The environmental pillar of ESG and financial performance: A portfolio analysis

Published Version:
Elettra Agliardi, Thomas Alexopoulos, Kleanthis Karvelas (2023). The environmental pillar of ESG and financial performance: A portfolio analysis. ENERGY ECONOMICS, 120, 1-16 [10.1016/j.eneco.2023.106598].

Availability:
This version is available at: https://hdl.handle.net/11585/919902 since: 2024-05-23

Published:
DOI: http://doi.org/10.1016/j.eneco.2023.106598

The Environmental Pillar of ESG and Financial Performance: A Portfolio Analysis

Abstract

Our research uses the environmental pillar of ESG as a proxy for environmental corporate social responsibility. We examine the performance of environmentally clustered portfolios by using simple quantitative investment strategies with optimum asset rotation. Post-hoc, sample-split analysis with non-parametric tests is performed. The results suggest that both environmental status and dynamic environmental performance are key characteristics of divergent financial behaviors. We show that environmentally low-rated companies present better financial performance, while environmental leaders are less risky and show more resilience. Assets with a dynamic environmental profile outperform on average in terms of returns and risk. Furthermore, supporting evidence of positive spillovers in high-rated environmental clusters is identified after the Paris Agreement. We evaluate the resilience of the environmental clusters during the COVID-19 crisis and the Russia-Ukraine war effect.

Keywords: Corporate Environmental Responsibility, ESG, Portfolio Rebalancing, ART Anova, Post-hoc Analysis, Paris Agreement, COVID-19 crisis, Russia-Ukraine war effect

JEL **Classification:** G11, G30, Q01, Q56

1. Introduction

The incorporation of Environmental, Social, and Governance (ESG) factors into the investment process and its financial impact are still controversial issues among the academia and investment community. To some, ESG remains vague and there is no consensus on this matter. Larcker et al. (2021) analyse seven common myths related to ESG and find that there is no conclusive association between ESG and financial performance.

In recent years, significant investments have been made in ESG stocks, sustainable funds and green bonds.[1] Following the Paris Agreement, which was adopted in December 2015 (entered into force in November 2016), with 190 countries plus the EU that agreed to reduce their carbon emissions and limit the global temperature increase, various studies have attempted to answer further questions regarding the impact of the Paris Agreement on assets' financial performance, and how to integrate carbon risk into financial products (Delis et al. 2019). The decision of the US administration to re-join the Paris Agreement in 2021 brought carbon risk back on top of the climate change agenda. In this path, the investment community monitors more closely firms which incorporate carbon emission decisions into their investment process. In February 2020, COVID-19 pandemic outbreak caused an unprecedented crisis in the financial markets since the global financial crisis of 2008-2009. The magnitude of the market crash was depicted in the 34% decline of the S&P 500 index in March 2020. The effect and the resilience of sustainable investments during times of crisis such as the COVID-19 pandemic have been investigated in recent studies (Albuquerque et al., 2020; Ding et al., 2020; Bae et al., 2021; Demers et al., 2021; Engelhardt et al., 2021). extraordinary

[1] A recent report by US SIF Foundation shows that investments in ESG and sustainable funds have increased from $1 trillion in 1998 to $8.4 trillion in 2022. See US SIF Foundation, "Report on US Sustainable and Impact Investing Trends," (2022).

Whether ESG criteria can be integrated successfully into investment decisions and to what degree these criteria affect financial performance is a constant debate in the literature.[2] And yet, despite its rapidly grown attention, the impact of ESG criteria remains unclear, especially in times of crises.

Our motivation is triggered by the inconclusive answers on this matter, as well as the recent developments with respect to global warming. Therefore, we focus solely on the Environmental pillar of ESG and aim to shed more light on its financial impact on portfolio performance and how this can be implemented. In this paper we explore whether and how environmental information feeds into the financial behaviour of a sample of S&P 500 stock portfolios. We aim to raise the level of consciousness of portfolio investors to an exogenous to them second-order problem, that of environment and climate change.

The contribution of our paper to the literature is threefold. First, we adopt the idea of clustering (i.e. clustering based on the environmental scores of ESG) instead of sorting, which has been broadly used in the literature, as sorting is often more vulnerable to short-term changes (Giese et al., 2019), "quality" factors (Bruno et al., 2021), and in some cases, to data driven results. In the first stage of our empirical analysis we employ seven straightforward criteria for optimum portfolio rebalancing. The use of clustering, instead of sorting, makes our rotation strategies more efficient, as all stocks belonging to a cluster are given the same chances of inclusion in a portfolio after rebalancing, while keeping at the same time the concept of environmental performance. Second, in order to be statistically consistent and answer the question as to whether the environmental cluster or the applied strategy drive our results,[3] we examine the significance of the effect of each factor

[2] The relationship between environmental and financial performance has been extensively examined in the literature, since Hamilton et al., 1993; Diltz, 1995; Cohen et al., 1997; Yamashita et al., 1999; Derwall et al., 2005; Kempf and Osthoff, 2007, to more recent studies on the impact of carbon emissions to financial performance and fixed income portfolio construction techniques, i.e. Bolton and Kacperczyk 2021, Görgen et al. 2020, Dorfleitner et al. 2020, Fabozzi et al. 2021.

[3] Many studies examine regional, sectoral, or applied strategy exposures (Bruno et al., 2021; Giese et al., 2020; Nagy et al., 2016).

separately and of their interaction. We expand our research by including a contrast hypothesis and a post-hoc analysis, in an attempt to identify environmental clusters with potential divergent behaviour. Third, we select the Paris Agreement as a shock event, a milestone event against climate change, and we examine if any positive spillovers can be identified in our portfolios. We expand our analysis after the COVID-19 crisis unfolded, to assess the effects of the pandemic on the relationship between ESG and financial performance, as well as the implications of the war in Ukraine.

We aim to identify long-term conclusions free from short-term variations, as well as to provide robust evidence attributed by the environmental performance and not by the selected strategy and/or parametrization. With clustering our empirical analysis examines the long-term relationship between environmental and financial performance. Unlike other studies that show an environmental premium being biased by short-term-driven results and other quality factors, we suggest that environmentally low-rated companies present higher returns. In contrast, companies with a dynamic environmental profile should be an optimal option, providing evidence of high risk-adjusted returns. Environmental leaders are the risk-averted option for an investor and show more resilience, indicating positive spillovers after the Paris Agreement. We evaluate the resilience of the environmental clusters during the COVID-19 crisis and the Russia-Ukraine war effect. Our results explained by the environmental performance and not by the applied strategy.

The rest of the paper is structured as follows. Section 2 reviews the literature on the relationship between environmental and financial performance and the ESG index. In Section 3 we present our empirical methodology and the data used in the paper. In Section 4 our empirical results are discussed. Section 5 gives concluding comments and offers some extensions of the current work.

2. The State-of-the-Art on ESG investing

The attractiveness of socially responsible investing (SRI), or sustainable investing, has considerably grown for more than two decades,[4] along with asset managers' attention to ESG data. This is reflected in the annual growth of 26.1% of new signatories of the Principles for Responsible Investment (PRI) in 2021, representing more than $110 trillion assets under management (AUM).[5] PRI signatories are committed to integrating ESG data when making investment decisions.

There is a widely held belief that sustainable investments pay off. However, it seems to be a continuously controversial issue in the literature (Renneboog et al., 2008; Larcker et al., 2021; Liang and Renneboog, 2021). Contradicting results started to appear in early studies too. Hamilton, Jo and Statman (1993), using data from socially responsible mutual funds with mainly environmental criteria, find that there is not SRI premium and social responsibility elements do not affect stock returns. This is in line with the basic finance principle, where such elements that are not proxies for risk do not affect the stock returns. Cohen, Fenn, and Konar (1997) report that there is neither an environmental premium nor a penalty for investing in low-polluter companies, having constructed industry-clustered portfolios with environmental responsibility criteria to investigate the financial performance difference between low-polluter and high-polluter companies in the United States (US).

Studies have attempted to explain also the effect of sustainable investments during times of financial crises and shock events. The relation between ESG and financial performance of US commercial banks over 2003-2010 was examined by Cornett et al. (2016). They find positive evidence related to ESG performance. Especially after the financial crisis, banks gain from engaging in sustainability. Similarly, Lins et al. (2016), study ESG factors of non-financial

[4] United Nations, "Financing for Sustainable Development Report 2019" (2019)

[5] United Nations Principles for Responsible Investment, "PRI Quarterly signatory update Q2 2021" (June 2021).

companies during 2008-2009 and showed that high-ESG rated firms had significantly higher stock returns, profitability and growth than low rated ones.

The Paris Agreement as a shock event, was a milestone against climate change. Recent studies investigate whether the Paris Agreement can raise the awareness of investors. Monasterolo and de Angelis (2020) analyze a sample of "green" and "brown" indices from the US and EU (1999-2018), and argue that low carbon emissions are associated with low risk following the Paris Agreement. Bolton and Kacperczyk (2021) seem to agree with Monasterolo and de Angelis (2020). They examine the impact of carbon emissions in relation to stock returns over 2015-2017 and provide evidence of a carbon premium, with companies which generate higher carbon emissions demonstrating significantly higher stock returns. Seltzer et al. (2021) study companies with poor environmental performance and demonstrate results of lower credit ratings, over the 2009-2017 period in the US, especially in states where the environmental regulations are stricter. They show that Paris Agreement as a shock event affected negatively "polluter" companies.

The implications of the COVID-19 pandemic in the financial markets has drawn attention in the recent literature, showing ambiguous results. The resiliency of the high-ESG rated firms during the COVID-19 crisis is documented by Albuquerque et al. (2020). They analyze Environmental and Social ratings and stock returns from 2017 to 2019 in the US and show that high rated firms exhibited lower volatility after the COVID-19 outbreak. On the same page, Ding et al. (2020) evaluate 6,000 firms globally and argue that firms with better CSR performance demonstrated resilience during the pandemic. Lower idiosyncratic stock volatility with abnormal returns for high-ESG rated European firms is evidenced by Engelhardt et al. (2021). They analyze a sample of European firms in 2019-2020. Bae et al. (2021) and Demers et al. (2021) contribute to the literature with similar mixed results. Both studies utilize ESG data from 2018-2019 and analyze US stock returns following the COVID-19 outbreak. They show no relation between ESG performance and stock returns and thus CSR does not immunize firms in uncertain times.

On the selection of the appropriate ESG portfolio construction techniques, numerous studies have provided deviate evidence on whether ESG and environmental investments outperform (Clark et al., 2015), while others tend to blame other "quality" factors, such as high profitability and conservative investment (Bruno et al., 2021). Two widely used SRI approaches (Clark et al., 2015), are the simple exclusionary approach, in which companies from controversial industries are excluded, and the inclusion strategy, in which inclusion screens are applied.[6] One of the most prevalent portfolio construction techniques is sorting. Sorting stocks based on certain criteria, has been broadly applied in investment strategies, and basically in the momentum strategy, i.e. buying the best-in-class stocks and selling the laggards, or other alternatives. In such way it has been applied in portfolio construction techniques utilizing ESG criteria. An ESG momentum strategy applied by Nagy et al. (2016) over the period 2007-2015, advantaging stocks with increasing ESG scores. Their results showed that the ESG momentum strategies outperform the market. They provided also evidence of factor and industry contributions, showing a tilt towards mid-cap companies. A similar combined approach is followed by Verheyden et al. (2016), who apply as an additional step a preliminary ESG screening, excluding stocks with the lowest ESG scores. Their findings show improved returns, lower volatility by excluding low-rated companies with an ESG momentum strategy. Giese et al. (2019) sort companies using the ESG aggregate score, utilizing data of 1600 stocks during 2007-2017. They show a link between ESG rating and financial performance and provide evidence of outperformance for the higher ESG-rated companies. They assess the ESG momentum strategy analogous to Nagy et al. (2016) and find significant financial impact on stock returns. They evidence though, that ESG momentum works only in the short-term. Using 1600 stocks from developed markets from 2006 to 2019, Giese et al. (2021) extend their previous research by analyzing the aggregate ESG score, ESG factors individually and their sub-

[6] Exclusion and inclusion strategy have been widely applied in the literature (Alessandrini and Jondeau, 2020; Dimson et al., 2020; Trinks et al., 2017; Nagy, Kassam and Lee, 2016; Clark et al., 2015; Hong and Kacperczyk, 2009; Diltz, 1995).

components. They show a positive return and positive alpha for higher ranked companies. However, environmental and carbon emission indicators show weakness in long-term results. Early studies also, using similar sorting techniques show that ESG criteria can be a substantial gauge in the investment process (Derwall et al., 2005; Yamashita et al., 1999; Kempf and Osthoff, 2007; Carhart 1997).

