ESG and Stock Return Prediction

Advanced Programming Project – 2025

CHEKARA Ishaq
ishac.chekara@unil.ch
ID: 20440046

January 11, 2026

Contents

Abstract

This project investigates whether Environmental, Social, and Governance (ESG) scores provide incremental predictive power for monthly stock returns of S&P 500 constituents. We build a fully automated machine learning pipeline combining ESG data with historical stock returns. Predictive performance is compared between a baseline model using lagged returns and an ESG-enhanced model incorporating ESG scores and categorical sector and industry variables. Model performance is evaluated out-of-sample using RMSE, MAE, and R^2 metrics. The results suggest that ESG information provides limited incremental predictive power for short-term returns, consistent with the view that ESG characteristics are more closely related to long-term risk and stability than to short-term return dynamics.

1 Introduction

Predicting stock returns is notoriously difficult due to market efficiency and the high level of noise inherent in financial data. As a result, the objective of this project is not to achieve high predictive accuracy, but rather to assess whether ESG information provides incremental explanatory power beyond standard return-based predictors.

This project examines monthly stock returns for S&P 500 firms and compares:

- a **baseline model** based solely on lagged returns,
- an **ESG-enhanced model** augmenting lagged returns with ESG scores and firm characteristics.

The entire analysis pipeline is fully automated and reproducible, from data collection to model evaluation.

2 Research Question

The central research question of this project is:

> Does ESG information improve the prediction of future monthly stock returns once historical return information is taken into account?

This question is addressed by comparing the performance of baseline (only historical returns) and ESG-enhanced models.

3 Motivation

The motivation behind this project is partly personal. From the outset, focusing on a financial topic appeared to be a natural choice, and the progression toward ESG scores as the main focus followed shortly thereafter. ESG scores represent a distinctive metric, and their widespread use today is equally singular, as they aim to capture dimensions of corporate behavior that go beyond purely financial performance. Such an initiative was necessary to create incentives toward more responsible economic development. The strength of ESG scores lies primarily in their ability to aggregate social and governance dimensions alongside environmental considerations, rather than focusing solely on ecological indicators, which have often been used for greenwashing purposes. Evaluating firms based on their relationships with employees and the quality of their governance structures adds a layer of coherence to the ESG framework and aligns it more closely with its intended purpose: assessing a form of corporate responsibility and accountability among large firms.

4 Data

4.1 ESG Dataset

The ESG dataset contains annual ESG risk ratings for S&P 500 companies over multiple years. Each firm appears once per year.

The key variables include:

- Firm identifier: `ticker`, `year`
- ESG total score (`esg`)
- Environmental, Social, and Governance sub-scores (`e`, `s`, `g`)
- Sector and industry classifications

4.2 Price Data and Monthly Returns

Stock price data are obtained from Yahoo Finance using the `yfinance` library. Prices are resampled to month-end frequency and monthly returns are computed as percentage changes.

The final dataset is constructed by merging monthly returns with ESG data using firm ticker and calendar year.

4.3 Annual Returns (Exploratory Analysis)

An alternative specification using annual returns was explored. However, the annual dataset contained significantly fewer observations, resulting in higher estimation variance. Since results were qualitatively similar to the monthly analysis, annual returns are excluded from the final presentation.

5 Feature Engineering

5.1 Target Variable

The target variable is the one-month-ahead stock return.

5.2 Baseline Features

The baseline model includes lagged returns to capture short-term persistence:

- `ret_lag_1` to `ret_lag_6`

Lagged returns represent past values of a stock's return shifted backward in time and are commonly used in financial modeling to capture short-term dynamics in asset prices. For instance, a one-month lag corresponds to the return observed in the previous month, while higher-order lags capture returns further in the past. These variables allow the model to account for persistence, momentum, and mean-reversion effects that are frequently documented in the empirical finance literature. In this project, lagged returns form the core baseline information, providing a benchmark that reflects what can be inferred from historical price behavior alone. ESG variables are therefore evaluated only in terms of their incremental contribution beyond these established return-based predictors.

5.3 ESG Features

The ESG-enhanced specification augments the baseline with:

- ESG total score and E/S/G sub-scores
- Sector and industry dummy variables (one-hot encoded)

This design allows a clean comparison between models with and without ESG information.

6 Methodology

6.1 Preprocessing

Numerical features are normalized using Min-Max scaling. Categorical variables are encoded using one-hot encoding. Observations with missing target or ESG values are excluded to ensure consistent feature availability.

