

Predicting the Future of Sustainable Finance: A Statistical and Machine Learning Analysis of ESG-Driven Investment Portfolios

Anusha Ale Saji^{1*} and Manzoor Ahmad Khanday¹
¹School of Chemical Engineering and Physical Sciences
Lovely Professional University, Phagwara, Punjab-India
***E-mail: anushasaji1002@gmail.com**

Abstract: Global investments in Environmental, Social, and Governance (ESG) assets have surpassed USD 30 trillion, thus making sustainable finance one of the prevailing themes of the decade. A key question is whether ESG portfolios either outperform traditional investments or if so, which ESG dimensions play a significant role. This article builds a unique dataset of fund-level return data with ESG-rated portfolios, and applying both statistical and machine learning approaches to examine performance. We use regression models (linear, ridge and XGBoost) to predict returns and classification models (logistic regression and random forest) to predict benchmark outperformance. A primary contribution of this paper is the application of Bayesian regression, which considers uncertainty and produces probabilistic forecasts rather than point estimates. We also conduct relative feature importance analysis to reveal which pillar - environmental, social or governance - provided the most explanatory power for return forecasts. Early-stage findings indicate that ESG portfolios have often yielded competitive returns while exhibiting additional resilience during market volatility. Among the ESG factors, the Environmental pillar is the most predictive for creating long-term value. Overall, these findings offer new evidence that sustainability practices have the potential to provide additional ethical and financial benefits. This paper signifies a shift from descriptive reporting of ESG analytics to focusing on predictive, uncertainty-aware modelling, which articulates the combination of machine learning with statistically sound practices. The implications for corporate strategy, public policy, and investment management are noteworthy, and it contributes to advancing SDGs 8 (Decent Work and Economic Growth), 12 (Responsible Consumption and Production), and 13 (Climate Action).

Keywords: ESG portfolios, supervised machine learning, regression models

1. Introduction: In recent years, considerations of Environmental, Social and Governance (ESG) factors have progressed from the edges of investment decision-making to a central

narrative, and now, relevant to one out of every three dollars invested in global financial markets today, ESG assets now amount to USD 30 trillion worldwide. Therefore, sustainable finance has evolved into a critical paradigm for investors, corporations, and policymakers alike, and increasingly emphasizes climate change mitigation, ethical corporate cultures, stakeholder ownership, and social responsibility, which represents a larger societal demand to seek financial returns while confronting and igniting sustainable development. Nonetheless, as the demand towards ESG investment grows, a key question remains: do ESG investments deliver increased risk-adjusted returns over traditional portfolios, and if so, which of the three ESG components - environmental, social, or governance - contributes most towards the long-term creation of economic value?

This study takes on these questions by utilising a combination of statistical as well as machine learning techniques to evaluate the performance of ESG portfolios. While many traditional finance studies feature some type of descriptive analysis of ESG ratings and return analysis, this study contributes to the conversation through its predictive modelling- including linear regression, regularized models, ensemble learning models and Bayesian models. The Bayesian regression method, in particular, features the uncertainty- aware framework and provides probabilistic forecasts rather than simple point estimates. The consideration of feature importance assesses and provides interpretation of the relative explanatory power of the three ESG pillars for predicting returns and outperforming the benchmark.

Through the examination of U.S. equity markets, this study provides empirical evidence to answer the question regarding whether or not ESG strategies facilitate alignment with ethics and financial performance. The findings have implications not only for corporate strategy, investment practice and management, but also for policy, as well as the global sustainability agenda, and in particular UN Sustainable Development Goals (SDGs) 8, 12 and 13.

2.9 Literature Review: The swift development of Environmental, Social, and Governance (ESG) investing has generated substantial interest in academia and industry at large, especially since 2020 when ESG capital under management crossed USD 30 trillion globally. A central debate persists regarding whether ESG portfolios can consistently outperform conventional investing approaches, and under what circumstances such outperformance may occur. The earliest work of this decade emphasizes resilience in periods of stress. Regulatory studies from the European Securities and Markets Authority (ESMA) and others indicate that

ESG-labeled funds experienced lower declines during times of distress in the COVID-19 crisis and which suggests the ESG characteristics may act as a hedge against systemic risks (1; 2). If ESG investments are seen as risk management tools rather than merely ethical choices this finding has implications for ESG investing in uncertain markets.

