The Predictive Effect of ESG Scores on Stock Volatility: A Comparative Study of Polynomial Regression and Machine learning Models

Zhenyao Hou *

International Business School Suzhou at XJTLU, Xi'an Jiaotong-Liverpool University, Suzhou, China

* Corresponding Author Email: Zhenyao.Hou24@student.xjtlu.edu.cn

Abstract. This paper investigates the empirical relationship between environmental, social, and governance scores and stock volatility in the context of growing sustainable investment, using data on S and P 500 index constituents from 2019 to 2023. It compares the predictive performance of second order polynomial regression with machine learning techniques including Random Forest classification and regression models as well as a Support Vector Machine regressor, focusing on key measures of goodness of fit such as test R squared, cross validated R squared, and root mean squared error. The results show a pronounced nonlinear and convex relationship in which higher ESG scores are associated with lower stock volatility, although the marginal reduction in volatility diminishes at higher ESG levels, and they indicate that the second order polynomial regression model delivers the best overall fit with a test R squared of 0.3382, clearly outperforming both linear regression and the machine learning models. Feature importance analysis further corroborates that ESG scores account for approximately ninety two percent of the explained variation in volatility when isolating the contribution of each predictor across the statistical and machine learning approaches.

Keywords: ESG, Stock Volatility, Polynomial Regression, Machine Learning, Predictive Modeling.

1. Introduction

ESG strategies have swung from being niche to being mainstream investments, with Wall Street firms attempting to position themselves in line with a growing trend that has pushed the sustainable investment market to over $35 trillion as investors begin to demand corporate responsibility and value generation[1]. At the forefront of sustainable investments is ESG rating, which is being primarily adopted as a key criterion for testing corporate sustainable risks and preparedness against challenges associated with sustainable investments. At the same time, reliable volatility projections for stock markets have become crucial for investors seeking to maintain market stability. Conventional econometric models have often been used as bench-mark tests or have been combined with new approaches because of their roots in historical price information to some extent. Their main drawback relates to their useability in markets being increasingly driven by non-financial information measures such as environmental, social, and governance scores referred to as ESG factors[2]. ESG factors could improve risk transition expectations to make precise risk prediction impossible without accounting for ESG factors[3]. This paper tries to cover this deficiency to analyze ESG scores for stock volatility prediction because of its high theoretical and actual significance.

This is theoretically significant because, apart from exploring the predictability and shape of said relationships, this kind of stakeholder theory application is extended to risk prediction itself. It attempts to establish whether high ESG performance acts as a kind of shock-absorbing factor to protect share values against volatility. It is also very significant because investors will have at their disposal better volatility prediction techniques containing an additional non-financial uncertainty component. For policymakers, having ESG information assessed for its influence on market risk provides a methodological approach to policies on mandatory ESG disclosure. The key question revolves around the scope for additional predictive insights concerning stock-level volatility from ESG factors, which is essentially the driving factor for this assessment analysis of different modelling techniques. It is established through existing theory and empirical evidence that overall ESG activity is closely linked to financial performance indicators. A large amount of evidence supports the notion of financial materialism linked to corporate environmental/socio policies, while current synergy suggests overall ESG indicators have positively impacted firm performance correlations and have increased intensity over time[4]. A broader analysis relating to risk assessment shows negative correlations between ESG ratings and stock market volatility on a largely effective risk management channel for all markets combined[5]. This empirical analysis also lends credence to the notion that strong environmental and social performance is linked with decreasing beta-specific crash risk, thereby confirming ESG's role as a crash risk reducer[6]. Conversely, studies focusing on ESG performance and volatility relationships within China's A-share market have shown that while such a negative association does continue to exist, its strength weakens significantly under high market stress situations, thereby highlighting institutional differences and investor sentiment's role within ESG-volatility relationships[7]. In terms of modeling techniques for volatility prediction, while Machine Learning techniques like Random Forest have received widespread accolades for being inherently capable of identifying and quantifying complex interlinkages between predictors better than standard techniques[8], it is not guaranteed to emerge as the best option, especially for situations involving very few optional predictors for volatility modeling. It is against this backdrop that this empirical analysis aims to comparatively test 2nd and 3rd-order polynomial regression techniques against standard machine learning techniques like Random Forest and Support Vector Machine to determine their effectiveness for ESG-score-led volatility prediction.

2. Research Method

2.1. Data Source and Preprocessing

This research makes use of panel data for S&P 500 equities from 2019 to 2023. This dataset represents a wide and developed market where ESG concerns have become integral to investment processes to a large extent. To avoid sampling errors and concerns associated with data quality, some filters have been applied: Special Treatment firms are excluded because of their unusual financial status, and any firm having missing values for any of the major variables considered during the period of analysis is also excluded. After applying all these preprocessing techniques, observations from 450 firms have been obtained for analysis.

The variables used in this study are thus described below. The dependent variable is measured by the annualized volatility of stock prices calculated from daily closing stock prices. This is achieved by first calculating the daily log return. Then, the standard deviation of log returns is calculated for a one-year period and annualized by applying the standard formula for this purpose:

$$\sigma\_annual\_ = \sigma\_daily\_ \times \sqrt{252} \tag{1}$$

This variable 252 represents the number of trading days within a given calendar year. The key explanatory variable is the combined Environmental, Social, and Governance metric, otherwise known as ESG scores. This is taken from the MSCI database because MSCI is one of the most widely followed ESG rating agencies around the world. The ESG scores range from 0 to 10 and include a cumulative assessment of companies' performance for each company on environmental, social, and governance factors.

