Advanced Programming 2025

ESG Scores And Stock Returns: Explanatory and Predictive Evidence from the S&P 500

Final Project Report

Jeyarajah Niruban
Niruban.Jeyarajah@unil.ch
Student ID: 21422019

December 18, 2025

Abstract

In this paper, the research aims to examine the relationship between the ESG (Environment, Social and Governance) score and stock returns, as well as the extent to which ESG scores contain predictive information for future returns. Using a panel of S&P 500 companies from 2015 to 2023, we combine traditional econometric methods—including cross-sectional OLS regressions and Fama–MacBeth estimations-with machine learning techniques such as linear regression, Lasso and Random Forest models. The econometric results reveal a modest but statistically significant negative association between ESG scores and realized stock returns. The machine learning analysis shows that ESG contributes little to out-of-sample predictive performance, with traditional firm characteristics (size, momentum, profitability, valuation and market beta) dominating forecasting accuracy. Overall, the findings suggest that while ESG scores may capture important aspects of corporate sustainability, they provide limited value for short-term stock return prediction.

Keywords: data science, Python, machine learning, ESG score, stock returns

Contents

1 Introduction

Climate change has become one of the defining challenges of our century. Firms across the world are increasingly expected to operate in a more socially and environmentally responsible manner. In response, investors and regulators rely on ESG (Environmental, Social and Governance) scores to quantify a company's sustainability practices and overall non-financial performance. These scores aim to capture how firms manage environmental impact, social responsibility, and governance quality, and they often influence portfolio allocation, regulatory assessments, and public perception.

However, despite the widespread adoption of ESG ratings, their relationship with financial performance remains controversial. While some studies suggest that strong ESG practices reduce risk or improve long-term stability, others argue that ESG scores add little information beyond traditional firm characteristics. In particular, it is unclear whether ESG provides *incremental predictive power* for future stock returns, beyond well-established financial factors such as size, valuation, profitability, or momentum.

The objective of this study is to investigate the link between ESG scores and stock performance using a panel of the U.S. S&P 500 companies from 2015 to 2023. More specifically, we examine (i) whether ESG scores are associated with realized returns across sectors, and (ii) whether ESG information improves the out-of-sample prediction of next-year returns when combined with standard financial characteristics. To address these questions, we combine classical econometric tools (descriptive analysis, cross-sectional regressions, Fama–MacBeth estimations) with modern machine learning methods (Lasso regression and Random Forest).

Finally, the structure of this report follows with the Section 2 which reviews related work on ESG and return predictability. Section 3 describes the dataset and the methodology used. Section 4 presents empirical and machine learning results, including performance metrics and visualizations. Section 5 discusses the implications and limitations of our findings. Section 6 concludes and outlines directions for future research.

2 Literature Review / Related Work

A growing body of research has examined the relationship between ESG performance and financial returns, yet the evidence remains highly mixed. Several recent studies find that higher ESG scores are associated with lower short-term stock performance. For instance, Escobar-Saldívar et al. (2025) document a negative relationship between ESG ratings and U.S. stock returns using a large cross-sectional sample of publicly traded firms.

This finding is consistent with the global evidence reported by Alves, Krüger and van Dijk (2024), who compare ESG ratings from multiple providers across more than 16,000 stocks in 48 countries. Their results show little to no systematic relationship between ESG ratings and global stock returns, highlighting the difficulty of obtaining consistent predictive signals from ESG metrics.

Other studies suggest that the impact of ESG may depend on the time horizon considered. da Cunha et al. (2025) show that ESG performance tends to exhibit weak or even negative associations with short-term returns, whereas long-term financial performance appears more positively related to sustainability practices. This points to a possible mismatch between the investment horizon of most empirical studies and the inherently long-term nature of ESG initiatives.

In contrast to evidence of weak or negative relationships, some studies find positive ESG effects in specific institutional contexts. Yin et al. (2023), using an unbalanced panel of listed firms in China, report that ESG performance positively influences stock returns for non–state-owned enterprises and for firms located in eastern provinces. The authors argue that state-owned firms face greater political and social obligations, which may weaken the financial relevance of ESG performance.

Overall, the existing literature does not provide a clear consensus. The relationship between ESG and stock performance appears to vary across markets, time horizons, and firm types. This ambiguity motivates the present study, which revisits the question using S&P 500 companiy's data from 2015–2024 and evaluates both explanatory and predictive relationships.

3 Methodology

3.1 Data Description

All firm-level data used in this study were obtained from the Refinitiv Eikon database. The sample consists of companies included in the S&P 500 index, covering the period from January 2015 to December 2023. For each firm-year observation, the dataset contains a wide range of financial and non-financial variables, including market capitalization, beta, ESG Global Score, debt-to-equity ratio, price-to-earnings ratio (P/E), annual total stock return, sector and industry classification, and return on equity (ROE).

