Graphical Abstract

Identifying Risk Variables From ESG Raw Data Using A Hierarchical Variable Selection Algorithm

Zhi Chen, Zachary Feinstein, Ionuţ Florescu

Highlights

Identifying Risk Variables From ESG Raw Data Using A Hierarchical Variable Selection Algorithm

Zhi Chen, Zachary Feinstein, Ionuţ Florescu

- We propose a novel Hierarchical Variable Selection (HVS) algorithm to identify important risk variables from raw, high-dimensional ESG data.
- The HVS algorithm illustrates superior explanatory power for corporate risk compared to models using traditional, pre-aggregated ESG scores.
- HVS achieves superior explanatory power compared to traditional variable selection methods while simultaneously selecting a more parsimonious set of variables.
- We illustrate the applicability of the HVS algorithm for identifying sets of risk variables for companies from different industry sectors.

Identifying Risk Variables From ESG Raw Data Using A Hierarchical Variable Selection Algorithm

Zhi Chen[a,1,*], Zachary Feinstein[a], Ionuţ Florescu[a,b]

[a]*School of Business, Stevens Institute of Technology, 1 Castle Point Terrace , Hoboken, 07030, NJ, USA*
[b]*Hanlon Financial Laboratories, School of Business, Stevens Institute of Technology, 1 Castle Point Terrace, Hoboken, 07030, NJ, USA*

Abstract

Environmental, Social, and Governance (ESG) factors aim to provide non-financial insights into corporations. In this study, we investigate whether we can extract relevant ESG variables to assess corporate risk, as measured by logarithmic volatility. We propose a novel Hierarchical Variable Selection (HVS) algorithm to identify a parsimonious set of variables from raw data that are most relevant to risk. HVS is specifically designed for ESG datasets characterized by a tree structure with significantly more variables than observations. Our findings demonstrate that HVS achieves significantly higher performance than models using pre-aggregated ESG scores. Furthermore, when compared with traditional variable selection methods, HVS achieves superior explanatory power using a more parsimonious set of ESG variables. We illustrate the methodology using company data from various sectors of the US economy.

Keywords: ESG, Financial Risk, Variable Selection, Hierarchical Data, Sustainable Finance
JEL: G320, C550, C520, Q560, M140, G200

1. Introduction

Environmental, Social, and Governance (ESG) factors represent non-financial information used when analyzing corporations and investment strate-

*Corresponding author
Email address: zchen100@stevens.edu (Zhi Chen)

gies (Giese et al., 2019; Verheyden et al., 2016; Yalew et al., 2020). Environmental criteria assess a company's resource management, emissions, and environmentally-friendly innovations. Social criteria focus on how a company manages its relationships with employees, communities, suppliers, and customers. Governance reflects a company's leadership structure and decision-making processes.

To assess ESG factors, data providers, such as LSEG,[1] Bloomberg, and MSCI collect raw ESG variables from corporations and often design a hierarchical structure to aggregate this data into a single overall score (Halbritter and Dorfleitner, 2015; Shen et al., 2023; Wang et al., 2023). The availability of these scores coincided a period of rapid expansion in ESG investing. Foley et al. (2024) reports that sustainably managed assets in the United States market grew from \$6.57 trillion in 2014 to \$17.08 trillion by 2020, an increase of more than 160% . However, after this initial growth, literature has raised questions about the efficacy of ESG investing as well as the ESG scores. For example, a review by (Pérez et al., 2022) shows that ESG scores are poor predictors of asset returns. Additionally, each data provider has its own way to aggregate raw variables, resulting in different rating agencies providing inconsistent scores for the same company (Park, 2024). These doubts led to a subsequent downturn in ESG investment, with assets decreasing to \$8.40 trillion by 2022 (Anne Tucker, 2024; Alliance, 2022; Gibson Brandon et al., 2021).

Previous studies have found that ESG factors are not significant in relation to corporate equity returns (e.g., Pérez et al. (2022); Engelhardt et al. (2021); Deng and Cheng (2019); Fauver et al. (2018)). On the other hand, recent work shows a potentially stronger relationship between ESG variables and risk-related measures of a company's performance. Specifically, Engelhardt et al. (2021) reports an R-squared value around 0.04 when regressing a company's historical volatility using the overall ESG score. Furthermore, this relationship becomes stronger when using more granular scores. Alsaadi et al. (2017) finds a value of 0.15 using individual E, S, and G component scores. Fauver et al. (2018) uses an even finer set of variables and achieve a value of 0.20 in its relationship with historical volatility. Bonacorsi et al. (2024) constructs credit risk variables (Z-score) and regresses them on ESG raw variables. The authors report an R-squared value around 0.40. In our

[1]The London Stock Exchange Group (LSEG), was formerly known as Refinitiv.

work, we build on these ideas by creating a model that relates logarithmic volatility to raw variables.

We propose a Hierarchical Variable Selection algorithm (HVS), which analyzes the entire set of raw ESG data, to select a parsimonious set of variables that are most relevant to the logarithmic volatility of returns. The HVS algorithm is specifically designed for ESG datasets that contain many variables with a limited number of observations. Data providers have been expanding the set of ESG raw variables considered within their datasets. For instance, Refinitiv advertises the number of raw variables collected as 400 in 2017 and 600 in 2023. Increasing the number of features collected can potentially attract more clients. However, past data may not be augmented easily with values for the new variables introduced in later years. The resulting database is very sparse. This sparsity leads to issues like multicollinearity among available variables. To address these challenges, the HVS algorithm uses the tree structure of the LSEG dataset to select relevant variables iteratively.

Compared with classical variable selection algorithms such as Stepwise regression, Lasso regression, and PCA, HVS performs well while simultaneously improving the model's explainability. Crucially, we show that HVS is robust for out-of-sample data.

Since ESG dataset captures some non-financial information, a natural question is whether the company's value is affected by this information. To investigate this, we combine the HVS selected factors with the annual factors from the Fama-French 3-factor model as a proxy for financial information. Our results confirm that HVS selected variables bring little to no contribution to change in returns. However, they are highly significant for logarithmic volatility of returns even in the presence of financial market factors.

We apply the HVS algorithm to companies from various sectors of the US economy. This case study is motivated by the data providers practice to aggregate raw variables on a per-sector basis (Eikon, 2022). Our algorithm directly identifies important risk variables (as measured by logarithmic volatility) for each sector from the data. Our finding reveals that the most relevant risk variables are significantly different for each sector. As such, the HVS algorithm can be of use to, e.g., the sustainability accounting standards board (SASB), an organization aiming to identify ESG issues that affect a corporation's financial materiality within each sector (Madison and Schiehll, 2021; GawÄ et al., 2022; Behl et al., 2022; Miranda et al., 2023). We also apply the HVS algorithm to large- and small-capitalization companies. The results indicate that relevant ESG risk variables are significantly different for small companies than the relevant ones for large companies.

The remainder of this work is organized as follows. Section 2 explains our rationale for using logarithmic volatility rather than the standard form of volatility. Section 3 introduces our Hierarchical Variable Selection (HVS) algorithm and illustrates each step's results using the Energy sector as an example. In section 4, we evaluate the algorithm's robustness and benchmark its performance. Finally, section 5 presents case studies: we identify relevant risk variables for each industry sector, as well as for both large- and small-capitalization companies by sector.

2. Logarithmic volatility as response variable

Many studies attempt to link the ESG overall score with the return of a corporation, but they do not find a significant correlation between them (Engelhardt et al., 2021; Li et al., 2021; Halbritter and Dorfleitner, 2015). Due to the lack of correlation, other studies consider alternative return-related response variables such as earnings per share, ROA, or Tobin's Q (Deng and Cheng, 2019; Fauver et al., 2018; Alareeni and Hamdan, 2020), but these studies also reveal no significant correlation. In this paper, we corroborate that ESG is not a reliable factor for future returns (see Appendix B). However, recent research suggests that ESG is more informative when assessing corporate risk. Specifically, in Engelhardt et al. (2021), authors regress ESG overall scores on volatility. They find a statistically significant negative correlation with volatility. The significant negative correlation between ESG overall scores and other risk metrics has also been highlighted in the literature (Zhou and Zhou, 2021; Lööf et al., 2022).

We began our analysis using historical volatility as the response variable to measure risk.

$$s = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \bar{r})^2} \tag{1}$$

where: N is the number of trading days in a year; r_t is the daily return of day t; $\bar{r}$ is the average return over N days. However, diagnostic checks reveal that the model's residuals violate the normality assumption for a linear regression. To address this, we apply the Box-Cox transformation to identify an appropriate functional form for the response variable.

$$y^{(\lambda)} = \begin{cases} \frac{y^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0 \\ \log(y), & \text{if } \lambda = 0 \end{cases} \tag{2}$$

The results suggest that a logarithmic transformation ($\lambda = 0$) of volatility better satisfies the model assumptions, as illustrated in Table 1. Additionally, using logarithmic volatility leads to an increase in the R-squared value of the HVS algorithm, indicating improved model fit and explanatory power.

Table 1: Comparison of regular volatility and logarithmic volatility

Volatility Form	The p-value of Jarque-Bera Test For Residuals	R^2
Historical Volatility	0.0000	0.39
Logarithmic Volatility	0.1389	0.51

As a result, to quantify risk we calculate the annual logarithmic volatility using daily return observations for each company in the dataset using the following formula:

$$\log(s) = \log\left(\sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(r_t - \bar{r})^2}\right) \tag{3}$$

This quantity is going to be our response variable throught this study.