6.2 Models

The following regression models are considered:

- Linear Regression
- Ridge Regression
- Lasso Regression
- Random Forest Regressor
- Gradient Boosting Regressor

All models are evaluated using identical training and testing splits.

7 Evaluation Strategy

Models are evaluated using a fixed train-test split, which is kept identical across all specifications to ensure fair comparison.

Performance metrics include RMSE, MAE, and R^2. The marginal value of ESG information is assessed using:

$$\Delta R^2 = R^2_{\text{ESG-enhanced}} - R^2_{\text{Baseline}}.$$

8 Results

8.1 Baseline Models (Lagged Returns Only)

We first evaluate baseline models that rely exclusively on past return information through lagged returns (up to six months). Table 1 summarizes the predictive performance on the monthly dataset.

Table 1: Baseline model performance (monthly returns)

Model	RMSE	MAE	R^2
Linear Regression	≈ 0.096	≈ 0.072	≈ 0.007
Ridge Regression	≈ 0.096	≈ 0.072	≈ 0.007
Lasso Regression	≈ 0.095	≈ 0.071	≈ 0.007
Random Forest	≈ 0.091	≈ 0.067	≈ 0.064
Gradient Boosting	≈ 0.093	≈ 0.070	≈ 0.037

Overall, baseline models achieve modest predictive performance, with R^2 values ranging from approximately 0.00 for linear models to 0.06 for Random Forest. This result is consistent with the financial literature, as predicting short-horizon stock returns is known to be extremely challenging.

Tree-based models (Random Forest and Gradient Boosting) clearly outperform linear models in terms of RMSE and R^2. This suggests that return dynamics exhibit nonlinear patterns and interaction effects that linear specifications fail to capture.

8.2 ESG-Enhanced Models

We then augment the feature space with ESG-related variables (overall ESG score, E/S/G sub-scores) as well as sector and industry indicators. Table 2 reports the results for ESG-enhanced models.

Table 2: ESG-enhanced model performance (monthly returns)

Model	RMSE	MAE	R^2
Linear Regression	≈ 0.096	≈ 0.072	≈ 0.003
Ridge Regression	≈ 0.096	≈ 0.072	≈ 0.006
Lasso Regression	≈ 0.095	≈ 0.071	≈ 0.009
Random Forest	≈ 0.092	≈ 0.067	≈ 0.09
Gradient Boosting	≈ 0.093	≈ 0.069	≈ 0.04

Reported values are rounded for readability. Exact metrics are available in the project repository.

Adding ESG variables yields heterogeneous effects across models. For linear regressions, ESG information has virtually no impact on predictive performance. For tree-based models, ESG features sometimes lead to small improvements.

8.3 ESG Contribution Analysis

To isolate the marginal contribution of ESG information, we compute:

$$\Delta R^2 = R^2_{\text{ESG}} - R^2_{\text{Baseline}}.$$

The results indicate that:

- ΔR^2 is close to zero for linear models,
- ΔR^2 can be slightly positive for tree-based models,
- ESG variables do not consistently improve out-of-sample performance across specifications.

This suggests that ESG scores contain, at best, a weak and unstable signal for short-term return prediction once past returns and sector effects are accounted for.

8.4 Exploratory Graphical Analysis

In addition to the quantitative evaluation of predictive models, several exploratory graphical analyses were conducted to provide complementary insights into the relationship between ESG scores and stock returns. All figures are reported in the Appendix for reference.

First, the distribution of monthly stock returns exhibits a bell-shaped form centered around zero, with noticeable fat tails. This confirms the well-known stylized facts of financial returns and highlights the high level of noise inherent in short-term return prediction.

Second, pairwise correlations between ESG scores (total ESG and E, S, and G sub-scores) and monthly returns are close to zero. This indicates the absence of a strong linear relationship between ESG characteristics and short-horizon returns, which is consistent with the weak predictive performance observed in the regression models.

Finally, a portfolio-style analysis compares the cumulative performance of firms in the top 20% and bottom 20% of the ESG score distribution. At each date, firms are ranked by ESG score and grouped accordingly. The resulting cumulative return paths show that high-ESG portfolios tend to slightly outperform low-ESG portfolios over the sample period. However, the performance gap remains modest and is characterized by significant volatility and periods of convergence.

Overall, these graphical results reinforce the main empirical findings of the project. While firms with higher ESG scores may exhibit marginally better long-term performance, ESG information does not appear to provide a strong or stable signal for predicting monthly stock returns. The observed differences are small relative to the overall variability of returns, suggesting that ESG characteristics are more relevant as long-term firm attributes rather than short-term return predictors.

