sup>6</sup>This is a necessary restriction, since according to Golany and Roll (1989) the number of comparable units in DEA data sets should be twice the number of in- and outputs in order to ensure proper behavior of the efficiency model.

Industry Sector (GICS Code)	# of Stocks		$\mu$ (%)	$\sigma$ (%)	$\theta$	MV (M)
Energy (10)	33	Mean	2.89%	14.17%	63.04	19,941.16
		Median	3.39%	15.37%	64.00	19,267.45
		Max	35.38%	18.82%	78.47	29,447.12
		Min	-35.53%	6.24%	41.01	11,784.01
Materials (15)	34	Mean	3.28%	15.50%	62.75	15,890.62
		Median	3.74%	16.66%	64.91	14,872.29
		Max	42.79%	21.64%	77.63	27,456.85
		Min	-37.88%	7.99%	40.60	7,986.21
Industrials (20)	60	Mean	3.17%	15.96%	59.74	30,058.72
		Median	3.01%	17.36%	61.31	28,927.55
		Max	48.90%	22.95%	76.57	48,406.33
		Min	-37.79%	7.72%	37.69	15,773.71
Consumer Discretionary (25)	71	Mean	2.58%	16.05%	59.31	23,400.83
		Median	2.98%	17.38%	60.32	22,206.24
		Max	45.34%	22.02%	74.78	39,032.47
		Min	-39.33%	8.06%	40.70	10,746.08
Consumer Staples (30)	20	Mean	2.89%	13.59%	55.09	11,154.87
		Median	3.61%	14.78%	56.61	10,851.75
		Max	37.35%	17.60%	73.27	18,305.22
		Min	-35.76%	7.00%	32.65	5,459.13
Health Care (35)	41	Mean	2.65%	14.66%	60.01	25,634.75
		Median	2.78%	16.14%	61.65	25,154.05
		Max	46.80%	20.39%	75.25	41,174.05
		Min	-35.70%	6.51%	40.57	13,888.31
Financials (40)	87	Mean	2.89%	15.92%	61.29	19,321.14
		Median	3.06%	17.19%	63.27	18,437.54
		Max	46.93%	22.67%	76.21	32,301.47
		Min	-37.59%	8.18%	40.78	10,006.32
Information Technology (45)	55	Mean	3.10%	14.98%	71.15	29,972.41
		Median	3.26%	16.19%	72.90	28,329.11
		Max	44.96%	20.31%	83.60	49,261.40
		Min	-35.00%	8.05%	53.35	16,304.03
Utilities (55)	27	Mean	2.77%	16.03%	58.51	14,443.97
		Median	2.96%	17.67%	60.42	13,565.16
		Max	47.63%	21.17%	76.10	23,419.84
		Min	-38.77%	7.33%	34.67	7,882.75

**Table 2:** This table shows the descriptive statistics of the data set, sorted according to the General Industry Classification Standard (GICS). In column 2 the number of stocks contained in each sector are given, while the columns 4 - 7 show the sector-specific mean, median, maximum and minimum values for return ( $\mu$ ), volatility ( $\sigma$ ), ESG Score ( $\theta$ ) and the market capitalization (MV) given in Mio. USD.

and with the selected and well established benchmarks. These are given by a standard Markowitz portfolio optimization and a naive  $1/N$  portfolio of all assets available. The benchmarks are chosen since they represent standard approaches to active as well as passive portfolio management, respectively. Then we analyze the differences in performance between the three MEA models in order to understand the effects of additional variables in the multi-directional efficiency analysis on portfolio performance.

For every sector a number of out-of-sample performance measures related to financial and social performance are computed, i.e. average return, volatility, Sharpe Ratio, 95% VaR, 95% expected shortfall, average ESG score ( $\bar{\theta}$ ) and Delta Ratio. Furthermore we compute p-values to evaluate sector-specific and joint significances based on various tests, such as the JKM-test (Ledoit and Wolf, 2008) for Sharpe Ratios, a paired t-test for the means of excess returns and ESG scores and a paired f-test for the variance of excess returns. For Sharpe Ratio significance testing, we refer to joint JKM tests, as the power of the sector-specific tests (JKM) is low, since we are only able to observe 32 quarters of out-of-sample data.

All models are implemented based on the full sample period.<sup>7</sup> The first eight quarters are used to load the input and output variables, which enter the MEA models. We end up with 32 quarters of out-of-sample data with rebalancing being done every quarter. The results of the out-of-sample performance of the benchmarks are given in Table 3, while the results of each MEA model are shown in Tables 4 to 6.

With respect to the benchmark results, we see a strong outperformance of the equally-weighted portfolio compared to the mean-variance portfolio with regards to financial performance, both in terms of returns as well as volatility. This is in line with our expectation,

---

<sup>7</sup>We also implemented a shorter period in which only the period after the financial crisis is considered. We only report the full-period results in this section, since the post-crisis results have shown to be similar and do not change our findings. We therefore conclude, that our results are stable and unaffected by major economic changes and circumstances. The post-crisis results can be received from the authors upon request.

Mean-variance portfolio							
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)
55	-1.60%	13.77%	-0.17	58.95	4.28	-0.26	-0.33
15	0.79%	12.25%	0.01	68.43	5.59	-0.21	-0.29
25	0.88%	11.77%	0.02	54.28	4.61	-0.20	-0.26
35	2.55%	10.66%	0.18	68.36	6.41	-0.17	-0.22
45	0.26%	10.01%	-0.05	69.98	6.99	-0.17	-0.20
20	1.88%	15.69%	0.08	76.27	4.86	-0.26	-0.35
40	1.40%	13.74%	0.05	83.26	6.06	-0.24	-0.31
10	-1.09%	10.89%	-0.17	95.01	8.72	-0.21	-0.26
30	0.05%	12.34%	-0.05	67.09	5.44	-0.21	-0.28
Mean	0.57%	12.35%	-0.01	71.29	5.88	-0.21	-0.28
Median	0.79%	12.25%	0.01	68.43	5.59	-0.21	-0.28