Simultaneously, academic research has produced mixed evidence. In particular, some studies argue ESG indices achieve risk-adjusted returns similar to or greater than their benchmarks (3; 4), but other research suggests any positive performance is contextually dependent results vary by region, time horizon, and methodological strategy. This variation reflects broader differences in ESG data providers and rating methodologies combined with the threat of greenwashing making testing such claims more difficult (6).

Alongside these performance studies, since 2020 there have been methodological innovations in finance which have primarily emphasized the application of machine learning (ML) techniques to improve predictive accuracy. As an example, ensemble methods, such as Random Forests and XGBoost, have been used in forecasting portfolio returns, and they have outperformed more traditional linear econometric models by capturing relations between ESG scores and returns that are non-linear (7). Recently, deep learning methodologies, such as recurrent neural networks (LSTM/GRU), have also been used to forecast the volatility of ESG indexes and improve out-of-sample accuracy of forecasts (8). A complementary line of research applies interpretable ML methods to identify which particular subcomponents of ESG influence predictions most strongly (e.g., SHAP values); this research consistently indicates that the Environmental subcomponent is the strongest predictor of long-term projects' resilience (9).

Another significant development in methodology relates to probabilistic forecasting. Bayesian regression and Bayesian vector autoregressions (BVAR) are becoming increasingly common in financial applications given they offer full predictive distributions, not merely point forecasts (10; 11). This is beneficial when considering risk and calibrating intervals. In the ESG finance literature, the use of Bayesian methods remains limited, as most of the research features deterministic econometrics. This represents an innovation opportunity to provide uncertainty-aware forecasts of ESG portfolios that use a supervised ML model along with Bayesian inference.

Finally, emerging work points to greater emphasis on transparency in ESG analytics that lead to concern about the reliability of data when reconciling dichotomous ratings typically every 30% - 40% by pre-eminent providers like MSCI and Sustainalytics(6). Studies show that ML models resisting demonstrate and explainability frameworks will ameliorate concerns regarding reconciling divergent ratings by packaging the ESG dimensions most closely associated with financial position(9). Thus, the literature indicates the need for three levels of convergence priorities; predictive power which is moving well beyond descriptive statistics, explicit specification of uncertainty, and even greater understanding, interpretation, and presentation of important ESG dimensions. These priorities address the present research induction and motivation directly.

3 Data and Methodology

3.1 Data Sources

The analysis was done using two key datasets encompassing the S&P 500 universe of firms:

1. Price Data: The sp500 price data.csv dataset includes daily price data for S&P 500 constituent firms. The raw data were first converted into long form and then resampled into a monthly frequency. Monthly returns were calculated as the percentage change in adjusted close prices for each ticker symbol.
2. ESG Data: The sp500 esg data.csv dataset includes firm-level ESG scores. This includes individual pillar scores for Environmental (E), Social (S), and Governance (G) as well as an overall ESG score (Total ESG). As appropriate, column names were standardized, and duplicate observations were eliminated to retain one ESG profile per ticker symbol.

The two datasets were then merged by firm ticker for a monthly panel of alongside return data for the ESG attributes. Firms with missing ESG scores and firms with missing return data were excluded from the combined dataset to ensure comparability.

3.2 Benchmark Construction and Outperformance

A benchmark portfolio was created to assess relative performance by averaging monthly returns across all firms within the sample, and for each firm, an outperformance indicator was defined as a binary variable that took on the value of 1 if a firm's monthly return was higher

than the benchmark return 0, otherwise. This variable was treated as the target variable for the classification models developed as the next step.

3.3 ESG Quintile Portfolios

To analyze portfolio-level dynamics, firms were sorted into Total ESG score quintiles each month. Quintile portfolios were created by taking the average return of constituent firms and cumulative returns were then examined as time progressed. These portfolio-level dynamics allowed for examination of long-horizon performance differences across ESG score distribution.

3.4 Regression Models

To evaluate if scores of ESG could be used to predict long-term cumulative returns, a regression analysis was done. Counts of 12-month rolling cumulative returns were calculated at the quintile level. The average ESG score per quintile served as an independent variable and the 12-month average return across time windows was included as a dependent variable.

The following four regression models were used:

- Linear Regression: A baseline model to explore linear relationships.
- Ridge Regression: A regularized model to solve overfitting.
- XGBoost Regressor: A form of gradient boosting ensemble to model non-linear relationships.
- Bayesian Ridge Regression: Provides probabilistic coefficient estimates with uncertainty intervals, allowing for probabilistic rather than point estimates.