In addition to the raw data extracted from Refinitiv, several variables were created to support the empirical analysis. First, a momentum measure was constructed to capture past stock performance, defined as the previous year's return. Second, a dummy variable (*HighImpact*) was created to identify firms operating in environmentally intensive sectors. This classification follows the intuition that ESG considerations may affect these industries differently. Third, an interaction term between the ESG Global Score and the sector dummy was introduced to examine whether the relationship between ESG and returns varies across industries.

3.2 Approach

3.2.1 Descriptive Analysis

The analysis begins with a descriptive exploration of all variables included in the study. This step serves three purposes: (i) to understand the global structure of the dataset, (ii) to identify potential outliers and data inconsistencies, and (iii) to document preliminary patterns between ESG scores and stock returns across sectors. Given the presence of extreme values in financial ratios, winsorization at the 1% and 99% quantiles (sector-specific) was applied to mitigate the influence of outliers.

3.2.2 Cross-Sectional Regression Framework

To estimate the contemporaneous relationship between ESG performance and realized returns, several cross-sectional regression models were implemented.

Model 1: ESG Only. The baseline specification evaluates the simple association between ESG scores and annual returns:

$$R_{i,t} = \alpha + \beta_0 ESG_{i,t} + \varepsilon_{i,t}.$$

Model 2: ESG with Controls. To account for standard firm characteristics widely used in empirical asset pricing, the following multiple regression is estimated:

$$R_{i,t} = \alpha + \beta_0 ESG_{i,t} + \beta_1 Size_{i,t} + \beta_2 PE_{i,t} + \beta_3 Momentum_{i,t} + \beta_4 ROE_{i,t} + \beta_5 DebtEq_{i,t} + \varepsilon_{i,t}.$$

Model 3: Sector Interaction. Because ESG considerations may be more financially relevant in environmentally intensive industries, the third specification includes a high-impact sector dummy and its interaction with ESG:

$$R_{i,t} = \alpha + \beta_0 ESG_{i,t} + \beta_1 HighImpact_{i,t} + \beta_2 (ESG \times HighImpact)_{i,t} + \gamma' Controls_{i,t} + \varepsilon_{i,t}.$$

This model tests whether the ESG–return relationship differs across sectors with greater environmental externalities.

Fama–MacBeth Regressions. Finally, Fama–MacBeth (1973) two-step regressions were implemented to examine whether the ESG effect is stable over time. This method runs a cross-sectional regression for every year, then computes the time-series average of the coefficients and their standard errors, thereby assessing the persistence of ESG effects.

3.2.3 Predictive Analysis: Machine Learning Models

The second part of the methodology evaluates whether ESG scores contain incremental predictive information for future returns. Three models of increasing flexibility were tested using an expanding-window design (2015–2021 train → 2022 test; 2015–2022 train → 2023 test):

- **Linear regression**, serving as a benchmark predictive model.
- **Lasso regression**, which introduces L1 regularization to perform variable selection and identify the most relevant predictors.
- **Random Forest**, a non-linear ensemble method capable of capturing interactions and complex patterns that linear models may miss.

For each model, the target variable is the one-year-ahead stock return, constructed using a forward shift. Only firms for which the next-year return is available were included in the predictive exercise. The performance of each model is evaluated using out-of-sample R^2 and Mean Squared Error (MSE). To ensure a fair comparison, all linear models use standardized predictors, while Random Forest is trained on the raw features.

3.2.4 Summary of the Approach

The combined use of econometric and machine learning methods allows us to examine both the explanatory power of ESG (in-sample associations) and its predictive power (out-of-sample forecasting). Cross-sectional and Fama–MacBeth regressions assess whether ESG systematically correlates with realized returns, while machine learning models test whether ESG adds forecasting value beyond traditional financial characteristics.

3.3 Implementation

All analyses were implemented in Python using a fully reproducible workflow. Data manipulation and preprocessing were performed with `pandas` and `numpy`, econometric estimations with `statsmodels`, and predictive modelling with `scikit-learn`. The entire pipeline follows a time-consistent design to avoid look-ahead bias and ensure valid out-of-sample evaluation.

The raw datasets obtained from Refinitiv were merged using firm identifiers and reporting dates to construct a balanced firm–year panel. After merging, the data were cleaned to remove inconsistencies, and missing values in financial ratios were imputed using sector-level medians to preserve cross-sectional structure. Financial variables exhibiting extreme observations were winsorized at the 1% and 99% quantiles within sectors.

Feature engineering steps included the construction of one-year momentum (lagged return), a high-impact sector dummy, and the interaction term between ESG score and sector classification. The target variable for prediction, next-year return, was constructed through a forward shift within each firm to ensure strict temporal ordering.

For the econometric analysis, cross-sectional regressions were implemented with `statsmodels`, and Fama–MacBeth estimates were computed by running a separate cross-sectional regression for each year and averaging the resulting coefficients over time.

