

RESEARCH ARTICLE OPEN ACCESS

Sustainable Finance Disclosure Versus Performance: A Clustering Approach

Veronica Distefano^{1,2} | Vincenzo Gentile¹ | Paolo Antonio Cucurachi¹ | Sandra De Iaco^{1,3,4} 

¹Department of Economic Sciences, University of Salento, Complesso Ecotekne, Lecce, Italy | ²European Centre for Living Technology (ECLT), Ca' Foscari University of Venice, Venice, Italy | ³National Centre for HPC, Big Data and Quantum Computing, Bologna, Italy | ⁴National Biodiversity Future Centre, Palermo, Italy

Correspondence: Sandra De Iaco (sandra.deiaco@unisalento.it)

Received: 26 January 2025 | **Revised:** 20 April 2025 | **Accepted:** 8 May 2025

Funding: This study was supported by the National Recovery and Resilience Plan (NRRP), the Mission 4 Component 2 Investment 1.4—Call for Tender no. 3138 of 16 December 2021, rectified by decree no. 3175 of 18 December 2021 of the Italian Ministry of University and Research funded by the European Union—Next Generation EU, award number: project code CN-00000033, concession decree no. 1034 of 17 June 2022 adopted by the Italian Ministry of University and Research: CUPF87G22000290001, project title “National Biodiversity Future Center—NBFC”.

Keywords: cluster analysis | dimension reduction | ESG | SFDR | structural topic modeling | sustainable finance

ABSTRACT

The Sustainable Finance Disclosure Regulation (SFDR) imposes on investment companies the publication of sustainability related profiles of the portfolio creation process. This obligation seeks to enhance the information provided to customers on financial products with regard to the Environmental, Social and Governance (ESG) criteria. In order to understand to which extent this information is relevant for investment decisions, the reliability of the classification among funds with a different ESG focus (articles 6, 8, and 9) is studied from a financial perspective and verified through a hierarchical cluster analysis on principal components of the performance measures of various funds distributed in Europe.

1 | Introduction

In recent years, the increasing attention to environmental, social, and governance (ESG) issues has led to a rapid spread of financial instruments as well as investment decision-making processes linked to those activities or projects that take care of the environment and are beneficial to society.

Accordingly, specialized rating agencies have grown and have developed their own ESG scores algorithms through which issuers and financial investment instruments are classified and labeled at a certain degree of proximity to an ESG oriented approach. The purpose of rating agencies' ESG scores is to assist investors in more accurately assessing the non-financial performance of investment funds (Hartzmark and Sussman 2019). In this context, the Sustainable Finance Disclosure Regulation (SFDR) was designed in March 2021 to enhance transparency

in the market for sustainable investment products. SFDR represents a critical regulatory development within the European Union's broader agenda to promote sustainable economic growth and environmental responsibility. As part of the European Green Deal and the EU Sustainable Finance Action Plan, the SFDR seeks to enhance the transparency, comparability, and accountability of sustainability-related disclosures in the financial sector. The SFDR applies to a wide range of financial market participants and financial advisers, including asset managers, pension funds, insurance companies, and investment firms operating within the EU. Its primary purpose is to address the information asymmetry between investors and financial institutions regarding environmental, social, and governance (ESG) considerations in investment decision-making processes.

The regulation mandates standardized disclosures at both the entity and product levels. At the entity level, financial firms are

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Business Strategy and the Environment* published by ERP Environment and John Wiley & Sons Ltd.

required to disclose how they integrate sustainability risks into their investment or advisory processes. At the product level, the SFDR has harmonized rules for funds classification by indicating modalities and explicitly providing information necessary for a clear taxonomy based on the declaration regarding the integration of ESG factors at the asset manager level. In particular, Article 6 Funds are those whose objectives are not aligned with ESG's themes, Article 8 Funds's interest is directed toward ESG themes but do not represent the main criteria of the fund, while Article 9 Fund's main objective is to comply with ESG issues.

By imposing clear disclosure obligations, the SFDR aims to prevent greenwashing and ensure that financial products marketed as “sustainable” are held to objective and consistent criteria. This increased transparency is intended to empower investors to make more informed decisions aligned with their sustainability preferences.

Moreover, the SFDR operates alongside other key regulatory initiatives such as the EU Taxonomy Regulation and the Corporate Sustainability Reporting Directive (CSRD), forming a cohesive framework to reorient capital flows toward sustainable activities. Collectively, these measures seek to integrate sustainability more deeply into the financial system and support the EU's commitment to achieving climate neutrality by 2050.

Recently, some authors have analyzed the impact of SFDR over different perspectives finding that investor attention to SFDR has strong predictive power over equity market prices especially during bearish markets (Birindelli et al. 2023); moreover, a better ESG label leads to larger fund net inflows (Becker et al. 2022; Scherer and Hasaj 2023; Duncombe et al. 2023), and investment funds with sustainability objectives have higher financial incentives (Cremasco and Boni 2022; Wang 2024). Although there are various contributions which examined the determinants of financial performance, including comparisons between socially responsible (SR) and non-SR funds (Choi and Lee 2024; Hart and Zingales 2017), only few works have analyzed the alignment between investors' interest to performance and investment sustainability orientation in the light of the new SFDR classification, as in Cosma et al. (2023) and Scheitza and Busch (2024) that shows the outperformance in the very short run of Article 9 funds over conventional funds during the two recent crises (Covid and Russia-Ukraine war) even if with a lack of performance persistence over the long run. Additionally, the study highlights the underperformance of light green funds compared to dark green ones.

To the best of our knowledge, the present study is the first to address the SFDR classification by reversing the traditional approach. Traditionally, research in this field of study starts by examining existing classifications and then analyzes how these predefined categories relate to differences in performance. In contrast, this paper takes a different and more exploratory approach. Rather than relying on existing labels from the outset, it first identifies groups of financial products that naturally cluster together based on their performance characteristics. These newly formed clusters are then compared to the regulatory classifications to assess how well the current regulatory framework

reflects actual performance patterns. This shift in methodology offers a meaningful contribution to the existing body of knowledge. It can also be used to assess the reliability and accuracy of current regulatory classifications in meeting their intended goals and introduces a systematic way to evaluate both the original regulatory groupings and the newly derived clusters based on a desired impact or characteristics. By applying statistical measures to both sets of classifications, the study provides a clearer understanding of whether the regulatory categories are truly aligned with observable achievement of a desired target. The purpose of this paper is to contribute to the wide and still open debate on the alignment of investors' performance objectives and investment ESG orientation (Gantchev et al. 2024; Buchetti et al. 2025); in other terms, it aims at providing valuable information for investors' approach to ESG investments, asset managers' strategies correct implementation, and policymakers' evaluations about the effectiveness of the new regulation. A clustering analysis has been developed to discover inherent cluster structure in data, in order to evaluate the reliability of the classification among the prior (regulatory-based) and posterior (performance-based) groups. The choice of this approach is justified since it allows for a deeper understanding of the underlying structure of the data. The novelty of this paper concerns the use of unsupervised methods in order to explain that there is no link between the sustainable fund orientation and their financial performance; namely, there is no evidence of a possible trade-off between the two domains. Therefore, in this context, an innovative approach has been proposed taking into account the use of a classification based on asset managers declaration instead of external rating agencies. This is in contrast with the ESMA (European Securities and Markets Authority) mainstream position assuming that ESG funds outperform non-ESG funds (ESMA 2022a, 2022b). Note also that the heterogeneity of sustainable funds is well-documented in existing literature, as highlighted by key references such as Sandberg et al. (2009) and Townsend (2020). To control for potential biases introduced by differing investment styles, which could influence both performance outcomes and classification results, a homogeneous sample was constructed, since only European equity large cap blend style funds distributed in Europe have been used (Cortez et al. 2009, 2012; Fatemi et al. 2024). The analysis exclusively considers mutual funds characterized by similar portfolio structures. In particular, the sample was restricted to large-cap equity funds with no pronounced value or growth bias (within a $\pm 10\%$ range) and a sectoral composition consistent with their respective benchmarks. The only distinction remaining in our sample is between active and passive funds that is taken into account by using appropriate risk measures and risk adjusted performance measure.

Cluster methodology together with an unsupervised dimensionality reduction technique has been applied to examine the return, risk, and efficiency of funds subject to the SFDR classification using daily data, gathered from Refinitiv database, over the course of 2 years since the adoption of the SFDR. This is an innovative approach for testing whether funds with a certain ESG classification exhibit common patterns of behavior in terms of return, risk, and efficiency; in particular, this approach deductively tests whether the data analysis provides homogeneous groups of funds characterized by the same ESG classification. The paper is organized as follows. After

presenting the research hypothesis focused on studying the relationship between ESG orientation and funds performance (Section 2), a brief description of the methodology is presented (Section 3). This paper proposes an innovative two-stage analysis to evaluate the validity of a specific hypothesis focused on verifying through cluster analysis on principal components (PCs), whether the prior SFDR classification provides useful information concerning performance expectations. Then, the dataset is described and the variables extraction method together with the hierarchical clustering algorithm (HCA) is provided (Section 4). At the end, a comparison between the prior classification and classification based on clustering model is discussed (Section 5).

2 | Theoretical Analysis and Research Hypothesis

In this section, firstly, the more general literature on the performance of sustainable versus conventional funds has been analyzed, then several contributions focusing on SFDR, which have emerged in the recent period, have been discussed.