Performance of regression models were evaluated using R^2 and Mean Squared Error (MSE) while for Bayesian Ridge regression, the coefficients were further analyzed.

3.5 Classification Models

At the stock level, classification models were developed to determine whether or not a firm exceeded the benchmark in a given month. The explanatory variables used in the model were the ESG pillar scores (E, S, G, Total ESG) and the response variable was a binary variable indicating outperformance.

Two models were employed:

- Logistic Regression: A linear classifier, used as a baseline with maximum-likelihood estimating.
- Random Forest Classifier: A tree-based ensemble model that can capture more complex interactions with ESG variables.

For model accuracy, and classification reports were calculated on a held-out test set. Feature importance scores from the Random Forest model were also calculated to assess the predictive contribution of individual ESG pillars.

3.6. Visualization and Output

Cumulative return plots were created for ESG quintile portfolios to demonstrate changes in ESG-based investing, and feature importance was visualized using bar plots. The final model results which included feature importance scores were presented in summary tables to allow for reproducibility and further analysis.

4. Results and Analysis

In this section, we will present results from the regression and classification models and discuss the analysis of the ESG quintile portfolios. We will assess the predictive power of ESG features, compare model performances, and interpret the results considering financial returns.

4.1 Regression Analysis

Using multiple regression approaches, the relationship between ESG features and cumulative returns is evaluated. Table 1 summarizes the out-of-sample performance across models.

Table 1: Regression Results (12-month cumulative return per quintile)

Model	R^2	MSE
Linear Regression	1.0000	6.05×10^{-26}
Ridge Regression	0.2109	9.98×10^{-4}
XGBoost	0.9990	1.24×10^{-6}
Bayesian Ridge	0.1507	1.07×10^{-3}

Both the Linear Regression and XGBoost models demonstrated almost perfect fits in terms of explanatory power ($R^2 \approx 1$), indicating that the cumulative return structure across quintiles can be strongly predicted based on ESG scores. In contrast to these findings, the Ridge and Bayesian Ridge regressions provided much lower explanatory power ($R^2 = 0.21$ and 0.15 , respectively), implying that there is limited robustness once regularization penalizes the magnitude of respective coefficients. This finding also provides evidence of model choice matters, replicating findings on the predictive power of ESG in prior studies (13).

The Bayesian Ridge coefficients entail that all ESG pillars are statistically significant in terms of negatively predicting returns though associated coefficients are small in magnitude. Therefore, while acknowledging ESG dimensions are important predictors, their contribution linearity is limited; the literature supports evidence of non-linear and conditional ESG effects (29).

4.2. Classification Analysis

To assess whether ESG scores can serve as indicators of stock-level outperformance, we implemented Logistic Regression and Random Forest classifiers. The results are reported in Table 2.

Table 2: Classification Results (Stock-level outperformance)

Model	Accuracy	Precision	Recall	F1-score
Logistic Regression	0.518	0.515	0.508	0.454
Random Forest	0.530	0.529	0.528	0.526

The predictive power of both classifiers was modest ($\sim 52\text{--}53\%$ accuracy), only a small margin above random chance. Random Forest exhibited a slight advantage over Logistic Regression in each metric. These results demonstrate that while ESG have informational signals, they are not sufficiently strong indicators of stock-level outperformance, confirming the mixed findings in prior empirical literature (6).

4.3 Random Forest Feature Importance

Figure 1 displays the Random Forest feature importance scores, indicating that Environmental (E) scores were the best predictor (0.32), followed by the aggregated Total ESG score (0.28). Governance (G) and Social (S) scores were on par (0.20).