For the predictive analysis, each machine learning model was fit using an expanding-window procedure: the model was trained on all available data up to year $t-1$ and evaluated on year t. Linear models (OLS and Lasso) used standardized predictors, with the scaler fitted exclusively on the training set. Random Forest models were trained on unscaled features. Model performance was quantified using out-of-sample R^2 and Mean Squared Error (MSE).

This implementation ensures that all steps respect the chronological structure of the data, that no information from the test set leaks into the training process, and that both econometric and machine learning models are evaluated under strictly comparable conditions.

```python
splits = [(2015, 2021, 2022), (2015, 2022, 2023)]

for start, end, test_year in splits:
    train = data[(data['Year'] >= start) & (data['Year'] <= end)]
    test  = data[data['Year'] == test_year]

    X_train = train[features].values
    y_train = train['Next_Year_Return_w'].values

    X_test = test[features].values
    y_test = test['Next_Year_Return_w'].values
```

Listing 1: Expanding-window construction for out-of-sample prediction

4 Results

4.1 Descriptive Analysis

The analysis begins with a descriptive exploration of the variables included in the dataset. Correlation results indicate that none of the explanatory variables exhibit strong linear relationships with one another, suggesting limited multicollinearity and justifying their joint inclusion in multivariate regressions.

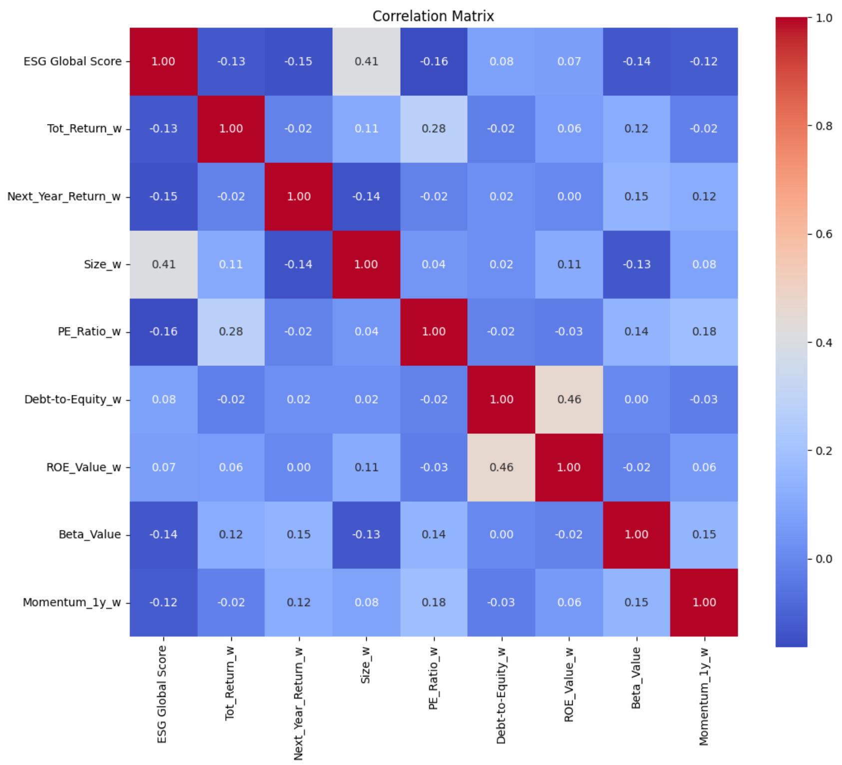

Figure 1: Correlation Heatmap

Summary statistics further confirm the presence of substantial dispersion in financial ratios, particularly the P/E ratio, ROE and the debt-to-equity ratio, which motivated the use of sector-level winsorization at the 1% and 99% quantiles.

Table 1: Summary Statistics

Statistic	ESG	Tot_Return	Next_Year_Return	Size	PE	DebtEq	ROE	Beta
count	4475	4451	3976	4466	4312	4180	4477	4312
mean	61.26	15.59	10.44	28.14	49.45	19.10	49.44	0.97
std	17.12	36.34	37.29	53.47	1287.43	886.56	56.05	0.42
min	4.76	-88.83	-71.40	7.31	0.74	0.00	-2485.00	-0.07
25%	49.57	-4.58	-3.47	10.10	40.83	8.65	0.66	0.66
50%	64.41	13.60	14.90	19.98	74.52	15.39	0.99	0.99
75%	74.29	31.06	32.96	28.17	131.14	26.97	1.22	1.22
max	93.30	743.44	743.44	1791.58	42210	31560	3.68	3.68

Figure 2 presents ESG scores across sectors. ESG performance varies meaningfully between industries: basic materials, consumer defensive, consumer staples, healthcare and real estate display the highest average ESG scores, whereas communication services and financials rank substantially lower.