Studies though, have provided evidence of applied strategy, sector, equity or geographic factors biases. Bruno et al. (2021) assess several studies using sorting and ESG momentum techniques and provide evidence of sector biases results, and exposures to high profitability and equity style factors. Similarly, Auer and Schuhmacher (2016), show that ESG rated companies outperform only when applying various portfolio screens with geographic and industry criteria. This study yields several combinations of ESG, geographic and industries conditions that should be considered in order to provide risk-adjusted performance.

Sorting techniques have shortcomings. They are driven by short-term variations and lead to data driven results. Increased attention to ESG leads to inflated short-term performance (Bruno et al., 2021), while in the long-term, lower expected returns (Cornell, 2020; Pastor et al., 2021). Likewise, Giese et al. (2019) evidence that ESG momentum strategy shows outperformance in the short-term. Long-term positions can be easily overestimated when considering ESG short-term driven results.

Hence, in an attempt to avoid such shortcomings, we utilize the clustering technique with environmental scores of ESG. In order to overcome biases associated with the applied strategy, in our empirical analysis we employ seven criteria for optimum portfolio rebalancing. Our rotation criteria have no memory of past values after a particular time shift. More specifically, while the lookback concept keeps past information for the next rebalances, we accept that after several rebalances, the shift in time is bigger than the lookback period, thus keeping no past information compared to the first rebalance. This recursive procedure happens many times, since the lookback period is much smaller than the entire period of analysis of 2003-2022. Therefore, there will be enough such realizations with zero or low memory from the past, or in other words, enough restarts of the chances of inclusion of an asset in a portfolio. Clustering methods is a common practice for portfolio optimization techniques (Tola et al., 2008; Chen and Huang, 2009). Our aim is to provide robust results explained by the environmental performance and not dependent on the selected specification, as explained more in depth in Section 3. Finally, our analysis considers also shock events such as the Paris Agreement, the COVID-19 period, as well as the Russia-Ukraine war, and examines if any spillovers effects can be identified in our portfolios.

3. Data, Clusters analysis & Empirical Methodology

3.1. Data

The two main data categories concern the daily returns of the S&P 500 companies and their environmental score of ESG index, respectively. All data have been acquired from Thompson Reuters – Eikon (2022). Thomson Reuters database provides ESG scores of approximately 9,000 companies globally, which measure companies' ESG performance based on over 400 different metrics under 10 categories, across the three Environmental, Social and Governance pillars.[7] The timespan of our sample is from 2003 until 2022, in daily frequency for returns and in annual frequency regarding environmental scores. As mentioned above, the ESG framework, sets four different clusters from A to D, based on the ESG score of each year (each quartile, starting from zero score until one hundred, represents one cluster from D to A respectively), with their respective subdivisions, plus [+], normal and minus [-]. The same cluster notation can be used in the disaggregate level based only on one ESG index component, E, S, or G score. This analysis uses only the E score to allocate the assets to clusters "A" to "D". To that end, we use the highest

[7] Thomson Reuters ESG data are widely used in the CSR literature (Albuquerque et al., 2020; Bae et al., 2021; Demers et al., 2021). The relevant metrics are grouped into 10 categories: resource use, emissions and innovation (Environmental pillar), workforce, human rights, community, and product responsibility (Social pillar) and management, shareholders and CSR strategy (Governance pillar). For an overview and analysis of the ESG report and ratings providers, please refer to Huber and Comstock (2017).

frequency in years from the entire period an asset remains in a cluster to make the final allocation. Nevertheless, there is the case where a cluster remains approximately equally in two or more clusters, and hence the aforementioned criterion loses useful information. Usually this happens when high change rates of ESG scores exist. In this case, we classify our assets to a new cluster, called "T" for transition. To avoid discretionary criteria for including an asset to cluster T, we use the geometric mean of the absolute changes of the environmental performance (i.e. the "E" component of the ESG scores) of the examined period, and select only the assets whose means are bigger from the 3rd quartile (Q_3). We selected the geometric mean as a measure of central tendency, since the samples of "E" scores of many assets include outliers, resulting in potential over or underestimation of the measure. Quoting Clark-Carter's (2010) words, "The geometric mean has an advantage over the arithmetic mean in that it is less affected by extreme values in a skewed distribution", which is our case here. Hence, we have:

$$Cluster_i = \begin{cases} Cluster\ T, & Geom.Mean_{\Delta(E)} \geq Q_3 \\ Cluster_i(\max \text{Years}), & Geom.Mean_{\Delta(E)} < Q_3 \end{cases} \quad \text{(eq.1)}$$

where $\text{Geom.Mean}_{\Delta(E)} = \sqrt[n]{(1+|\Delta E_{2-1}|)(1+|\Delta E_{2-1}|)\dots(1+|\Delta E_{n-(n-1)}|)} - 1$

Table (1) below presents examples of cluster allocation before and after the use of the logical statement. As shown, there are cases where the highest frequency in years does not set the cluster, but the high change rate of environmental performance does through its geometric mean.

[Insert Table 1, about here]

To estimate the computational time of cluster allocation, we use the big O notation for complexity classes. Eq.1 is a conditional statement; hence we take the maximum runtime from the two statements included. The upper statement is a basic assignment operation, while the lower is a nested condition of basic operations as well. Although the total time is a function of the size n of all assets, the overall complexity class remains a constant time O(1) class, given we know the number of assets in advance, which remains unchanged. As shown in the following sections, our annual and daily data, of E scores and returns respectively, are not used in the same estimation block together, avoiding us to use aggregation or disaggregation techniques for mixed frequencies. While assets allocations in clusters, as explained above require only the annual data of the E scores, the clusters dependence analysis, portfolio rebalancing and its post hoc analysis use only daily returns.

3.2. Clusters dependence analysis

To have a comprehensive insight of the features of the clusters, which might allow us to explain better some of our empirical results, we make an initial analysis of their structure dependencies. We do this on two levels, first on their linear relationships using principal component analysis (PCA), and second on the extreme values of their performance using the extreme value theory. It is well known that PCA aims to reproduce original variance structure of the initial dataset, here each cluster's assets, using a relatively small number of eigenvectors. Our focus is to assess the level of homogeneity of the assets of its cluster. The more variance explained with fewer principal components after eigenvector-eigenvalue decomposition, the more homogenous the behaviour of the cluster in terms of stronger linear dependencies among assets, as eigenvalue λ_i represents the amount of variability between the assets of a cluster explained by the respective eigenvector.

Figure (1) and Table (2) below present the results from the PCA Analysis. As shown, the best environmental clusters exhibit larger values of their first eigenvalues, and reach the same level (e.g. 0.6) of explained variances using less eigenvectors compared to less environmentally friendly clusters and Cluster T as well. These results suggest stronger ties or linear dependencies in assets belonging to clusters with higher environmental performance and the opposite for clusters D or T. In other words, assets in higher environmental status look more towards the same direction than assets on the rest of the clusters.

[Insert Table 2, about here]

[Insert Figure 1, about here]

We further explore dependence structures, within each cluster, in their extreme values. Towards that direction, we implement a multivariate extreme value test within each cluster based on max-stability. Also, as a measure of tail fatness, we estimate the Obesity Index (Cooke et al., 2011). Before testing, we address the issue of *i.i.d.* variables using McNeil et Frey (2000) two-step approach. By default, an ARMA(p1,q1) process with GARCH(p2,q2) errors is assumed. Nevertheless, after testing our series for autocorrelation and partial autocorrelation[8], we see that while several autocorrelations seem significant under the *i.i.d.* hypothesis, the autocorrelations are within the confidence bands of the GARCH hypothesis, as explained in Proberts et Boshnakov work (n.a.). Given this, we use a pure *GARCH(p, q)* model as below:

$$X_t = \mu_0 + \sigma_t Z_t, \qquad \text{(eq.2)}$$

$$\sigma_t^2 = a_0 + \sum_{i=1}^{p} a_i (X_{t-i} - \mu_o)^2 + \sum_{j=1}^{q} \beta_j \, \sigma_{t-j}^2$$

where $(Z_t)_{t \in \mathbb{Z}}$ a strict white noise with mean 0 and variance σ^2, independent of $(X_s)_{s<t}$ for all t. Lags p, and q are identified with the AIC Information criterion.

The standardised residuals of eq.(2) replace our initial log return series in the following analysis of extreme values. Also, we separate return series for positive and negative returns to address the non-symmetric behaviour financial returns usually shown. Figure (2) below presents autocorrelations and partial autocorrelations for an asset's log return series (up) and for the standardised residuals after the GARCH filter (middle). Also, Q-Q plots for them (returns are in left and residuals in right) are depicted respectively.

[Insert Figure 2, about here]

[8] All test results are in the data file

Regarding the Obesity Index (*OI*), it is a scalar measure and it is defined for a random variable X as:

$$OI(x) = P(X_1 + X_4 > X_2 + X_3 | X_1 \leq X_2 \leq X_3 \leq X_4\,, \quad X_i \;\; are\; i.i.d.\,, \qquad \text{(eq.3)}$$

It is expected that this probability is higher in the case of heavy-tailed distributions than in thin-tailed ones. Given the joint density of n order statistics:

$$f_{1,2,\dots,n}(x_1, x_2, \dots, x_n) = \begin{cases} n! \prod_{i=1}^{n} f(x_i), & x_1 < x_2 < \cdots < x_n, \\ 0, & otherwise \end{cases} \qquad \text{(eq.4)}$$

And the restrictions $x_1 + x_4 > x_2 + x_3$ & $x_1 < x_2 < x_3 < x_4$

$OI(X)$ equals the following integral:

$$OI(X) = 24 \int_{-\infty}^{\infty} f(x_1) \int_{x_1}^{\infty} f(x_2) \int_{x_2}^{\infty} f(x_3) \int_{x_3+x_2-x_1}^{\infty} f(x_4)\, dx_4 dx_3 dx_2 dx_1 \qquad \text{(eq.5)}$$

In our data, we estimate *OI* by bootstrapping. We run 2^8 trials of samples of 4 points for each asset return series, in every cluster[9]. Although *OI* does not completely reflect the tail index or the extreme value index, since it is applied to the whole distribution, it is important as it does not require estimating a parameter of an ex-ante hypothetical distribution whose moments may or may not be infinite. *OI* follows the idea that larger values are further apart or that the sum of two samples is driven by the larger of the two. Assuming our returns for each cluster, are identically distributed, then the higher the *OI* value the heavier the tail of our series. Table (3) below, includes the estimates of *OI* for all the clusters and their 95% confidence bounds, separately for positive and negative returns.