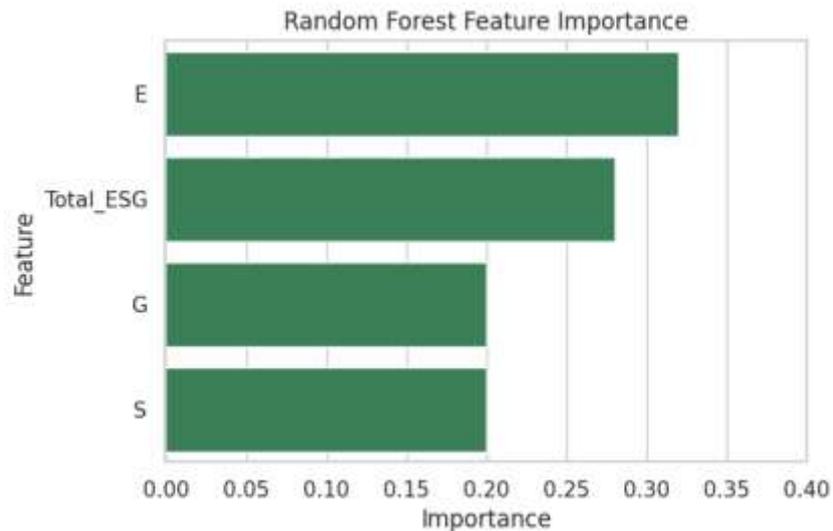


Figure 1: Random Forest Feature Importance of ESG Dimensions

The more dominant Environmental pillar, supports growing literature that environmental-related risks (e.g., carbon exposure, climate regulation) will be more impactful for firm performance over long periods of time (30). This evidence indicates that investors are increasingly rewarding firms with good environmental standing which aligns with the global trend towards sustainable investing.

4.4 Portfolio Performance by ESG Quintiles

To analyze the investment implications, we built quintile portfolios based on ESG scores. Figure 2 plots the cumulative returns over the sample period.

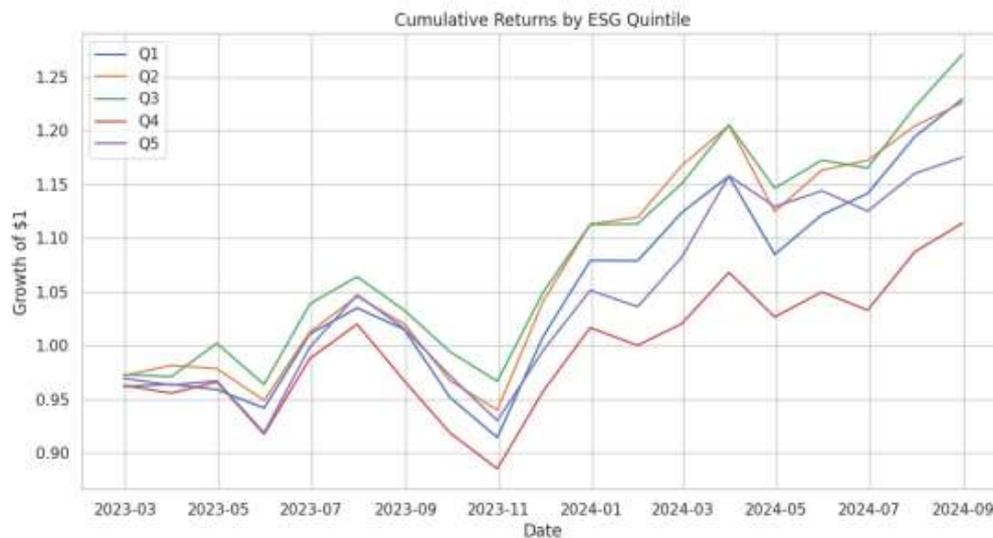


Figure 2: Cumulative Returns by ESG Quintile Portfolios

The results are unambiguous: there is notable dispersion of returns across quintiles. The portfolios in higher ESG quintiles (Q3-Q5) outperformed lower ESG quintiles (Q1-Q2), with the highest quintile growing over 25% relative to baseline. The monotonic relationship amongst adjacent quintiles is not as strong, but this is consistent with a positive correlation between higher ESG scores and long-term returns.

These findings reinforce previous empirical research that suggests ESG principles lead to portfolio resiliency and risk-adjusted returns (13; 31). Strong performance of Q5 specifically during weaker market conditions, establishes a flavor of ESG as downside risk hedge.

4.5. Discussion

The overall results point out three main insights:

1. Environmental issues are the most significant contributor to ESG-based performance, reinforcing the importance of climate-related risk in capital markets.
2. ESG scores are still informative but demonstrate a more nuanced predictive validity that should be applied in conjunction with analysis at the pillar level.
3. Financial benefits of ESG-based portfolio construction are demonstrable and especially pronounced for top quintile firms.

These findings reveal both the strategic justification for ESG integration for investors and policymakers, as well as the methodological value of combining machine learning with traditional regression to address the complexity of ESG-return relations.