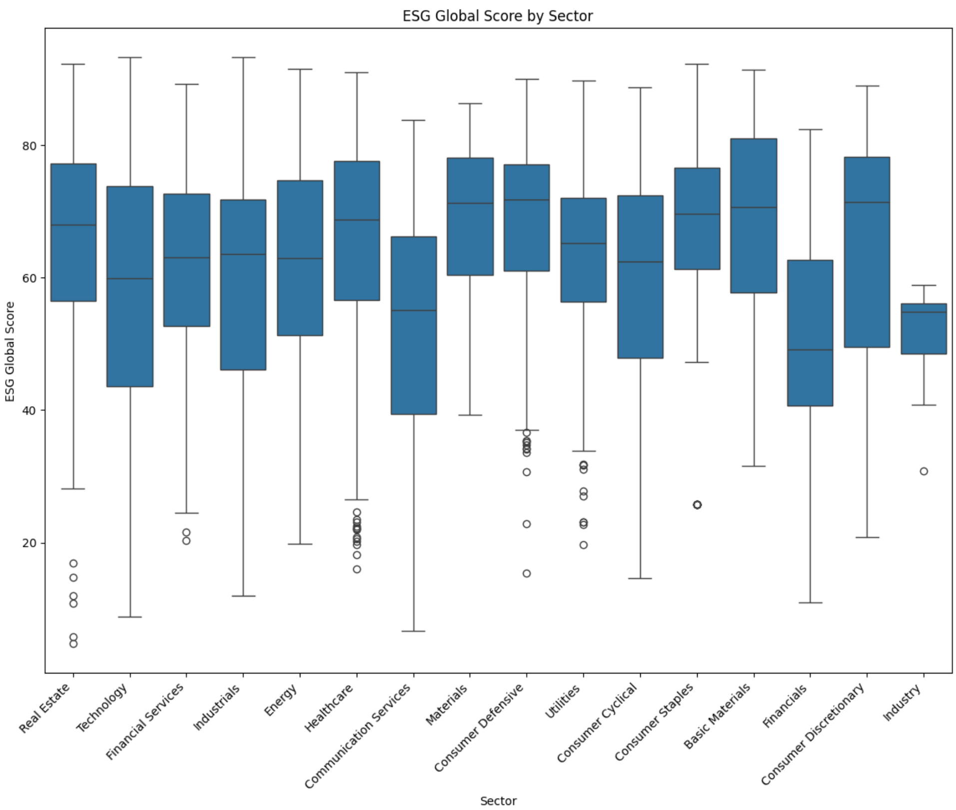

Figure 2: ESG by Sector

In contrast, Figure 3 shows that sector-level returns do not follow the same pattern. Technology, industrials, basic materials, financials and healthcare exhibit the highest average returns, whereas consumer discretionary, utilities and real estate underperform.

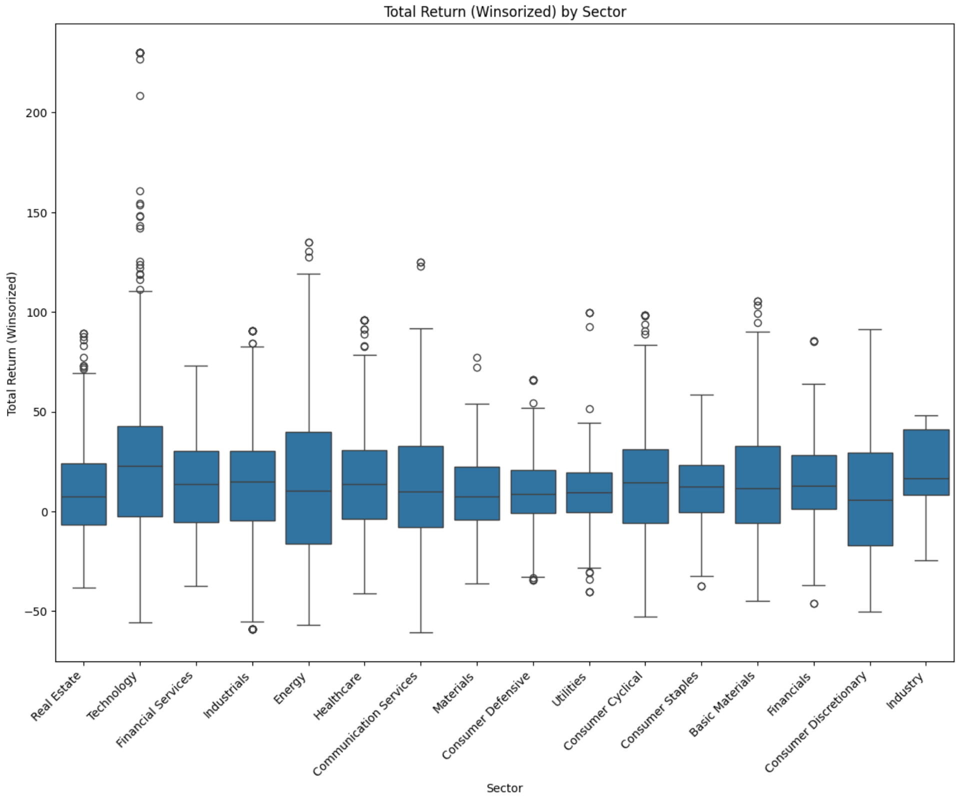

Figure 3: Return by Sector

A comparison between high-impact and low-impact sectors reveals that ESG scores are remarkably similar across both groups. However, stock performance differs markedly: low-impact sectors outperform high-impact sectors on average. This suggests that sector characteristics, rather than ESG alone, may drive performance differences.