[Insert Table 3, about here]

[9] Code file "Obesity Index.R" is available in the data file

As shown, negative tails are heavier compared to positive ones, for all the clusters they are referring to, but the T where the opposite is true. This is usual in financial returns as there is asymmetry between negative and positive disturbances, with bad news resulting in deeper losses compared to returns from good news. The second fact is that the worse the environmental cluster the heavier the tail, while for Cluster T, it seems to be closer to the environmentally good clusters A, B than to the bad ones, C and D., Given these two facts, it can be argued that Cluster T, performs rather not so conventionally compared to the other clusters, as it suffers less from bad news and earns more from good news, while it is not affected so much from extreme values like clusters C or D do.

Moving to tail dependences, our focus is to examine dependence structures of extreme values in a portfolio of assets. To that end, we skip bivariate tests of extreme value copulas, despite their broad application in financial analyses, and proceed with multivariate extreme tests, that correspond to a portfolio of assets. Hence we apply the multivariate test of Kojadinovic et al. (2011) for a d-dimensional random vector $\boldsymbol{X}$ of *i.i.d.* elements and continuous marginal cumulative distributions F_1,.., F_d.

It is well known that, based on Sklar (1959), the c.d.f. of this vector is,

$$F(\boldsymbol{x}) = C\{F_1(x_1), \dots, F_d(x_d)\}, \;\; \boldsymbol{x} \in \mathbb{R}^d, \qquad \text{(eq.6)}$$

where $C: [0,1]^d \to [0,1]$ is a copula capturing the dependence of vector's X components. If, additionally, the following applies to copula C:

$$C(\boldsymbol{u}) = \left[C\left(u_1^{\frac{1}{r}}, \dots, u_d^{\frac{1}{r}}\right)\right]^r, \forall \boldsymbol{u} \in [0,1]^d, \forall r > 0 \qquad \text{(eq.7)}$$

The unknown copula C belongs to the class of extreme-value copulas. Keeping Kojadinovic et al. notation the test for the hypothesis $H_o: C \in \textit{extreme value}$ is based on processes of the form:

$$\mathbb{D}_{r,n}(u) = \sqrt{n}\left[\left(C_n\left(u_1^{1/r}\right)\right)^r - C_n(u)\right], u \in [0,1]^d, r > 0 \qquad \text{(eq.8)}$$

$$\mathbb{D}_{r,n}(u,t) = \sqrt{n}\left[\left(C_n\left(u_1^{1/r}, t\right)\right)^r - C_n(u,t)\right], u \in [0,1]^d, r > 0$$

which can be used for the hypothesis H_o. Thus, the proposed tests for multivariate extreme value dependence, considering Cramer-Von Mises functionals, are:

$$S_{r,n} = \int_{[0,1]^d} \{\mathbb{D}_{r,n}(\boldsymbol{u})\}^2 d\boldsymbol{u} \quad \& \quad T_{r,n} = \int_{[0,1]^d} \{\mathbb{D}_{r,n}(\boldsymbol{u})\}^2 dC_n(\boldsymbol{u}) \qquad \text{(eq.9)}$$

Based on these Test statistics, we run a sufficient number of random samples (portfolios) of assets of length of six, ten and fifteen, just like the portfolios in our methodology that follows. We examine again positive and negative returns, separately per cluster. Table (4) includes, the times the Null hypothesis is not rejected, that is the times the portfolios extreme values belong to the family of extreme value copulas. Moreover, Figure (3), shows the p-values of these tests for an overview.

[Insert Table 4, about here]

[Insert Figure 3, about here]

As we move from Cluster A to Cluster T, it becomes more probable that the unknown portfolio copula belongs to the class of extreme-value copulas. Also, this is the case as the portfolio's length increases, regardless the cluster it belongs to.

All in all, given the PCA analysis, the Obesity index and the multivariate extreme value test dependencies, the results sheds light to the amphoteric role of Cluster T. On one hand, this cluster has the lowest homogeneity, in terms of linear dependency, resembling more clusters C and D, on the other thought, its tails are closer to these of clusters A and B regarding their fatness. In addition, it responds inversely compared to all the other clusters, with good news having stronger spillover effects on its right tail compared to bad news asymmetries on the left. Last, if seen in a portfolio performance context, its extreme dependencies seem to belong more often to extreme-value copulas than with the other clusters making its performance less stable. The results in this analysis, stresses the volatile character of the low ranked clusters C and D, both when normal (with PCA) and extreme value conditions are examined (with obesity and *ev* tests), while Cluster T has mixed results.

3.3. Methodology

The proposed methodology is structured in four levels. In the first level we classify the S&P 500 companies to an environmental cluster (EC). We select the clustering instead of sorting as the latter inserts significant short-term variations, whereas clustering with environmental criteria can identify a long-term relationship between environmental and financial performance. Bruno et al. (2021) show that increased attention to ESG leads to inflated short-term performance, while in the long-term to lower expected returns (Cornell 2020; Pastor et al., 2021). The second level includes standard but straightforward criteria for optimum portfolio rebalancing on monthly basis in the clusters of the S&P 500 companies that we have already created. We utilize various criteria, regarding the applied optimization strategy and the portfolio parametrization, in order to overcome the problem of contradictory results that might be affected by the one or the other selected specification. In the third level, in order to provide robust conclusions attributed only by the environmental performance and not by other factors such as the selected strategy, we feed the results out of the second stage and apply a non-parametric factorial analysis of variance[10], to

[10] Results from a permutational multivariate analysis of variance are also available in the following doi: 10.17632/2jsb9z3d69.4.

identify true factors of variation among the S&P 500 stocks. The two factors of interest here are the environmental cluster and the applied strategy. We examine the significance of their effects separately and the effect of their interaction. Last, in the fourth level, we run appropriate nonparametric hypothesis tests to reach in out-of-sample conclusions and answer our main research question.

Hence, following the clustering of the S&P 500 companies into environmental clusters as described in the previous section, we start our analysis by applying seven standard rotation criteria that work as strategies[11]. These are: the Arithmetic Mean (AM), the Mean-Variance[12](U), the Minimum Volatility (S), the Skewness (SK), the Kurtosis (KR), the Sharpe-Ratio (SR) and the Normality Test (NT). Constant rebalancing, on a monthly basis, works as optimum asset rotation of the portfolios of the five different environmental clusters. We choose a number of rotation criteria to prevent biased errors and false conclusions spurring from one only strategy.

Towards this direction, we compute the respective sample measures of the four central moments plus the mean-variance and the Normality measures that as follows:

$$\mu_{R_{t,i}} = E(R_i) = \sum_{j=t-m}^{j=t} r_{j,i}\, P\big[R_i = r_{j,i}\big], i = 1, \dots, n \qquad \text{(eq.10)}$$

$$\sigma_{R_{t,i}} = \sqrt{E[(R_i - E[R_i]^2} = \sqrt{\sum_{j=t-m}^{j=t}(r_{j,i} - E[R_i]^2)\, P\big[R_i = r_{j,i}\big]} \qquad \text{(eq.11)}$$

$$U_{R_{t,i}} = \mu_{R_{t,i}} - \lambda {\sigma_{R_{t,i}}}^2 \qquad \text{(eq.12)}$$

$$SR_{R_{t,i}} = \mu_{R_{t,i}} / \sigma_{R_{t,i}} \qquad \text{(eq.13)}$$

$$SK_{R_{t,i}} = [(R_i - E(R_i))^3]/\sigma_{ti}(m)^3 = \sum_{j=t-m}^{j=t}\big(r_{j,i} - \mu_{R_{t,i}}\big)^3 P\big[R_i = r_{j,i}\big]/\, \sigma_{ti}(m)^3 \qquad \text{(eq.14)}$$

[11] Standard rotation criteria used in previous research (Alexopoulos, 2018), addressing different research questions and assets.
[12] We use the mean variance utility U, with λ equals to unit allowing a relative low degree of risk aversion.

$$KR_{t,i} = E[(R_i - E(R_i))^4]/\sigma_{ti}(m)^4 = \sum_{j=t-m}^{j=t} \left(r_{j,i} - \mu_{R_{t,i}}\right)^4 P\left[R_i = r_{j,i}\right]/\, \sigma_{ti}(m)^4 \qquad \text{(eq.15)}$$

$$NT_{t,i}(m) = \quad {SK_{R_{t,i}}}^3 + \left(KR_{t,i} - 3\right)^2 \qquad \text{(eq.16)}$$

where *R* is the returns variable and it is assumed a discrete random variable, *m* stands for the look-back period, and *i* for the asset. Accordingly, we rank all the stocks of each environmental cluster based on the applied criterion each time, and keep the top *k* of them, for the next month's portfolio return.

We select the first top *k* stocks for inclusion in our portfolio after having them sorted in descending when the criteria of AM, U, SR and NT are used and in ascending order for the rest S, SK and KR criteria. To avoid data driven results, we apply, for each strategy and environmental cluster, a number of simulations with different lookback periods *m* and portfolio lengths *l* each time, that work as a sort of sensitivity analysis. To that end, we examine three lookback periods and three portfolio lengths, making it a total of 9 different runs per strategy and per group. In aggregate we simulate 63 cases per each of the five environmental clusters. For each case-portfolio, besides its arithmetic mean, we estimate standard performance measures, such as the geometric mean, the volatility, the Sharpe Ratio, the percentage of periods of positive returns, and indicators of the downside risk, such as the Maximum Drawdown, the 0.95VaR, and so on. A selection from the above measures is used as response variables in the next level of our research. The use of clustering instead of sorting, based on the environmental scores of ESG, makes our rotation strategies more efficient, as all stocks belonging in a cluster are given the same chances of inclusion in a portfolio after rebalancing while keeping at the same time the concept of environmental performance. On the contrary, if the sorting technique had been applied, this might not be possible as we might consider each time the same top environmental performers for rotation in a portfolio.

Proceeding in our analysis, we examine the significance of the main and the interaction effects of our examined factors, i.e. the environmental cluster and the applied strategy. The positive answer out of this analysis is basic, for the following post hoc clustering analysis to be valid. Therefore, we first test for the validity of anova assumptions. The response variables are by default independent as they stem from the different ESG groups of the S&P 500 stocks. We examine the existence of significant outliers, of normality and of homogeneity of variance for each combination of the groups of the two factors. For normality we use the Shapiro-Wilk Test (1965) and the D'Agostino-Pearson test (1973). For homoscedasticity of variance we select the Levene's Test (1961). As it is shown, in the next section, the test results answer in negative on the non-violation of the assumptions of Anova analysis. To overcome this, we select to use the non-parametric factorial analysis of variance approach proposed by Wobbrock et al. (2011) based on the concept of the aligned ranked data transformation.

Although rank transformations have appeared in statistics for years,[13] there is criticism about non robust results when interaction effects are examined in a factorial design (Salter and Fawcett, 1993). In our two-factor analysis of environmental cluster (EC) and strategy (ST), we examine the significance of the effect of each factor (EC & ST) and of their interaction (EC*ST) on a response variable of interest, stripping from it each time, all effects but one.

To achieve this, we use the Aligned Ranked Transformation (ART) of Wobbrock et al. (2011), which includes a pre-processing step that first aligns the data for both single and interaction effects, before assigning the ranks. This alignment helps in estimating the effects as marginal means, which then can be isolated from the effects of the rest factors and their interactions as well. Hence, ART transformation is twofold, first it does not require normality of our data, and second it estimates the true effect of each factor and of its interaction, stripped from the other factors effects.

The aligned transformation from the initial response variable is estimated as follows:

[13] Conover and Iman (1981), Sawilowsky (1990)

$$Y_{aligned} = Y_{raw} - cell\ mean + estimated\ effect \quad \text{(eq.17)}$$

Cell mean is the average of all responses with the same levels of EC and ST factors. In our data is the average of 9 responses per response variable. The estimated effect (EE) is for the main effect for EC and ST:

$$EE_{EC} = \overline{EC_\iota} - \mu \ \text{ and } EE_{ST} = \overline{ST_\iota} - \mu \quad \text{(eq.18)}$$

and for the interaction effect of EC*ST:

$$EE_{EC*ST} = \overline{EC_\iota * ST_J} - \overline{EC_\iota} - \overline{ST_\iota} + \mu \quad \text{(eq.19)}$$

The $Y_{aligned}$ responses are then ranked, with the smallest value receiving rank 1, the next smallest rank 2, and so on.[14] The transformed results are fed a standard factorial Anova in order to examine the significance, or not, of each factor (EC and ST) and of their interaction. F- statistics and p-values are presented in respective tables for all the transformed response variables.