2.1 | Performance of Sustainable Versus Conventional Funds

The financial performance of ESG funds has stimulated a lively debate with respect to its connection with ESG issues that has not yet reached a consensus and has developed three different theories that can be summarized in negative, positive, or absence of relationship. In the literature, there is a substantial body of work examining the determinants of financial performance, including comparisons between socially responsible (SR) and non-SR funds, with notable studies (Brammer et al. 2005; Leite and Cortez 2014). However, all these papers were not focused on SFDR classification that started in 2021. SFDR classification is a self-statement of the investment companies, while many studies on SRI funds were based on SRI ratings provided by third parties.

In general, many studies show no statistically significant differences in performance between ethical and conventional funds. This finding has been found in several markets, including the United States, Europe, and Australia. Some specific studies that support this neutrality in performance include Bauer et al. (2007) for Germany, the United Kingdom, and the United States over the period 1990–2001, after controlling for investment style, and Kreander et al. (2005) for European funds. Using a matched-pairs approach, Schröder (2004) and Bauer et al. (2007) also concluded that socially responsible funds do not underperform their unfiltered counterparts. Cortez et al. (2009) found that European socially responsible funds generally exhibit neutral performance relative to conventional and socially responsible benchmarks. Gregory (1997), using a size-adjusted benchmark, found no significant differences between the returns of ethical and nonethical funds in the United Kingdom. Other studies have also highlighted that the difference in performance between conventional funds and socially responsible funds is irrelevant or not significant (Goldreyer and Diltz 1999; Hamilton et al. 1993; Shank et al. 2005; Statman 2000; Rathner 2013).

However, there is also evidence of underperformance of ethical (or SRI) funds in some studies and contexts.

Works that support the theory of a negative relationship between ESG orientation and performance find their foundation on the argument that ESG funds are portfolios less efficient since the reduction of the universe of possible investments implies a lower diversification capable of excluding companies characterized by returns above the average (Hamilton et al. 1993; Adler and Kritzman 2008). Similarly, other authors have demonstrated that ESG funds perform worse than conventional ones because of the higher screening activity that implies higher monitoring or agency costs (Bauer et al. 2007; Zeidan 2022; Chegut et al. 2011; Fabozzi et al. 2008; Statman and Glushkov 2008; Capelle-Blancard and Monjon 2014; Leite and Cortez 2014; Yue et al. 2020).

Renneboog et al. (2008) examined SRI funds in 17 countries across Europe, North America, and Asia, finding that investors pay a “price for ethics”, with SRI funds underperforming their domestic benchmarks. Likewise, Jones et al. (2008) found an underperformance of SRI funds relative to the Australian market. Renneboog et al. (2011) found that SRI funds with “Sin/Ethical” or environmental filters significantly underperform conventional funds globally. Ibikunle and Steffen (2017) found an underperformance of European green funds relative to conventional funds over the overall period analyzed. A significant study in this regard is that of Renneboog et al. (2008) who found that fund returns decrease with the intensity of screening on social and corporate governance criteria; the specific type of responsible investing criteria used also appears to influence financial performance. A crucial point highlighted is the heterogeneity of Socially Responsible Investing (SRI) funds, which apply different levels of intensity in their screening criteria. This variation in intensity is critical to understanding its relationship to financial performance. In addition, Renneboog et al. (2011) examined how screening intensity affects the relationship between money flows and performance. They found that more “Sin/Ethical” screening significantly reduces sensitivity to past negative returns compared to conventional funds. More environmental screening increases the sensitivity of money flows to positive performance, while more social screening reduces the sensitivity of flows to both positive and negative performance.

Barnett and Salomon (2011) hypothesized and found a curvilinear (U-shaped) relationship between social screening intensity and financial performance. Authors found that community relations screening was associated with higher financial performance, while environmental and employee relations screening was associated with lower performance. This study suggests that the choice is not simply between being an SRI fund or not, but how socially responsible to be, managers should carefully consider the effects of their screening strategies on performance.

In contrast, some studies have reported an outperformance of ethical funds. Luther et al. (1992) had found weak evidence of outperformance for UK ethical trusts. Mallin et al. (2006) concluded that ethical funds tended to outperform matched nonethical funds, although they did not directly control for bias toward small businesses in ethical portfolios. Varma and Nofsinger (2014) found that SRI funds outperformed their

conventional counterparts in periods of crisis and underperformed in other periods.

The idea of a positive relationship between financial performance and ESG funds orientation is supported by the possibility of obtaining higher returns by ESG investing in comparison to conventional investing since funds using SRI criteria might actually select companies with more sustainable profitability and better long-term prospects (Gil-Bazo et al. 2010; Becchetti et al. 2014; Braeken and Assen 2016; Neves et al. 2023). In this perspective, the commitment to social and environmental issues implies that screening techniques provide valuable information to fund managers about firms cost of debt and default risk (Atif and Ali 2021). Further, Helliari et al. (2022) proved that SRI funds are more diversified than conventional funds, whereas Klinkowska and Zhao (2023) focused on the association between SRI fund flows and their performance and showed that SRI funds with greater performance attract more investments. The differences in performance persistence between SRI and non-SRI funds were investigated by Lean et al. (2015) who found evidence of the SRI funds' performance persistence.

2.2 | Theoretical Implication of SFDR Classification System

The SFDR classification system has significant theoretical implications across regulatory design, investor behavior, and the understanding of sustainability in finance. By introducing a three-tiered classification system for financial products, Articles 6, 8, and 9 aim to increase transparency and mitigate greenwashing by providing a framework for disclosing sustainability-related information. However, the SFDR is neutral regarding specific product design, not mandating the exact nature of sustainable investment objectives for Article 9 funds. This has led to discussions about whether it effectively functions as a labeling regime. Scheitza and Busch (2024) highlight that while Article 9 funds are often termed “dark green” or impact-related, a substantial portion follow general ESG strategies rather than focusing on transformative impact. A key theoretical implication explored is the concept of “category fuzziness”, as investigated by Cremasco and Boni (2022). Applying organizational category theory, they found that despite the SFDR's attempt to create distinct categories based on sustainability goals, the financial incentives of fund managers in Articles 6 and 9 funds appear similar, influenced by both financial and sustainability performance. This suggests that the regulatory distinctions might not always translate into distinct internal behaviors and challenges the SFDR's effectiveness in creating clear, self-discriminating categories. From the perspective of investor behavior, the SFDR intends to provide more transparent and comparable information, enabling investors to make more informed decisions aligned with their sustainability preferences. The theory is that increased transparency should incentivize fund managers to sustainably invest more and allow investors to allocate capital accordingly. Research by Becker et al. (2022) indicates that funds subject to the SFDR saw an increase in their sustainability ratings post-announcement, and funds with better ESG labels significantly experienced higher net fund flows. This suggests that investors do value ESG alignment and react to the introduced

labels. However, Ferriani (2024), comparing SFDR labels with Morningstar sustainability ratings, found that investors generally rely more on industry-developed sustainability scores for portfolio decisions, with the notable exception of Article 9 funds, which did attract positive inflows. This implies that investors might perceive the Article 9 label as a more credible signal of a primary focus on sustainability. It has also been theorized that the increased investment in sustainability-labeled funds could lead to ESG asset overvaluation. Furthermore, investors might focus on the label itself rather than the actual degree of ESG integration. The study by Abouarab et al. (2025) adds to this by showing that Article 9 funds exhibited a reduction in their greenwashing index after the SFDR's implementation, unlike Article 8 funds, suggesting the regulation might be effective in guiding genuine decarbonization efforts for funds with explicit sustainability objectives.

The SFDR attempts to codify and standardize the understanding of sustainability within financial markets by creating distinct categories based on the level of sustainability integration. Article 9's requirement for a “sustainable investment objective” theoretically distinguishes funds with a primary sustainability focus from those that merely consider ESG factors (Article 8) or do not (Article 6). However, the findings of Scheitza and Busch (2024), indicating that not all Article 9 funds align with a strong impact-generating purpose, raise questions about the practical application of these distinctions and the interpretation of a “sustainable investment objective”. Similarly, Abouarab et al. (2025)'s finding that only Article 9 funds showed a significant reduction in greenwashing suggests that the SFDR's impact on genuinely embedding sustainability practices might be more pronounced for funds with the highest level of declared commitment. Cruciani and Santagiustina (2023)'s analysis of Key Investor Information Documents (KIIDs) reveals that sustainability information is discussed across multiple sections, consistently with the SFDR classification. This has implications for how sustainability is framed and communicated to investors, suggesting that it is an integrated aspect of fund disclosure rather than a separate theme. This work reverts the analysis of the performances of ESG funds versus non-ESG funds using the SFDR classification based on the asset manager's declaration. This approach overcomes most of the limits that affect sustainability ratings in the classification of funds (Berg et al. 2022; Martinez-Meyers et al. 2024) and permits to verify whether funds that belong to the same SFDR categories, as Articles 6, 8 and 9 exhibit similar risk-adjusted performances. Thus, the research hypothesis can be formulated as follows.

H1: SFDR Classification Provides Information on the Performance Metrics of Mutual Funds.

However, different performances within the same category of funds raise some doubts on the reliability of the research hypothesis. Indeed, this evidence highlights important issues in the open debate on the performance of ESG funds and leaves space to claim that the relationship between ESG orientation and performance might be absent or weak. Thus, the paper aims to assess whether the SFDR classification system, which categorizes funds based on their ESG orientation, has practical value in guiding investment decisions from a financial perspective.