Naive portfolio							
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)
55	2.58%	12.06%	0.15	61.31	5.08	-0.19	-0.26
15	2.64%	10.56%	0.18	64.50	6.11	-0.17	-0.24
25	2.21%	11.31%	0.13	60.84	5.38	-0.18	-0.25
35	2.68%	11.65%	0.17	61.67	5.29	-0.18	-0.23
45	2.88%	10.95%	0.20	71.46	6.53	-0.17	-0.24
20	2.89%	11.25%	0.19	61.66	5.48	-0.18	-0.25
40	2.60%	11.33%	0.17	63.24	5.58	-0.18	-0.24
10	2.68%	10.94%	0.18	65.66	6.00	-0.17	-0.24
30	2.31%	10.50%	0.15	57.16	5.44	-0.17	-0.22
Mean	2.61%	11.17%	0.17	63.06	5.66	-0.18	-0.24
Median	2.64%	11.25%	0.17	61.67	5.48	-0.18	-0.24

**Table 3:** This table shows the out-of-sample results for the benchmark portfolios.  $\bar{\mu}$  represents the average return,  $\sigma$  portfolio volatility, SR denotes the Sharpe Ratio and  $\bar{\theta}$  stands for the average ESG Score. The Delta Ratio (DR) is computed by  $\frac{\bar{\theta}}{\sigma}$ . VaR and ES represent the Value-at-Risk and Expected Shortfall of the 95% quantile, respectively.

since we use quarterly data and therefore have a limited number of observations, which is known to strongly impact the performance of mean-variance portfolios compared to an equally-weighted asset allocation (DeMiguel et al., 2009).

Interestingly, we observe higher mean ESG Scores in the mean-variance portfolios. However when comparing Delta Ratios - the tradeoff between ESG Score and financial risk - we find no significant differences between the benchmarks. Given the data set used these results suggest that the naive portfolio is overall the preferred choice for investors.

## 5.1 Efficiency based on Volatility and Returns - Model 1

In this model, only volatility and expected return are considered in the multi-directional efficiency analysis. We find that the efficiency-weighted (Strategy 1) and the efficiency-screened (Strategy 3) portfolios strongly outperform the mean-variance portfolio in terms of financial performance and the combination of both (Strategy 2) weakly outperforms it. The efficiency-weighted portfolios show the statistically significantly highest out-of-sample Sharpe Ratios on average out of all asset allocation strategies. The naive portfolio shows a statistically significantly higher Sharpe Ratio, but the difference (15%) is comparatively small with respect to Strategy 1. It is interesting to note at this point that the superior performance of the efficiency-weighted portfolio in contrast to the mean-variance benchmark is actually achieved by using less information than what is used for the benchmark, since covariances are not considered in the MEA. These results are supported by our joint t-tests. It is also interesting to note that the f-tests on the variance of excess returns confirm weakly significantly lower financial risk of Strategy 2 against the naive portfolio benchmark, while underpinning all other results detailed above.

In terms of ESG Scores  $\theta$ , the mean-variance portfolio significantly outperforms all asset allocation strategies. However this is not surprising, since the ESG Score is not considered as an output variable in this MEA model, hence no attempt was made here to actively improve the social performance of the portfolios. In fact, the ESG Scores are not significantly different, when the strategies are compared among each other.<sup>8</sup> The naive portfolio shows statistically significantly smaller ESG scores compared to the efficiency-weighted portfolio in almost all sectors. Interestingly, when social performance is evaluated while considering financial risk, the various portfolio strategies under MEA Model 1 are highly competitive. Both under Strategy 1 and 2, the Delta Ratios outperform both the

---

<sup>8</sup>We conduct all significance tests for comparing the models with the benchmarks, but also comparing the models among each other. For the sake of readability we only report the detailed test results for the benchmark comparison and refer to the model comparison in the discussion of the results.



mean-variance and the naive benchmarks. The mean and median results of the Delta Ratio under Strategy 3 indicate a performance roughly on par with both benchmarks.

## 5.2 Efficiency based on Volatility, Returns and ESG scores - Model 2

In addition to volatility and expected return, also ESG scores are considered as output in Model 2. First, all portfolio strategies now strongly outperform the mean-variance portfolio in terms of financial performance. Again, the efficiency-weighted portfolios show on average the highest Sharpe Ratios (0.16) out of all asset allocation strategies, however, the drop-off to Strategy 2 (0.12) and 3 (0.12) is much lower than compared to Model 1. Strategy 1 shows the highest average out-of-sample returns, while volatility is on average lower in Strategy 2. In Model 2 the Sharpe Ratios of Strategies 1, 2 and 3 are not significantly different compared to the naive portfolio at the 5% level, as indicated by the joint t-tests.<sup>9</sup> The f-tests on the variance of excess returns indicate highly significantly lower financial risk of Strategy 2 against the naive portfolio benchmark, while again underpinning all other results detailed above.

With a view to social responsibility, i.e. ESG scores, the superior performance of the mean-variance benchmark portfolio observed in Model 1 is now by far less pronounced. While Strategy 1 (efficiency-weighting,  $\bar{\theta} = 66.58$ ) and Strategy 3 (efficiency-screening,  $\bar{\theta} = 69.37$ ) are again outperformed with high and low statistical significance respectively, Strategy 2 (efficiency-weighting & efficiency-screening,  $\bar{\theta} = 78.15$ ) now clearly (and highly significantly) surpasses the benchmark ( $\bar{\theta} = 71.29$ ). The naive portfolio shows a significantly smaller mean ESG score compared to all three portfolio strategies. When evaluating social performance together with financial risk, the three portfolio strategies under MEA

---

<sup>9</sup>In Strategy 3 we observe more variability in the results of the sector portfolios with regards to financial performance. Therefore the mean and median results offer a less intuitive interpretation in combination with the significance tests when compared to the other strategies.