The last step in our empirical methodology includes a complex contrast hypothesis and a post-hoc analysis. Contrast tests ask a specific question by framing a "custom" hypothesis, as opposed to the general anova null against alternative hypothesis. In our analysis, we attempt to identify potential environmental clusters whose financial performance seems to be more distant or in other words more "islanded" in comparison with the other ECs. To that end we examine for each response variable a number of custom contrast tests, we call "Diffs of Diffs". Of all the possible combinations of the factors EC and ST, we ask if the difference EC(i)-EC(k) is different with ST(m) compared to ST(l). To that end, we run appropriate F-tests on the following:

Null hypothesis:

$$H_0: [\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(m)} - [\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(l)} = 0$$

against the alternative

[14] All the aligned and ranked transformations for all the response variables are available in the following doi: 10.17632/2jsb9z3d69.4.

H_1: $[\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(m)} - [\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(l)} \neq 0$

where $i \neq k$, $i, k = 1, \dots ,5$, & $m \neq l$, $m, l = 1, \dots ,7$.

Hence, we count the number of the rejected H_0 per cluster and response variable:

$$\sum_{R=1}^{5} \left[\sum_{H_0=Rejected} 1 \right]_R$$

(eq.20)

where R stands for the response variable.

The EC_i with the highest number in significant differences among all the response variables has the most islanded performance between them.

The second path of the post-hoc analysis involves a standard non-parametric Mann-Whitney (MW) U test, for two independent samples. For each response variable and rotation strategy, we have five independent samples based on the EC. The samples are populated from all the possible combinations of the two sets listed below.

First set for the lookback period, LB : $\{12,18,24\}_{\text{months}}$

Second set for the portfolio length, PL: $\{6,10,15)_{\text{assets}}$

In each response variable we organize our hypothesis of MW tests in a pair-wise mode between the five different environmental groups. We use an upper-tailed test since we sort them in descending order based on their sample mean $\bar{X}$. That is:

$\bar{X}_{sample(i)} > \bar{X}_{sample(i+1)} > \cdots > \bar{X}_{sample(i+4)}$

Therefore, the null hypothesis is:

H_0: $\mu_{sample(i)} - \mu_{sample(i+1)} = 0$

Against the alternative

H_1: $\mu_{sample(i)} - \mu_{sample(i+1)} \geq 0$

Where i=1,…,5 stand for the different ECs

According to the MW-U test results, we calculate a cumulative ranking for each cluster by adding all the individual rankings for all the applied strategies and response variables, as shown in eq (14). The smallest cumulative ranking corresponds to the best performance, the next one to the second best and so on.

$$Total\ Ranking_{EC(i)} = \sum_{rv=1}^{5} \sum_{st=1}^{7} p_{EC(i),rv,st} \quad \text{(eq. 21)}$$

Where $p_{EC(i),r,st}$ is the individual ranking of the i environmental cluster, in *rv* response variable and with *st* rotation strategy. In case of non-rejection of the null hypothesis, the two examined ECs are given an equal rank. For the response variables of Total Return, Sharpe Ratio, and Geometric Mean the highest rank, i.e. 1st, is assigned to the biggest significant $\mu_{sample(i)}$, on the opposite for the response variables of MaxDD*Vol and 0,95VaR, it is assigned to the smallest one. We assume equal weighting for all the response variables.

Last, for the sake of completeness, we consider the impact, if any, of significant shock events. Upon that, we select the Paris Agreement adopted in December 2015, which is considered a milestone event against climate change and examine if potential positive spillovers exist in our portfolios of all the ECs. We expand our analysis following the COVID-19 outbreak in February 2020, to assess the effects of the pandemic on the relationship between environmental and financial performance, as well as the consequences of the war in Ukraine in 2022. We study the year before and after the shock events, by contrasting the different relative changes in annual performances scaled to the annual relative changes of the market's index, SPY (i.e. elasticities) as shown in eq. 22 that follows.

$$D_{EC_i} = \left| \left[{(RV_{i,t} - RV_{i,t-1})}\Big/{RV_{i,t-1}} \right] \Big/ \left[{(RV_{SPY,t} - RV_{SPY,t-1})}\Big/{RV_{SPY,t-1}} \right] \right| \quad \text{(eq. 22)}$$

Where RV stands for the response variables as described before, i is for the EC, and t=2016 for Paris Agreement, t=2020 for COVID-19 and t=2022 for war in Ukraine. The hypothesis of this out-of-sample comparison is if the changes on the annual performances (before and after the shock event) are significantly different, among all the ECs or if they are all moving in the same response path in magnitude and sign after the shock event, regardless their environmental status. To that end, we apply the same nonparametric MW-U test in pair wise mode between ECs as before, using upper-tail tests after having ordered their sample means in descending order. For each cluster we estimate a sample of 7 different values of elasticities, based on the respective strategy each time. All in all, Figure (4) below presents, the applied framework of our methodology.

[insert Figure 4, about here]

4. Empirical results and discussion

Table (5) presents the rebalancing results of the top-10 and bottom-10 portfolios from all the environmental clusters. Given all the possible parametrizations for rotation criteria, lookback periods and portfolio sizes, the results are supportive of a critical view. To assess the portfolios performance, for each parametrization we estimate a set of return and risk related variables.[15] A selection of them is used, as response input variables in the next levels of our analysis.

[Insert Table 5, about here]

Prima facie, we observe that the low-rated environmental companies in cluster "D", as well as those in cluster "T" with a dynamic environmental performance produce higher returns on average, contrary to the best environmental performers in "A" and "B" with lower on average performance. Our findings are consistent with Bolton and Kacperczyk (2021), Monasterolo and de Angelis (2020) and Dorfleitner et al. (2020), as well as with other empirical studies[16], supporting the argument that low environmental performance is associated with higher returns, often with highly diversified portfolios in global scale. On the other hand, a positive link between environmental

[15] Ar. Mean, Gm. Mean, Volatility, Sharpe Ratio, Total Return, Min, Max, Skewness, Kurtosis, %+ Periods, Ar. Mean of %+, Vol. of %+, %- Periods, Ar. Mean of %-,Vol. of %-, MaxDD, Vol*MaxDD, 0.95-VaR. All results are available in the following doi: 10.17632/2jsb9z3d69.4.

[16] Hong and Kacperczyk, 2009; Blitz and Fabozzi, 2017; Dunn et al, 2018.

leaders and reduced risk is suggested with companies in clusters "A" and "B". In line with Hoepner et al's (2018) findings, best environmental performers are less risky in contrast to others with low-rates ones. Hoepner et al. (2018) find that ESG engagement reduces firms' downside risk, as well as their exposure to downside risk factor. Similarly, Martiradonna et al. (2022) construct portfolios including and excluding green bonds, using various allocation strategies, to assess their risk performance while focusing on the COVID-19 crisis. They suggest that green portfolios outperform in terms of risk, while providing diversification benefits, including the period of COVID-19. Baker et al (2022) focus on green corporate bonds and suggest also that AAA-rated bonds present significantly low risk.

Our results evidence improved risk-adjusted returns for companies with a dynamic environmental performance in clusters "T", indicating a potential diversification effect, advantaging the strengths of both the environmental leaders and laggards. In the literature, Nagy, Kassam and Lee (2016) construct two investment strategies considering the ESG current and dynamic rating, and demonstrate evidence of different financial performance between the portfolios. The "ESG tilt" strategy, which is more long-term in nature, with an alpha signal related to the current ESG score, and the "ESG momentum" strategy which overweights firms with dynamic ESG status, utilizing an alpha signal related to the change in the score. Both portfolios outperform the benchmark. They exhibit that the ESG tilt portfolio lean towards less volatile stocks, while ESG momentum towards more profitable stocks. A qualitative representation of our remarks are shown in Figure (5).

[Insert Figure 5 about here]

Figure (6) shows the time series and the histogram of returns for the portfolio of environmental cluster "D" and of "A" respectively. The larger width in the time series band, the fatter tails and large spread of the histogram, as well as the estimated statistics on the right, support the previously stated argument for cluster "D" compared to "A" accordingly.

[insert Figure 6, about here]

The results, for normality, and equality of variance, in Table (6) answer in negative to the non-violation of the standard Anova assumptions, justifying the adoption of the ART anova. In most of the cases there is neither normality nor homoscedasticity of variance.[17]

[Insert Table 6, about here]

Table (7) presents the results of the ART factorial nonparametric anova.[18] For each response variable, we examine the significance of the main and the interaction effects of the factors EC and ST. Based on the F-statistic values, it can be induced, that portfolios listed in the S&P 500, behave differently depending on their environmental status and or the applied optimization strategy.[19] Although in the literature there are further analyses examining the causal effects between environmental factors and financial performance,[20] the ART anova and the post hoc analysis results could be used here as an initial guide or a map for efficient asset selection depending each time on the investors needs either for returns or stability..

[Insert Table 7, about here]

Given the significant main and interaction effects, that EC and ST factors have on portfolios, the next question to be answered is if there is an environmental cluster with an idiosyncratic financial performance compared to the others, or if all clusters are in the same behavioural "neighbourhood". To answer this, we created a specific contrast test, called "Diffs of Diffs", which we applied

[17] Results for symmetry are included in the data file with the respective box-plot figures are available in the following doi: 10.17632/2jsb9z3d69.4.

[18] All the Aligned and ranked transformations are available in the following doi: 10.17632/2jsb9z3d69.4

[19] For robustness purposes, we additionally run a permutational multivariate analysis of variance given all the response variables. The results are on the same page with the ART factorial analysis, supporting the argument of significant effects of the environmental status.

[20] The causal relationship between corporate environmental & social performance and financial performance is investigated thoroughly in the literature (Preston & O'Bannon, 1997; Orlitzky et al., 2003; Waddock & Graves, 1997; Makni et al., 2008; Nelling & Webb, 2009; Hang et al., 2018 etc.)

recursively for all possible comparisons among clusters and strategies. The results in Table (8) show the number of all the significant differences of eq.20 of an environmental cluster in pair-wise mode with the others, given all the different criteria.[21] In total there are 84 different contrast tests per cluster. As shown by the same number of significant differences between clusters in pair wise mode, clusters "D", "C" and "T" belong to the same behavioural "neighbourhood", in contrast with environmental leaders "A" and "B" which have also common characteristics. Cluster "D" nevertheless, has more significant differences with the others, suggesting a more distinctive (financial) behaviour.

[Insert Table 8, about here]

We visualize the results of Table (8) in Figure (7), where the test difference between two clusters is measured in a vertical black line, (e.g. A-B). Comparisons for significant differences, given the applied criteria, can be deduced from the length of these lines. In case of a significant difference the two examined lengths should be of unequal sizes, the opposite suggests non-significant differences. In Figure, we have drawn the line for all the "A-B" differences.

[Insert Figure 7, about here]

Table (9) includes the results of the MW non-parametric upper-tail tests, implemented in ordered and independent samples, as explained in the methodology. The results align with the previous in-sample findings of the first-level portfolio rebalancing. In more detail, we see that the highest rankings (smaller values), in terms of return, are in "D" and "T" while based on the cumulative ranking from the two risk measures, it is clusters "A", "B" and "T" respectively. Out of these in-

[21] All contrast test results with their respective p-values are available in the following doi: 10.17632/2jsb9z3d69.4.

and out-of-sample results, we conclude that cluster T assets work well both in return and risk terms. [Insert Table 9, about here]

That is to say, in case an investor wishes to diversify their portfolio towards profit, then the investor should first look for potential replacements or additions, with stocks from “D” and “T” environmental pools, while if one aims at less volatile environments, then they should mix their portfolios with stocks belonging to “A” and “B” clusters. Stocks listed in cluster “T” should be considered as an optimal option demonstrating best risk-adjusted returns.