8.5 Why Do Tree-Based Models Perform Better?

The superior performance of Random Forest and Gradient Boosting can be explained by several factors:

- **Nonlinearity:** Financial return dynamics are highly nonlinear. Tree-based models naturally capture nonlinear relationships without requiring explicit functional form assumptions.
- **Interaction effects:** Tree ensembles can exploit interactions between lagged returns, ESG scores, and sector membership, which linear models ignore.
- **Robustness to noise:** Random Forests average across many decorrelated trees, reducing variance and overfitting in noisy financial data.
- **Threshold effects:** ESG variables may matter only beyond certain thresholds (e.g., very high or very low ESG scores), a pattern better captured by decision trees.

However, even for tree-based models, the absolute level of R^2 remains low, reflecting the inherent difficulty of predicting equity returns and the limited incremental information content of ESG scores at a monthly frequency.

9 Discussion

Several factors can explain the limited contribution of ESG scores:

- ESG scores evolve slowly relative to monthly returns.
- ESG-related information may already be partially incorporated into prices.
- The signal-to-noise ratio in return prediction is inherently low.

Beyond predictive accuracy, it is important to emphasize that ESG scores are not primarily designed to forecast short-term stock returns. ESG indicators evolve slowly over time and are more closely related to firms' long-term sustainability, governance quality, and exposure to non-financial risks. Consequently, the limited predictive contribution of ESG scores for monthly returns observed in this project is not unexpected, as short-term return dynamics are largely driven by market-wide shocks, macroeconomic conditions, and momentum effects.

From a performance measurement perspective, these results closely relate to the concepts discussed in the Investment course. In particular, the ESG-enhanced model can be interpreted as an attempt to generate incremental alpha relative to a baseline benchmark based on lagged returns. Alpha represents the component of performance that cannot be explained by systematic risk factors or benchmark exposures and is commonly interpreted as evidence of abnormal returns. As emphasized in the course, apparent performance improvements must be evaluated relative to an appropriate benchmark and assessed out-of-sample. Our findings indicate that once standard return-based predictors and sector exposures are controlled for, ESG variables do not generate a persistent or robust source of alpha, but instead resemble a style dimension rather than genuine selection skill.

While ESG characteristics do not appear to generate a robust return premium, they may nonetheless remain economically relevant through other channels. In particular, ESG scores may be more informative for predicting measures of financial risk rather than expected returns. Firms with stronger ESG profiles may exhibit lower volatility, reduced downside risk, or greater resilience during periods of market stress. Therefore, the absence of short-term return predictability does not imply that ESG information lacks financial relevance.

Nevertheless, recent developments in financial markets illustrate that even major asset managers are reevaluating their ESG-branded products. For example, BlackRock, the world's largest asset manager, announced the closure of certain ESG-focused funds and adjustments to its sustainable finance offerings.

The underlying causes are multifaceted. Investor interest in ESG funds has declined, but this trend cannot be explained solely by weaker performance relative to other funds. In the United States, the political shift toward conservative positions has limited the expansion of initiatives such as ESG, despite their relevance beyond purely financial considerations.

Furthermore, the Securities and Exchange Commission amended its "name rule," requiring that at least 80% of a fund's assets be invested in accordance with the characteristics suggested by its name. This regulatory change particularly targeted ESG funds and further illustrates that such products are no longer at the peak of their popularity in the United States, despite the country being one of the main markets where ESG investing has developed and expanded over the past two decades.

Such market behavior aligns with our empirical finding that ESG scores offer limited incremental predictive power for short-term stock returns.

10 Conclusion

This project developed an end-to-end machine learning pipeline to assess whether ESG scores improve the prediction of monthly stock returns for S&P 500 firms.

The empirical results indicate that while lagged returns contain modest predictive information, ESG scores provide limited incremental value for short-term return prediction. Once sector and industry effects are properly taken into account, the numerical ESG scores (E, S, and G) do not exhibit robust out-of-sample predictive power for monthly returns.

These findings suggest that ESG information should not be interpreted primarily as a return-predictive signal at short horizons. Instead, ESG characteristics are more plausibly related to firms' long-term risk profiles, resilience, and exposure to non-financial risks. Therefore, the absence of a strong ESG return premium does not contradict the economic relevance of ESG considerations.

Future research could extend this framework by examining alternative financial targets such as volatility, downside risk, drawdowns, or performance during market stress periods, where ESG information may play a more meaningful role.