Results								p-values			
Strategy 1: Efficiency-weighting								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	2.18	11.24	0.13	62.25	5.54	-0.18	-0.25	0.06	0.90	0.05	0.00
15	2.11	9.66	0.14	66.20	6.85	-0.16	-0.23	0.26	0.88	(0.06)	0.00
25	1.86	10.71	0.11	62.25	5.81	-0.18	-0.24	0.32	0.83	0.00	0.00
35	2.57	10.95	0.17	60.90	5.56	-0.17	-0.22	0.52	0.46	(0.00)	(0.00)
45	2.50	10.01	0.18	71.79	7.17	-0.16	-0.21	0.14	0.79	0.13	0.02
20	2.59	10.66	0.18	62.21	5.84	-0.17	-0.23	0.31	0.80	(0.00)	0.03
40	2.48	10.88	0.16	64.53	5.93	-0.17	-0.23	0.33	0.64	(0.00)	0.00
10	2.61	10.53	0.18	66.79	6.34	-0.17	-0.23	0.08	0.47	(0.00)	0.00
30	2.00	9.80	0.13	60.17	6.14	-0.16	-0.20	0.15	0.76	(0.00)	0.00
Mean	2.32	10.49	0.15	64.12	6.13	-0.17	-0.23	Joint Tests			
Median	2.48	10.66	0.16	62.25	5.93	-0.17	-0.23	Joint Tests			
JKM								0.01	(0.02)		
T-Test								0.01	(0.00)	(0.00)	0.00
F-Test								0.01	0.14		
Strategy 2: Efficiency-weighting & Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	0.46	11.54	-0.02	68.33	5.92	-0.21	-0.27	0.17	(0.04)	0.00	0.00
15	0.72	10.58	0.00	69.18	6.54	-0.18	-0.27	0.51	0.86	0.27	0.00
25	1.04	10.91	0.03	74.67	6.84	-0.19	-0.26	0.45	0.76	0.00	0.00
35	2.15	10.01	0.14	56.44	5.64	-0.16	-0.19	0.56	0.59	(0.00)	(0.00)
45	1.89	9.05	0.13	68.72	7.59	-0.14	-0.19	0.15	0.66	0.82	(0.03)
20	1.34	11.11	0.06	57.90	5.21	-0.19	-0.23	0.55	0.89	(0.00)	(0.03)
40	1.36	10.37	0.06	70.39	6.79	-0.17	-0.21	0.48	0.80	(0.00)	0.00
10	2.49	10.15	0.18	63.88	6.29	-0.16	-0.20	0.04	0.51	(0.00)	0.78
30	1.53	10.17	0.08	71.07	6.99	-0.16	-0.20	0.18	0.70	0.00	0.00
Mean	1.44	10.43	0.07	66.73	6.42	-0.17	-0.23	Joint Tests			
Median	1.36	10.37	0.06	68.72	6.54	-0.17	-0.21	Joint Tests			
JKM								0.08	(0.01)		
T-Test								0.09	(0.01)	(0.00)	0.00
F-Test								0.00	0.08		
Strategy 3: Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	1.06	13.32	0.03	59.89	4.50	-0.23	-0.30	0.11	0.88	0.40	0.73
15	2.08	9.73	0.14	56.59	5.82	-0.14	-0.21	0.19	0.59	(0.00)	(0.00)
25	0.26	11.38	-0.04	66.87	5.88	-0.21	-0.28	0.67	0.87	0.00	0.02
35	0.74	12.01	0.00	63.59	5.29	-0.21	-0.27	0.76	0.82	0.88	0.31
45	2.21	9.90	0.16	66.13	6.68	-0.14	-0.18	0.13	0.59	(0.10)	(0.01)
20	3.20	13.87	0.18	67.54	4.87	-0.21	-0.27	0.23	0.52	(0.02)	0.09
40	1.48	11.54	0.07	69.62	6.03	-0.19	-0.23	0.47	0.77	(0.00)	0.04
10	2.33	13.40	0.12	49.01	3.66	-0.23	-0.29	0.05	0.63	(0.00)	(0.00)
30	2.36	12.83	0.13	72.26	5.63	-0.20	-0.29	0.12	0.55	0.00	0.00
Mean	1.75	12.00	0.09	63.50	5.37	-0.19	-0.26	Joint Tests			
Median	2.08	12.01	0.12	66.13	5.63	-0.21	-0.27	Joint Tests			
JKM								0.04	(0.06)		
T-Test								0.04	(0.08)	(0.00)	0.35
F-Test								0.36	0.85		

**Table 4:** This table shows the out-of-sample results for the MEA Model 1, which uses volatility as input and expected return as output.  $\bar{\mu}$  represents the average return,  $\sigma$  portfolio volatility, SR denotes the Sharpe Ratio and  $\bar{\theta}$  stands for the average ESG Score. The Delta Ratio (DR) is computed by  $\frac{\bar{\theta}}{\sigma}$ . VaR and ES represent the Value-at-Risk and Expected Shortfall of the 95% quantile, respectively. P-values in brackets indicate results significantly lower than the respective benchmarks' results starting at a 10% level.

Model 2 perform very well. The Delta Ratios of Strategies 1 through 3 clearly outperform the mean-variance and the naive benchmarks.

### **5.3 Efficiency based on Different Risk Measures, Returns and ESG scores - Model 3**

In Model 3, additional risk measures (i.e. VaR 95% and ES 95%) were added as input variables to the MEA model to evaluate whether portfolio risk can be further positively influenced. We observe improvements with regards to average returns, volatilities, ESG Scores, Sharpe Ratios and Delta Ratios in all strategies compared to Model 2, while average portfolio VaR and ES remain virtually unchanged.<sup>10</sup> However the differences in these values are too small to translate into statistically significant changes in our tests compared to the previous results. One probable explanation for this result could be the observed negative sample correlation between ESG Scores and stock volatility. This would imply that most of the risk improvement potential has already been realized as a result of considering ESG Scores as additional variable, which is an interesting finding. Under these circumstances additional risk variables do not seem to provide further significant information value to the asset selection and thus, to the resulting portfolio performance.

Nevertheless, Model 3 also shows that the MEA results are robust and not easily affected by the choice of the variables included in the linear programming system.

### **5.4 Discussion of Results**

The above mentioned results are also visualized in Figures 1 and 2, which illustrate the out-of-sample Sharpe Ratio and Delta Ratio sector mean results for each portfolio strategy and each MEA model, respectively. The top figure of Figure 1 shows these sector results of Strategy 1 of all MEA models. It can be seen that these are clustered very closely,

---

<sup>10</sup>The average ESG Score of Strategy 3 slightly decreases compared to Model 2 as the only exception.