It is worth pointing out that the research question is suggested by ESMA position on the outperformance of ESG funds and by the evidence provided by the extensive literature before the introduction of the SFDR classification (ESMA 2022a, 2022b). Consequently, this raises the important issue on the possible influence of the SFDR classification on the financial performance prediction, which represents essentially the position of the European regulator.

3 | Methodology

This paper proposes an innovative two-step analysis to investigate the hypothesis that a priori funds classification of SFDR provides useful information concerning performance expectations. First of all, principal components analysis (PCA) has been used to find the most significant orthogonal linear combinations of funds characteristics, which are called PCs (Jolliffe 2002). Then, these PCs have been included as new variables in the HCA which has been performed by using the Ward criterion (Strauss and von Maltitz 2017). The choice of this variables reduction is justified, since it aims at identifying a smaller set of PCs that capture the most relevant information in the original dataset with minimal loss of information. The agglomerative hierarchical clustering based on PCs can help to improve the interpretation of the underlying data structure as well as can help to lead to a more stable clustering (Everitt et al. 2011). In the following, a brief review of the PC analysis and the hierarchical clustering, used in this paper, has been proposed, and external and internal clustering validation measures computed to evaluate the performance of the clustering algorithm on the ESG dataset (Rousseeuw 1987; Maulik and Bandyopadhyay 2002) have been clarified (S.3).

3.1 | Variables Extraction With PCA

The PCA, also known as feature extraction method, is an unsupervised dimensionality reduction method and enhances the performance of the modeling process with minimum information loss. It was introduced by Pearson (1901) and then developed by Hotelling (1933).

More specifically, this procedure uses an orthogonal transformation to transform a set of cases of correlated variables into a set of values of linearly uncorrelated variables, known as PCs. The PC components are used in order to interpret the dependence among the selected variables over the PCs as well as to evaluate the relationships among observations. Thus, the PCs are built so that they are ordered with respect to the portion of explained variance. Note that a few uncorrelated variables (the first few PCs) are deemed to be useful. In the following, the PCA is presented.

Formally, given a vector x of p variables, the aim is to find a linear combination of x that maximizes the variance. Thus, it turns out that for $k = 1, 2, \dots, p$, the k th PC is given, as follows:

$$z_k = \alpha'_k x \quad (1)$$

where α'_k is an eigenvector of the variance \sum , corresponding to its k th largest eigenvalue λ_k . To derive the form of PCs, the constraint used is that α_k has been chosen to have unit length

$\alpha'_k \alpha_k = 1$, then $\text{var}(z_k) = \lambda_k$, where $\text{var}(z_k)$ denotes the variance of z_k . It is worth noting that to maximize $\alpha'_1 \sum \alpha_1$ subject $\alpha'_1 \alpha_1 = 1$, the technique of Lagrange multipliers is used. Thus, the function to be maximize is as follows:

$$\alpha'_1 \sum \alpha_1 - \lambda(\alpha'_1 \alpha_1 - 1), \quad (2)$$

where λ is a Lagrange multiplier; the maximum is obtained by setting equal to zero the first-order derivative with respect to α_1 , that is,

$$\sum \alpha_1 - \lambda \alpha_1 = 0 \quad (3)$$

Thus, λ is an eigenvalue of \sum and α_1 is the corresponding eigenvector. Note that to avoid the influence of scale and dimensionality, the original variables are standardized before performed the PCA.

3.2 | Clustering Approach

Cluster analysis is an unsupervised method with the aim to identify groups of similar observations according to a set of particular criteria. In the, there exist several types of clustering approaches which have been developed for different structure of data. The main categories of clustering algorithms are hierarchical and partitive algorithms. Mainly, hierarchical approaches are divided in agglomerative and divisive. The agglomerative allows to find the structures of data in which initially each observation is considered as an individual cluster and then, at each step, the most similar pairs of clusters are merged; this is known as bottom-up processes. On the other hand, the divisive algorithms are top-down processes. The choice of clustering method is related to the aim of the study, which can involve pattern classification, image segmentation and data mining. The definition of similarity between two objects implies a notion of distance. Similarity between clusters is generally measured by linkage methods, and in this study, the Ward-type linkage has been considered. In particular, Ward's linkage as the distance between two clusters corresponds to the minimum within-cluster variance.

In this paper, the agglomerative hierarchical clustering for analysis ESG data set has been selected. The hierarchical clustering produces nested sequence of partitions between one cluster at the top and clusters of individual points at the bottom. The output of hierarchical clustering is a tree and can be visualized as a tree in the form of a dendrogram.

3.3 | Performance Metrics of Classification

In this section, a brief description of the metrics adopted to evaluate the performance of clustering algorithm on the ESG data set is provided. In particular, these metrics have been considered to assess if there is a significant difference between the prior SFDR classification and the classification of funds obtained by the cluster procedure. The validity measures consist in two main categories: external clustering validation and internal clustering validation measures. The internal

validation measures evaluate the classification derived by the cluster procedure, based on the information within the data. On the contrary, the external measures evaluate the partitions derived by the cluster procedure, based on available supervised information, namely, information that has not been used for clustering.

Formally, let $C = C_1, \dots, C_K$ be a particular clustering partition of the n observations into K disjoint clusters; the internal (Connectivity, Average Silhouette Width and Dunn indices) and external (adjusted Rand index, ARI) are defined in the following.

- The connectivity index is defined as follows:

$$Conn(C) = \sum_{i=1}^n \sum_{j=1}^L \delta(i, nn_{i(j)}), \quad (4)$$

where $nn_{i(j)}$ is the j -th nearest neighbor of the i -th observation, L is the number of nearest neighbors of i observations, $\delta(i, nn_{i(j)})$ is equal to zero if i -th and j -th are observations in same cluster, and $\delta(i, nn_{i(j)})$ is equal to $\frac{1}{j}$ if they belong to different clusters. The connectivity index lies in the range $[0, \infty)$: low values of the connectivity index are considered optimal. This index measures the connectedness how relates to what extent units are located in the same cluster in other terms, their nearest neighbors in the space.

- The average silhouette width reflects the compactness and separation of clusters and takes the form as follows:

$$s(k) = \frac{1}{n} \sum_{i=1}^k \sum_{x \in C_i} \frac{b(x) - a(x)}{\max(a(x), b(x))}, \quad (5)$$

where $a(x) = \frac{1}{n} \sum_{n_i y \in C_i, y \neq x} d(x, y)$ and $b(x) = \min_{1 \leq j \leq k, j \neq i} \frac{1}{n_j} \sum_{n_j y \in C_j} d(x, y)$.

The values of the average silhouette width are in the range $[-1, 1]$ and higher values indicate better cluster result in terms of compactness and separation of clusters.

- The Dunn index is defined as follows:

$$Dunn(C) = \min_{1 \leq j \leq k} \left\{ \frac{\min_{x \in C_j, y \in C_l} d(x, y)}{\max_{1 \leq l \leq k} \left(\max_{x, y \in C_l} d(x, y) \right)} \right\}, \quad (6)$$

where the Dunn index is the ratio of smallest distance between units in different clusters (intercluster separation) and the largest intracluster compactness among all clusters. The Dunn index assumes values in the range $[0, \infty)$, where high values of Dunn index are considered optimal.

- The ARI is a measure of agreement between two data partitions: one derived by the clustering procedure C_1 and the other is the true clustering of the samples of ESG dataset based on domain knowledge C_0 . Therefore, in this paper, the ARI index has been used to compare and evaluate how similar are the partitions of funds related to ESG with respect to the funds classification obtained by using hierarchical clustering based on PCA. The ARI is defined as follows:

$$ARI = \frac{2(ad - bc)}{(a + b)(b + d) + (a + c)(c + d)}, \quad (7)$$

where a, b, c , and d indicate, respectively, the number of sample pairs belonging to the same cluster in both C_0 and C_1 , the number of sample pairs belonging to the same cluster in C_0 but to different clusters in C_1 , the number of sample pairs belonging to different clusters in C_0 but to the same clusters in C_1 , and the numbers of pairs belonging to different clusters in both C_0 and C_1 . The ARI index assumes values in the range $[0, 1]$ where values of index close to 0 indicate a random agreement between C_0 and C_1 partitions, while ARI values close to 1 indicate that C_0 and C_1 clustering both represent the same clusters.