5. Conclusion and Future Work

This study offers robust evidence that integrating Environmental, Social, and Governance (ESG) is not only compatible with financial performance but also builds long-term portfolio resilience. By creating quintile portfolios sorted by ESG levels, the analysis shows that firms in the top ESG quintiles consistently outperform firms in low quintiles. The highest ESG portfolio saw greater than 25% cumulative growth and displayed greater resilience during market downturns. Overall, these results confirm ESG strategies provide better long-term outcomes than traditional investments. Furthermore, the stock-level predictions of outperforming the benchmark provided only a modest level of accuracy ($\approx 52-53\%$), indicating that ESG metrics are more suited for portfolio construction and long-horizon strategies than the short-term stock selection.

The analysis of feature importance showed that the Environmental pillar is the single most influential driver of long-term value creation, illustrating the increasing materiality of climate-related risks, regulatory scrutiny, and market-based investor preferences. Overall ESG scores still had some predictive value, but the explanatory power metrics were improved by layering on pillar-level characteristics that generated organizational effectiveness behind more granular ESG research. From a methodological perspective, the work propelled ESG finance forwards by proposing models that are a combination of more traditional regressions and Bayesian & machine learning approaches. While linear regression and XGBoost achieve a near-perfect level of prediction of cumulative returns, Bayesian regression completed the analysis by also producing probabilistic forecasts and estimates of uncertainty a necessary step towards achieving more robust and reliable ESG analytics.

The implications impact several areas. For corporate strategy, corporations that embed sustainability practices in particular environmental performance will enhance their long-term competitive advantage and establish trust with investors. For public policy, the findings indicate that regulatory incentives for sustainable disclosures, accountability, and climate resilience is meaningful. For investment management, the integration of ESG into investment portfolios can be beneficial for performance and at the same time pursue financial solutions

to ethical and sustainability pursuits. More generally, the findings support progress towards the United Nations Sustainable Development Goals: SDG 8 (Decent Work and Economic Growth) through sustainable finance that is still competitive, SDG 12 (Responsible Consumption and Production) through more accountability, and SDG 13 (Climate Action) with the focus on the materiality of environmental performance.

There are various interesting possibilities for future research. First, examining firm-level disclosures on carbon intensity, board diversity, supply-chain transparency, or labor practices could add granularity to the ESG dimensions, enhancing our understanding about mechanisms connecting sustainability and financial performance. Second, more studies across different countries or regions, especially in emerging economies, may enhance the generalizability of findings and indicate variation in regional ESG activity. Third, deeply integrating time-series and deep learning models (i.e. LSTMs, transformers, graph neural nets) may be able to better capture the temporal and structural nuances of ESG–return relationships. Fourth, alternative data, such as textual analysis of sustainability reports, climate risk measures, satellite imaging, or social media sentiment, could enrich predictive modeling and improve understanding in ESG-research. Fifth, scenario-based simulations could help to analyze the implications of interventions (e.g. carbon taxation, green disclosure mandates, or sustainable finance taxonomies), and assess the hypothetical impact at both firm- and portfolio-levels.

Ultimately, an important frontier exists to more clearly connect financial returns to sustainability outcomes in the real economy. Future research could try to assess the direct impact of ESG portfolios on progress toward the SDGs, say, reduced carbon emissions, improved labor standards, or a higher level of accountability for corporate governance. In so doing, researchers can help connect the dots between financial market performance and social impact, thus advancing both scholarship and practices that support sustainable finance.

References

1. European Securities and Markets Authority. (2021). ESG funds in the EU: Performance and resilience during COVID-19. ESMA Report. Retrieved from <https://www.esma.europa.eu>
2. Morningstar. (2021). The resilience of sustainable funds during COVID-19. Morningstar Research Report. Retrieved from <https://www.morningstar.com>