Results								p-values			
Strategy 1: Efficiency-weighting								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	2.25	11.17	0.14	64.58	5.78	-0.18	-0.24	0.05	0.77	0.00	0.00
15	2.29	9.58	0.16	69.72	7.28	-0.16	-0.23	0.24	0.76	0.19	0.00
25	2.03	10.43	0.13	64.75	6.21	-0.17	-0.23	0.29	0.61	0.00	0.00
35	2.50	10.74	0.16	64.88	6.04	-0.17	-0.22	0.52	0.54	(0.00)	0.00
45	2.55	9.93	0.18	73.31	7.38	-0.16	-0.20	0.14	0.70	0.02	0.00
20	2.53	10.79	0.17	64.33	5.96	-0.17	-0.23	0.32	0.90	(0.00)	0.00
40	2.52	10.80	0.17	66.37	6.15	-0.17	-0.23	0.32	0.49	(0.00)	0.00
10	2.42	10.53	0.16	69.25	6.58	-0.17	-0.23	0.09	0.72	(0.00)	0.00
30	2.25	9.64	0.16	62.03	6.43	-0.15	-0.19	0.11	0.39	(0.00)	0.00
Mean	2.37	10.40	0.16	66.58	6.42	-0.17	-0.22				
Median	2.42	10.53	0.16	64.88	6.21	-0.17	-0.23				
JKM								0.01	(0.09)		
T-Test								0.01	(0.01)	(0.00)	0.00
F-Test								0.00	0.11		
Strategy 2: Efficiency-weighting & Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	1.59	11.35	0.08	76.11	6.71	-0.16	-0.22	0.05	0.80	0.00	0.00
15	0.95	9.11	0.03	81.32	8.93	-0.16	-0.23	0.45	0.88	0.00	0.00
25	1.76	9.00	0.12	84.44	9.38	-0.15	-0.21	0.25	0.55	0.00	0.00
35	2.07	9.66	0.14	74.77	7.74	-0.15	-0.20	0.56	0.62	0.00	0.00
45	2.25	9.03	0.17	79.48	8.80	-0.14	-0.16	0.12	0.58	0.00	0.00
20	1.27	11.08	0.05	73.18	6.60	-0.19	-0.24	0.56	(0.04)	(0.08)	0.00
40	2.39	10.44	0.16	81.84	7.84	-0.17	-0.21	0.30	0.53	0.83	0.00
10	1.80	11.82	0.09	76.26	6.45	-0.20	-0.25	0.08	0.76	(0.00)	0.00
30	2.66	9.21	0.21	75.98	8.25	-0.14	-0.17	0.04	0.28	0.00	0.00
Mean	1.86	10.08	0.12	78.15	7.86	-0.16	-0.21				
Median	1.80	9.66	0.12	76.26	7.84	-0.16	-0.21				
JKM								0.02	(0.06)		
T-Test								0.02	(0.02)	0.00	0.00
F-Test								0.00	0.03		
Strategy 3: Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	0.96	14.14	0.02	64.30	4.55	-0.23	-0.30	0.10	0.89	0.07	0.10
15	2.52	9.86	0.19	66.13	6.71	-0.13	-0.19	0.11	0.49	(0.08)	0.18
25	0.94	11.50	0.02	71.67	6.23	-0.20	-0.26	0.48	0.74	0.00	0.00
35	1.20	8.35	0.06	67.25	8.05	-0.13	-0.19	0.70	0.72	0.65	0.03
45	2.31	8.97	0.18	71.65	7.99	-0.13	-0.15	0.09	0.53	0.24	0.46
20	3.13	13.49	0.18	75.53	5.60	-0.21	-0.25	0.24	0.53	0.55	0.01
40	3.07	11.51	0.21	76.13	6.61	-0.18	-0.23	0.21	0.36	(0.00)	0.00
10	1.43	14.11	0.05	55.80	3.95	-0.24	-0.29	0.09	0.79	(0.00)	(0.00)
30	2.76	11.94	0.17	75.85	6.35	-0.17	-0.24	0.06	0.46	0.00	0.00
Mean	2.04	11.54	0.12	69.37	6.23	-0.18	-0.23				
Median	2.31	11.51	0.17	71.65	6.35	-0.18	-0.24				
JKM								0.01	0.85		
T-Test								0.01	0.84	(0.08)	0.00
F-Test								0.17	0.68		

**Table 5:** This table shows the out-of-sample results for the MEA Model 2, which uses volatility as input and expected return and ESG Score as output.  $\bar{\mu}$  represents the average return,  $\sigma$  portfolio volatility, SR denotes the Sharpe Ratio and  $\bar{\theta}$  stands for the average ESG Score. The Delta Ratio (DR) is computed by  $\frac{\bar{\theta}}{\sigma}$ . VaR and ES represent the Value-at-Risk and Expected Shortfall of the 95% quantile, respectively. P-values in brackets indicate results significantly lower than the respective benchmarks' results starting at a 10% level.