Overall, the descriptive analysis highlights two key insights:

- ESG scores differ substantially across industries, with some sectors systematically exhibiting stronger sustainability practices.
- Stock returns do not correlate with these patterns; sectors with low environmental impact achieve stronger financial performance.

This suggests that raw ESG levels may not directly translate into sector-level return differences.

4.2 Cross-Sectional Regression Results

The first regression specification considers ESG as the sole explanatory variable. Results indicate a negative and statistically significant association between ESG scores and annual returns. Firms with higher ESG scores tend to earn lower contemporaneous returns. The following results are on the appendix, table 2.

The second specification incorporates standard firm characteristics. After controlling for size, valuation, momentum, profitability, and leverage, the coefficient on ESG remains negative and statistically significant. In addition:

- larger firms exhibit higher returns (positive size effect);
- higher P/E ratios are associated with stronger performance (growth premium);

- profitability (ROE) is positively related to returns;
- firms with higher leverage earn lower returns;
- momentum is negatively related to returns, indicating reversal effects.

In the third model, a high-impact sector dummy and an interaction term between ESG and high-impact industries are added. ESG remains significantly negative, and high-impact sectors underperform low-impact sectors. However, the interaction term is not statistically significant, suggesting that the ESG–return relationship does not differ meaningfully across environmentally intensive sectors.

4.3 Fama–MacBeth Results

To test whether the ESG effect is persistent over time, Fama–MacBeth regressions were conducted. In the ESG-only specification, the average coefficient on ESG is –0.217 with a t-statistic of –2.95, indicating a significant negative relationship across years.

In the specification with controls, ESG remains significantly negative (average coefficient –0.236, t = –4.12) and high-impact sectors underperform on average. These results confirm that the ESG effect is robust to standard firm characteristics.

Finally, the interaction between ESG and high-impact sectors is negative and significant, indicating that the ESG penalty is stronger in environmentally intensive industries.

Overall, the Fama–MacBeth results confirm that the negative ESG–return relation is stable and persistent over time.

4.4 Machine Learning Results

Out-of-sample predictive performance was evaluated using expanding-window forecasts. Linear regression models achieve out-of-sample R^2 values ranging from 3% to 8%, indicating limited predictive power.

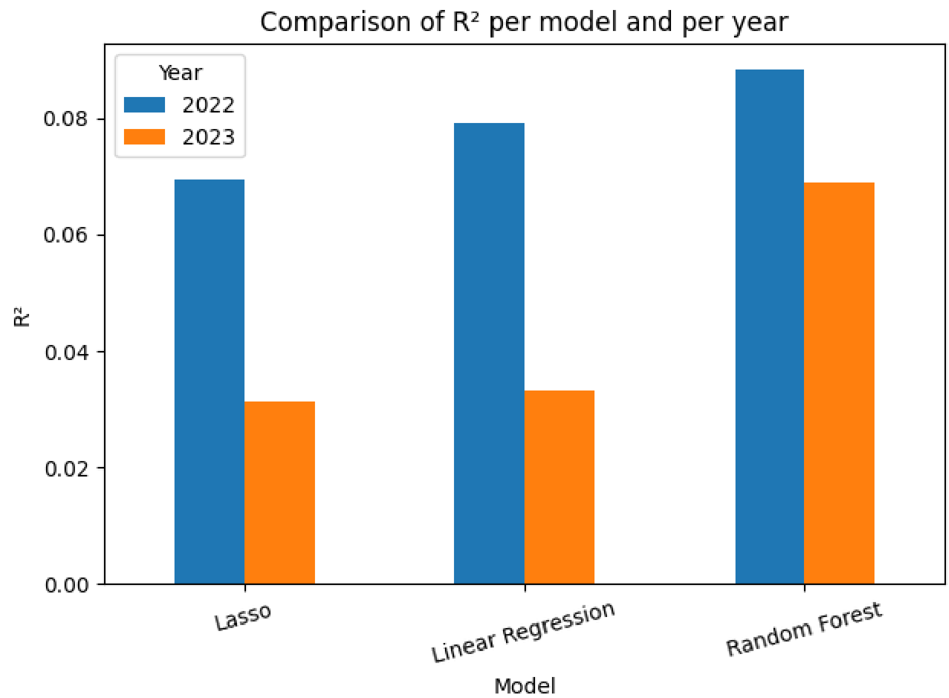

Figure 4: R^2 Comparaison

Lasso regression selects $\alpha \approx 0.21$ in both windows and achieves similar R^2 values, confirming that regularization does not materially improve predictive accuracy. However, Lasso provides useful information on variable importance, shrinking several coefficients toward zero while retaining traditional risk factors as key predictors.