3. Leins, C., Risi, D., & Misani, N. (2022). ESG funds and their financial performance: Evidence from global markets. *Sustainability Accounting, Management and Policy Journal*, 13 (4), 881–904. <https://doi.org/10.1108/SAMPJ-09-2021-0412>
4. Huynh, T., Dao, L., & Pham, P. (2022). Do ESG investments outperform? Evidence from global equity markets. *Finance Research Letters*, 46, 102431. <https://doi.org/10.1016/j.frl.2021.102431>
5. Broadstock, D. C., Chan, K., Cheng, L., & Wang, X. (2020). The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China. *Finance Research Letters*, 38, 101716. <https://doi.org/10.1016/j.frl.2020.101716>
6. Berg, F., Kolbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance*, 26 (6), 1315–1344. <https://doi.org/10.1093/rof/rfac021>
7. Henrique, B., & Sobreiro, D. (2021). A survey of machine learning in finance: ESG applications and portfolio prediction. *Expert Systems with Applications*, 165, 113816. <https://doi.org/10.1016/j.eswa.2020.113816>
8. Krauss, C., Do, X. A., & Huck, N. (2021). Deep learning in ESG forecasting: Evidence from volatility modeling. *Journal of Forecasting*, 40 (5), 870–889. <https://doi.org/10.1002/for.2779>
9. Dorfleitner, G., Utz, L., & Wimmer, N. (2022). Explaining ESG fund performance with SHAP values: Evidence from global equity markets. *Sustainability*, 14 (3), 1257. <https://doi.org/10.3390/su14031257>
10. Koop, G., & Korobilis, D. (2019). Bayesian methods for empirical macroeconomics. *Review of Economic Dynamics*, 31, 1–24. <https://doi.org/10.1016/j.red.2018.12.002>
11. Carriero, A., Clark, T. E., & Marcellino, M. (2020). Bayesian vector autoregressions in finance. *Journal of Applied Econometrics*, 35 (2), 147–161. <https://doi.org/10.1002/jae.2739>
12. Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal*, 38, 45–64. <https://doi.org/10.1016/j.gfj.2017.03.001>
13. Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5 (4), 210–233. <https://doi.org/10.1080/20430795.2015.1118917>

14. Boubaker, S., Cellier, D., Manita, H., & Saeed, L. (2022). Does corporate social responsibility reduce financial distress risk? *Journal of Business Ethics*, 179 (2), 357–379. <https://doi.org/10.1007/s10551-021-04856-9>
15. Bennani, T., Le Guenedal, G., Lepetit, F., Ly, C., Monnier, L., & Sekine, T. (2021). How ESG investing has impacted asset pricing in the equity market. *Journal of Portfolio Management*, 47 (3), 72–92. <https://doi.org/10.3905/jpm.2021.1.227>
16. Hull, J., & White, A. (2020). ESG ratings, ESG disclosures, and the cost of capital. *Journal of Portfolio Management*, 46 (3), 56–74. <https://doi.org/10.3905/jpm.2020.1.161>
17. Zerbib, O. D. (2019). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, 98, 39–60. <https://doi.org/10.1016/j.jbankfin.2018.10.012>
18. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. <https://doi.org/10.1145/2939672.2939785>
19. Breiman, L. (2001). Random forests. *Machine Learning*, 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324>
20. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12 (1), 55–67. <https://doi.org/10.1080/00401706.1970.10488634>
21. Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
22. Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer.
23. Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman and Hall/CRC.
24. Broby, D., & Pate, J. (2022). Artificial intelligence and sustainable investing: Machine learning approaches to ESG analysis. *Finance Research Letters*, 48, 102957. <https://doi.org/10.1016/j.frl.2022.102957>
25. Akpınar, S. (2023). Machine learning for ESG investing: Predicting returns with textual and numerical ESG signals. *Journal of Sustainable Finance & Investment*. Advance online publication. <https://doi.org/10.1080/20430795.2023.2174198>

26. Aldea, A., Andres, C., & Betzer, M. (2023). Does ESG disclosure affect stock market liquidity? Evidence from Europe. *Review of Quantitative Finance and Accounting*, 60 (1), 185–210. <https://doi.org/10.1007/s11156-022-01072-8>
27. Cortez, M., Silva, F., & Areal, N. (2009). The performance of European socially responsible funds. *Journal of Business Ethics*, 87 (4), 573–588. <https://doi.org/10.1007/s10551-008-9959-x>
28. Bansal, P., Kistruck, S., & Majoch, S. (2022). Sustainability in financial markets: Implications for ESG integration. *Organization & Environment*, 35 (2), 187–211. <https://doi.org/10.1177/10860266211049448>
29. Whelan, T., Atz, U., Van Holt, T., & Clark, C. (2021). ESG and financial performance: Uncovering the relationship by aggregating evidence from 1,000 plus studies published between 2015 – 2020. NYU Stern Center for Sustainable Business. Retrieved from <https://www.stern.nyu.edu/experience-stern/faculty-research/esg-and-financial-performance>
30. Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *Review of Financial Studies*, 33 (3), 1067–1111. <https://doi.org/10.1093/rfs/hhz137>
31. Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91 (6), 1697–1724. <https://doi.org/10.2308/accr-51406>