Results								p-values			
Strategy 1: Efficiency-weighting								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	2.13	10.75	0.13	65.22	6.07	-0.17	-0.23	0.05	0.80	0.00	0.00
15	2.33	9.59	0.17	69.63	7.26	-0.16	-0.22	0.23	0.71	0.21	0.00
25	2.01	10.48	0.12	64.87	6.19	-0.17	-0.23	0.29	0.65	0.00	0.00
35	2.53	10.78	0.17	64.79	6.01	-0.17	-0.22	0.52	0.49	(0.00)	0.00
45	2.60	9.90	0.19	73.38	7.41	-0.16	-0.20	0.13	0.63	0.02	0.00
20	2.57	10.82	0.17	64.07	5.92	-0.17	-0.23	0.32	0.87	(0.00)	0.00
40	2.49	10.80	0.16	66.51	6.16	-0.17	-0.23	0.32	0.55	(0.00)	0.00
10	2.45	10.56	0.16	69.55	6.59	-0.17	-0.23	0.08	0.69	(0.00)	0.00
30	2.15	9.63	0.15	62.07	6.45	-0.15	-0.19	0.12	0.55	(0.00)	0.00
Mean	2.36	10.37	0.16	66.68	6.45	-0.17	-0.22				
Median	2.45	10.56	0.16	65.22	6.19	-0.17	-0.23				
JKM								0.01	(0.09)		
T-Test								0.01	(0.01)	(0.00)	0.00
F-Test								0.00	0.10		
Strategy 2: Efficiency-weighting & Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	1.58	9.81	0.09	75.08	7.65	-0.15	-0.21	0.03	0.75	0.00	0.00
15	1.10	9.03	0.04	81.42	9.02	-0.16	-0.22	0.41	0.85	0.00	0.00
25	1.76	9.06	0.12	83.39	9.20	-0.15	-0.21	0.25	0.55	0.00	0.00
35	2.04	9.45	0.14	73.91	7.82	-0.15	-0.20	0.56	0.62	0.00	0.00
45	2.36	8.67	0.19	78.50	9.05	-0.13	-0.16	0.11	0.52	0.00	0.00
20	1.51	11.12	0.07	71.35	6.42	-0.19	-0.24	0.51	(0.07)	(0.01)	0.00
40	2.18	10.04	0.15	81.29	8.10	-0.17	-0.21	0.33	0.59	(0.08)	0.00
10	1.86	11.75	0.10	76.67	6.52	-0.20	-0.24	0.08	0.75	(0.00)	0.00
30	2.31	8.98	0.18	75.89	8.45	-0.14	-0.17	0.06	0.39	0.00	0.00
Mean	1.86	9.77	0.12	77.50	8.03	-0.16	-0.21				
Median	1.86	9.45	0.12	76.67	8.10	-0.15	-0.21				
JKM								0.01	(0.07)		
T-Test								0.02	(0.02)	0.00	0.00
F-Test								0.00	0.01		
Strategy 3: Efficiency-screening								SR		$\theta$	
Sector	$\bar{\mu}$ (%)	$\sigma$ (%)	SR	$\bar{\theta}$	DR	VaR (%)	ES (%)	MV	Naive	MV	Naive
55	1.15	12.52	0.04	62.27	4.97	-0.21	-0.29	0.07	0.84	0.15	0.32
15	2.77	9.84	0.21	66.93	6.80	-0.13	-0.19	0.08	0.42	0.82	0.06
25	0.79	10.56	0.01	70.38	6.66	-0.18	-0.24	0.53	0.79	0.00	0.00
35	1.46	8.54	0.09	65.68	7.69	-0.13	-0.19	0.65	0.66	0.82	0.08
45	2.46	8.38	0.22	70.41	8.40	-0.12	-0.14	0.07	0.45	0.43	0.72
20	3.85	14.12	0.23	70.42	4.99	-0.21	-0.27	0.18	0.40	0.85	0.07
40	2.77	10.78	0.19	75.16	6.97	-0.18	-0.22	0.24	0.41	(0.00)	0.00
10	1.92	14.44	0.09	54.89	3.80	-0.24	-0.29	0.06	0.72	(0.00)	(0.00)
30	2.25	12.47	0.12	76.12	6.10	-0.19	-0.28	0.10	0.56	0.00	0.00
Mean	2.16	11.29	0.13	68.03	6.27	-0.18	-0.23				
Median	2.25	10.78	0.12	70.38	6.66	-0.18	-0.24				
JKM								0.00	0.79		
T-Test								0.01	0.78	(0.01)	0.00
F-Test								0.09	0.54		

**Table 6:** This table shows the out-of-sample results for the MEA Model 3, which uses volatility, VaR 95% and ES 95% as input and expected return and ESG Score as output.  $\bar{\mu}$  represents the average return,  $\sigma$  portfolio volatility, SR denotes the Sharpe Ratio and  $\bar{\theta}$  stands for the average ESG Score. The Delta Ratio (DR) is computed by  $\frac{\bar{\theta}}{\sigma}$ . VaR and ES represent the Value-at-Risk and Expected Shortfall of the 95% quantile, respectively. P-values in brackets indicate results significantly lower than the respective benchmarks' results starting at a 10% level.

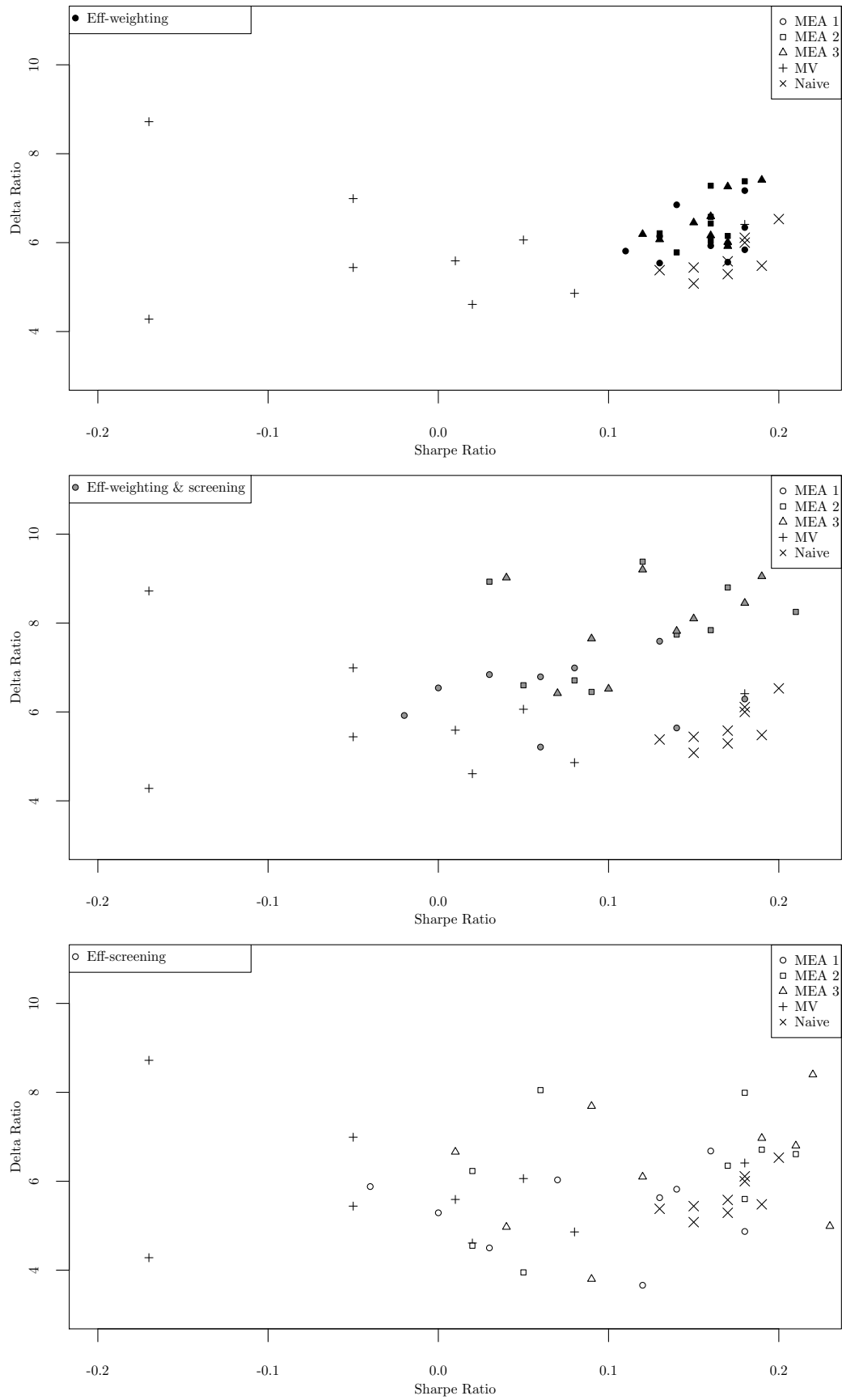
compared to Strategies 2 and 3 presented in the middle and bottom figure and are situated above the cluster of naive portfolios in the two-dimensional plane. Figure 2 presents all portfolios of each MEA model, separately. It shows that the inclusion of the social variable does lead to higher Delta Ratio results and that there is no visible difference between Model 2 in the middle figure and Model 3 in the bottom figure.