Table (10) results, concerning the Paris Agreement, show in general environmental leaders in cluster “A” have a better response after its ratification. We recognize that their response could be the aggregate effect of other underlying factors, therefore potentially decreasing the actual efficacy of the Agreement in the financial sector, with a negative sign for D- and B-rated companies, having underperformed in return and risk measures, based both on the in- and out- of sample results.

[Insert Table 10, about here]

Applying in the same way a MW non-parametric upper-tail test, we assess the impact of the COVID-19 crisis. Table (11) indicates no significant differences in clusters’ relative changes after the COVID-19 outbreak. Our findings are consistent with Bae et al. (2021) and Demers et al. (2021) who analyze the immunity of sustainable investing during COVID-19 pandemic and show no relation between ESG performance and stock returns. The results from MW non-parametric test in Table (12) show analogous results following the Russia-Ukraine war, with clusters “B”, “D” and “C” appearing slight negative effects.

[Insert Table 11, about here]

[Insert Table 12, about here]

All in all, our results have created an initial guide for portfolio management based on the concept of corporate environmental responsibility reflected through the ESG rankings and values. Our research identifies the environmental status of a firm as a key characteristic for its financial performance and a parameter not to be overlooked in a portfolio formation or update.

5. Conclusions

The focus of the paper is to examine whether the concept of environmental corporate social responsibility has divergent effects on the financial behaviour of portfolios of stocks, and if so, how this could be exploited by potential investors. Towards this direction, we used the environmental pillar of the ESG ranking, of all the stocks of S&P 500, and clustered them in five environmental groups based on a proposed logical statement. To avoid biased conclusions derived from a specific strategy or rotation criterion for portfolio rebalancing, we employ seven straightforward criteria, working, so to speak, as a sensitivity analysis. To reach in, out of sample conclusions and be statistically consistent, we implemented the ART factorial analysis of variance. The results suggested that environmental status is a significant factor for diversification on a portfolio's performance. Given that, we expanded our research and attempted to answer if stocks belonging to an environmental cluster have more distant behaviour from the other clusters, and last which cluster should be the most suitable "pool", as an investor's first choice for portfolio management. Based on standard rankings of return and risk measures, and the employment of the lower tail non-parametric Mann-Whitney U test, we conclude that low environmental-rated companies in cluster "D" produce higher returns, while "A" and "B" clusters are the risk averted option for an investor. Our findings suggest that companies with a dynamic environmental performance "T" should be an optimal option for asset selection, providing evidence of high risk-adjusted returns. Our results suggest positive spillovers of the environmental leaders following the Paris Agreement in 2016. We expand our research to examine the implication of COVID-19 pandemic and evidence no significant differences in clusters' relative changes. Finally, similar results are shown following the war in Ukraine, with clusters “B”, “D” and “C” showing insignificant negative effects. In terms of future research, it would be fruitful to investigate whether the nexus of environmental and financial performance differs in other stocks listed in European and Asian markets and how this is translated for good and bad environmental status.

Data references

Agliardi, E., Alexopoulos, T., and Karvelas, K. 2021. “Empirical Results "The Environmental Pillar of ESG and Financial Performance: A Portfolio Analysis"”, Mendeley Data, V4, doi: 10.17632/2jsb9z3d69.4

References

Albuquerque, R., Koskinen, Y., Yang, S., and Zhang, C. 2020. Resiliency of environmental and social stocks: An analysis of the exogenous COVID-19 market crash. *The Review of Corporate Finance Studies 9*(3): 593–621. https://doi.org/10.1093/rcfs/cfaa011

Alessandrini, F., and Jondeau, E. 2020. ESG Investing: From Sin Stocks to Smart Beta. *The Journal of Portfolio Management 46*(3): 75-94. doi:10.3905/jpm.2020.46.3.075

Alexopoulos, T. 2018. To trust or not to trust? A comparative study of conventional and clean energy exchange-traded funds. *Energy Economics 72*: 97-107.

Auer, B. R., and Schuhmacher, F. 2016. Do socially (ir)responsible investments pay? New evidence from international ESG data. *The Quarterly Review of Economics and Finance 59*: 51-62. doi:10.1016/j.qref.2015.07.002

Bae, K.-H., El Ghoul, S., Gong, Z. (J., and Guedhami, O. 2021. Does CSR matter in times of crisis? evidence from the COVID-19 pandemic. *Journal of Corporate Finance 67*. https://doi.org/10.1016/j.jcorpfin.2020.101876

Blackrock and Ceres. 2015. 21st Century Engagement, Investor Strategies for Incorporating ESG Considerations into Corporate Interactions. Ceres archive. https://www.ceres.org/resources/reports/21st-century-engagement-investor-strategies-incorporating-esg-considerations. Accessed 10 July 2021.

Blitz, D., and Fabozzi, F. J. 2017. Sin Stocks Revisited: Resolving the Sin Stock Anomaly. *The Journal of Portfolio Management* https://doi.org/10.3905/jpm.2017.2017.1.070

Bolton, P., and Kacperczyk, M. 2021. Do investors care about carbon risk? *Journal of Financial Economics* 142(2): 517-549. doi:10.1016/j.jfineco.2021.05.008

Bruno, G., Esakia, M., and Goltz, F. 2021. “Honey, I shrunk the ESG alpha”: Risk-adjusting ESG portfolio returns. *The Journal of Investing*. https://doi.org/10.3905/joi.2021.1.215

Carhart, M. M. 1997. On Persistence in Mutual Fund Performance. *The Journal of Finance 52*(1): 57-82. doi:10.1111/j.1540-6261.1997.tb03808.x

Chen, L.-H., and Huang, L. 2009. Portfolio optimization of equity mutual funds with fuzzy return rates and risks. *Expert Systems with Applications*, 36(2), 3720-3727. https://doi.org/10.1016/j.eswa.2008.02.027

Clark, G. L., Feiner, A., and Viehs, M. 2015. From the Stockholder to the Stakeholder: How Sustainability Can Drive Financial Outperformance. *SSRN Electronic Journal*. doi:10.2139/ssrn.2508281

Clark-Carter, D. (2010). Measures of Central Tendency. International Encyclopedia of Education, 264–266. doi:10.1016/b978-0-08-044894-7.01343-9

Cohen, M. A., Fenn, S., and Konar, S. 1997. *Environmental and financial performance: are they related?*. Washington, DC: Investor Responsibility Research Center, Environmental Information Service.

Conover, W. J., and Iman, R. L. 1981. Rank Transformations as a Bridge between Parametric and Nonparametric Statistics. *The American Statistician* 35(3): 124-129. doi:10.2307/2683975

Cooke R.M., Nieboer D., and Misiewicz J. 2011. Fat-Tailed Distributions: Data, Diagnostics and Dependence. *Resources for the future*. https://www.rff.org/publications/working-papers/fat-tailed-distributions-data-diagnostics-and-dependence/ Accessed 14 November 2022.

Cornell, B. 2020. ESG preferences, risk and return. *European Financial Management*, 27(1), 12-19. https://doi.org/10.1111/eufm.12295

D'Agostino, R., and Pearson, E. S. 1973. Tests for departure from normality. Empirical results for the distributions of b2 and b1. *Biometrika* 60(3): 613-622. doi:10.1093/biomet/60.3.613

Delis, M. D., De Greiff, K., and Ongena, S. R. 2019. Being stranded with fossil fuel reserves? climate policy risk and the pricing of Bank Loans. *SSRN Electronic Journal*. doi:10.2139/ssrn.3451335

Demers, E., Hendrikse, J., Joos, P., and Lev, B. 2021. ESG did not immunize stocks during the COVID-19 crisis, but investments in intangible assets did. *Journal of Business Finance & Accounting 48*(3-4): 433–462. https://doi.org/10.1111/jbfa.12523

Derwall, J., Guenster, N., Bauer, R., and Koedijk, K. 2005. The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal* 61(2): 51-63. doi:10.2469/faj.v61.n2.2716

Diltz, J. D. 1995. Does Social Screening Affect Portfolio Performance? *The Journal of Investing* 4(1): 64-69. doi:10.3905/joi.4.1.64

Dimson, E., Marsh, P., and Staunton, M. 2020. Exclusionary screening. *The Journal of Impact and ESG Investing* 1(1): 66-75. doi:10.3905/jesg.2020.1.1.066

Dorfleitner, G., Kreuzer, C., and Sparrer, C. 2020. ESG controversies and controversial ESG: About Silent Saints and Small Sinners. *Journal of Asset Management* 21(5): 393-412. doi:10.1057/s41260-020-00178-x

Dunn, J., Fitzgibbons, S., and Pomorski, L. 2018. Assessing risk through environmental, social and governance exposures. *Journal of Investment Management*, 16(1), 4-17.

Fabozzi, F. J., Ng, P. W., and Tunaru, D. E. 2021. The impact of corporate social responsibility on corporate financial performance and credit ratings in Japan. *Journal of Asset Management* 22(2): 79-95. doi:10.1057/s41260-021-00204-6

Fortado, L. 2017. Why activists are cheerleaders for corporate social responsibility. *Financial Times* December 26.

Giese, G., Lee, L.-E., Melas, D., Nagy, Z., and Nishikawa, L. 2019. Foundations of ESG investing: How ESG affects equity valuation, risk, and performance. *The Journal of Portfolio Management*, 45(5), 69–83. https://doi.org/10.3905/jpm.2019.45.5.069

Giese, G., Nagy, Z., and Lee, L.-E. 2021. Deconstructing ESG ratings performance: Risk and return for E, S, and G by Time Horizon, sector, and weighting. *The Journal of Portfolio Management*, 47(3), 94-111. https://doi.org/10.3905/jpm.2020.1.198

Görgen, M., Nerlinger, M., and Wilkens, M. 2020. Carbon risk. *SSRN Electronic Journal*. doi:10.2139/ssrn.2930897

Hamilton, S., Jo, H., and Statman, M. 1993. Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds. *Financial Analysts Journal* 49(6): 62-66. doi:10.2469/faj.v49.n6.62

Hang, M., Geyer-Klingeberg, J., and Rathgeber, A. W. 2018. It is merely a matter of time: A meta-analysis of the causality between environmental performance and financial performance. *Business Strategy and the Environment* 28(2): 257-273. doi:10.1002/bse.2215

Hoepner, A. G., Oikonomou, I., Sautner, Z., Starks, L. T., and Zhou, X. 2018. ESG Shareholder Engagement and Downside Risk. *SSRN Electronic Journal*. doi:10.2139/ssrn.2874252

Hong, H., and Kacperczyk, M. 2009. The price of sin: The effects of social norms on markets. *Journal of Financial Economics* 93(1): 15-36. doi:10.1016/j.jfineco.2008.09.001

Huber, B., and Comstock, M. 2017. ESG Reports and Ratings: What They Are, Why They Matter. https://corpgov.law.harvard.edu/2017/07/27/esg-reports-and-ratings-what-they-are-why-they-matter/. Accessed 25 July 2021.

Kempf, A., and Osthoff, P. 2007. The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management* 13(5): 908-922. doi:10.1111/j.1468-036x.2007.00402.x

Kojadinovic, I., Segers, J., & Yan, J. 2011. Large-sample tests of extreme-value dependence for multivariate copulas. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 39(4): 703–720. http://www.jstor.org/stable/41304494

Larcker, D. F., Tayan, B., and Watts, E. M. 2021. Seven Myths of ESG. *SSRN Electronic Journal*.