3. The Hierarchical Variable Selection (HVS) Algorithm

Within this section we introduce a novel algorithm that selects a parsimonious set of ESG variables relevant to the logarithmic volatility of returns. Our algorithm is specifically designed to use the hierarchical structure of the LSEG dataset. We will first describe this structure and then detail the steps of the algorithm. We illustrate the proposed Hierarchical Variable Selection (HVS) algorithm using data from companies in the Energy sector.

Notably, our methodology can be used to support the principles outlined by the Sustainability Accounting Standards Board (SASB). SASB outlines ESG topics which they call the “materiality map” likely to affect a company’s performance in a given sector. HVS can provide an empirical complement to this expert-defined mapping. Additionally, while the SASB materiality map typically outlines broad material topics or concepts, our algorithm pinpoints to specific raw variables within these broader areas that have the most relevance to a corporation's risk profile.

3.1. LSEG dataset and its structure

We chose to use the LSEG dataset because it provides access to a large number of raw variables. Figure 1 describes the structure of this dataset. Specifically, the raw variables are aggregated hierarchically through several tiers: from the base data to 15 category scores, then combined into 3 pillar scores, and ultimately into a single overall score. In this aggregation process, LSEG groups raw variables into 15 categories, as illustrated by the arrows in Figure 1. Many of the raw variables in the dataset contain missing values. We assign 0 to missing values for boolean variables according to Clarkson et al. (2008); Cheng et al. (2014). We then remove any numeric variables that have less than 80% data availability within each sector. Finally, any company-year observation that still contains missing values after these steps is removed entirely. The details of data processing steps are provided in Appendix A. Taking the Energy sector as an example, the initial dataset had 617 variables and 695 company/year data points. After the data processing steps, the final dataset for the Energy sector contains 255 variables and 422 observations.

3.2. Steps of the Hierarchical Variable Selection Algorithm

Algorithm 1: Hierarchical Variable Selection (HVS) Algorithm

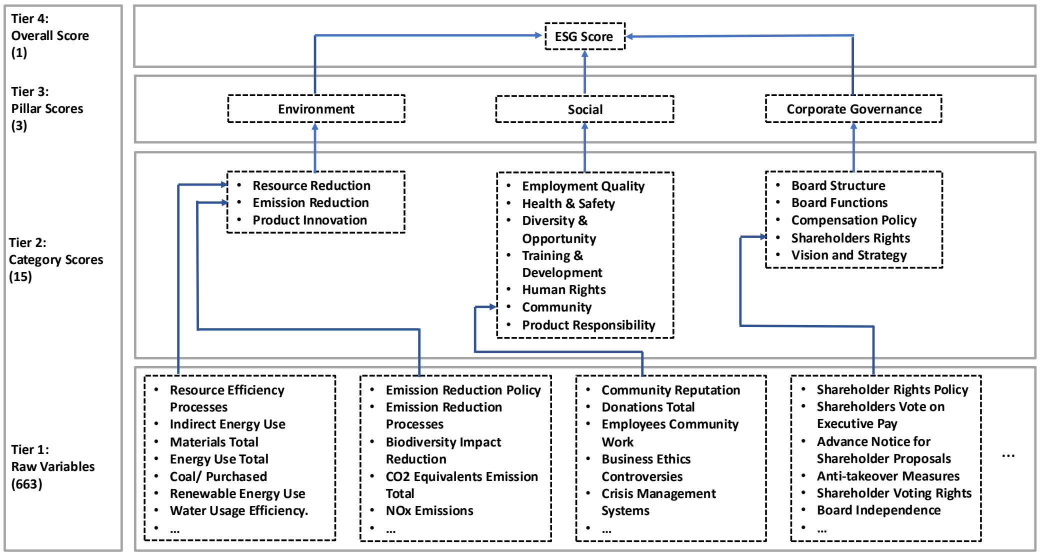
Step 1: Perform a Stepwise regression with raw variables within each category. Use the AIC criterion for model selection;

Step 2: Perform a Stepwise regression with all significant variables identified in Step 1 across categories. Use the AIC criterion for model selection;

Step 3: Perform a Ridge regression with the selected variables identified in step 2;

Result: A set of selected variables and weights

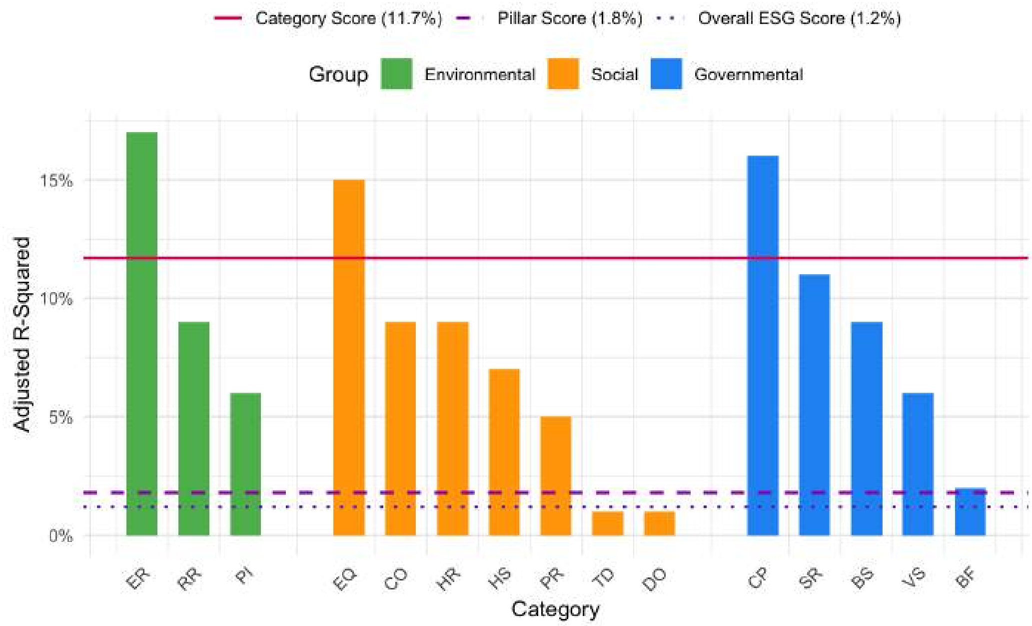
Figure 1: Hierarchical Structure of the LSEG ESG Dataset



Step 1. The first step of the HVS algorithm aims to select a comprehensive set of raw variables within each of the 15 categories (Tier 2 in Figure 1) that are most relevant to risk (as measured by logarithmic volatility). To achieve this, we perform a Stepwise regression using the AIC criterion for each categories, which produces 15 such sets of variables. For the Energy sector example, these sets combine to a total of 55 selected variables from an initial 255. The performance of these 15 regressions for the Energy sector is visualized in Figure 2 on page 8. The bars are the Adjusted R-squared values for each regression. We use Adjusted R-squared values for visualization, as they penalize a model for having more factors. For comparison, the plot also includes three horizontal lines representing the Adjusted R-squared values from benchmark models, each using one of LSEG's aggregated score (overall, pillar, or category) as predictors.

Looking at the three lines in Figure 2, the Adjusted R-squared increases from 1.2% using the single overall ESG score, slightly to 1.8% with the more detailed pillar scores, and then rises substantially to 11.7% when using the even more granular category scores. This trend aligns with prior literature, which shows that using more detailed ESG scores consistently leads to better model performance (Bonacorsi et al., 2024; Fauver et al., 2018; Alsaadi et al., 2017). Notably, models built using selected raw variables from cer-

Figure 2: This bar chart illustrates the Adjusted R-squared values of using selected raw variables within each category to regress logarithmic volatility of returns. The dashed lines indicate the Adjusted R-squared values obtained from linear regression models using aggregated scores provided by LSEG as regressors (i.e., 15 category scores, 3 pillar scores, single overall score).



tain specific ESG categories—'Emission Reduction' from the E pillar, 'Employment Quality' from the S pillar, and 'Compensation Policy' from the G pillar—outperform the model that uses all 15 pre-aggregated category scores. This is particularly striking given that these 15 category scores are aggregated from the entire set of raw variables. This finding suggests that the aggregation process itself may obscure or dilute the impact of the most critical risk-relevant variables.

Step 2. The second step of the HVS algorithm further refines the variables selected in Step 1. We aggregate all variables selected by the 15 regressions in the previous step. However, simply adding them together could introduce multicollinearity. To mitigate this, HVS performs a second Stepwise regression with these selected variables.

The Adjusted R-squared value from this consolidated model is significantly higher than the values achieved by any of the models using selected variables from a single category (Figure 3). This indicates that variables from different categories provide unique and complementary contributions to the ESG risk model. To our knowledge, this Adjusted R-squared value is also larger than those reported in other studies (Engelhardt et al., 2021; Zhou and Zhou, 2021; Lööf et al., 2022). In addition to increasing performance, Step 2 of the algorithm creates a parsimonious set of variables. For example, in the Energy sector the number of variables is reduced from 51 to 31 and at the same time the Adjusted R-squared value increases from around 17% to 48%.