Figure 3 illustrates the out-of sample Sharpe Ratio and Delta Ratio overall mean results for all three MEA models and all three portfolio strategies.

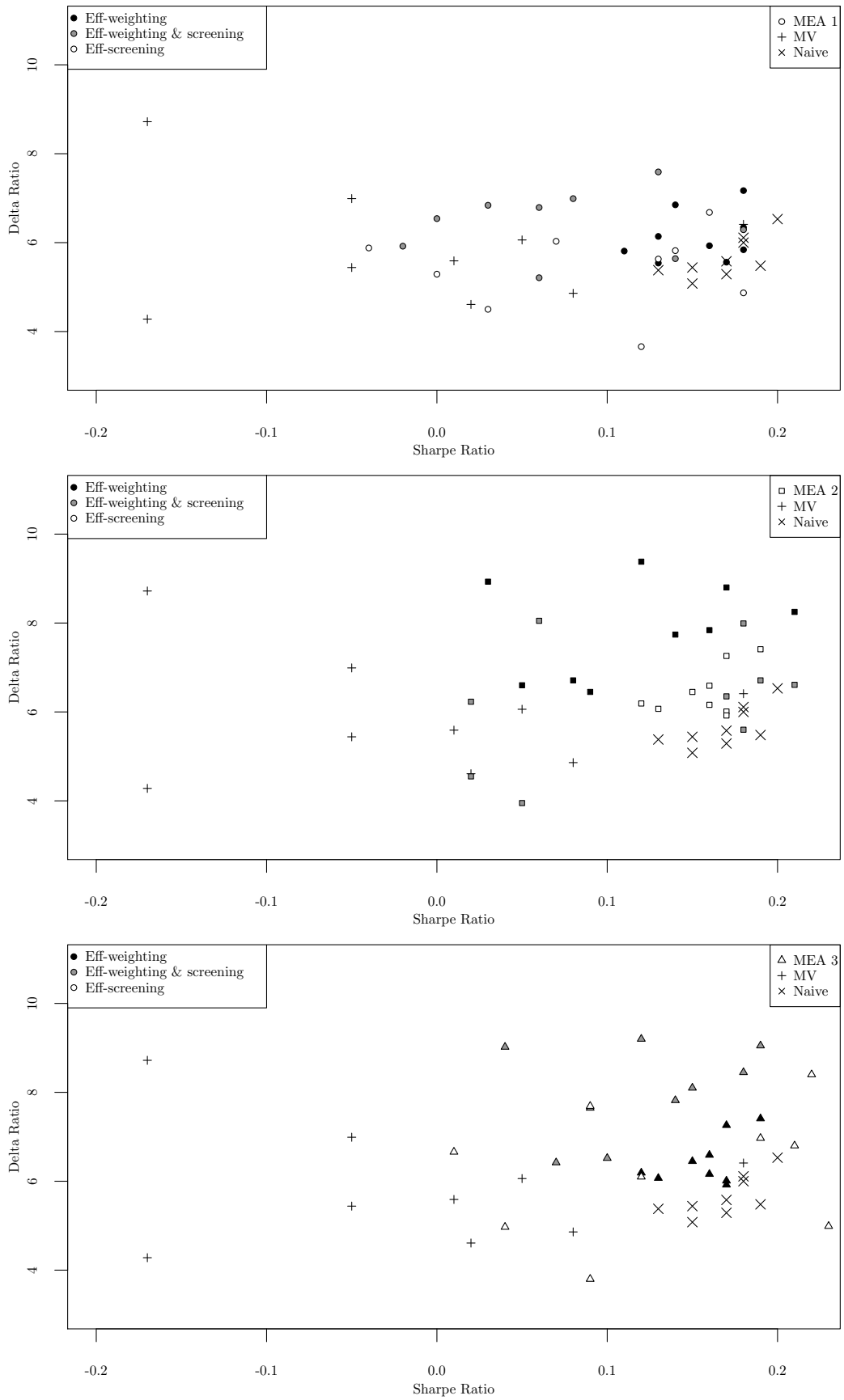
Overall, it can be concluded that all but one of the tested strategies succeed in outperforming the mean-variance portfolio in terms of financial and social performance measured by Sharpe and Delta Ratios, with Strategy 1 (efficiency-weighting) achieving on average the highest financial performance of all strategies.

Including ESG Scores in the multi-directional efficiency analysis (i.e. Model 2) has a positive effect across all three strategies for both Sharpe and Delta Ratios. The effect is strongly significant and most pronounced in Strategy 2, which is the combination of efficiency-weighting and efficiency-screening. This strategy delivers the highest increase in Sharpe Ratio, ESG Score and Delta Ratio compared to Model 1 and overall yields on average the highest ESG Score and Delta Ratio of all tested strategies.