Levene, H. 1961. Robust tests for equality of variances. *Contributions to probability and statistics. Essays in honor of Harold Hotelling* 279-292.

Liang, H., and Renneboog, L. 2021. Corporate Social Responsibility and Sustainable Finance. *Oxford Research Encyclopedia of Economics and Finance*. doi:10.1093/acrefore/9780190625979.013.592

Makni, R., Francoeur, C., and Bellavance, F. 2008. Causality between Corporate Social Performance and Financial Performance: Evidence from Canadian Firms. *Journal of Business Ethics* 89(3): 409-422. doi:10.1007/s10551-008-0007-7

Martiradonna, M., Romagnoli, S., and Santini, A. 2022. The Beneficial Role of Green Bonds as a New Strategic Asset Class Dynamics: Dependencies, Allocation and Diversification Before and During the Pandemic Era. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4049419

McNeil, A. J., & Frey, R. 2000. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance* 7(3-4), 271–300. https://doi.org/10.1016/s0927-5398(00)00012-8

Nagy, Z., Kassam, A., and Lee, L. 2016. Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies. *The Journal of Investing* 25(2): 113-124. doi:10.3905/joi.2016.25.2.113

Nelling, E., and Webb, E. 2009. Corporate social responsibility and financial performance: The “virtuous circle” revisited. *Review of Quantitative Finance and Accounting* 32(2): 197-209. doi:10.1007/s11156-008-0090-y

Orlitzky, M., Schmidt, F. L., and Rynes, S. L. 2003. Corporate Social and Financial Performance: A Meta-Analysis. *Organization Studies* 24(3): 403-441. doi:10.1177/0170840603024003910

Pastor, Ľ., Stambaugh, R. F., and Taylor, L. A. 2021. Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011

Pedersen, L.H., Fitzgibbons S., Pomorski L. 2020. Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics* (In Press)

Preston, L. E., and O'bannon, D. P. 1997. The Corporate Social-Financial Performance Relationship. *Business & Society* 36(4): 419-429. doi:10.1177/000765039703600406

Proberts J., and Boshnakov G., ARMA-GARCH modelling and white noise tests https://cran.r-project.org/web/packages/sarima/vignettes/garch_tests_example.pdf Accessed 14 November 2022.

Renneboog, L., Ter Horst, J., and Zhang, C. 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance* 32(9): 1723-1742. doi:10.1016/j.jbankfin.2007.12.039

Salter, K. C., and Fawcett, R. F. 1993. The art test of interaction: A robust and powerful rank test of interaction in factorial models. *Communications in Statistics - Simulation and Computation* 22(1): 137-153. doi:10.1080/03610919308813085

Sawilowsky, S. S. 1990. Nonparametric Tests of Interaction in Experimental Design. *Review of Educational Research* 60(1): 91-126. doi:10.3102/00346543060001091

Seltzer, L., Starks, L. T., and Zhu, Q. 2021. Climate regulatory risks and corporate bonds. *SSRN Electronic Journal*. doi:10.2139/ssrn.3563271

Shapiro, S. S., and Wilk, M. B. 1965. An analysis of variance test for normality, complete samples. *Biometrika* 52(3-4): 591-611. doi:10.1093/biomet/52.3-4.591

Sklar, A. 1959. Fonctions de Répartition à n Dimensions et Leurs Marges. *Publications de l’Institut Statistique de l’Université de Paris*. 8: 229-231.

Thompson Reuters Eikon, Refinitiv Eikon Online, 2020. http://eikon.thomsonreuters.com/index.html. Accessed 02 January 2021.

Tola, V., Lillo, F., Gallegati, M., and Mantegna, R. N. 2008. Cluster Analysis for portfolio optimization. *Journal of Economic Dynamics and Control*, 32(1), 235-258. https://doi.org/10.1016/j.jedc.2007.01.034

Trinks, P. J., Scholtens, L., Mulder, M., and Dam, L. 2017. *Divesting Fossil Fuels: The Implications for Investment Portfolios*. Groningen: University of Groningen, SOM research school.

United Nations, 2019. *Financing for Sustainable Development Report 2019.* Inter-agency Task Force on Financing for Development, New York

United Nations, Principles for Responsible Investment: PRI Quarterly signatory update Q2 2021, https://www.unpri.org/signatory-resources/quarterly-signatory-update/4609.article. Accessed 17 January 2021.

US SIF, 2022. *Report on Sustainable and Responsible Investing Trends 2022.* Forum for Sustainable and Responsible Investment, Washington DC

Verheyden, T., Eccles, R.G. and Feiner, A. 2016. ESG for All? The Impact of ESG Screening on Return, Risk, and Diversification. *Journal of Applied Corporate Finance*, 28(2) 47-55. https://doi.org/10.1111/jacf.12174

Waddock, S. A., and Graves, S. B. 1997. The Corporate Social Performance-Financial Performance Link. *Strategic Management Journal* 18(4): 303-319. doi:10.1002/(sici)1097-0266(199704)18:43.0.co;2-g

Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J. 2011. The aligned rank transform for nonparametric factorial analyses using only anova procedures. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*. doi:10.1145/1978942.1978963

Yamashita, M., Sen, S., and Roberts, M. C. 1999. The rewards for environmental conscientiousness in the US capital markets. *Journal of Financial and Strategic Decisions* 12(1): 73-82

Figures

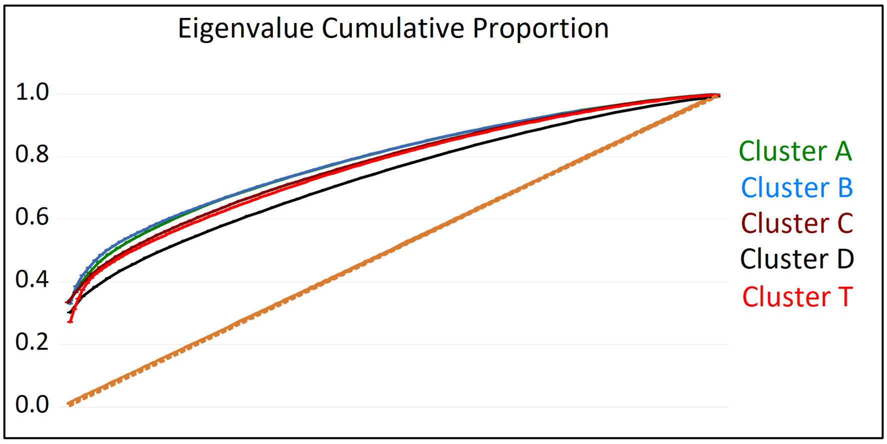


Figure 1. Eigenvalue Cumulative proportion for all the clusters

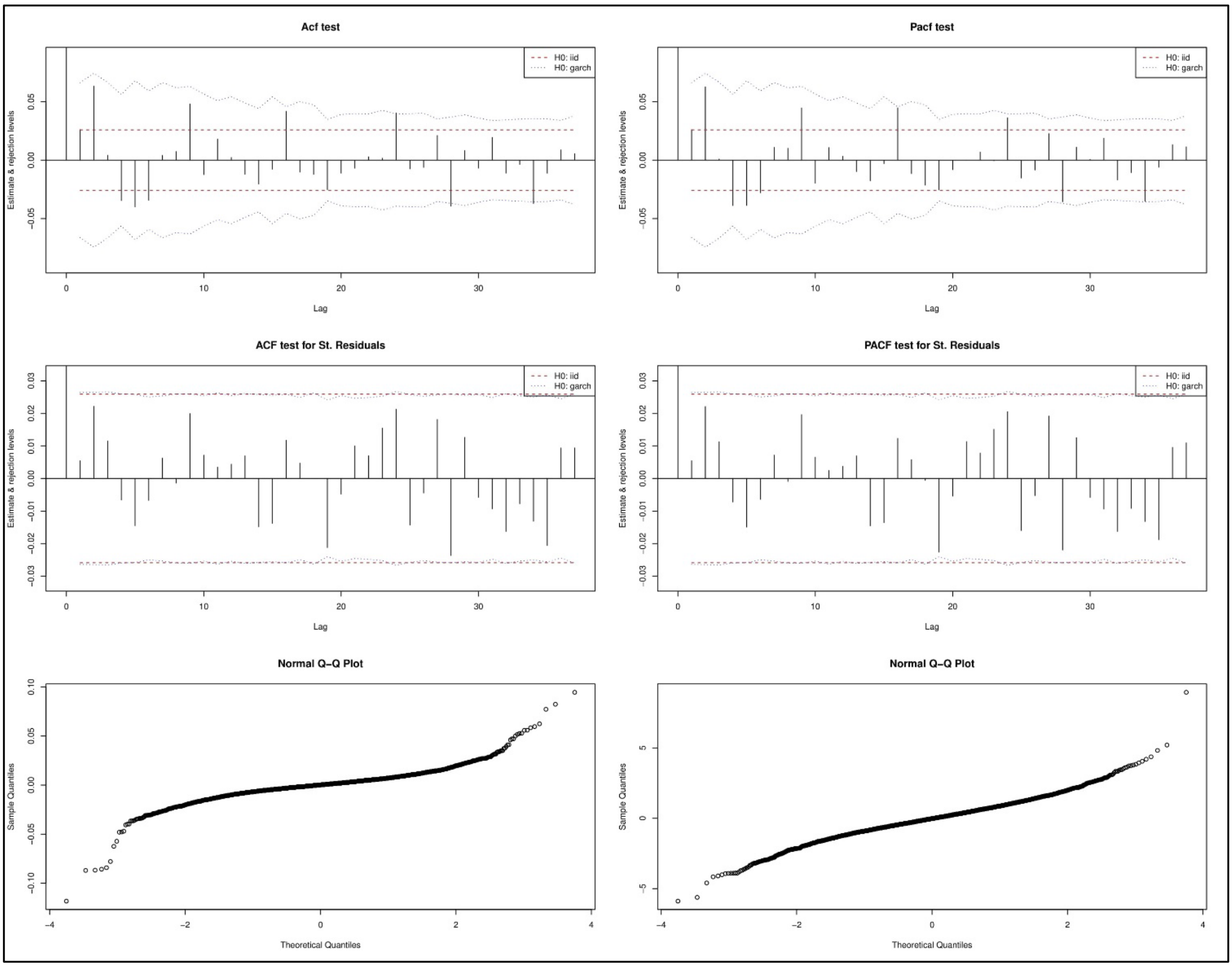


Figure 2. Partial- and autocorrelations for original return series and standardized residuals after GARCH filtering

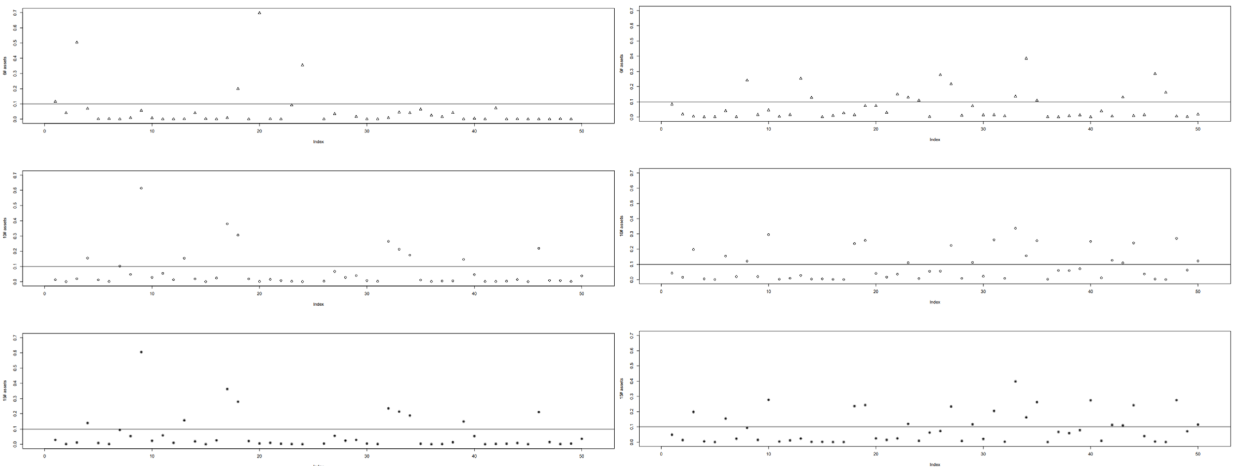


Figure 3. P-values of tests of multivariate extreme values copulas for negative returns of portfolios of Cluster A (left) and Cluster T (Right)