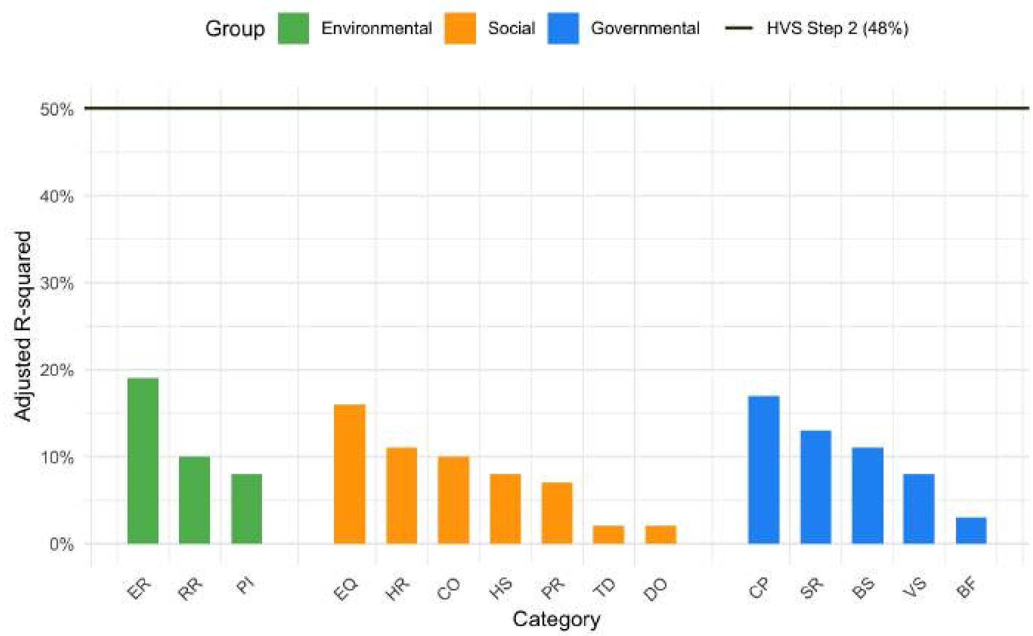


Figure 3: Adjusted R-squared values of HVS Step 1 vs. HVS Step 2 for the Energy Sector

Step 3. The objective of Step 3 is to enhance the model's robustness. Given that ESG datasets typically contain a large number of variables but relatively few observations, there is a high risk of overfitting. This step is focused on improving the model's out-of-sample performance. We use Ridge regression (Hoerl and Kennard, 1970) as it has a penalty term that reduces the risk of overfitting the dataset. We use a 5-fold cross validation technique to determine the regularization parameter. Additionally, we standardize all variables, as this will enable us to compare the effect of each variable on the response.

Table 2: Comparison of Model Fit Metrics after HVS Step 2 (Stepwise) and Step 3 (Ridge) for the Energy Sector

Step	No.Selected Variables	%dev	AIC	BIC
HVS Step 2	31	0.51	345.47	480.48
HVS Step 3	31	0.50	349.95	482.87

This final step retains the same set of variables selected in Step 2, but with weights modified through Ridge regression for enhanced robustness. Table 2 shows that for the Energy sector, the %dev[2] value obtained in Step 3 is nearly identical to the value obtained in Step 2. In fact all statistics are very similar. However, this final step is crucial as it creates a more robust methodology as we will show in the next section. As an added bonus this step allows us to assess the relative importance of selected variables for log volatility of annual returns, thus highlighting which ESG risks are more relevant. We provide the definitions of all variables identified as important for the Energy sector in Appendix D.

In Figure 4 we visualize the importance of the 15 ESG categories of the Energy sector. Let the final model from Step 3 be:

$$\text{Logarithmic Volatility} = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \epsilon \tag{4}$$

where k is the total number of selected variables, and the predictor variables (x_i) have been standardized. The coefficient β_i quantifies the impact of variable i on log volatility and since the variables are standardized they have comparable magnitudes. Each of the k selected variables belongs to one of the 15 specific ESG categories $(C_1, C_2, \ldots, C_{15})$. To determine the importance of each category we sum the absolute values of the β_i's of variables from that category. Specifically, if I_j denotes the set of indices for variables in category

[2]The Ridge regression cannot calculate an R^2 value. Instead, it calculates the Percent Deviance Explained %dev (Hoerl and Kennard, 1970) to refer to the proportion of deviance explained by the Ridge regression. In the case of multivariate regression %dev is exactly equal to the R^2 value.

C_j we have:

$$\text{Score}_{C_j} = \sum_{i \in I_j} |\beta_i| \tag{5}$$

For example, if variables x_1, x_5, and x_{12} belong to the Community category, its total influence score is calculated as $|\beta_1| + |\beta_5| + |\beta_{12}|$. This is done for all 15 categories. Finally, the importance of each category C_j is calculated as its relative score expressed as a percentage:

$$\% \text{ Importance}_{C_j} = \frac{\text{Score}_{C_j}}{\sum_{n=1}^{15} \text{Score}_{C_n}} \times 100\% = \frac{\sum_{i \in I_j} |\beta_i|}{\sum_{i=1}^{k} |\beta_i|} \times 100\% \tag{6}$$

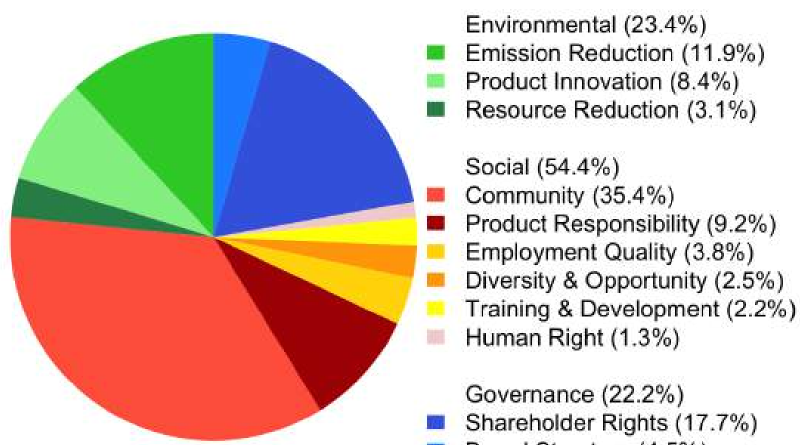
Figure 4 displays a pie chart of importance weights for each category in the Energy sector. The colors are chosen such that all environment (E) categories are shades of green, social (S) categories are shades of red, and governance (G) are shades of blue. Each category is assigned a fixed color. The pie chart only shows 11 categories as the other 4 did not contain any significant ESG raw variable.

Naively, one may expect environmental variables to dominate the other categories for companies in the Energy sector. However, we can clearly see that variables related to social categories account for the most significant part of ESG-related risk. This seemingly counterintuitive result will be addressed later in Section 5.3 when we study each sector factored by the size of the company.

4. HVS Algorithm validation

In this section we illustrate the robustness of the HVS methodology. As mentioned in the discussion of Step 3, the nature of ESG data (few observations, large number of variables) makes overfitting easy. To our knowledge, no prior work on ESG data has tested model performance on out-of-sample data. Additionally, since HVS selects ESG variables related to changes in logarithmic volatility, we also compare it with other factor selection methods. As in Section 3, we illustrate our results using the Energy sector. The analysis, conducted across all sectors, yields consistent findings. Comprehensive results for all other sectors are provided in Appendix C.

Figure 4: This pie chart illustrates the sum of absolute value of coefficients for selected raw variables within each category resulted from Step 3 for the Energy sector companies.



4.1. In-sample and out-of-sample analysis

In order to assess the robustness of the methodology we introduce two out-of-sample validation designs. The first design evaluates temporal robustness by testing whether the algorithm can learn a model from the historical data and use it to predict future logarithmic volatility of equity returns. In this design, we train the HVS model using data from a five-year historical window to predict the logarithmic volatility of corporate returns for the next year. The five-year window is rolled forward annually through the dataset. In Golbayani et al. (2020) the authors show that this temporal data split is much more relevant for financial data than the traditional cross-validation method.

The second validation design assesses cross-sectional robustness using leave-one-out-company validation within each sector. We iteratively train the HVS model on data from all companies within the Energy sector except one, and then use the trained model to predict the logarithmic volatility for the held-out company. This process is repeated for each company in the sector. For both the temporal and cross-sectional validation scenarios, model prediction accuracy is evaluated using Mean Squared Error (MSE).

To illustrate the results, we compare the performance of the full HVS algorithm with the intermediate model that terminates after Step 2 (i.e., without the final Ridge regression step). We also include a benchmark model that does not include any ESG factors, where the forecast is simply the average value of the in-sample log volatilities. Specifically, when forecasting the next year log volatility we use the past five-year average log volatility for the respective company. Table 3 presents the MSE values for both designs.

Table 3: Comparison of HVS with Benchmark Variable Selection Methods for the Energy Sector

Design	Model	IS MSE	OOS MSE
Cross-sectional	Mean of Response	0.1871	0.2438
	HVS Step 2	0.1134	0.2122
	HVS Step 3	0.1290	0.1839
Temporal	Mean of Response	0.1871	0.2438
	HVS Step 2	0.0931	0.2413
	HVS Step 3	0.1130	0.2132

We can see from Table 3 that the full HVS algorithm achieves lower out-of-sample MSE values than the intermediate model that terminates after Step 2. To test whether the improvement is statistically significant we perform matched pairs tests using every combination of models. The p-values of these tests are presented in Table 4. These results provide statistical evidence that applying the Ridge regression in Step 3 of the HVS algorithm is important for the robustness of the algorithm when assessing out-of-sample data.

4.2. HVS versus traditional variable selection methods

To assess whether the multistep variable selection procedure of the HVS algorithm is justified, we benchmark its performance against four classical methods.

1. We use a Principal Component Analysis (PCA) (PCA_1) and we select components that explain 80% of the total variance.
2. We use a PCA which is constrained to use the same number of components as the number of variables selected by HVS (PCA_2).