4 | Data and Classification Analysis

The dataset used in this study has been collected by Refinitiv Database from March 10, 2021 to March 10, 2023 and consists

TABLE 1 | Description and descriptive statistics of relevant performance measures.

| Sections | Performance measures | Formula | Mean | std | min | max |
|------------------------|----------------------|---|------|------|-------|------|
| Return | GM | $\sqrt[n]{\prod_{i=1}^n (1 + \bar{R}_{pi})} - 1$ | 0.11 | 0.10 | -0.26 | 0.61 |
| Indicators | ER | $ER = \bar{R}_p - \bar{R}_f$ | 0.07 | 0.04 | -0.12 | 0.26 |
| | Alpha | $Alpha = \bar{R}_p - \bar{R}_{bmk}$ | 0.03 | 0.04 | -0.16 | 0.22 |
| Risk | σ | $\sigma = Stdev(\bar{R}_p)$ | 0.17 | 0.02 | 0.07 | 0.29 |
| indicators | TE | $TE = Stdev(\bar{R}_p - \bar{R}_{bmk})$ | 0.12 | 0.03 | 0.02 | 0.31 |
| | MaxDD | $MaxDD = \frac{TroughValue - PeakValue}{PeakValue}$ | 0.24 | 0.06 | 0.11 | 0.51 |
| Risk adjusted | SR | $SR = \frac{\bar{R}_p - \bar{R}_f}{\sigma_p}$ | 0.41 | 0.26 | -0.63 | 1.33 |
| Performance indicators | IR | $IR = \frac{\bar{R}_p - R_{bmk}}{TE}$ | 0.28 | 0.35 | -0.82 | 1.23 |

where \bar{R}_p is the mean of the time series of the return of the mutual funds, \bar{R}_f is the mean of the time series of the return of the risk-free asset, and \bar{R}_{bmk} is the mean of the times series of the return of the benchmark.

of 440 observations of various performance metrics referred to funds distributed in Europe and specialized on the European stock market with benchmarks coherent with a large cap blend style and by a set of 8 performance measures (Table 1). Taking into account that the investment style may influence the performance and the classification results, a homogeneous sample, covering only mutual funds, characterized by a similar portfolio, has been considered for the analysis.

Looking at the composition of the portfolio of the mutual funds, the analysis has been limited only on large cap funds without a value-growth bias ($\pm 10\%$) and with a sectoral breakdown consistent with the composition of the benchmark. In order to provide a comprehensive assessment of the performance of each fund, a set of financial metrics has been calculated, encompassing various dimensions of return, risk, and risk-adjusted. For every fund in the sample, not only has the geometric mean of the returns been computed to capture the average performance over time, but additional indicators have been also included to evaluate both absolute and relative performance. Specifically, the analysis has incorporated the calculation of standard deviation of returns to assess the total volatility or dispersion of returns around the mean (Derwall and Koedijk 2009; Schröder 2007), as well as the alpha, which measures the excess return (ER) of a fund relative to a benchmark and serves as an indicator of manager skill (Edwards and Caglayan 2001). Furthermore, the tracking error (TE) volatility has been examined to quantify how closely a fund follows its benchmark index, thereby providing insight into active risk (Aber et al. 2009; Petajisto 2013). To evaluate downside risk, the maximum draw-down has been computed, which identifies the largest peak-to-trough decline in fund value over the observed period and provides a sense of vulnerability during adverse market movements (Heidorn et al. 2009; Brandon and Qing 2022). Additionally, two widely used risk-adjusted performance measures have been employed: the Sharpe ratio, which evaluates returns relative to total risk (Miralles-Quirós et al. 2019; López Prol and Kim 2022; Kourtis 2016), and the information ratio (IR), which refines this assessment by focusing on the fund's active return relative to its TE, useful in assessing the efficiency with which a manager generates ERs given the level of active risk taken (Abate et al. 2021).

In the following section, the performance measures are classified into three macro sections: Return, Risk and Efficiency (Buchetti et al. 2025; Shahrouf 2022).

The first section is composed of three indicators, and it is focused on performance indicators assessing the absolute and relative return of each fund: the geometric mean of returns (GM), the ER defined as the difference between the return of the fund, and a risk free asset and the alpha indicator described as the difference between the return of the fund and the return of the MSCI Europe Net Return. The second section is composed by three risk indicators: the standard deviation (σ) measure of dispersion of fund's returns around the average, the TE that is the deviation from the benchmark of a fund and computed as standard deviation of alpha), and then maximum drawdown (MaxDD). The last section is composed by two risk adjusted performance indicators assessing the ability to reward each unit of risk in addition to the risk free asset,

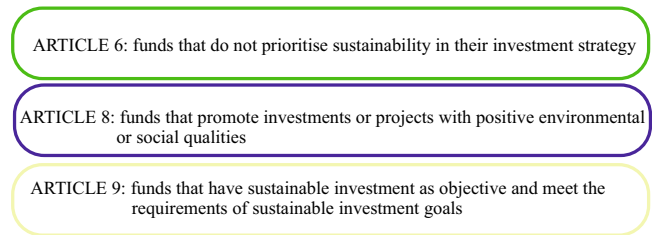


FIGURE 1 | SFDR classification by article and definitions.

namely, in the case of the Sharpe ratio (SR) and to the benchmark in the case of the IR.

In addition, the classification variable defined as $SFDR_0$ describes the rules for the classification of funds, as illustrated in Figure 1. More specifically, the $SFDR_0$ can be defined as the fund manager's attitude to a more or less sustainability-oriented portfolio composition and reporting.

It is worth highlighting that the variable of classification $SFDR_0$ and the financial indicators used in this paper can be considered a valid syntheses of ESG factors in the investment decision-making process. In order to increase the accuracy of learning approach, a preprocessing data analysis has been performed to detect the presence of outliers values and removing funds with missing data. After the preprocessing phase, the dataset consists of 440 funds for which 8 performance measures are taken with respect to the original dataset. Table 1 provides the name of financial indicators used in this paper jointly with a brief description of them and descriptive statistics. The descriptive statistics show positive risk adjusted performance indicators (namely, Sharpe ratio and IR) with a mean of 0.41 e 0.28 and a standard deviation of 0.26 and 0.35, respectively. That means that on average, the funds of the sample were able to outperform the risk free asset and the benchmark. This is consistent with the positive average of GM. The average value of standard deviation is consistent with the risk of the benchmark, even though the TE suggests an active strategy of the funds. The MaxDD is clearly affected by the negative performance of the European stock market in 2022.

4.1 | Variables Extraction Method for HCA in Financial Context

In this section, the PCA and the HCA have been applied to evaluate the specific hypothesis regarding the adequacy of the priori classification of funds ($SFDR_0$) as defined in the previous section. More specifically, the paper proposes to check the null hypothesis that the priori classification is suitable to provide the economic information necessary for a correct classification with respect to the financial characteristics of mutual funds. First of all, before proceeding with the study, the variable Alpha has been removed. The choice to remove this variable is justified by considering that the Alpha indicator is the numerator of the IR and therefore the ability to outperform the benchmark is already covered by this measure. Moreover, in some cases, Alpha indicator can be interpreted as a form of ER that also takes into account systematic risk (Berg et al. 2022). The PCA analysis has been performed

on the standardized values of the continuous original variables. Consequently, the PCA has been carried out on the correlation matrix, and the PCs with eigenvalues greater than one, based on Kaiser's rule (Kaiser 1958; Braeken and Assen 2016), have been included as covariates of HCA algorithm.

Table 2 presents the eigenvalues and the ratios of explained variances, from which it is evident as the PC variances are in progressive decline. From the eigenvalues in Table 2, it is worth to underline that the first two PCs can be selected, as the cumulative contribution rate exceeding the 84% is deemed to effectively reduce dimensionality while ensuring the preservation of sufficient information. Only the 15% of the variance remains unexplained.

Thus, the first PC of the data presents the highest eigenvalue (4.85); this component explains the greatest variation (69%) of the variables, making it the most significant. The influence is also equally distributed among all the variables. Similarly, the component with the second largest eigenvalue, (1.083) represents the second PC. This component explains an additional 16% of the variation of the variables.

After checking the results in Table 2, it is reasonable the hypothesis according to which the first component provides a measure related to the performance assessment, where SR and the IR indices are considered relevant for evaluating portfolios with normal distributions of absolute and relative returns in comparison with the risk free asset and the benchmark returns. In addition, the ER indicator allows investors to compare sets of funds against each other. On the other hand, the second PC provides a measures related to risks indicators described by the variables: σ , TE, and MaxDD. This second PC describes a relevant multidimensional aspect of risk management. Indeed, the PC component provides evidence regarding the dependence among the specific metrics that organizations use to monitor and assess potential risks with respect to the impact of the overall performance.

Furthermore, the PCA analysis results have been displayed on the factor plane by focusing on the relations between PCs and financial variables (Gabriel 1971), as illustrated in Figure 2 and Table 3, respectively. In particular, the interpretation of the graphical output allows detecting among the variables (identified by arrows) which ones contribute to the construction of the axes on a circle, called the correlation circle.

TABLE 2 | Eigenvalues and percentage of explained variance.

| PC | Eigenvalues | % of explained variance | Cumulative % of total variance |
|------|-------------|-------------------------|--------------------------------|
| PC 1 | 4.851867677 | 69.31239538 | 69.31240 |
| PC 2 | 1.082675030 | 15.46678615 | 84.77918 |
| PC 3 | 0.886083156 | 12.65833081 | 97.43751 |
| PC 4 | 0.103876709 | 1.48395299 | 98.92147 |
| PC 5 | 0.054288125 | 0.77554464 | 99.69701 |
| PC 6 | 0.018119349 | 0.25884785 | 99.95586 |
| PC 7 | 0.003089953 | 0.04414219 | 100.00000 |

From the Figure 2 and Table 3, it is evident that PC 1 best explains the variability of ER, GM, SR and IR variables. These values increase as PC 1 increase. On the contrary, PC 2 best explains the variability of MaxDD, σ and TE variables. However, among these, the MaxDD and the σ variables have been found to have higher weights than the one for the TE variable (vectors with longer lengths). Indeed, it is important to point out that if the arrows identify the variables, the angles between the respective vectors indicate the different degree of collinearity between the variables. Therefore, the arrows arrow indicate which variables account for most of PCs, while the length of the arrows corresponds to the intensity of the relationship.