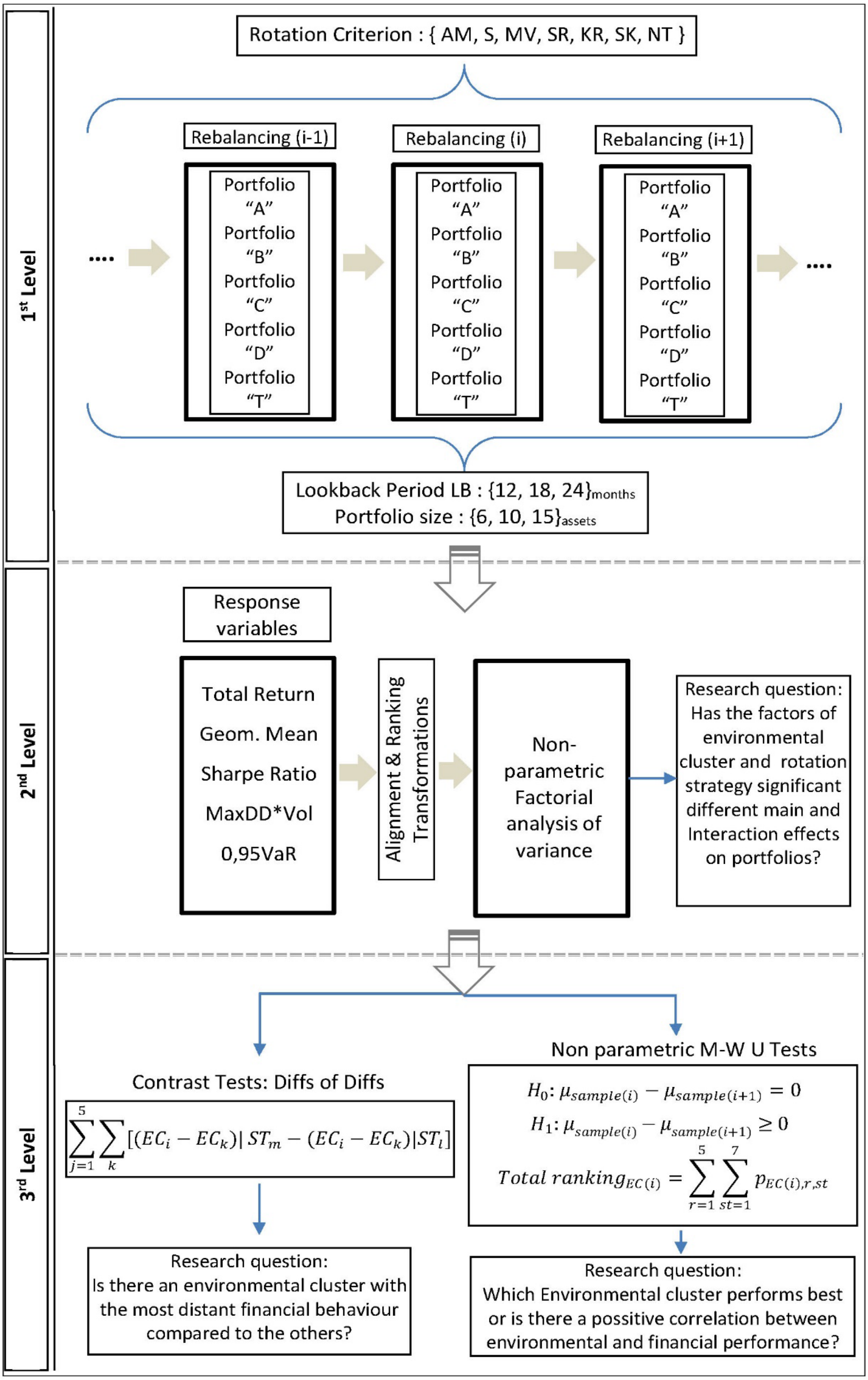
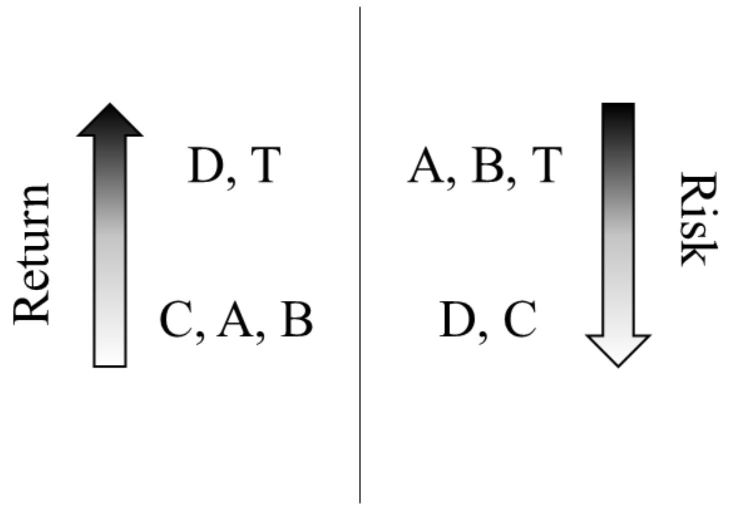
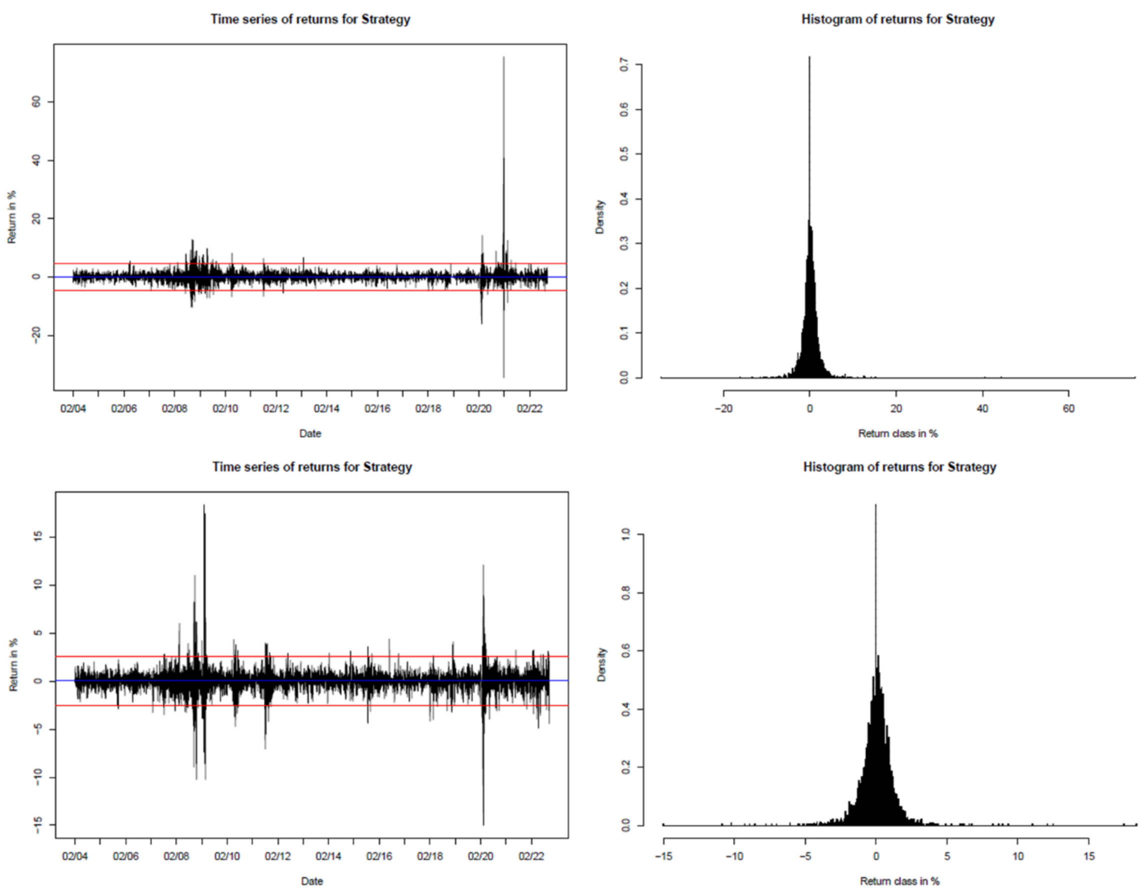


Figure 4. The overall framework of the Empirical methodology



* Dark shading represents the best performing clusters

Figure 5. A qualitative representation out of the rebalancing results for all the portfolios.



Up right and left: A Top-10, based on returns, environmental cluster "D" portfolio
Bottom right and left: A Top-10, based on risk, environmental cluster "A" portfolio

Figure 6. Time series and histogram of returns for the portfolios of environmental clusters "A" and "C"