Table 4: Statistical Test For Comparison between HVS step 2 and HVS step 3

Design	Comparison of MSE	IS P-Value	OOS P-Value
Cross-sectional	HVS Step 3 $<$ Mean of Response	0.0000	0.0000
	HVS Step 3 $<$ HVS Step 2	1.0000	0.0060
	HVS Step 2 $<$ Mean of Response	0.0000	0.0933
Temporal	HVS Step 3 $<$ Mean of Response	0.0000	0.0094
	HVS Step 3 $<$ HVS Step 2	1.0000	0.0332
	HVS Step 2 $<$ Mean of Response	0.0000	0.4594

3. The “Stepwise” benchmark uses Stepwise selection but instead of using the Hierarchical approach in HVS it selects all relevant variables in one step.
4. The “Lasso” benchmark uses all raw variables and imposes a LASSO type constraint.

Table 5: Comparison of HVS Performance with Benchmark Methods.

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA_1	81	0.40	450.66	790.24
PCA_2	31	0.30	471.84	610.95
Stepwise	60	0.56	288.11	537.68
Lasso	72	0.46	388.95	687.61
HVS	31	0.50	349.95	482.87

Table 5 presents model statistics for each of the selection methods. HVS selects a small number of relevant variables with a comparably high %dev value. Based on %dev and AIC, the Stepwise model has a better in sample model at the cost of doubling the number of selected variables. This typically means multicollinearity and overfitting.

This is evident in Table 6 where we present the MSE values for HVS and the Stepwise method for out-of sample data. While Stepwise regression shows a better fit to the in-sample data, its MSE values increase significantly for out-of-sample data.

Table 6: In-Sample and Out-of-Sample MSE for HVS and Stepwise Regression.

Design	Model	IS MSE	OOS MSE
Cross-sectional	Stepwise	0.0802	0.3594
	HVS	0.1296	0.1833
Temporal	Stepwise	0.0571	1.5900
	HVS	0.1235	0.2124

5. Case study

This section illustrates the applicability of the proposed methodology using two case studies. First, we examine which ESG risk factors are important for different industry sectors. We also explore whether these identified risk variables provide additional information beyond more traditional financial factors – the Fama-French factors. We analyze nine sectors as defined by the Refinitiv Business Classification (TRBC): Energy, Finance, Healthcare, Utilities, Food, Consumer, Mining, Industry, Real Estate. The remaining 4 sectors were excluded from our analysis due to the limited number of available observations in the database.

Second, we examine how relevant ESG risk variables differ by company size. For this analysis, we define large-cap companies as constituents of the S&P 500 and small-cap companies as constituents of the Russell 2000.[3] Notably, existing literature analyzes the impact of ESG factors for large capitalization companies only, for example those corporations included in S&P 500, China A-listed companies, and the STOXX Europe 600 (Alareeni and Hamdan, 2020; Deng and Cheng, 2019).

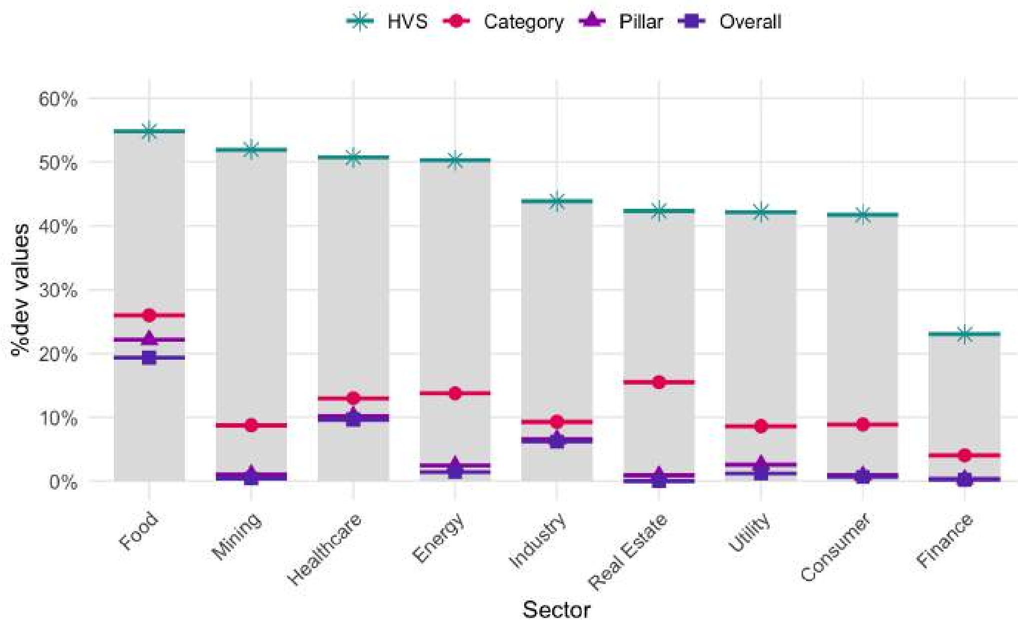
5.1. Sector analysis

Different industries operate under unique business models, face distinct regulatory environments, and are subject to varying stakeholder pressures. This inherent heterogeneity suggests that the most important ESG factors for financial risk as measured by log volatility are unlikely to be uniform across the entire market.

[3] We choose to use S&P 500 instead of Russell 1000 to avoid some mid-cap companies and thus illustrate the difference in factors better.

In Figure 5 we present the %dev values for HVS (dark cyan) and %dev values obtained from Ridge regressions of logarithmic volatility with LSEG's aggregated scores. This plot shows that HVS performance is universally better than using any of the aggregated scores. Among these sectors, the finance sector exhibits the worse performance. We will revisit these findings in Section 5.3.

Figure 5: A comparison of the %dev values for HVS with corresponding %dev values using aggregated LSEG scores.



In Figure 6 we visualize the relative importance of each pillar (E, S, and G) as well as categories for each sector. The pie charts illustrate this breakdown, with shades of green, red, and blue corresponding to the Environmental, Social, and Governance pillars, respectively. Each category's percentage weight is calculated using the coefficients of the final Ridge regression in HVS, as detailed in Section 3.2.

It is clear that the ESG risk variables are significantly different for each sector. The HVS algorithm allows us to provide a more in depth examination of important categories within each sector. Across all sectors, the Community category ("Tomato" shade) consistently shows significant relevance to ESG risk as measured by log volatility. Similarly, the Shareholder Rights ("Royal

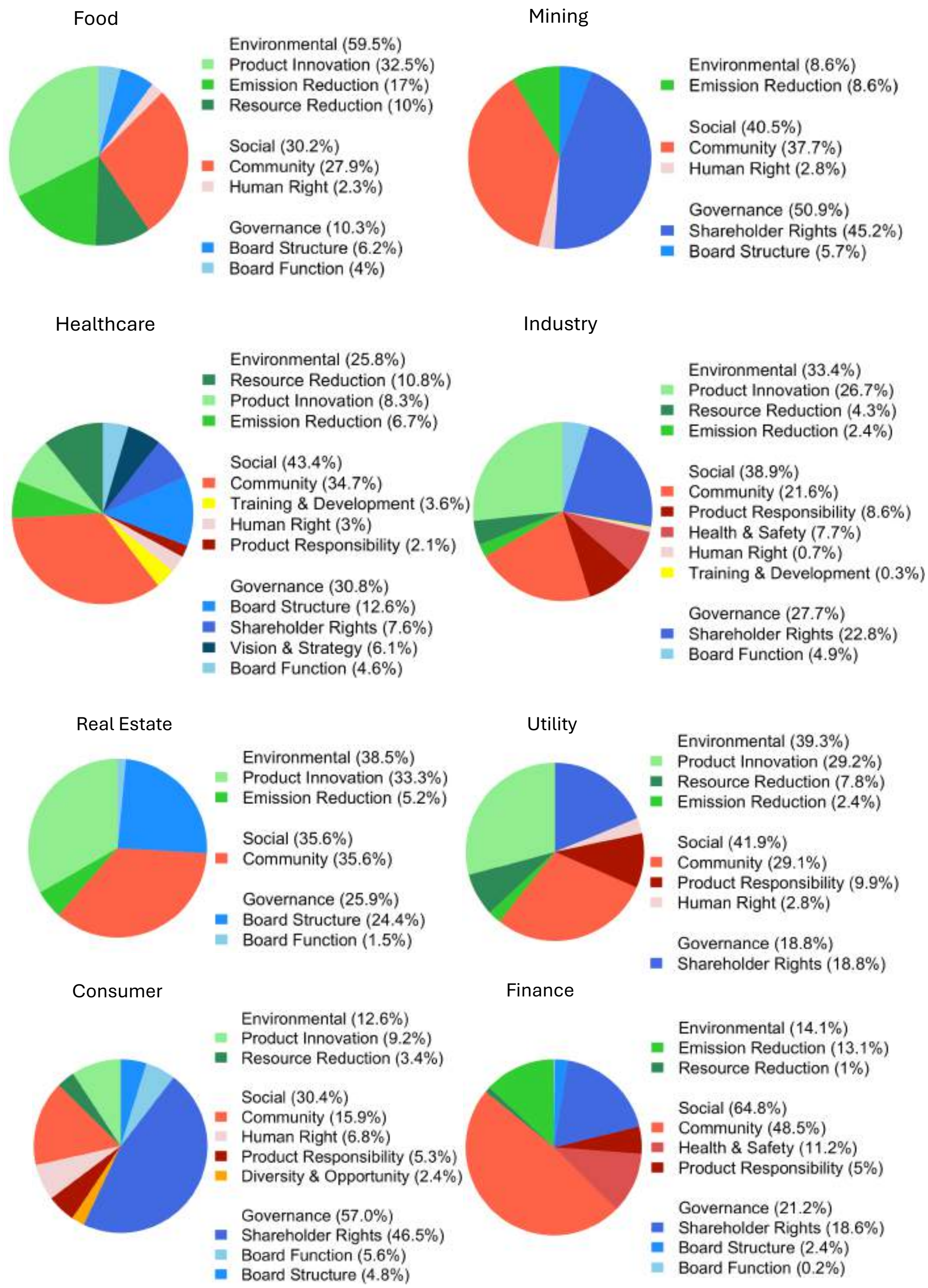


Figure 6: Relevant ESG factors for logarithmic volatility of returns grouped by sector.

blue" shade) is the other category that is prevalent throughout each sector.