4.2 | Financial Data Clustering Based on PCA

Within this context, the PCA has been used as a feature extraction to find the optimal subset of variables in terms of PCs ignoring redundant and irrelevant information. As a

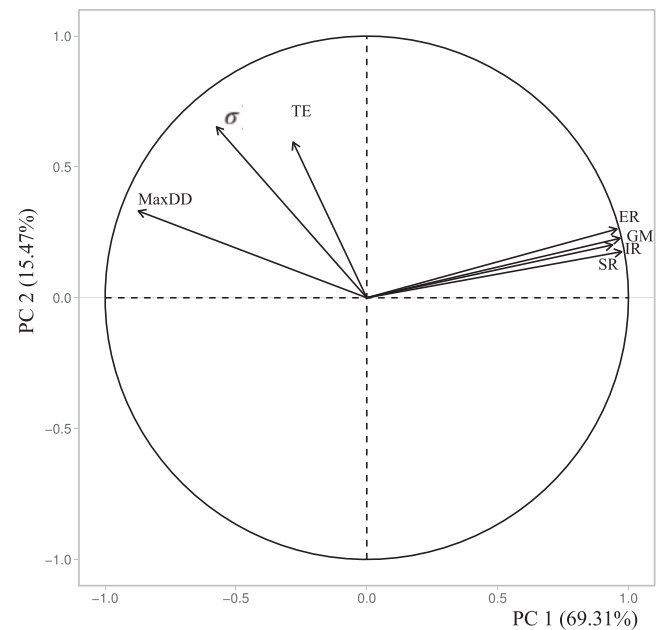


FIGURE 2 | Projection of the nine financial variables on the PC 1 and PC 2.

TABLE 3 | Pearson correlation coefficient between financial variables and the first two principal components (PCs).

| Financial variables | PC 1 | PC 2 |
|---------------------|--------|-------|
| GM | 0.968 | 0.227 |
| ER | 0.956 | 0.263 |
| σ | -0.574 | 0.653 |
| TE | -0.282 | 0.595 |
| MaxDD | -0.873 | 0.332 |
| SR | 0.974 | 0.176 |
| IR | 0.939 | 0.202 |

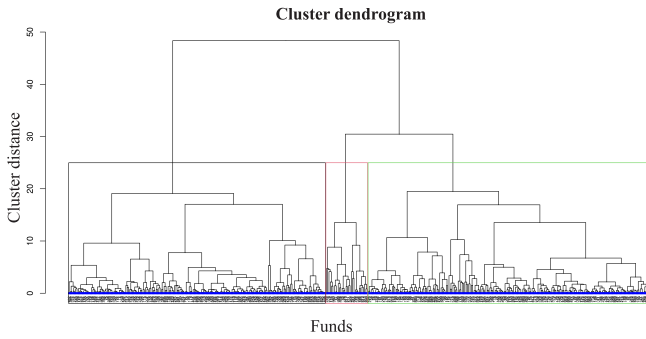


FIGURE 3 | Dendrogram of the funds and financial indicators based on Ward linkage method of hierarchical clustering.

consequence, these PCs have been included as covariates of an AHC which aims to identify groups of similar observations according to a set of particular criteria. As previously mentioned, the cluster analysis, conducted on the PCs, is characterized by the estimation of the dissimilarity matrix associated with the distances between the 440 funds (Euclidean distance) and similarity between clusters measured by linkage methods (Ward type linkage). In addition, the optimal number of clusters has been defined to 3 in accordance with several indices that combine information regarding intra and inter cluster compactness. The findings of cluster analysis has been depicted graphically in a diagram, known as cluster dendrogram, as given in Figure 3. It shows the graph generated from the dataset by using Ward's linkage of the multidimensional relationships among all of the samples, where the horizontal axis indicates the funds and the vertical axis indicates the distance (level of similarity) at which the clusters are merged. Furthermore, the partitions obtained have the following size: Cluster 1 (includes 213 funds), Cluster 2 (includes 32 funds), and Cluster 3 (includes 195 funds).

5 | Comparison Between the Prior Classification and Classification Based on Clustering Model

In this section, the hierarchical clustering based on PCA has been applied to assess the specific hypothesis, namely, that the SFDR classification provides a consistent information for the funds with respect to the classification obtained through the AHC.

In order to clarify the modeling findings, the priori classification of funds regulated under the SFDR that classifies financial products into three main categories according to their characteristics and sustainability is called $SFDR_0$. This priori classification of funds is called $SFDR_0$, and it is associated to the model “ $Model_0$ ”. On the other hand, the classification of funds under unlabeled categories based on the proposed approach is called $SFDR_1$, and it is associated to the model “ $Model_1$ ”. The choice to compare the partitions of $SFDR_0$, defined $Model_0$, and $SFDR_1$, denoted as $Model_1$, is justified considering the lively debate on the impact of ESG orientation on performance results. Thus, the misalignment of the distribution of funds in $Model_0$ and in $Model_1$ has been assessed. The validation criteria to assess the clustering partitions have been based on internal validation measures, as shown in Table 4. From the findings reported in Table 4, it is

TABLE 4 | Performance metrics for the prior classification $SFDR_0$ based on $Model_0$ and the classification driven by AHC approach $SFDR_1$ for $k = 3$ ($Model_1$).

| Model | Silhouette | Dunn | Connectivity |
|-----------|------------|------|--------------|
| $Model_0$ | 0.09 | 0.02 | 76.96 |
| $Model_1$ | 0.27 | 0.05 | 51.73 |

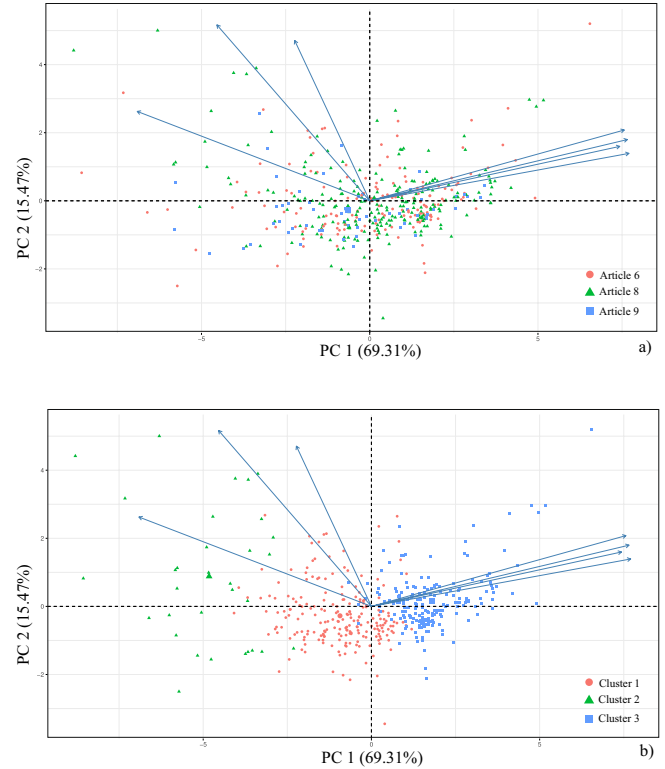


FIGURE 4 | PCA plot regarding the multivariate variation among 50 funds in terms of financial variables. The colormaps of the partitions correspond to the partitions achieved with $Model_0$ (a) with respect to the $Model_1$ (b).

evident that for all values, $Model_1$ outperforms the priori classification $Model_0$, in terms of separation and compactness.

In particular, the optimal values (for $k = 3$) for the average silhouette width and the Dunn index are 0.27 and 0.05, respectively. Similarly, the lowest value of the connectivity index equal to 51.73 suggests better clustering structure compared to other partition in terms of within-cluster variance. Figure 4 illustrates the colormaps of the partitions achieved with $Model_1$ with respect to $Model_0$. It is evident that the cluster partitions based on $Model_1$ provide a more compact result than those obtained considering only the priori information or $Model_0$.

On the other hand, a key aspect of this work has also been to compare the clustering procedure proposed in this paper with the priori classification (or priori knowledge) by using the external criteria. In this context, the ARI index has been computed in order to provide a measure of similarity between these two different partitions. The ARI index has been computed, and the score equal to 0.05 has revealed that two cluster solutions can be considered

independent; in other terms, the partitions provided by priori the classification $Model_0$ and the partitions provided by PCA based on AHC; $Model_1$ presents a low agreement of clusters, as also evident from the contingency table (Table 5) where there are the fund frequencies in relation to the partitions of the two models $SFDR_0$ and $SFDR_1$, respectively. Table 5 shows an evident misalignment of the distribution for the funds in $Model_0$ and $Model_1$, suggesting the poor predictive information about performance measures contained in the SFDR classification. Finally, the performance of the obtained clusters ($Model_1$) and the priori classification ($Model_0$) have also been evaluated with respect to the original financial data.