When investigating the performance of the strategies with regard to the naive portfolio, we find that for Model 2 and 3 there is no significant difference in financial performance at the 5% level, while all strategies surpass the naive portfolio in terms of social performance. This means that all strategies based on efficiency information that also incorporates a social variable are strictly preferred to the naive portfolio. Nevertheless Strategies 2 and 3 have a clear conceptual shortcoming compared to Strategy 1 since the asset universe is potentially reduced significantly, with only fully efficient firms being considered for the portfolio, which may result in unfavorable consequences, due to diversification losses.

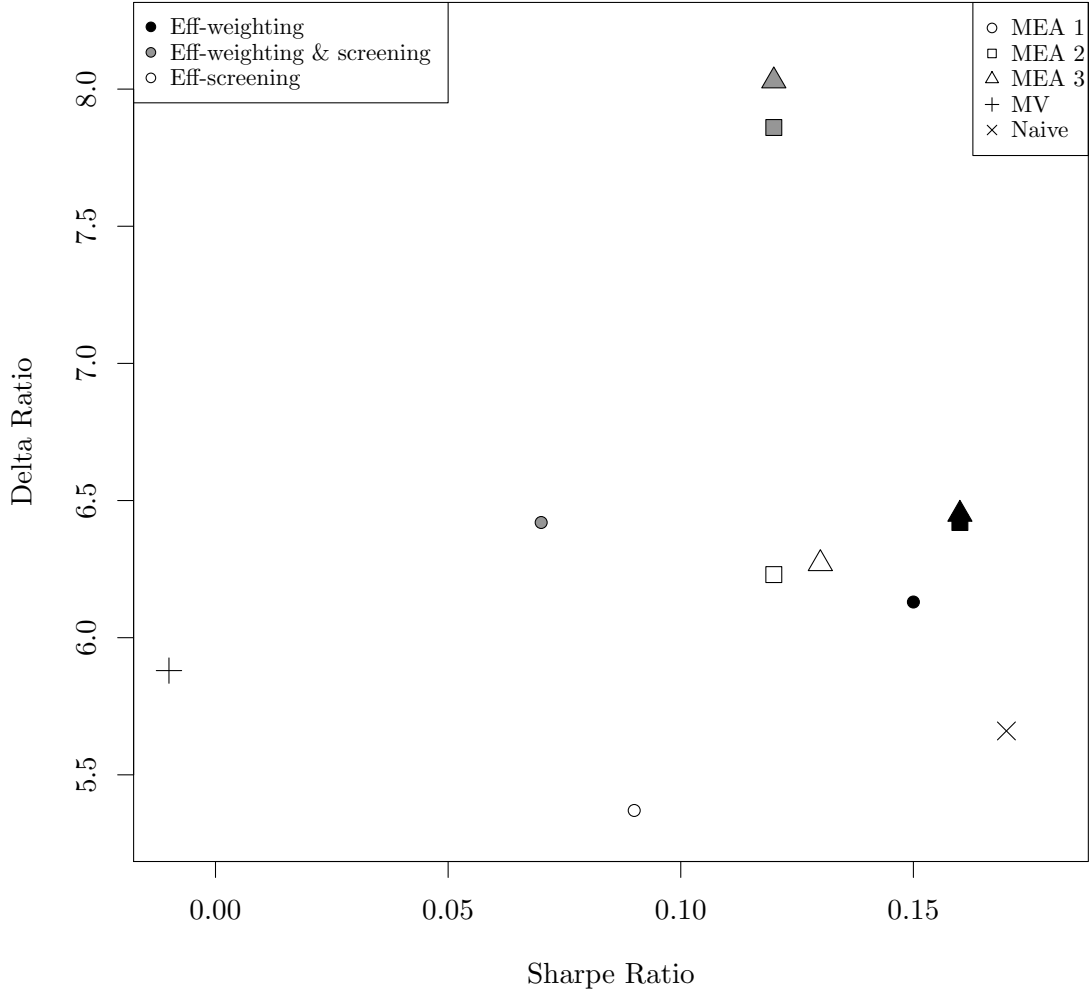


**Figure 1:** This Figure shows the average out-of-sample results for each sectors compared in terms of asset allocation strategies. Circles, squares and triangles designate the Sharpe and Delta Ratio mean results for the three respective MEA models, while black, gray and white filling indicates strategy 1 (Efficiency-weighting), strategy 2 (Efficiency-weighting & Efficiency screening) or strategy 3 (Efficiency screening) with the three MEA models. The symbols + and x reference the two portfolio benchmarks, i.e. the mean-variance portfolio and the naive portfolio.



**Figure 2:** This Figure shows the average out-of-sample results for each sectors compared in terms of model specifications. Circles, squares and triangles designate the Sharpe and Delta Ratio mean results for the three respective MEA models, while black, gray and white filling indicates strategy 1 (Efficiency-weighting), strategy 2 (Efficiency-weighting & Efficiency screening) or strategy 3 (Efficiency screening) with the three MEA models. The symbols + and x reference the two portfolio benchmarks, i.e. the mean-variance portfolio and the naive portfolio.





**Figure 3:** This Figure shows the average out-of-sample results over all sectors compared in terms of asset allocation strategies and model specifications. Circles, squares and triangles designate the Sharpe and Delta Ratio mean results for the three respective MEA models, while black, gray and white filling indicates strategy 1 (Efficiency-weighting), strategy 2 (Efficiency-weighting & Efficiency screening) or strategy 3 (Efficiency screening) with the three MEA models. The symbols + and x reference the two portfolio benchmarks, i.e. the mean-variance portfolio and the naive portfolio.

On the contrary, Strategy 1 preserves the initial asset universe, which should make it a potentially more interesting choice for investors.

Our results show that given the selected benchmarks it is possible for social responsible investors to achieve a superior level of social performance without compromising financial performance. If an even higher social performance is preferred, a comparably small sacrifice in financial performance has to be accepted. Furthermore, the proposed strategies are not

only preferred by social responsible investors looking for social performance, but are also a viable, more socially responsible option for traditional investors only caring for financial performance when given the choice between the strategies and the benchmarks.