It is also worth noting the surprisingly low importance of the environmental variables for the Mining sector. When looking at the Energy sector in Figure 4 we also saw that these variables are not dominating ESG factors. We will revise this fact in section 5.3.

ESG Variables as a complement to financial factors. We further investigate whether the selected ESG raw variables can contribute additional information beyond well-established financial factors. Our analysis compares three regression models for each sector, all of which use logarithmic volatility as the response variable. These consist of (1) a baseline yearly Fama-French 3-factor model, (2) a model containing only the HVS-selected ESG variables, and (3) an augmented model that integrates both the financial and ESG factors.

Table 7 presents the model fit statistics for this analysis. Augmenting the baseline model with HVS-selected variables substantially improves its performance across every sector, as shown by a significant increase in the %dev and a corresponding decrease in the BIC. This finding indicates that the ESG factors identified by HVS provide additional information for explaining the volatility of returns.

Table 7: Assessing the Incremental Explanatory Power of HVS-Selected ESG Variables Beyond the Fama-French 3-Factors, by Sector. The values represent the %dev from regressions on logarithmic volatility of returns.

	Fama-French 3 factors		Selected ESG Variables		Combined Model	
Sector	%dev	BIC	%dev	BIC	%dev	BIC
Food	0.12	374.74	0.55	233.35	0.68	115.34
Mining	0.13	194.00	0.52	165.15	0.62	129.59
Healthcare	0.11	802.67	0.51	629.65	0.61	481.00
Energy	0.13	561.75	0.50	482.87	0.63	371.34
Industry	0.17	733.48	0.44	646.76	0.58	419.72
Real Estate	0.36	219.08	0.42	241.01	0.72	117.51
Utility	0.32	343.69	0.42	434.81	0.73	106.65
Consumer	0.19	421.33	0.42	418.82	0.60	266.91
Finance	0.37	554.31	0.23	874.57	0.59	278.98

The Fama-French 3-factor model was designed to explain stock returns. For a comprehensive comparison, we repeat the previous analysis using corporate returns as the response variable. The results are available in Appendix B.

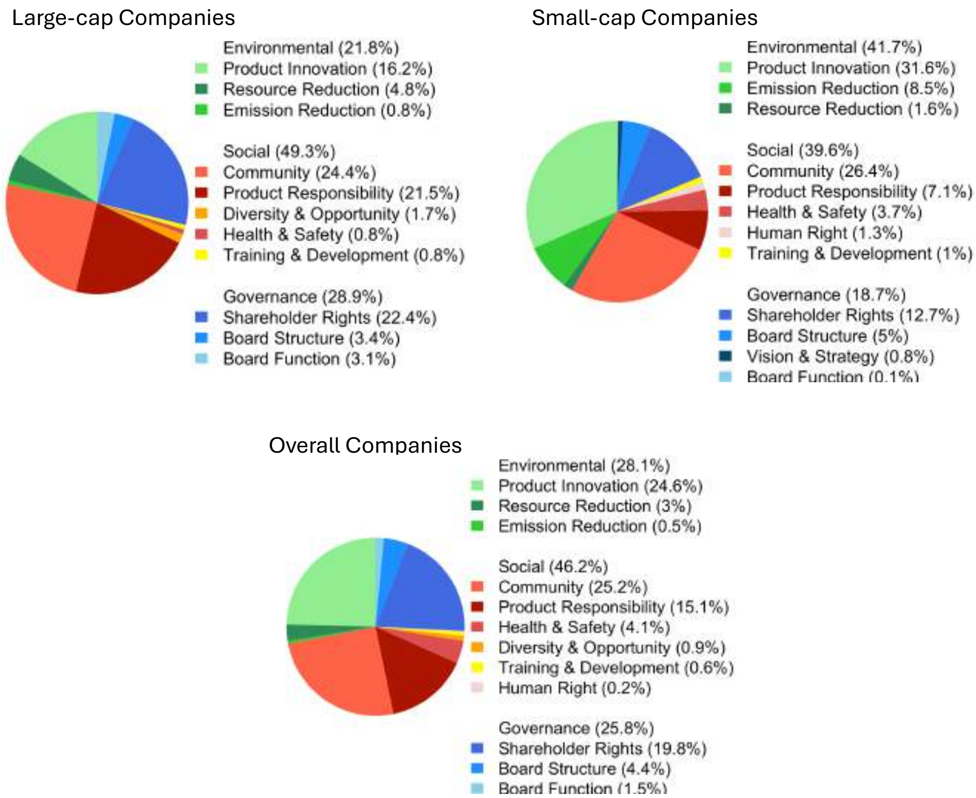
5.2. Size analysis

As far as the authors are aware, all prior ESG studies (Alareeni and Hamdan, 2020; Jain et al., 2019; Shen et al., 2023) focus on large capitalization companies defined as total market value in excess of $10 billion. At the time of writing this manuscript, this includes all S&P 500 companies. In this section we expand these studies by examining the difference in factors for large and small capitalization companies. We consider companies in the S&P 500 index to represent large cap companies and companies in the Russell 2000 index to represent small cap companies. Those companies considered to be sized "between" these defined large-cap and small-cap groups are generally referred to as mid-cap. We chose to avoid mid cap companies to see if there is a difference in ESG variables based on market cap and thus we focus on the two extremes to highlight potential differences.

We observe that the %dev value is higher for large-cap companies at 30% compared to 26% for small-cap companies. Next, in Figure 7 we present the important categories and pillars obtained for large versus small capitalization companies across all sectors.[4] The resulting pie charts reveal significant differences between the two groups. While the Social variables are similar, Environmental variables are much more important for small cap companies while Governance is more significant for large caps. Additionally, we may attempt to explain the observed large differences in percent importance for some categories. Product innovation is significantly more important for small cap companies' logarithmic volatility. Small cap companies could seek novel products to gain a competitive advantage. In contrast, Product Responsibility is more important for large caps as they are established corporations and product controversies could lead to large downside risks. Finally, Shareholder Rights is significantly more important for large-cap companies. This may be because the category measures protections for minority shareholders, and in large public companies, many investors fall into this group.

[4]It is important to note that the HVS algorithm, including the variable standardization in Step 3, is applied separately to the large-cap and small-cap datasets. Therefore, our analysis does not compare the absolute magnitude of risk between the two groups. Instead, we compares the relative importance of the ESG categories within each group

Figure 7: Important ESG risk variables for large-cap and small-cap companies across all sectors



5.3. Sector by size analysis

Next, we repeat the analysis for large and small companies but within each sector. Table 8 shows that in every sector analyzed, the %dev value for the large-cap group is substantially higher than for the small-cap group. This suggests that the ESG risk factors are more significantly related to the volatility of large-cap firms than to that of small-cap firms. The Finance sector provides the most stark illustration of the difference between large- and small-cap firms. Interestingly, the table also reveals a different phenomenon in sectors like Food and Healthcare, where the %dev for the Overall model, which combines both size groups, is higher than the %dev for either the large- or small-cap groups individually. This suggests that this split by market capitalization may not be the optimal way to segment companies for ESG risk analysis. Investigating alternative grouping methodologies that could capture these interaction effects and potentially yield even more powerful models presents a promising direction for future research.

Figures 8 to 10 present a sector-by-sector comparison of the relative importance of ESG categories for large- and small-cap companies. The importance of variables is very different for small versus large. First, for large-cap companies within the Energy sector, environmental variables are the overwhelmingly dominant risk drivers, with factors such as ‘Resource Reduction’, ‘Emission Reduction’, and ‘Product Innovation’ being particularly significant. These variables perhaps appear because the large companies for the Energy sector often face more intense environmental scrutiny from regulators, investors, and the public. Thus, they would generally face greater exposure to regulatory penalties, environmental litigation, and reputational risk associated with environmental performance. Small Energy companies, in contrast, tend to operate more locally or within sub-segments of the Energy industry, where social factors such as labor relations, community impact, and employee safety take precedence.

The pie charts also help identify the differing sources of ESG-related risk between large-cap and small-cap companies in other sectors. In the Healthcare sector, for example, social variables are dominant risk drivers for Large-cap companies. This makes sense as this sector focuses on patient outcomes, access to care, and product safety. However, a more unexpected pattern emerges within Small-cap companies of the Healthcare sector, where Environmental factors surprisingly are much more important. For large-cap companies in the Food, Real Estate, and Utility industries, Environmental variables are identified as relevant risk factors. In contrast, for small-cap companies within these same sectors, the algorithm found no Environmental variables to be significantly related to risk.