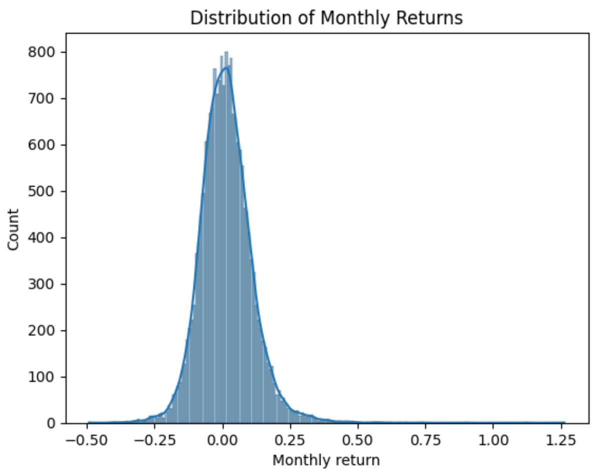
References

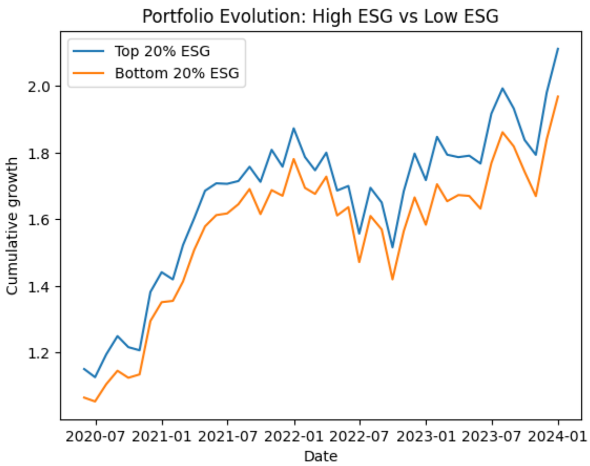
1. Yahoo Finance data accessed via the `yfinance` Python library.
2. Scikit-learn documentation on preprocessing, pipelines, and regression models.
3. S&P Global ESG Scores provided through CEDIF.

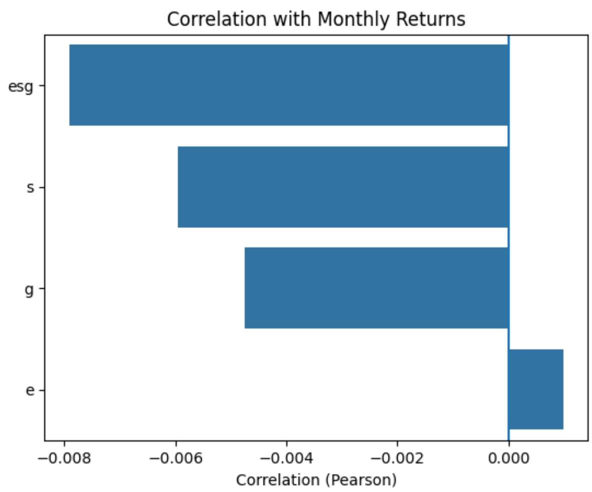
A Code Repository

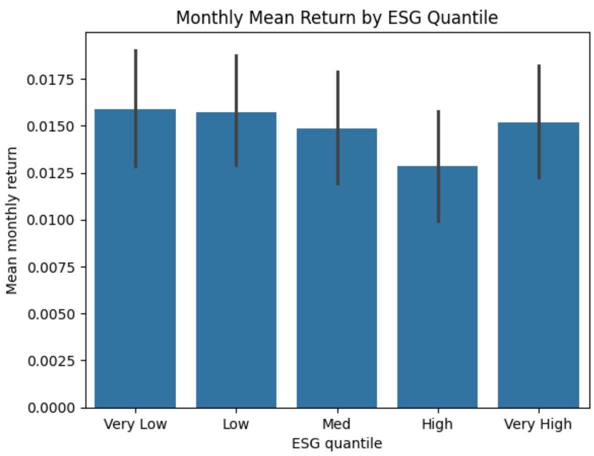
GitHub Repository: https://github.com/Sh4kk0/AP-ESG-Prediction-Project

B Graphs




Portfolio Evolution: High ESG vs Low ESG
Top 20% ESG
Bottom 20% ESG
Cumulative growth
2.0
1.8
1.6
1.4
1.2
2020-07
2021-01
2021-07
2022-01
2022-07
2023-01
2023-07
2024-01
Date


Correlation with Monthly Returns
esg
s
g
e
−0.008
−0.006
−0.004
−0.002
0.000
Correlation (Pearson)


Monthly Mean Return by ESG Quantile
Mean monthly return
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
Very Low
Low
Med
High
Very High
ESG quantile