Random Forest models yield the highest predictive performance, with out-of-sample R^2 between 7% and 9%. Feature importance analysis shows that Size, Momentum, Beta, ROE and P/E are the dominant predictors of future returns. ESG contributes only modestly (importance ≈ 0.104), and the interaction and high-impact indicators contribute almost none.

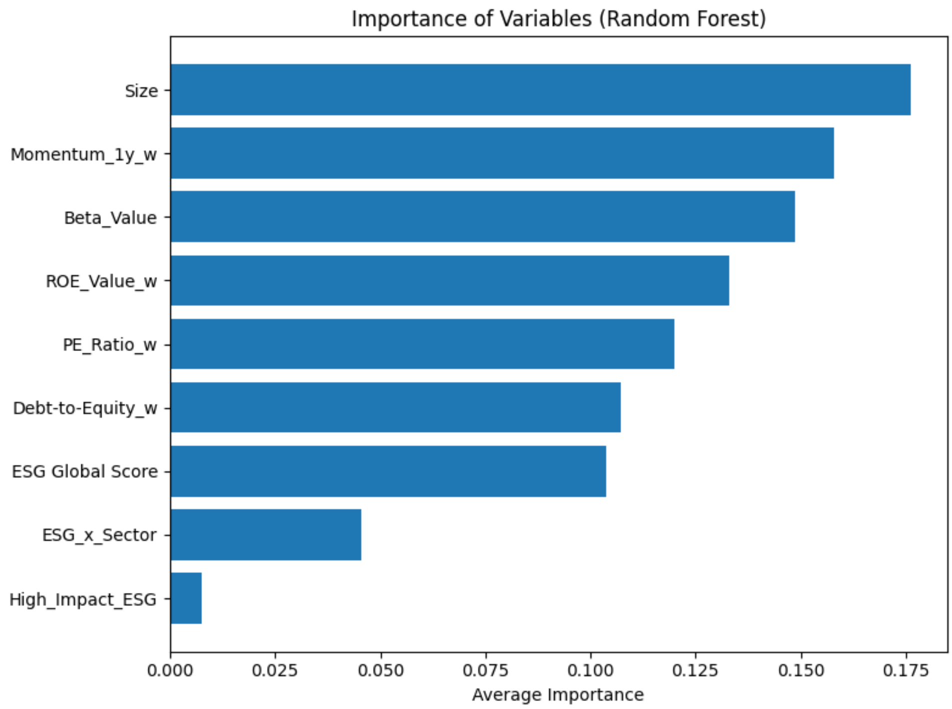

Figure 5: RF importances

Across all machine-learning experiments, ESG exhibits limited predictive power, while standard firm characteristics explain the majority of predictive performance. These findings align with the asset-pricing literature, where stock returns are notoriously difficult to forecast, and sustainability measures tend to influence long-term corporate risk rather than short-term predictability.

5 Discussion

The results obtained in this study align closely with the existing academic literature on ESG and asset pricing. Consistent with prior work, the analysis indicates that ESG scores have limited explanatory and predictive power for stock returns. Traditional firm characteristics—such as size, momentum, profitability, valuation and market beta—remain the dominant drivers of return variation.

From a data perspective, the availability of high-quality financial and ESG data from Refinitiv greatly facilitated the empirical work. However, missing data represented a notable challenge. In particular, the lack of complete information for several firms in the later years of the sample prevented reliable prediction for 2024. This highlights a broader limitation of ESG datasets: coverage and consistency issues remain common, especially for newly disclosed metrics.

A key limitation of this project is the presence of survivorship bias. Because the sample is restricted to firms that appear in the S&P 500 during the analysis window, companies that delisted, merged, or exited the index are excluded. This may overstate performance and distort the relationship between financial characteristics and returns. A more comprehensive approach would involve including the entire U.S. equity universe or using a global database to avoid index-selection effects.

Another limitation concerns the nature of the return predictability problem itself. Stock returns are notoriously difficult to forecast, and the low out-of-sample R^2 values obtained across all models are consistent with a vast body of evidence in empirical finance. Machine learning can capture non-linearities, but even advanced models like Random Forest achieve only modest improvements, suggesting that ESG metrics provide little incremental forecasting value relative to traditional factors.

Finally, ESG scores may exert their influence over longer horizons than those studied here. ESG policies often involve long-term strategic commitments, and their financial effects may materialize over multi-year periods, through risk reduction, cost improvements, or enhanced corporate resilience.

6 Conclusion and Future Work

6.1 Summary

This study examined whether ESG scores are associated with stock performance and whether they contain incremental predictive information for future returns. Across all econometric specifications—including simple cross-sectional regressions and Fama–MacBeth estimations—the ESG score consistently shows a negative but economically modest relationship with realized stock returns. These effects remain statistically significant in several models, yet their magnitude suggests that ESG is not a primary driver of short-term performance.