The Real Estate sector presents an interesting finding: the scope of the companies analyzed dramatically changes the selected risk variables. When all companies are analyzed together, Environmental variables are the most significant to ESG risk. However, when the dataset is split into large- and small-cap subgroups, Environmental variables are not selected for either group. This may be because the significant risk variables for the combined dataset are universal but masked by more dominant local variables within the narrowly defined subgroups. In the smaller, more homogeneous large- and small-cap groups, ESG variables, such as 'Shareholder Rights' for large-caps or 'Community' for small-caps, are overwhelmingly dominant. The HVS algorithm correctly identifies these strong, group-specific signals and prioritizes them. In the larger, more heterogeneous combined dataset, these powerful but group-specific effects may be less dominant overall. This allows the 'Product Innovation' variable, which has a relatively strong impact across both groups, to emerge as the significant risk factor for the overall companies.

Table 8: Comparison of %dev Values of HVS algorithm for Large- and Small-Cap Companies by Sector

	Large-cap	Small-cap	Overall
Food	0.50	0.38	0.55
Mining	0.55	0.33	0.51
Healthcare	0.37	0.32	0.50
Energy	0.48	0.39	0.50
Industry	0.42	0.39	0.43
Real Estate	0.47	0.37	0.42
Utility	0.36	0.24	0.42
Consumer	0.49	0.29	0.42
Finance	0.52	0.06	0.23

6. Conclusions

In this study we provide evidence that ESG variables are related to financial risk, as measured by logarithmic volatility. We developed the Hierarchical Variable Selection (HVS) algorithm specifically to understand ESG datasets. HVS achieves higher explanatory power than using aggregated ESG scores. Compared with traditional selection models, our algorithm achieves superior performance in modeling risk while identifying a more parsimonious set of variables.

Applying HVS to different sectors, we find that ESG risk variables differ across industries, highlighting the importance of conducting sector-specific analysis. Furthermore, when analyzing firms split into Large and Small, our results reveal that the relevant ESG risk factors differ between Large- and Small-cap companies, reflecting variations in public expectations, regulatory exposure, and strategic positioning.

Our results also illustrate the wide applicability of the proposed algorithm. Across multiple sectors and firm sizes, it consistently identifies risk-relevant ESG variables with high explanatory power. This robustness across

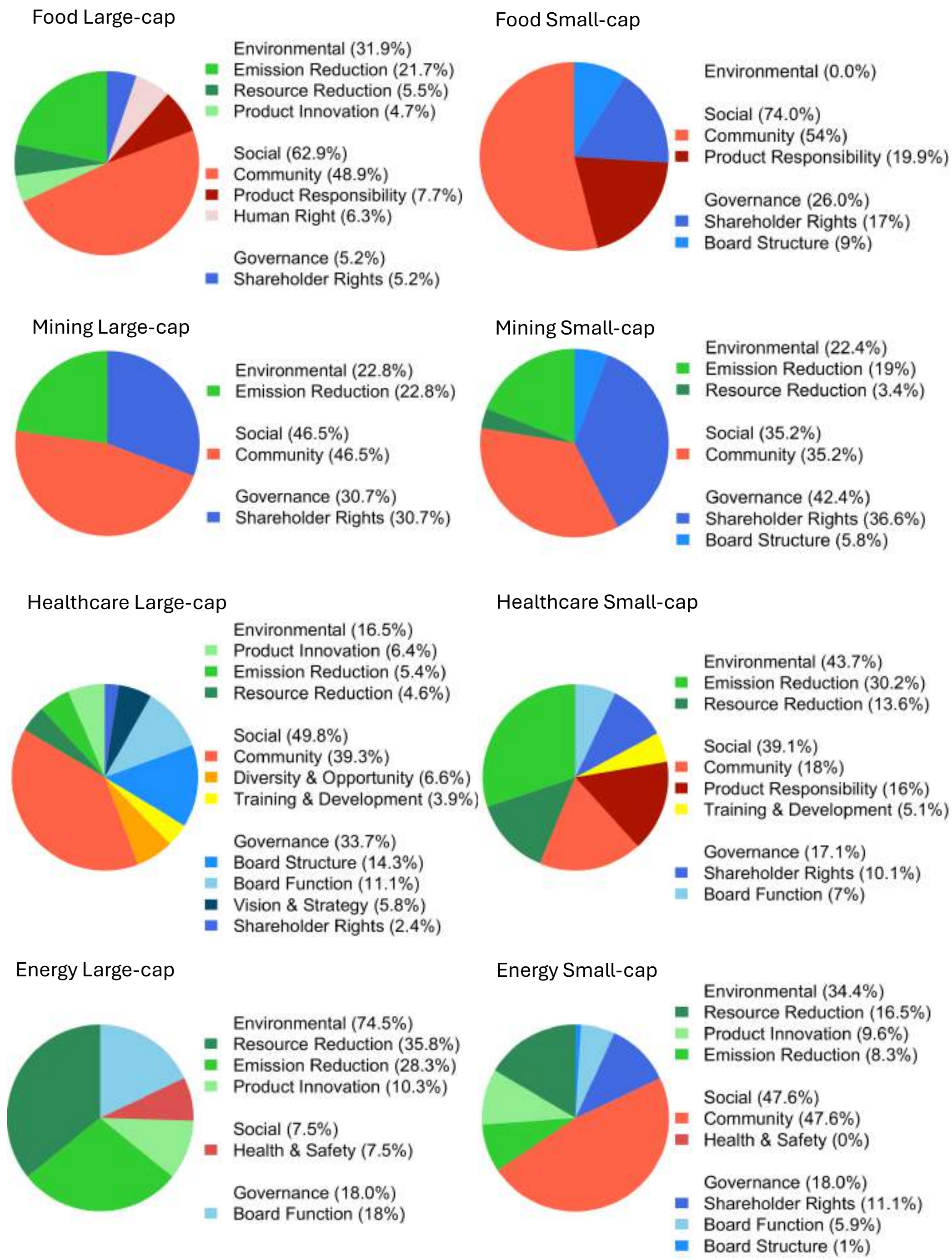


Figure 8: ESG variables within each sector for large and small cap companies.

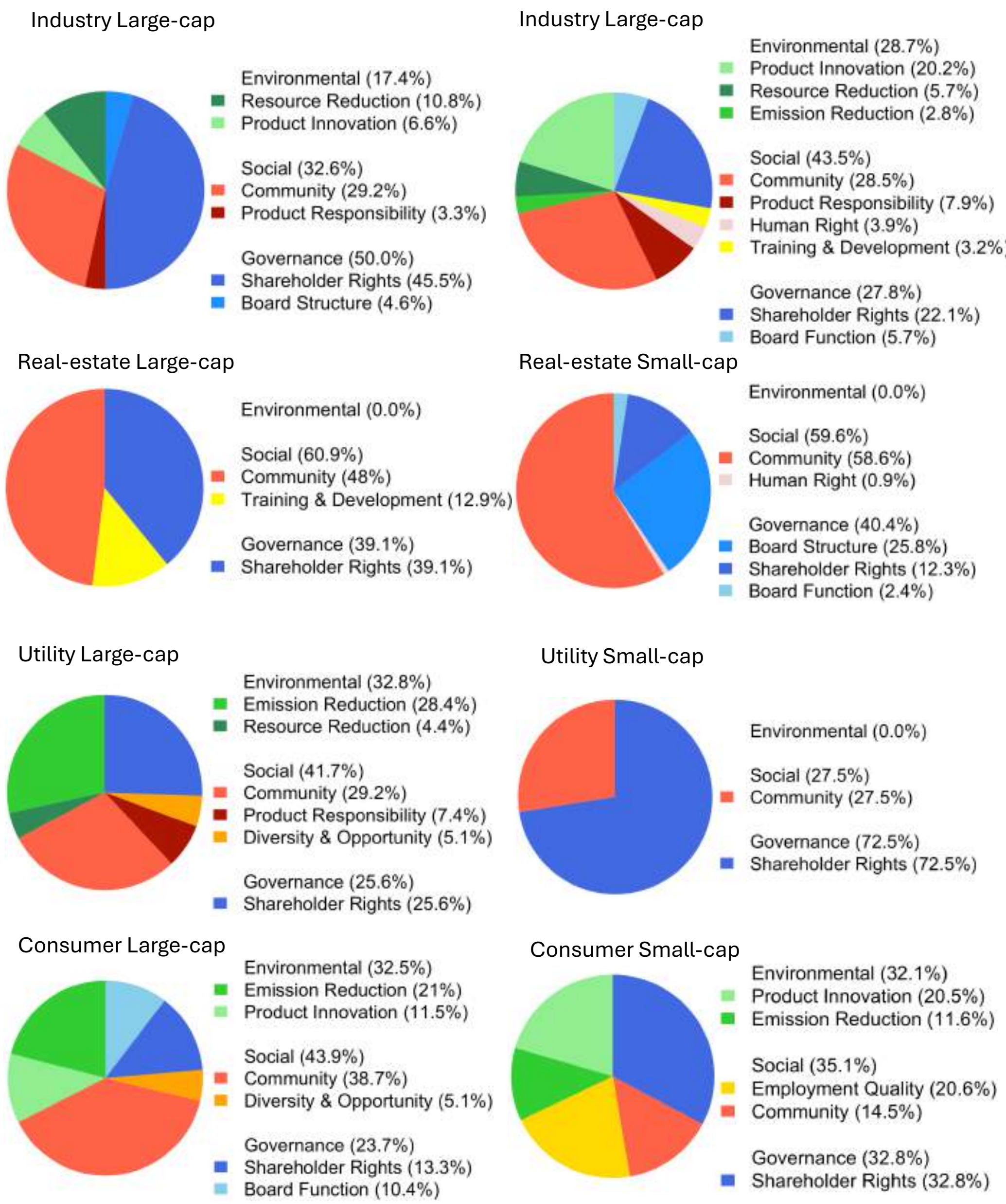
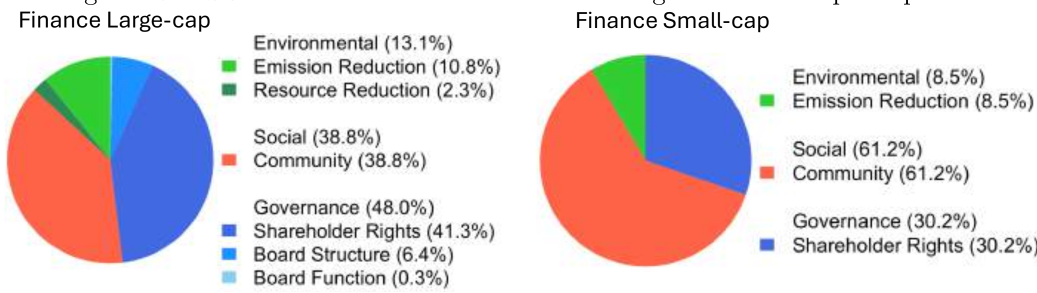


Figure 9: ESG variables within each sector for large and small cap companies.