Therefore, to this aim, the boxplots categorized according to the three obtained clusters and with respect to the $SFDR_0$ classification are shown in Figure 5. From Figure 5, it is evident that the classification obtained using the proposed method is able to identify clusters characterized by specific financial variables. In particular, according to the results obtained, the funds found in Cluster 3 are the most efficient (those with the greatest value on average

TABLE 5 | Contingency table used to determine which funds have been correctly assigned to null ($Model_0$) and alternate distributions ($Model_1$).

| | | $Model_1$ | | | Total |
|-----------|-----------|-----------|-----------|-----------|-------|
| | | Cluster 1 | Cluster 2 | Cluster 3 | |
| $Model_0$ | Article 6 | 65 | 6 | 86 | 157 |
| | Article 8 | 115 | 19 | 89 | 223 |
| | Article 9 | 33 | 7 | 20 | 60 |
| Total | | 213 | 32 | 195 | 440 |

with respect to the Risk Adjusted performance indicators, i.e., SR and IR), while Cluster 2 includes the most risky funds (that present on average higher values with regards to risks metrics, i.e., σ , TE and MaxDD). For what concerns the funds found in Cluster 1, it is important to point out that these funds are characterized by similar values in terms of TE to Cluster 3 but have lower average values of performance indicators compared to the latter.

On the other hand, the priori classification based on $SFDR_0$ does not highlight the presence of significant differences among the theoretical clusters in terms of financial variables. This result is justified by the idea that the performance of funds is strongly affected by other factors (i.e., management style and costs) that can explain the heterogeneity of the performance patterns of fund showing the same SFDR classification.

Finally, the averages associated to these financial variables have been compared with respect to both classifications by using the Kruskal–Wallis test. By fixing a significance level of 1%, this test checks for the validity of the null hypothesis that there is no difference among the partitions based on cluster hypothesis. From the results in Table 6, it is clear that the null hypothesis cannot be rejected for the partitions based on the prior classification $SFDR_0$, with respect to all the financial indicators, since the p -values are all greater or equal to the significance level, which means that there are no statistically significant differences among partitions. An opposite conclusion can be derived for the partition obtained by the proposed method, for which the null hypothesis can be rejected, highlighting that the differences among the funds arise when the partitions are obtained by releasing the a priori classification $SFDR_0$.

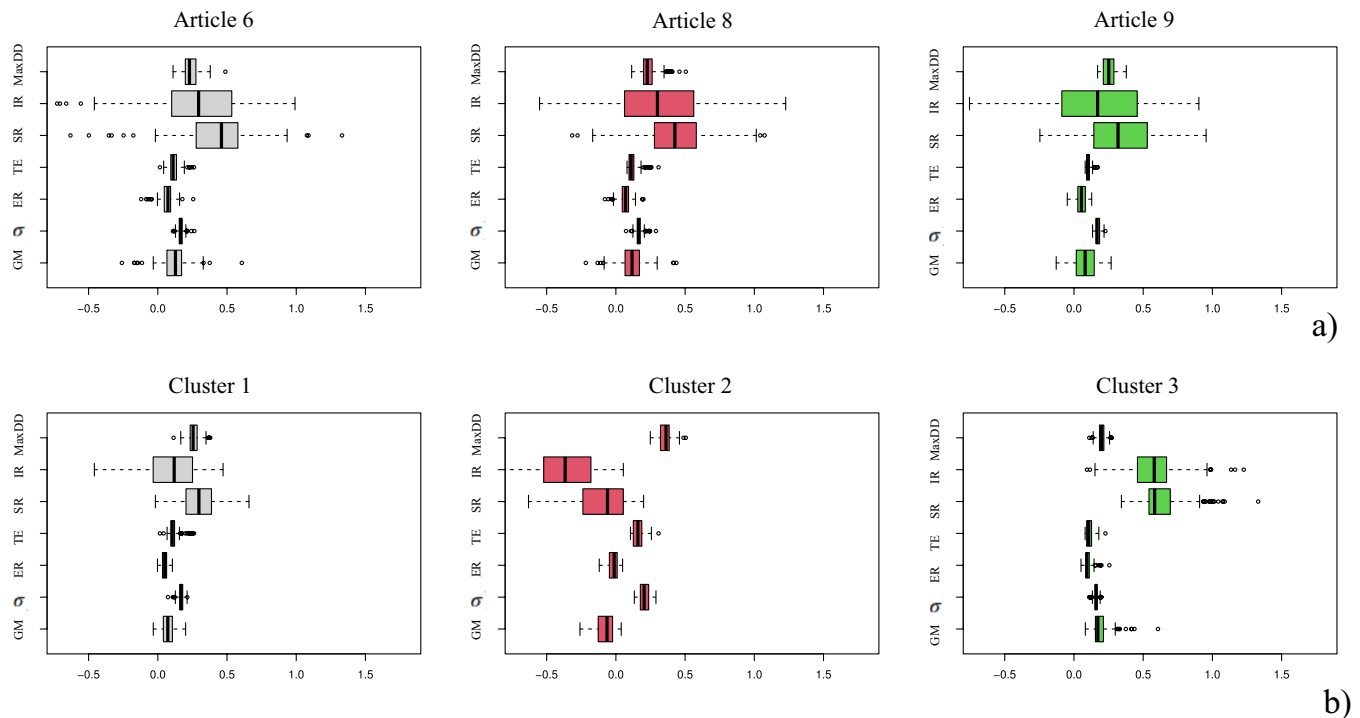


FIGURE 5 | Boxplots related to $p = 7$ financial variables classified by clusters with relation to a priori classification (a) and with respect to classification based on the AHC procedure (b).

TABLE 6 | Financial variables of the “prior classification” and based on AHC algorithm.

| Financial variables | “Prior classification” | | | AHC algorithm classification | | |
|---------------------|---------------------------|----|---------|------------------------------|----|---------|
| | Kruskal–Wallis statistics | df | p-value | Kruskal–Wallis statistics | df | p-value |
| GM | 8.324 | 2 | 0.02 | 323.458 | 2 | 0.00 |
| ER | 8.148 | 2 | 0.02 | 317.057 | 2 | 0.00 |
| σ | 4.414 | 2 | 0.11 | 74.780 | 2 | 0.00 |
| TE | 10.349 | 2 | 0.01 | 43.668 | 2 | 0.00 |
| MaxDD | 6.868 | 2 | 0.03 | 243.637 | 2 | 0.00 |
| SR | 8.612 | 2 | 0.02 | 322.245 | 2 | 0.00 |
| IR | 5.784 | 2 | 0.06 | 308.305 | 2 | 0.00 |

This demonstrates that even after the introduction of the SFDR (that represents a key step to market information since it shifts to asset managers the disclosure related to the sustainability orientation of the fund), it is not possible to make proven assumptions of over or underperformance associated with ESG’s funds orientation. This is also consistent with the position of ESMA, for which the evidence supporting the outperformance of ESG funds can be considered scarce; on the other hand, the general finding is that they tend to perform about the same—neither significantly better nor worse than conventional funds (ESMA 2022a, 2022b).

6 | Conclusions

The UE is leading the ESG deal at worldwide level and has decided to use the financial system as a transmission mechanism in order to improve the sustainability of the economic decisions. Also, the asset management industry is part of this project, and the investment companies have to disclose in the key information document the category of each mutual fund. The SFDR came into force in March 2021, and the first reaction of the market has been a strong shift of Assets Under Management (AUM) toward Articles 8 and 9 funds (Becker et al. 2022).

The academic research on this topic has been often based on the ESG ratings provided by different companies (Morningstar, Refinitiv) that unfortunately present a great divergence. The SFDR imposes an official classification of the mutual funds, but it is based on the declaration of the asset management companies and is exposed to the greenwashing risk (overstatement of the ESG orientation). Overall, the relationship between the intensity of responsible investing criteria and the financial performance of funds is far from straightforward and likely depends on the intensity of screening, the type of criteria used, the geographic context, and the time period analyzed. It is worth considering that the imposition of ethical constraints can be seen as a special case of restrictions on portfolio selection, which can lead to nonzero Jensen performance.

It is important to consider that the methodology used to compare performance can affect results. The use of appropriate benchmarks, controlling for factors such as fund and company size (small company effect), investment style, fees, and the use of multifactor models are crucial aspects. Gregory (1997) argued

that by accounting for the small company effect with a two-factor benchmark, the outperformance of the UK ethical funds found in previous studies disappears. Moreover, management fees may influence net performance for investors, if ethical funds have higher fees without superior performance, this would result in lower returns for investors. In summary, the evidence suggests that the relationship between the intensity of ESG criteria and the financial performance of mutual funds is not linear. Very low or very high screening intensity might be associated with better financial performance than moderate intensity. However, it is critical to consider the specific type of ESG criteria and the market context, as different intensities and types of screening can affect performance and investor behavior in distinct ways.

Finally, it is important to note that the definition of “ethical” and “unethical” can be vague and variable, making it difficult to establish a universally accepted list of “unethical” companies. The European Supervisory Authorities (ESAs) published a joint opinion on the SFDR, highlighting how the current provisions, while aimed at transparency, have generated complexity and potential greenwashing risks. The ESAs propose the introduction of a categorization system for financial products based on sustainability criteria and/or a sustainability indicator to simplify disclosures, especially for retail investors (www.eiopa.europa.eu/publications/joint-esas-opinion-assessment-sustainable-finance-disclosure-regulation-sfdr_en). Although the SFDR regulation has an objective of transparency and disclosure of ESG policies by investment companies without making value judgments about the effects of such guidelines, ESMA has stated on several occasions most recently (Securities and Authority 2025) that ESG funds have outperformed non-ESG funds particularly in the equity sector by promoting investment in funds classified in Articles 8 and 9. The story telling around the ESG benefits not only on the sustainability but also on better performances in terms of return, risk, and efficiency that have driven the tendency of many investors to select Articles 8 and funds. This study showed that the expectations of better performances only based on the SFDR classification are biased. The innovative approach of this study was to revert the analysis of the traditional performance metrics by creating clusters of funds with similar behaviors.