The machine learning analysis reinforces this conclusion. Using an expanding-window forecasting framework, linear regression, Lasso and Random Forest models all exhibit limited out-of-sample predictive power, with traditional firm characteristics (size, momentum, profitability, valuation and beta) overwhelmingly dominating prediction accuracy. ESG contributes very little additional information, and its predictive influence is substantially smaller than that of conventional risk factors.

Taken together, the evidence suggests that ESG considerations may play a role in long-term corporate sustainability and risk management, but they do not meaningfully enhance the prediction of short-horizon stock returns within the period studied.

6.2 Future Directions

Several avenues could improve and extend this research. First, expanding the sample beyond the S&P 500 would reduce survivorship bias and yield a more representative universe of U.S. firms. Including small-cap stocks or global markets may reveal different ESG–return dynamics. Second, this study relies on ESG scores from a single provider (Refinitiv); future work could compare multiple rating agencies to assess the robustness of results given the well-known divergence across ESG ratings.

Third, ESG effects may materialize over longer horizons than those considered here. Exploring multi-year cumulative returns, long-term portfolio strategies, or risk-adjusted measures could provide complementary insights. Finally, alternative sustainability metrics—such as individual E, S and G pillars, carbon emissions, or firm-level controversies—may capture aspects of corporate behavior not reflected in aggregate ESG scores.

Overall, while ESG remains an important indicator of corporate sustainability practices, the evidence from this study indicates that it offers limited value for short-term stock return prediction.

References

1. Escobar-Saldivar, L., Villarreal-Samaniego, D., Santillán-Salgado, R. (2025). *The Effects of ESG Scores and ESG Momentum on Stock Returns and Volatility: Evidence from U.S. Markets*. Journal of Risk and Financial Management, 18, 367.

2. Alves, R., Krüger, P., Van Dijk, M. (2024). *Drawing Up the Bill: Are ESG Ratings Related to Stock Returns Around the World ?* Journal of Corporate Finance, 93, 102768.

3. Da Cunha, Í., Santos Policarpo, R., De Oliveira, P., Abdala, E., Do Nascimento Rebelatto, D. (2025). *A Systematic review of ESG indicators and corproate performance: proposal for a conceptual framework*. Future Business Journal, 11, 106.

4. Yin, X-N., Li, J-P., Su, C-W. (2023). *How does ESG performance affect stock returns ? Empirical evidence from listed companies in China*. Journal of Heliyon, 9(5), e16320.

5. OpenAI. (2025). ChatGPT(Version 5.1)[Largue Language Model]. https://www.openai.com/chatgpt

A Appendix

A.1 Statistical Tests

Table 2: OLS Model 1 Results

Dep. Variable:	Tot_Return_w	**R-squared:**	0.017
Model:	OLS	**Adj. R-squared:**	0.016
Method:	Least Squares	**F-statistic:**	75.22
Date:	Mon, 24 Nov 2025	**Prob (F-statistic):**	5.81e-18
Time:	17:30:40	**Log-Likelihood:**	-21454.
No. Observations:	4435	**AIC:**	4.291e+04
Df Residuals:	4433	**BIC:**	4.293e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	29.3283	1.713	17.119	0.000	25.970	32.687
Q('ESG Global Score')	-0.2333	0.027	-8.673	0.000	-0.286	-0.181

Omnibus:	1039.002	**Durbin-Watson:**	2.073
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	5088.387
Skew:	1.041	**Prob(JB):**	0.00
Kurtosis:	7.816	**Cond. No.**	238.

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Table 3: OLS Model 2 Results

Dep. Variable:	Tot_Return_w	**R-squared:**	0.124
Model:	OLS	**Adj. R-squared:**	0.122
Method:	Least Squares	**F-statistic:**	81.96
Date:	Mon, 24 Nov 2025	**Prob (F-statistic):**	3.08e-96
Time:	18:32:19	**Log-Likelihood:**	-16615.
No. Observations:	3489	**AIC:**	3.324e+04
Df Residuals:	3482	**BIC:**	3.329e+04
Df Model:	6		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	-67.8962	10.963	-6.193	0.000	-89.391	-46.401
Q('ESG Global Score')	-0.3101	0.033	-9.483	0.000	-0.374	-0.246
Size	9.1848	1.115	8.235	0.000	6.998	11.372
PE_Ratio_w	0.3338	0.020	16.498	0.000	0.294	0.373
Momentum_1y_w	-0.1273	0.017	-7.459	0.000	-0.161	-0.094
ROE_Value_w	0.0495	0.009	5.420	0.000	0.032	0.067
Q('Debt-to-Equity_w')	-0.0076	0.003	-2.965	0.003	-0.013	-0.003

Omnibus:	492.729	**Durbin-Watson:**	1.817
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1944.355
Skew:	0.654	**Prob(JB):**	0.00
Kurtosis:	6.415	**Cond. No.**	6.12e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.12e+03. This might indicate that there are
strong multicollinearity or other numerical problems.