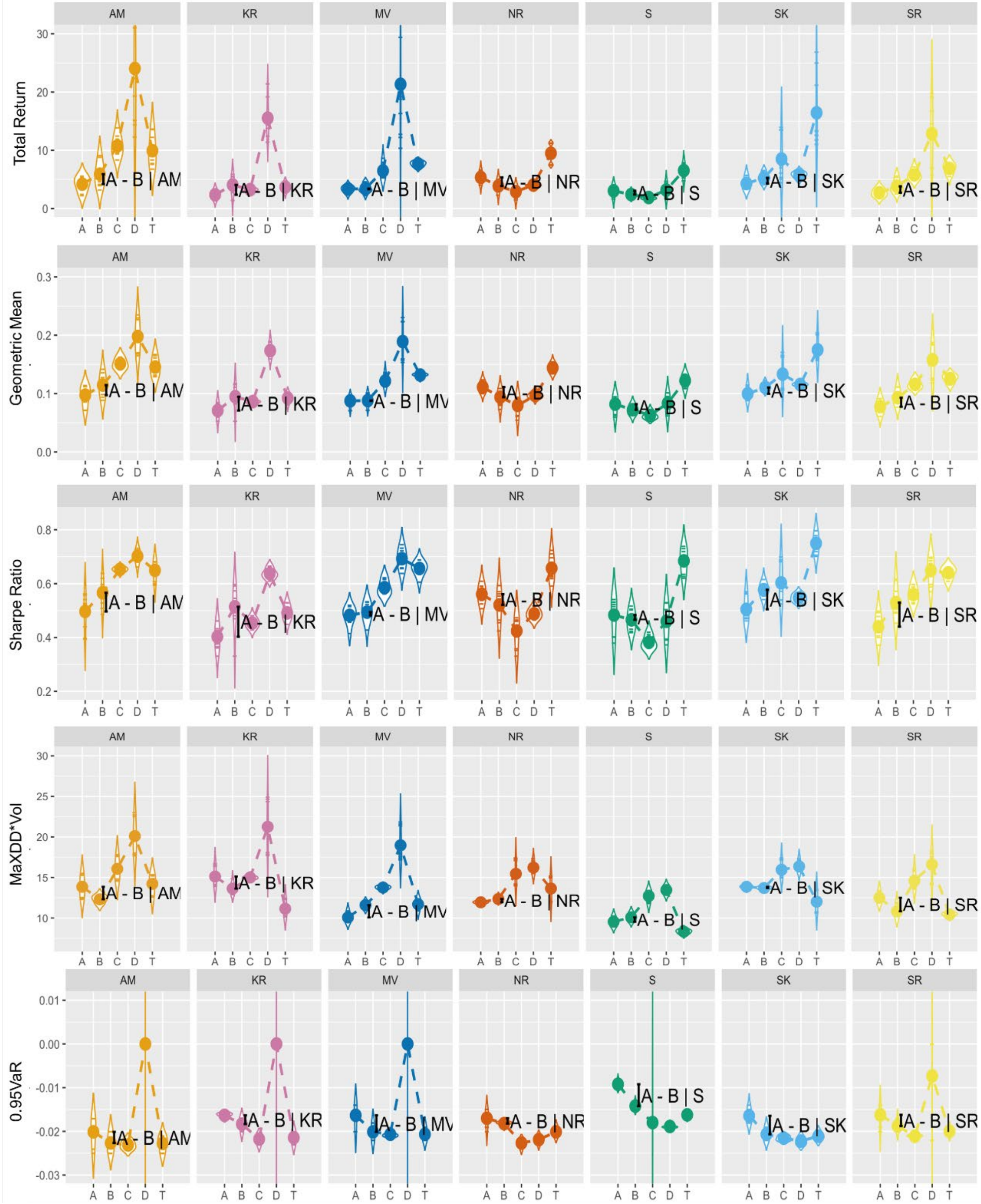


Figure 7. Visualization of the "diffs of diffs" contrast test results

Tables

Table 1. Companies' classification before and after the logical statement

Company	A	B	C	D	Classification based on the highest frequency of years	Geom.Mean of Δ(E) ≥ Q_3	Final Classification
Abiomed	10	0	4	3	A	False	A
Estee Lauder	4	6	4	5	B	True	T
Genworth Financial	0	9	4	4	B	False	B
Mckesson	1	6	10	0	C	False	C
Crown Castle	1	3	0	3	NA	True	T
Cit Group	0	2	3	13	D	False	D

*Values stand for the cumulative years in the specific ESG cluster

Table 2. First eigenvalue proportion and number of eigenvalues per Cumulative proportions

Cluster	1st Eigenvalue Proportion	Cumulative Proportion 50%	75%
A	0.34	#6	#26
B	0.33	#6	#34
C	0.33	#9	#33
D	0.30	#8	#25
T	0.27	#16	#63

Table 3. Estimate of the Obesity Index

	Positive tail		Negative tail	
Cluster	Obesity Index	Confidence Interval	Obesity Index	Confidence Interval
A	0.739	(0.737, 0.741)	0.748	(0.746, 0.750)
B	0.743	(0.741, 0.744)	0.752	(0.751, 0.754)
C	0.754	(0.752, 0.756)	0.754	(0.753, 0.756)
D	0.766	(0.764, 0.768)	0.778	(0.776, 0.780)
T	0.742	(0.741, 0.743)	0.736	(0.734, 0.736)

* The respective figures with the bootstrapped samples are available in the data file

Table 4. Percentages (%) of non-rejection of the null hypothesis
H_0: $Copula \in$ extreme value family

	Positive tail			Negative tail		
Cluster	#6 assets	#10 assets	#15 assets	#6 assets	#10 assets	#15 assets
A	12	22	22	16	24	22
B	16	20	18	18	16	12
C	10	28	30	18	20	20
D	16	48	48	12	28	28
T	34	46	48	28	38	36

Table 5. Top-10 & Bottom-10 portfolios

	Arithmetic Mean	Geometric Mean	Total Return	Sharpe Ratio	Volatility	MaxDD	0.95VaR
Top-10	0.275 (D)	0.235 (D)	39.866 (D)	0.797 (T)	0.186 (B)	-0.409 (T)	-0.008 (A)
	0.270 (D)	0.231 (D)	38.794 (D)	0.797 (T)	0.186 (B)	-0.409 (T)	-0.009 (A)
	0.270 (D)	0.230 (D)	36.480 (D)	0.779 (T)	0.188 (B)	-0.409 (T)	-0.009 (A)
	0.264 (D)	0.227 (D)	36.353 (D)	0.758 (T)	0.188 (B)	-0.416 (A)	-0.009 (A)
	0.259 (D)	0.224 (D)	31.047 (D)	0.745 (D)	0.188 (B)	-0.416 (A)	-0.009 (A)
	0.258 (D)	0.223 (D)	29.397 (D)	0.745 (T)	0.190 (B)	-0.416 (A)	-0.009 (A)
	0.236 (D)	0.205 (T)	26.880 (T)	0.744 (T)	0.190 (B)	-0.426 (T)	-0.010 (A)
	0.236 (D)	0.204 (T)	25.001 (T)	0.739 (T)	0.190 (B)	-0.426 (T)	-0.010 (A)
	0.233 (D)	0.201 (T)	23.092 (D)	0.733 (D)	0.192 (B)	-0.426 (T)	-0.010 (A)
	0.230 (D)	0.198 (D)	22.382 (D)	0.733 (D)	0.194 (T)	-0.430 (A)	-0.013 (B)
Bottom-10	0.079 (B)	0.053 (A)	1.397 (A)	0.332 (C)	0.382 (D)	-0.652 (D)	-0.025 (B)
	0.080 (A)	0.053 (B)	1.397 (B)	0.332 (A)	0.380 (D)	-0.652 (D)	-0.025 (B)
	0.081 (B)	0.055 (C)	1.493 (C)	0.332 (B)	0.378 (D)	-0.652 (D)	-0.025 (B)
	0.082 (A)	0.058 (C)	1.655 (C)	0.353 (C)	0.375 (D)	-0.650 (D)	-0.025 (T)
	0.082 (B)	0.059 (C)	1.664 (C)	0.359 (C)	0.374 (D)	-0.650 (D)	-0.025 (T)
	0.082 (C)	0.059 (C)	1.667 (A)	0.361 (C)	0.372 (D)	-0.650 (D)	-0.025 (T)
	0.082 (B)	0.059 (C)	1.684 (C)	0.362 (A)	0.359 (D)	-0.648 (D)	-0.025 (A)
	0.083 (C)	0.060 (A)	1.761 (D)	0.366 (C)	0.355 (D)	-0.648 (D)	-0.024 (A)
	0.083 (C)	0.060 (C)	1.773 (C)	0.367 (A)	0.351 (D)	-0.648 (D)	-0.024 (A)
	0.083 (A)	0.060 (C)	1.775 (A)	0.367 (C)	0.330 (D)	-0.648 (A)	-0.024 (C)
SPY	0.071	0.053	1.479	0.366	0.193	-0.537	-0.016

* Environmental clusters in parentheses

Table 6. Normality Tests & Equality of variance test of the Response variables

		Response Variable				
		TR	GM	SR	MaXDD*Vol	0.95VaR
Normality*	Shapiro-Wilk Test	Rejection=3 Non-Rejection=2	Rejection=3 Non-Rejection=2	Rejection=2 Non-Rejection=3	Rejection=3 Non-Rejection=2	Rejection=5 Non-Rejection=0
	D'Agostino-Pearson Test	Rejection=3 Non-Rejection=2	Rejection=1 Non-Rejection=4	Rejection=1 Non-Rejection=4	Rejection=1 Non-Rejection=4	Rejection=1 Non-Rejection=4
Equality of Variances	Levene's Test**	Rejection	Rejection	Rejection	Rejection	Rejection

* On the five Environmental Clusters A-T, Rejection of the Null Hypothesis
** All three versions of the test (mean, median, 10% trimmed mean) are used.

Table 7. Results from the Aligned Rank Transform (ART) non-parametric Anova

	F-values*		
Response Variable	Factor **EC**	Factor **ST**	Factor **EC*ST**
TR	137.80(***)	71.26(***)	28.37(***)
GM	209.13(***)	89.58(***)	30.00(***)
SR	159.11(***)	61.40(***)	24.99(***)
MaXDD*Vol	267.98(***)	88.76(***)	16.22(***)
0.95VaR	205.80(***)	78.29(***)	51.81(***)

* Significance in parentheses

Table 8. "Diffs of Diffs" Contrast test results

$$Contrast\ Test = [\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(m)} - [\mu_{EC(i)} - \mu_{EC(k)}]|_{ST(l)}$$

Response Variable	**Environmental Clusters**				
	EC (A)	**EC (B)**	**EC (C)**	**EC (D)**	**EC (T)**
Total Return	30	32	45	54	39
Geometric Mean	33	36	48	51	34
Sharpe Ratio	38	40	43	49	40
MaXDD*Vol	37	34	27	45	43
0.95VaR	19	19	20	60	16
Cumulative Sum	157	161	183	259	172

The values in cells correspond to the number of the rejected Null hypothesis of the contrast test.

Table 9. Clusters Rankings based on the non-parametric MW-U test

$$Ranking_{EC(i)} = \sum_{rv=1}^{5} \sum_{st=1}^{7} p_{EC(i),rv,st}$$

	Response Variable	EC (A)	EC (B)	EC (C)	EC (D)	EC (T)
		Environmental Clusters				
Measures of Return	Total Return	22	20	19	11	11
	Geometric Mean	24	20	19	11	11
Total Ranking_EC(i) on Returns		**46**	**40**	**38**	**22**	**22**
Measures of Risk	MaXDD*Vol	13	13	22	27	10
	0.95VaR	7	13	20	25	17
Total Ranking_EC(i) on Risks		**20**	**26**	**42**	**52**	**27**
Sharpe Ratio		**22**	**17**	**16**	**13**	**10**

Table 10. Environmental clusters responses after the Paris Agreement

	Response Variable	Rankings* 1st	2nd	3rd	4th	5th
Measures of Return	Total Return	A	C	T	D	B
	Geometric Mean	A	C	T	D	B
	Sharpe Ratio	T	A	C	B	D
Measures of Risk	MaXDD*Vol	D	T	C	A	B
	0.95VaR	A	D	B	C	T
	Best Performance	←				

*Based on their sample means
** Clusters included in a grey shaded box have equal population means, meaning non-rejection of the Ho: $\mu_{D_{EC(i)}} - \mu_{D_{EC(j)}} = 0$, or in other words having no significant differences in their responses after the Paris Agreement.

Table 11. Environmental clusters responses after the COVID-19 outbreak

	Response Variable	Rankings* 1st	2nd	3rd	4th	5th
Measures of Return	Total Return	D	A	B	C	T
	Geometric Mean	D	A	B	C	T
	Sharpe Ratio	C	T	A	B	D
Measures of Risk	MaXDD*Vol	D	A	B	C	T
	0.95VaR	A	T	C	B	D
	Best Performance	←				

*Based on their sample means
** Clusters included in a grey shaded box have equal population means, meaning non-rejection of the Ho: $\mu_{D_{EC(i)}} - \mu_{D_{EC(j)}} = 0$, or in other words having no significant differences in their responses after the COVID-19 outbreak.

Table 12. Environmental clusters responses after the war in Ukraine

Response Variable		Rankings*				
		1st	2nd	3rd	4th	5th
Measures of Return	Total Return	C	A	D	B	T
	Geometric Mean	C	A	T	D	B
	Sharpe Ratio	T	B	C	A	D
Measures of Risk	MaXDD*Vol	D	C	A	B	T
	0.95VaR	T	A	C	D	B
	Best Performance	←				

*Based on their sample means
** Clusters included in a grey shaded box have equal population means, meaning non-rejection of the Ho: $\mu_{D_{EC(i)}} - \mu_{D_{EC(j)}} = 0$, or in other words having no significant differences in their responses after the war in Ukraine.