Figure 10: ESG variables within each sector for large and small cap companies.



different market segments indicates that the algorithm may be effectively extended to other hierarchically structured datasets. In particular, HVS is well-suited for applications involving datasets that possess a hierarchical (tree-like) structure and are characterized by having a large number of variables with relatively limited observations.

Appendix A. Data Processing

First, we exclude demographic variables, such as CEO names, which fall outside the scope of this study. We then handle missing boolean variables and numeric variables. Following (Clarkson et al., 2008; Cheng et al., 2014), we assign 0 to missing values for all boolean variables. Clarkson et al. (2008) examine boolean variables that indicate whether companies issue policy, regulation, or measures related to ESG. These are coded "1" if present, "0" otherwise in our dataset. The authors infer that if companies issue ESG related policies, they are likely to report their efforts to the public, thus the missing values can be reasonably assessed as no such regulations are issued. Cheng et al. (2014) examine boolean variables that indicate whether companies are involved in events or practices considered adverse from an ESG perspective (e.g., weapons manufacturing, alcohol, etc.). They believe that if companies are involved in adverse practices, the public media is likely to report it, so the missing values can be considered as they are not involved. We carefully examined each variable description in the LSEG dataset, and we found that all boolean variables fit one of these two cases.

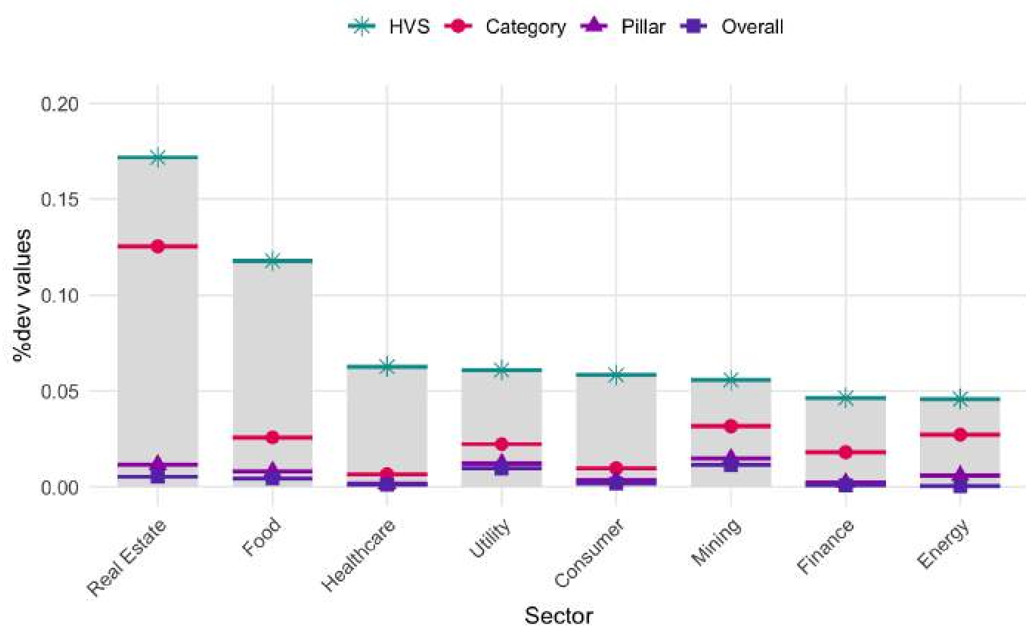
For numeric variables, we first consider "controversy variables" that quantify regulatory violations. Similar to our treatment of adverse event boolean variables, we assume missing controversy data implies no significant incidents occurred, and thus impute these missing values as "0". For the remaining numeric variables, we retain only those variables with at least 80% data availability within each sector across the entire analysis period. This step is performed separately with respect to each sector as we find data availability across industries is different.

Finally, any observation (i.e., company-year data point) still containing missing values after these steps is removed from the dataset. Taking the Energy sector as an example, the initial dataset had 617 variables and 695 company/year observations. After applying the data processing steps, the final dataset for the Energy sector contains 255 variables and 422 observations.

Appendix B. Applying HVS on Returns

When applying our algorithm to model returns, we find that the resulting %dev values are significantly lower compared to those obtained when modeling risk. The HVS algorithm identifies significantly fewer ESG variables as relevant for explaining returns than for risk. For example, in the Energy sector, it selects only six variables for returns compared to 31 for risk. These six variables are: Global Compact, Human Rights Processes in Policy Forced Labor, Sdg17PartnershipsToAchieveTheGoal, Veto Power or Golden share, Sdg13ClimateAction, Human Rights Process in Policy Freedom of Association.

Figure B.11: This bar chart illustrates the %dev of using selected raw variables within each category to regress returns.



In section 5, we find that the selected ESG raw variables can contribute additional information beyond Fama-French factors for logarithmic volatility of returns. We repeat the previous analysis using corporate returns as the response variable. Table B.9 shows that, as expected, the Fama-French factors generally provide a better fit for returns than they did for risk (as measured by logarithmic volatility). When we add the HVS-selected ESG variables, the explanatory power increases marginally. However, this slight improvement may be due to the introduction of new model parameters. This is indicated by the Bayesian Information Criterion (BIC) in Table B.9, as the value for the combined model is higher than for the Fama-French model alone. These findings illustrate that ESG factors are more related to risk rather than return, which confirms findings from Pérez et al. (2022); Engelhardt et al. (2021); Deng and Cheng (2019); Fauver et al. (2018).

Table B.9: Assessing the incremental Explanatory Power of HVS-Selected ESG Variables Beyond the Fama-French 3-Factors, by Sector. The values represent the %dev from regressions on return.

	Fama-French 3 Factors		Selected ESG Variables		Combined Model	
Sector	%dev	BIC	%dev	BIC	%dev	BIC
Food	0.15	82.92	0.12	137.04	0.21	109.55
Mining	0.21	315.15	0.06	346.21	0.22	322.22
Healthcare	0.16	732.12	0.06	832.45	0.19	743.89
Energy	0.22	456.43	0.05	565.24	0.27	467.21
Industry	0.27	412.50	–	–	0.27	412.50
Real Estate	0.32	40.27	0.17	86.85	0.50	8.56
Utility	0.24	-148.35	0.06	-43.52	0.25	-130.88
Consumer	0.22	512.32	0.06	604.81	0.24	527.56
Finance	0.43	-245.69	0.05	301.54	0.46	-242.00

Appendix C. Validation For Other Sectors

Table C.10: Comparison of HVS Performance with Benchmark Methods for Each Sector

Food

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	53	0.53	217.01	435.29
PCA	23	0.41	240.59	339.81
Stepwise	62	0.56	132.81	382.84
Lasso	52	0.51	224.70	437.04
HVS	23	0.55	136.10	233.35

Mining

Model	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	33	0.43	153.18	270.32
PCA	21	0.28	177.69	254.68
Stepwise	32	0.22	180.26	290.71
Lasso	39	0.50	134.90	274.14
HVS	21	0.52	89.52	165.15

Healthcare

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	66	0.41	620.36	930.22
PCA	40	0.37	621.00	812.39
Stepwise	67	0.55	351.60	661.47
Lasso	39	0.41	559.86	744.13
HVS	40	0.51	440.83	629.65

Industry

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	75	0.44	541.80	905.44
PCA	38	0.30	649.47	838.38
Stepwise	75	0.52	321.83	680.75
Lasso	65	0.43	520.81	834.51
HVS	38	0.44	460.58	646.76

Table C.11: Comparison of HVS Performance with Benchmark Methods for Each Sector

Real Estate

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	30	0.25	291.31	395.05
PCA	11	0.19	268.36	310.51
Stepwise	34	0.58	152.56	266.02
Lasso	2	0.07	271.95	283.68
HVS	11	0.42	200.11	241.01

Utility

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	56	0.37	400.61	641.1
PCA	30	0.27	411.57	544.26
Stepwise	56	0.42	300.24	536.58
Lasso	17	0.24	405.76	482.40
HVS	30	0.42	304.28	434.81

Consumer

Model	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	58	0.35	414.85	662.46
PCA	27	0.22	436.24	555.92
Stepwise	55	0.49	240.67	471.77
Lasso	37	0.35	369.37	528.19
HVS	27	0.42	301.27	418.82

Finance

Benchmark	**No.Selected Variables**	**%dev**	**AIC**	**BIC**
PCA	51	0.18	901.22	1160.8
PCA	20	0.06	972.38	1080.1
Stepwise	50	0.28	717.19	957.18
Lasso	4	0.11	881.80	908.29
HVS	20	0.23	769.72	874.57