The clustering procedure was based on a PCA analysis of the performance metrics that allow focusing only on the variables able to explain the features of the considered sample of funds. It is worth pointing out that the obtained partitions can be considered

better than the prior one in terms of homogeneity, and the contingency table shows a low correlation of the classifications based on ESG declaration and on performances. The main finding of this research supports the idea that using the SFDR classification to create expectations of better future performance could be misleading, since the different metrics seem to be affected by other factors. Our assumption is that the management style (active versus passive) and the cost of the funds should be taken into account and could explain better (or together) with the ESG orientation; the performance of mutual funds and future research on this topic is needed. For the time being, regulators and asset management companies should pay attention to feed the idea that ESG investing is not only able to help the environment, to support social programs and to lead to better governance but is also a good deal in terms of performance because the data show a different story.

Acknowledgments

The authors are grateful to the anonymous reviewers for their helpful and constructive comments on earlier versions of the manuscript. Open access publishing facilitated by Università del Salento, as part of the Wiley - CRUI-CARE agreement.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Data will be made available on request.

References

- Abate, G., I. Basile, and P. Ferrari. 2021. "The Level of Sustainability and Mutual Fund Performance in Europe: An Empirical Analysis Using ESG Ratings." *Corporate Social Responsibility and Environmental Management* 28: 1446–1455. <https://doi.org/10.1002/csr.2175>.
- Aber, J., D. Li, and L. Can. 2009. "Price Volatility and Tracking Ability of ETFs." *Journal of Asset Management* 10: 210–221.
- Abouarab, R., T. Mishra, and S. Wolfe. 2025. "Spotting Portfolio Greenwashing in Environmental Funds." *Journal of Business Ethics* 197: 811–839. <https://doi.org/10.1007/s10551-024-05783-z>.
- Adler, T., and M. Kritzman. 2008. "The Cost of Socially Responsible Investing." *Journal of Portfolio Management* 35: 52–56.
- Atif, M., and S. Ali. 2021. "Environmental, Social and Governance Disclosure and Default Risk." *Business Strategy and the Environment* 30: 3937–3959.
- Barnett, M., and R. Salomon. 2011. "Does It Pay to Be Really Good? Addressing the Shape of the Relationship Between Social and Financial Performance." *Strategic Management Journal* 33: 1304–1320. <https://doi.org/10.2307/41679849>.
- Bauer, R., J. Derwall, and R. Otten. 2007. "The Ethical Mutual Fund Performance Debate: New Evidence From Canada." *Journal of Business Ethics* 70: 111–124.
- Becchetti, L., R. Ciceretti, A. Dalò, and S. Herzel. 2014. "Socially Responsible and Conventional Investment Funds: Performance Comparison and the Global Financial Crisis." *Applied Economics* 47: 2541–2562.
- Becker, M. G., F. Martin, and A. Walter. 2022. "The Power of ESG Transparency: The Effect of the New SFDR Sustainability Labels on Mutual Funds and Individual Investors." *Finance Research Letters* 47: 102708.

- Berg, F., J. Koumlbel, and R. Rigobon. 2022. "Aggregate Confusion: The Divergence of ESG Ratings." *Review of Finance* 26: 1315–1344. <https://doi.org/10.1093/rof/rfac033>.
- Birindelli, G., H. Chiappini, and R. N.-U.-D. Jalal. 2023. "SFDR, Investor Attention, and European Financial Markets." *Finance Research Letters* 56: 104135.
- Braeken, J., and M. Assen. 2016. "An Empirical Kaiser Criterion." *Psychological Methods* 22: 450–466.
- Brammer, S., C. Brooks, and S. Pavelin. 2005. "Corporate Social Performance and Stock Returns: UK Evidence From Disaggregate Measures." *Environmental Economics* 35: 97–116.
- Brandon, R. T., and Y. Qing. 2022. "Maximum Drawdown as Predictor of Mutual Fund Performance and Flows." *Financial Analysts Journal* 78(4): 59–76. <https://ssrn.com/abstract=4150870>.
- Buchetti, B., F. R. Arduino, and S. Perdichizzi. 2025. "A Literature Review on Corporate Governance and ESG Research: Emerging Trends and Future Directions." *International Review of Financial Analysis* 97: 103759. <https://doi.org/10.1016/j.irfa.2024.103759>.
- Capelle-Blancard, G., and S. Monjon. 2014. "The Performance of Socially Responsible Funds: Does the Screening Process Matter?" *European Financial Management* 20: 494–520. <https://doi.org/10.2139/ssrn.1734764>.
- Chegut, A. M., H. Schenk, and B. Scholtens. 2011. "Assessing SRI Fund Performance Research: Best Practices in Empirical Analysis." *Sustainable Development* 19: 77–94.
- Choi, S. U., and W. J. Lee. 2024. "Financial Statement Comparability and Environmental, Social, and Governance (ESG) Performance." *Sustainability* 16, no. 18: 7993. <https://doi.org/10.3390/su16187993>.
- Cortez, M. C., F. Silva, and N. Areal. 2012. "Socially Responsible Investing in the Global Market: The Performance of US and European Funds." *International Journal of Finance & Economics* 17: 254–271.
- Cortez, M., F. Silva, and N. Areal. 2009. "The Performance of European Socially Responsible Funds." *Journal of Business Ethics* 87: 573–588. <https://doi.org/10.1007/s10551-008-9959-x>.
- Cosma, S., P. Cucurachi, V. Gentile, and G. Rimo. 2023. "Sustainable Finance Disclosure Regulation Insights: Unveiling Socially Responsible Funds Performance During Covid 19 Pandemic and Russia Ukraine War." *Business Strategy and the Environment* 33: 3242–3257. <https://doi.org/10.1002/bse.3650>.
- Cremasco, C., and L. Boni. 2022. "Is the European Union (EU) Sustainable Finance Disclosure Regulation (SFDR) Effective in Shaping Sustainability Objectives? An Analysis of Investment Funds' Behaviour." *Journal of Sustainable Finance & Investment* 2022: 1–19.
- Cruciani, C., and C. R. M. A. Santagiustina. 2023. "The Present and Future of Sustainability Disclosure in Equity Investment Funds Pre-Contractual Documents: Mapping ESG Discourse Through STM." *Finance Research Letters* 58: 104033. <https://doi.org/10.1016/j.frl.2023.104033>.
- Derwall, J., and K. Koedijk. 2009. "Socially Responsible Fixed-Income Funds." *Journal of Business Finance & Accounting* 36: 210–229. <https://doi.org/10.1111/j.1468-5957.2008.02119.x>.
- Duncombe, S., M. Park, M. Tarsalewska, and G. Trojanowski. 2023. "ESG Positioning in Private Infrastructure Fundraising." *International Review of Financial Analysis* 90: 102924. <https://doi.org/10.1016/j.irfa.2023.102924>.
- ESMA. 2022a. "The Drivers of the Costs and Performance of ESG Funds." *ESMA Report* 50-165: 2146.
- ESMA. 2022b. "Performance and Costs of EU Retail Investment Products Annual Statistical Report." *ESMA Report* 50-165: 1677.
- Edwards, F., and M. Caglayan. 2001. "Hedge Fund Performance and Manager Skill." *Journal of Futures Markets* 21: 1003–1028. <https://doi.org/10.2139/ssrn.281524>.