Table 4: OLS Model 3 Results

Dep. Variable:	Tot_Return_w	**R-squared:**	0.127
Model:	OLS	**Adj. R-squared:**	0.125
Method:	Least Squares	**F-statistic:**	63.11
Date:	Mon, 24 Nov 2025	**Prob (F-statistic):**	7.72e-97
Time:	18:32:35	**Log-Likelihood:**	-16609.
No. Observations:	3489	**AIC:**	3.324e+04
Df Residuals:	3480	**BIC:**	3.329e+04
Df Model:	8		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	-58.3920	11.296	-5.169	0.000	-80.539	-36.245
Q('ESG Global Score')	-0.3371	0.042	-7.979	0.000	-0.420	-0.254
High_Impact_ESG	-8.2043	3.892	-2.108	0.035	-15.835	-0.574
Q('ESG Global Score'):High_Impact_ESG	0.0815	0.060	1.351	0.177	-0.037	0.200
Size	8.5700	1.140	7.519	0.000	6.335	10.805
PE_Ratio_w	0.3349	0.020	16.448	0.000	0.295	0.375
Momentum_1y_w	-0.1311	0.017	-7.676	0.000	-0.165	-0.098
ROE_Value_w	0.0496	0.009	5.408	0.000	0.032	0.068
Q('Debt-to-Equity_w')	-0.0072	0.003	-2.822	0.005	-0.012	-0.002

Omnibus:	485.024	**Durbin-Watson:**	1.821
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	1903.499
Skew:	0.645	**Prob(JB):**	0.00
Kurtosis:	6.381	**Cond. No.**	6.34e+03

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 6.34e+03. This might indicate that there are strong multicollinearity or other numerical problems.

Table 5: Fama-MacBeth Model 1 Results

	beta_mean	fm_se	fm_t
Intercept	28.391117	6.491943	4.373285
Q('ESG Global Score')	-0.217102	0.073525	-2.952777

Table 6: Fama-MacBeth Model 2 Results

	beta_mean	fm_se	fm_t
Intercept	-28.330229	14.907813	-1.900361
Q('ESG Global Score')	-0.236400	0.057353	-4.121833
Size	7.465663	2.579295	2.894459
PE_Ratio_w	0.254388	0.047204	5.389151
Momentum_1y_w	-0.107027	0.078113	-1.370150
ROE_Value_w	0.127973	0.065230	1.961879
Q('Debt-to-Equity_w')	-0.004735	0.002355	-2.010890
High_Impact_ESG	-28.330229	14.907813	-1.900361

Table 7: Fama-MacBeth Model 3 Results

	beta_mean	fm_se	fm_t
Intercept	-28.330229	14.907813	-1.900361
Q('ESG Global Score')	-0.118200	0.028677	-4.121833
High_Impact_ESG	-28.330229	14.907813	-1.900361
Q('ESG Global Score'):High_Impact_ESG	-0.118200	0.028677	-4.121833
Size	7.465663	2.579295	2.894459
PE_Ratio_w	0.254388	0.047204	5.389151
Momentum_1y_w	-0.107027	0.078113	-1.370150
ROE_Value_w	0.127973	0.065230	1.961879
Q('Debt-to-Equity_w')	-0.004735	0.002355	-2.010890

B Code Repository

GitHub Repository: `https://github.com/Niru802/Niruban-project-repo.git`
All files required to run the project are organised in a clear and reproducible directory structure. The full codebase, data and results are contained in the project folder and can be executed directly using `python main.py` after creating the conda environment.

- **main.py** — The entry point of the project. Running this script loads the dataset, executes all machine learning models (Linear Regression, Lasso, Random Forest) using an expanding-window design, and saves the results to the `results/` folder.
- **src/** — Contains all Python modules used in the analysis:
  - `data_loader.py`: loads the cleaned dataset from `data/final_panel.csv`.
  - `models.py`: implements all predictive models and the expanding-window procedure.
  - `evaluation.py`: saves model outputs (metrics, feature importances) to disk.
- **data/** — Contains the cleaned modelling dataset used in the project:
  - `final_panel.csv`: final panel dataset with yearly firm-level variables (ESG score, financial ratios, sector classification, returns, engineered features).

- **results/** — Stores all outputs generated by `main.py`:
  - `ml_metrics.csv`: out-of-sample MSE and R^2 for 2022 and 2023.
  - `feature_importances.csv`: Random Forest feature importance scores.
- **environment.yml** — Conda environment specification containing all dependencies needed to reproduce the analysis.
- **notebooks** — Contains the notebooks used for creating the main dataset, descriptive analysis, statistical analysis and testing machine learning models for the project.
- **README.md** — Instructions on how to recreate the environment and run the project.
- **Niruban_project_report.pdf** — The final written report documenting the methodology, results and conclusions of the study.

To reproduce the project, users simply need to create the conda environment using `environment.yml`, activate it, and run the command:

```
python main.py
```

This reproduces all model results and regenerates the output files in the `results/` directory.