Appendix D. Definition of Risk Variables

Table D.12: Definition of Energy Risk Variables

Variable Name	Definition
PolicyCyberSecurity	Does the company have policies on protecting personal information, ensuring secure data handling practices, and mitigating information security risks?
Total Senior Executives Compensation	The total compensation paid to all senior executives (if total aggregate is reported by the company).
Board Size	The total number of board members at the end of the fiscal year.
Compensation Policy Elements/Policy Executive Retention	Does the company have a compensation policy to attract and retain executives?
Management Training)	Does the company claim to provide regular staff and business management training for its managers?
Staff Transport Impact Reduction Initiatives	Does the company report on initiatives to reduce the environmental impact of transportation used for its staff?
Armaments 5% Revenues	Are revenues generated from armaments larger than 5% of the total net revenues?
Estimated CO2 Equivalents Emission Total	The estimated total CO2 and CO2 equivalents emission in tonnes.
External Consultants	Does the board or board committees have the authority to hire external advisers or consultants without management's approval?
Limited Shareholder Rights to Call Meetings	Has the company limited the rights of shareholders to call special meetings?
Balanced Board Structure Policy Elements/Policy Board Size	Does the company have a policy regarding the size of its board?
CEO Compensation Link to Total Shareholder Return	Is the CEO's compensation linked to total shareholder return (TSR)?
Written Consent Requirements	Does the company permit actions to be taken without meeting by written consent?

Continued on next page

Table D.12 – *Continued from previous page*

Variable Name	**Definition**
e-Waste Reduction Initiatives	Does the company report on initiatives to recycle, reduce, reuse, substitute, treat or phase out e-waste?
Environmental Products	Does the company report on at least one product line or service that is designed to have positive effect on the environment or which is environmentally labeled and marketed?
Critical Countries Controversies	Number of controversies published in the media linked to activities in critical, undemocratic countries that do not respect fundamental human rights principles.
OilAndGasProducerFlag	Indicator used to identify companies involved in oil and gas production
CSR Sustainability External Audit	Does the company monitor its integrated strategy through belonging to a specific sustainability index? AND Does the company monitor its integrated strategy through conducting external audits on its reporting?
Renewable Energy Use	Does the company make use of renewable Energy?
Shareholder Rights Policy Elements/Policy Shareholder Engagement	Does the company have a policy to facilitate shareholder engagement, resolutions or proposals?
Management Departures	Has an important executive management team member or a key team member announced a voluntary departure (other than for retirement) or has been ousted?
Corporate Responsibility Awards	Has the company received an award for its social, ethical, community, or environmental activities or performance?
Fundamental Human Rights ILO or UN	Does the company have a policy to guarantee the freedom of association universally applied independent of local laws? AND Does the company have a policy for the exclusion of child, forced or compulsory labour?
Whistleblower Protection	Does the company describe the implementation of its community policy through a public commitment from a senior management or board member? AND Does the company describe the implementation of its community policy through the processes in place?
Unlimited Authorized Capital or Blank Check	Does the company have unlimited authorized capital or a blank check?

Continued on next page

Table D.12 – *Continued from previous page*

Variable Name	Definition
Elimination of Cumulative Voting Rights	Has the company reduced or eliminated cumulative voting in regard to the election of board members?
OECD Guidelines for Multinational Enterprises	Does the company have a policy to improve customer satisfaction or to strive to be a fair competitor?
Confidential Voting Policy	Does the company have a confidential voting policy (i.e., management cannot view the results of shareholder votes)?
Diversity and Opportunity Processes/Policy Diversity and Opportunity	Does the company have a policy to drive diversity and equal opportunity?
Poison Pill	Does the company have a poison pill (shareholder rights plan, macaroni defense, etc.)?
Agrochemical Products	Does the company produce or distribute agrochemicals like pesticides, fungicides or herbicides?

Declaration Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interests

The authors declares that he has no competing interests.

References

Alareeni, B.A., Hamdan, A., 2020. Esg impact on performance of us s&p 500-listed firms. Corporate Governance: The International Journal of Business in Society 20, 1409–1428.

Alliance, G.S.I., 2022. Global sustainable investment review 2022. URL: https://www.gsi-alliance.org/members-resources/gsir2022/.

Alsaadi, A., Ebrahim, M.S., Jaafar, A., 2017. Corporate social responsibility, shariah-compliance, and earnings quality. Journal of financial services research 51, 169–194.

Anne Tucker, Dana Brakman Reiser, Y.X., 2024. Is 2024 past peak esg? URL: https://corpgov.law.harvard.edu/2024/09/16/is-2024-past-peak-esg/#:~:text=After%20years%20of%20ESG%20(and,%249%20billion%20outflow%20in%202023.

Behl, A., Kumari, P.R., Makhija, H., Sharma, D., 2022. Exploring the relationship of esg score and firm value using cross-lagged panel analyses: Case of the indian energy sector. Annals of Operations Research 313, 231–256.

Bonacorsi, L., Cerasi, V., Galfrascoli, P., Manera, M., 2024. Esg factors and firms' credit risk. Journal of Climate Finance 6, 100032.

Cheng, B., Ioannou, I., Serafeim, G., 2014. Corporate social responsibility and access to finance. Strategic management journal 35, 1–23.

Clarkson, P.M., Li, Y., Richardson, G.D., Vasvari, F.P., 2008. Revisiting the relation between environmental performance and environmental disclosure: An empirical analysis. Accounting, organizations and society 33, 303–327.

Deng, X., Cheng, X., 2019. Can esg indices improve the enterprises' stock market performance?—an empirical study from china. Sustainability 11, 4765.

Eikon, R., 2022. Environmental, social and governance scores from refinitiv. London: Refinitiv Eikon .

Engelhardt, N., Ekkenga, J., Posch, P., 2021. Esg ratings and stock performance during the covid-19 crisis. Sustainability 13, 7133.

Fauver, L., McDonald, M.B., Taboada, A.G., 2018. Does it pay to treat employees well? international evidence on the value of employee-friendly culture. Journal of Corporate Finance 50, 84–108.

Foley, A.M., Heffron, R.J., Al Kez, D., Del Rio, D.D.F., Mcinerney, C., Welfle, A., 2024. Restoring trust in esg investing through the adoption of just transition ethics. Renewable and Sustainable Energy Reviews 199, 114557.

GawÄ, A., et al., 2022. Esg rating and market valuation of the firm: Sector approach. European Journal of Sustainable Development 11, 91–91.

Gibson Brandon, R., Krueger, P., Schmidt, P.S., 2021. Esg rating disagreement and stock returns. Financial Analysts Journal 77, 104–127.

Giese, G., Lee, L.E., Melas, D., Nagy, Z., Nishikawa, L., 2019. Foundations of esg investing: How esg affects equity valuation, risk, and performance. The Journal of Portfolio Management 45, 69–83.

Golbayani, P., Wang, D., Florescu, I., 2020. Application of deep neural networks to assess corporate credit rating. arXiv preprint arXiv:2003.02334 .

Halbritter, G., Dorfleitner, G., 2015. The wages of social responsibility—where are they? a critical review of esg investing. Review of Financial Economics 26, 25–35.

Hoerl, A.E., Kennard, R.W., 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12, 55–67. doi:10.1080/00401706.1970.10488634.

Jain, M., Sharma, G.D., Srivastava, M., 2019. Can sustainable investment yield better financial returns: A comparative study of esg indices and msci indices. Risks 7, 15.

Li, T.T., Wang, K., Sueyoshi, T., Wang, D.D., 2021. Esg: Research progress and future prospects. Sustainability 13, 11663.

Lööf, H., Sahamkhadam, M., Stephan, A., 2022. Is corporate social responsibility investing a free lunch? the relationship between esg, tail risk, and upside potential of stocks before and during the covid-19 crisis. Finance Research Letters 46, 102499.

Madison, N., Schiehll, E., 2021. The effect of financial materiality on esg performance assessment. Sustainability 13, 3652.

Miranda, B., Delgado, C., Branco, M.C., 2023. Board characteristics, social trust and esg performance in the european banking sector. Journal of Risk and Financial Management 16, 244.

Park, M., 2024. An empirical analysis on performance inconsistency among environmental, social and governance components of esg ratings. Asia-Pacific Journal of Business 15, 33–44.

Pérez, L., Hunt, V., Samandari, H., Nuttall, R., Biniek, K., 2022. Does esg really matter—and why. McKinsey Quarterly 60.

Shen, H., Lin, H., Han, W., Wu, H., 2023. Esg in china: A review of practice and research, and future research avenues. China Journal of Accounting Research , 100325.

Verheyden, T., Eccles, R.G., Feiner, A., 2016. Esg for all? the impact of esg screening on return, risk, and diversification. Journal of Applied Corporate Finance 28, 47–55.

Wang, D., Chen, Z., Florescu, I., Wen, B., 2023. A sparsity algorithm for finding optimal counterfactual explanations: Application to corporate credit rating. Research in International Business and Finance 64, 101869.

Yalew, S.G., van Vliet, M.T., Gernaat, D.E., Ludwig, F., Miara, A., Park, C., Byers, E., De Cian, E., Piontek, F., Iyer, G., et al., 2020. Impacts of climate change on energy systems in global and regional scenarios. Nature Energy 5, 794–802.

Zhou, D., Zhou, R., 2021. Esg performance and stock price volatility in public health crisis: evidence from covid-19 pandemic. International Journal of Environmental Research and Public Health 19, 202.