- Everitt, B. S., S. Landau, M. Leese, and D. Stahl. 2011. *Cluster Analysis*. John Wiley & Sons. <https://doi.org/10.1002/9780470977811>.
- Fabozzi, F. J., K. C. Ma, and B. J. Oliphant. 2008. "Sin Stock Returns." *Journal of Portfolio Management* 35: 82–94.
- Fatemi, A., I. Fooladi, Y. Zhao, and Z. Ma. 2024. "On the Superior Performance of SRI Funds." *International Review of Economics & Finance* 93: 567–581.
- Ferriani, F. 2024. "The Importance of Labels for Sustainable Investments: SFDR Versus Morningstar Globes." *Applied Economics Letters* 31, no. 18: 1813–1819. <https://doi.org/10.1080/13504851.2023.2208326>.
- Gabriel, K. R. 1971. "The Biplot Graphic Display of Matrices With Application to Principal Component Analysis." *Biometrika* 58: 453–467.
- Gantchev, N., M. Giannetti, and R. Li. 2024. "Sustainability or Performance? Ratings and Fund Managers' Incentives." *Journal of Financial Economics* 155: 103831.
- Gil-Bazo, J., P. Ruiz-Verdú, and A. Santos. 2010. "The Performance of Socially Responsible Mutual Funds: The Role of Fees and Management Companies." *Journal of Business Ethics* 94: 243–263. <https://doi.org/10.1007/s10551-009-0260-4>.
- Goldreyer, E., and J. D. Diltz. 1999. "The Performance of Socially Responsible Mutual Funds: Incorporating Sociopolitical Information in Portfolio Selection." *Managerial Finance* 25: 23–36.
- Gregory, A. 1997. "Ethical Unit Trust Financial Performance: Small Company Effects and Fund Size Effects." *Journal of Business Finance & Accounting* 24: 705.
- Hamilton, S., H. Jo, and M. Statman. 1993. "Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds." *Financial Analysts Journal - FINANC ANAL J* 49: 62–66. <https://doi.org/10.2469/faj.v49.n6.62>.
- Hart, O., and L. Zingales. 2017. "Companies Should Maximize Shareholder Welfare Not Market Value." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3004794>.
- Hartzmark, S., and A. Sussman. 2019. "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and FUND FLOWS." *Journal of Finance* 74, no. 6: 2789–2837.
- Heidorn, T., D. G. Kaiser, and C. Roder. 2009. "The Risk of Funds of Hedge Funds: An Empirical Analysis of the Maximum Drawdown." *Journal of Wealth Management* 12: 89–100. <https://api.semanticscholar.org/CorpusID:154983378>.
- Helliar, C., B. Petracchi, and N. Tantisantiwong. 2022. "Comparing SRI Funds to Conventional Funds Using a PCA Methodology." *Journal of Asset Management* 23: 581–595. <https://doi.org/10.1057/s41260-022-00264-2>.
- Hotelling, H. 1933. "Analysis of a Complex of Statistical Variables Into Principal Components." *Journal of Educational Psychology* 24: 498–520.
- Ibikunle, G., and T. Steffen. 2017. "European Green Mutual Fund Performance: A Comparative Analysis With Their Conventional and Black Peers." *Journal of Business Ethics* 145: 337–355. <https://doi.org/10.1007/s10551-015-2850-7>.
- Jolliffe, I. 2002. *Principal Component Analysis for Special Types of Data*. Springer.
- Jones, S., S. Van der Laan, G. Frost, and J. Loftus. 2008. "The Investment Performance of Socially Responsible Investment Funds in Australia." *Journal of Business Ethics* 80: 181–203. <https://doi.org/10.1007/s10551-007-9412-6>.
- Kaiser, H. F. 1958. "The Varimax Criterion for Analytic Rotation in Factor Analysis." *Psychometrika* 23, no. 3: 187–200. <https://doi.org/10.1007/BF02289233>.
- Klinkowska, O., and Y. Zhao. 2023. "Fund Flows and Performance: New Evidence From Retail and Institutional SRI Mutual Funds." *International Review of Financial Analysis* 87: 102596. <https://doi.org/10.1016/j.irfa.2023.102596>.
- Kourtis, A. 2016. "The Sharpe Ratio of Estimated Efficient Portfolios." *Finance Research Letters* 17: 72–78. <https://doi.org/10.1016/j.frl.2016.01.009>.
- Kreander, N., R. Gray, D. Power, and C. D. Sinclair. 2005. "Evaluating the Performance of Ethical and Non-Ethical Funds: A Matched Pair Analysis." *Journal of Business Finance & Accounting* 32: 1465–1493. <https://doi.org/10.1111/j.0306-686X.2005.00636.x>.
- Lean, H. H., W. R. Ang, and R. Smyth. 2015. "Performance and Performance Persistence of Socially Responsible Investment Funds in Europe and North America." *North American Journal of Economics and Finance* 34: 254–266.
- Leite, P., and M. Cortez. 2014. "Style and Performance of International Socially Responsible Funds in Europe." *Research in International Business and Finance* 30: 248–267. <https://doi.org/10.1016/j.ribaf.2013.09.007>.
- López Prol, J., and K. Kim. 2022. "Risk-Return Performance of Optimized ESG Equity Portfolios in the NYSE." *Finance Research Letters* 50: 103312. <https://doi.org/10.1016/j.frl.2022.103312>.
- Luther, R., J. Matatko, and D. Corner. 1992. "The Investment Performance of UK "Ethical" Unit Trusts." *Accounting, Auditing & Accountability Journal* 5: 1–1. <https://api.semanticscholar.org/CorpusID:155025121>.
- Mallin, C., B. Saadouni, and R. J. Briston. 2006. "The Financial Performance of Ethical Investment Funds." *Journal of Business Finance & Accounting* 22: 483–496. <https://doi.org/10.1111/j.1468-5957.1995.tb00373.x>.
- Martinez-Meyers, S., I. Ferrero-Ferrero, and M. J. Muñoz-Torres. 2024. "The European Sustainable Finance Disclosure Regulation (SFDR) and Its Influence on ESG Performance and Risk in the Fund Industry From a Multi-Regional Perspective." *Journal of Financial Reporting and Accounting*.
- Maulik, U., and S. Bandyopadhyay. 2002. "Performance Evaluation of Some Clustering Algorithms and Validity Indices." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, no. 12: 1650–1654.
- Miralles-Quirós, J. L., M. M. Miralles-Quirós, and J. Nogueira. 2019. "Diversification Benefits of Using Exchange-Traded Funds in Compliance to the Sustainable Development Goals." *Business Strategy and the Environment* 28: 244–255. <https://doi.org/10.1002/bse.2253>.
- Neves, M. E., A. Martins, M. Gouveia, and C. Pinho. 2023. "Are Green and Socially Responsible Funds More Efficient Than Conventional Funds? A DEA Approaches." *Journal of Economic Studies* 51: 261–286. <https://doi.org/10.1108/JES-12-2022-0677>.
- Pearson, K. 1901. "LIII. On Lines and Planes of Closest Fit to Systems of Points in Space." *Philosophical Magazine Series 1* 2: 559–572.
- Petajisto, A. 2013. "Active Share and Mutual Fund Performance." *Financial Analysts Journal* 69: 73–93. <https://doi.org/10.2139/ssrn.1685942>.
- Rathner, S. 2013. "The Influence of Primary Study Characteristics on the Performance Differential Between Socially Responsible and Conventional Investment Funds: A Meta-Analysis." *Journal of Business Ethics* 118: 349–363.
- Renneboog, L., J. Horst, and C. Zhang. 2008. "Socially Responsible Investments: Institutional Aspects, Performance, and Investor Behavior." *Journal of Banking & Finance* 32: 1723–1742. <https://doi.org/10.1016/j.jbankfin.2007.12.039>.
- Renneboog, L., J. Horst, and C. Zhang. 2011. "Is Ethical Money Financially Smart? Nonfinancial Attributes and Money Flows of Socially Responsible Investment Funds." *Journal of Financial Intermediation* 20: 562–588. <https://doi.org/10.1016/j.jfi.2010.12.003>.

- Rousseuw, P. 1987. "Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis." *Journal of Computational and Applied Mathematics* 20, no. 1: 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- Sandberg, J., C. Juravle, T. M. Hedesstrm, and I. Hamilton. 2009. "The Heterogeneity of Socially Responsible Investment." *Journal of Business Ethics* 87: 519–533. <https://doi.org/10.1007/s10551-008-9956-0>.
- Scheitza, L., and T. Busch. 2024. "SFDR Article 9: Is it All About Impact?" *Finance Research Letters* 62: 105179. <https://doi.org/10.1016/j.frl.2024.105179>.
- Scherer, B., and M. Hasaj. 2023. "Greenlabelling: How Valuable Is the SFDR Art 9 Label?" *Journal of Asset Management* 24, no. 7: 541–546.
- Schröder, M. 2004. "The Performance of Socially Responsible Investments: Investment Funds and Indices." *Financial Markets and Portfolio Management* 18: 122–142. <https://doi.org/10.1007/s11408-004-0202-1>.
- Schröder, M. 2007. "Is There a Difference? The Performance Characteristics of SRI Equity Indices." *Journal of Business Finance & Accounting* 34: 331–348. <https://doi.org/10.1111/j.1468-5957.2006.00647.x>.
- Securities, E., and M. Authority. 2025. *Costs and Performance of EU Retail Investment Products 2024 ESMA Market Report*. European Securities and Markets Authority. <https://doi.org/10.2856/3163056>.
- Shahrour, M. H. 2022. "Measuring the Financial and Social Performance of French Mutual Funds: A Data Envelopment Analysis Approach." *Business Ethics, the Environment & Responsibility* 31: 398–418.
- Shank, T. M., D. K. Manullang, and R. P. Hill. 2005. "Doing Well While Doing Good? Revisited: A Study of Socially Responsible Firms' Short-Term Versus Long-Term Performance." *Managerial Finance* 31: 33–46.
- Statman, M. 2000. "Socially Responsible Mutual Funds." *Financial Analysts Journal - FINANC ANAL J* 56: 30–39. <https://doi.org/10.2469/faj.v56.n3.2358>.
- Statman, M., and D. Glushkov. 2008. "The Wages of Social Responsibility." *Financial Analysts Journal* 65: 33–46.
- Strauss, T., and M. von Maltitz. 2017. "Generalising Ward's Method for Use With Manhattan Distances." *PLoS ONE* 12: e0168288.
- Townsend, B. 2020. "From SRI to ESG: The Origins of Socially Responsible and Sustainable Investing." *Journal of Impact and ESG Investing* 1: 10–25.
- Varma, A., and J. Nofsinger. 2014. "Socially Responsible Funds and Market Crises." *Journal of Banking & Finance* 48: 180–193. <https://doi.org/10.2139/ssrn.2142343>.
- Wang, H. 2024. "ESG Investment Preference and Fund Vulnerability." *International Review of Financial Analysis* 91: 103002. <https://doi.org/10.1016/j.irfa.2023.103002>.
- Yue, X.-G., Y. Han, D. Teresien, J. Merkyte, and W. Liu. 2020. "Sustainable Funds' Performance Evaluation." *Sustainability* 12: 8034.
- Zeidan, R. 2022. "Why Don't Asset Managers Accelerate ESG Investing? A Sentiment Analysis Based on 13,000 Messages From Finance Professionals." *Business Strategy and the Environment* 31: 3028–3039.