## 6 Conclusion

In this paper we analyse the usefulness of Data Envelopment Analysis for application to portfolio management with respect to socially responsible investments. Its flexibility in the selection of relevant parameters makes this method a perfect fit for analyzing investors preferences that are driven by more than only risk and return. For the application in socially responsible investing, we rely on multi-directional efficiency analysis, which features some useful properties compared to other DEA models. Herein, we compare three different MEA models, which include up to five variables, both financial and social, to compute relative efficiency scores. We further implement three asset allocation strategies and evaluate their out-of-sample performance among each other and with respect to selected and well-established benchmark portfolios.

We find that in all MEA models a superior financial and social performance can be achieved compared to a traditional mean-variance portfolio. Explicitly considering a social responsibility variable in the MEA model has a positive effect on all asset allocation strategies and these strategies are able to match the financial performance of the naive portfolio, while surpassing it in terms of social performance. Therefore, our results clearly outline the benefits of MEA portfolio modeling, not only for socially responsible investors but also traditional investors.

## References

- Asmild, M., Hougaard, J. L., Kronborg, D., and Kvist, H. K. (2003). Measuring inefficiency via potential improvements. *Journal of Productivity Analysis*, 19(1):59–76.
- Bahrani, R. and Khedri, N. (2013). Evaluation of relative efficiency and performance of companies using data envelopment analysis (dea) approach. *Elixir Finance Management*, 56:13299–13304.
- Ballesteros, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., and Plà-Santamaria, D. (2012). Socially responsible investment: A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research*, 216(2):487 – 494.
- Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9):1078–1092.
- Basso, A. and Funari, S. (2016). Dea performance assessment of mutual funds. In *Data Envelopment Analysis*, pages 229–287. Springer.
- Bilbao-Terol, A., Arenas-Parra, M., and Cañal-Fernández, V. (2012). Selection of socially responsible portfolios using goal programming and fuzzy technology. *Information Sciences*, 189:110–125.
- Bilbao-Terol, A., Arenas-Parra, M., Cañal-Fernández, V., and Bilbao-Terol, C. (2013). Selection of socially responsible portfolios using hedonic prices. *Journal of Business Ethics*, 115(3):515–529.
- Bogetoft, P. and Hougaard, J. L. (1999). Efficiency evaluations based on potential (non-proportional) improvements. *Journal of Productivity Analysis*, 12(3):233–247.

- Branda, M. and Kopa, M. (2014). On relations between dea-risk models and stochastic dominance efficiency tests. *Central European Journal of Operations Research*, 22(1):13–35.
- Charnes, A., Cooper, W. W., and 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.
- Chen, Z. and Lin, R. (2006). Mutual fund performance evaluation using data envelopment analysis with new risk measures. *Or Spectrum*, 28(3):375–398.
- Debreu, G. (1951). The coefficient of resource utilization. *Econometrica: Journal of the Econometric Society*, pages 273–292.
- DeMiguel, V., Garlappi, L., and Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953.
- Dia, M. (2009). A portfolio selection methodology based on data envelopment analysis. *INFOR: Information Systems and Operational Research*, 47(1):71–79.
- Edirisinghe, N. and 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.
- Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3):253–290.
- Gasser, S. M., Rammerstorfer, M., and Weinmayer, K. (2017). Markowitz revisited: Social portfolio engineering. *European Journal of Operational Research*, 258(3):1181–1190.
- Golany, B. and Roll, Y. (1989). An application procedure for dea. *Omega*, 17(3):237–250.

- Gregoriou, G. N. (2006). Optimisation of the largest us mutual funds using data envelopment analysis. *Journal of Asset Management*, 6(6):445–455.
- Hamilton, S., Jo, H., and Statman, M. (1993). Doing well while doing good? the investment performance of socially responsible mutual funds. *Financial Analysts Journal*, pages 62–66.
- Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M., and Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research*, 61(1):169–183.
- Ismail, M. K. A., Rahman, N. M. N. A., Salamudin, N., and Kamaruddin, B. H. (2012). Dea portfolio selection in malaysian stock market. In *Innovation Management and Technology Research (ICIMTR), 2012 International Conference*, pages 739–743. IEEE.
- Koopmans, T. (1951). Cl,analysis of production as an efficient combination of activities. *Cap. III of Koopmans, TC (ed), Activity Analysis of Production and Allocation. Cowles Commission Monograph*.
- Ledoit, O. and Wolf, M. (2008). Robust performance hypothesis testing with the sharpe ratio. *Journal of Empirical Finance*, 15(5):850–859.
- Lim, S., Oh, K. W., and 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.
- Lin, R. (2009). Mutual fund performance dynamic evaluation using data envelopment windows analysis. In *Management and Service Science, 2009. MASS'09. International Conference*, pages 1–4. IEEE.
- Lovell, C. K. (1993). Production frontiers and productive efficiency. *The Measurement of Productive Efficiency: Techniques and Applications*, pages 3–67.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

- Pätäri, E., Leivo, T., and Honkapuro, S. (2012). Enhancement of equity portfolio performance using data envelopment analysis. *European Journal of Operational Research*, 220(3):786–797.
- Pätäri, E. J., Leivo, T. H., and Samuli Honkapuro, J. (2010). Enhancement of value portfolio performance using data envelopment analysis. *Studies in Economics and Finance*, 27(3):223–246.
- Pendaraki, K. (2012). Mutual fund performance evaluation using data envelopment analysis with higher moments. *Journal of Applied Finance and Banking*, 2(5):97.
- Powers, J. and McMullen, P. (2000). Using data envelopment analysis to select efficient large market cap securities. *Journal of Business Management*, 7:31–42.
- Thomson Reuters (2012). Asset4 environmental, social and corporate governance data. <http://extranet.datastream.com/data/ASSET4%20ESG/>. Last Accessed on Mar. 25, 2015.
- Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3):498–509.
- US SIF (2016). Report on sustainable and responsible investing trends 2016. [http://ussifnewsite.membershipsoftware.org/files/SIF\\_Trends\\_16\\_Executive\\_Summary\(1\).pdf](http://ussifnewsite.membershipsoftware.org/files/SIF_Trends_16_Executive_Summary(1).pdf).
- Zhao, X. and Shi, J. (2010). Evaluation of mutual funds using multi-dimensional information. *Frontiers of Computer Science in China*, 4(2):237–253.