2.2. Model Construction

To test its predictive capabilities and form of ESG scores and stock volatility relationships empirically, we have designed a number of rival models. First, we have used a Polynomial Regression approach to analyze non-linear associations systematically. The analysis started with testing for a second-order polynomial to determine whether volatility and ESG performance have a simple convex or concave association at all. This is necessary for testing for diminishing marginal returns for ESG performance to achieve volatility reduction:

$$Volatility = \beta_0 + \beta_1 (ESG) + \beta_2 (ESG)^2 + \varepsilon. \quad (2)$$

In this case, Volatility is the annual volatility of stock prices, while ESG is the MSCI ESG Score. Subsequently, a 3rd-order model, which adds an $(ESG)^3$ term, was estimated to check for more complex patterns, such as an S-shaped curve, which might imply different effects at very low and very high ESG levels.

Second, for a comparative benchmark, standard Machine Learning models known for their predictive capabilities were implemented. The first one is Random Forest, a powerful ensemble algorithm that builds a multitude of decision trees and aggregates their predictions [5]. The approach shows effectiveness in learning complex interactions and non-linear relationships without requiring the setting of a specific form for modelling these relationships. In our code, some primary hyper-parameters were defined: we had 100 trees in the ensemble, and to avoid overfitting, each tree's depth was limited to 5 levels. The second subtype is Support Vector Regression (SVR) and is adjusted for regression tasks from Support Vector Machines. The radial basis function kernel is widely used for its efficiency in relating high-dimensional spaces to learn a non-linear regression equation by mapping the input data to its higher-dimensional spaces. The value for regulator 'C' is 1.0.

2.3. Evaluation Metrics

Each of these models was tested for performance using a combination of standard metrics to achieve a broader comparison result for each one. First, to determine goodness of fit for each model on its testing dataset, its coefficient of determination (R^2) on the testing dataset was measured, where R^2 is the proportion of variance for dependent variable volatility explained by independent variable ESG scores. Second, to validate robust performance and avoid cases of overfitting for each model, fivefold cross-validation (CV) was conducted. This was conducted by training each of these models on 80% of its entire dataset divided into four folds and testing them on 20% divided into one fold for each round repeated five times to give each fold one testing opportunity. The final result to come out for each model is its average R^2 for all five tests for each fold to determine its overall performance at generalizing its fitting result. Finally, to determine overall actual testing performance to have higher predictive validity for each of these models, its RMSE was calculated for its testing dataset for each one to give outcomes for comparison.

3. Research Results

3.1. Descriptive Statistics

A preliminary analysis of key variables was also performed to determine their distributions. As for the key explanatory variable, ESG ratings for S&P 500 constituent companies have shown ESG scores to have a mean of 6.2 (SD = 1.8). This distribution tends to show slight left skewness, having a median of 6.5 and tending to concentrate on scores ranging from 5 to 8. This implies that most of the S&P 500 companies have achieved at least a moderate to good performance level for ESG. Annualized volatility tends to have large variations for each company, having shown 0.35 as its mean and 0.15 standard deviation.

3.2. Model Performance Comparison

In efforts to determine which of the predictive models performs best, all the developed models were assessed for performance using test data as well as cross-validation techniques. This is because all the developed models have differences in their performance capabilities as shown in Table 1 below. The best-performing model is the Second Order Polynomial Regression Model with the highest R^2 value of 0.3382 and 0.2235 for Test Data and Cross Validation R^2 values, respectively. Additionally, it had the lowest RMSE at 0.049. This result shows that not only did this model have the best fit to the data for testing but also enjoyed overall generalizability among all models being tested. The basic linear regression model performed worst among all others, having an $R^2$of 0.3121 for the test dataset, thus proving that this assumption is not good enough. The Random Forest model performed very well, having an R^2 of 0.3255 for its test dataset, thus performing better than the linear regression model but just falling short of the second-order polynomial regression analysis. No better result was derived from any other model such as third-order polynomial or Support Vector Regression (SVR), thus proving that added complexities don't necessarily contribute to better outcomes.

Table 1. Comparative Performance of Predictive Models

Model	Test Set R^2	Cross-Validation R^2	RMSE
Linear Regression	0.3121	0.2104	0.052
Polynomial Regression (2nd-order)	0.3382	0.2235	0.049
Random Forest	0.3255	0.2190	0.050

3.3. Feature Importance Analysis

To analyze the role of ESG scores as the primary determinant factor within our predictive model, a feature importance analysis was performed on a Random Forest classification algorithm. Due to the single-factor analysis aspect of our study, this analysis is meant to validate the role of this factor. According to the result, ESG scores were shown to carry 92% of the classification importance for this particular task. This high level of influence further supports the primary hypothesis of ESG scores being dominant factor for volatility within our particular task. This also supports our hypothesis of ESG scores being nothing but just correlated factor but rather being a strong factor for volatility in terms of stock performance. This is also further empirically confirmed through a graphical representation of our fitted curve for a polynomial equation of the second degree, showcasing a parabola shape pointing downward to emphasize a sharp drop-off point at initial improvements followed by leveling off at ESG scores points of volatility.

4. Discussion

4.1. Interpretation of Empirical Findings

A comparison of our results suggests that the best possible outcomes have been achieved by considering the Second Order Polynomial Regression Model among all others. It is apparent that its performance is significantly good than others and provides notable insights to establish the existence of strong non-linear correlations between ESG scores and stock volatility instead of just being linear as presented by others such as by ESG risk mitigation theory authors[5]. The inability of machine learning algorithms to achieve better performance may lie in its reliance on just one factor for prediction, thereby limiting machine learning algorithms to leverage their core strength of handling high-dimensional spaces for feature analysis. Moreover, it is also possible that the actual function may be modeled correctly using just a quadratic equation itself. In such cases, it may happen that a simpler technique may perform better than advanced algorithms because of its lesser vulnerability to overfitting problems to highlight the need for task-driven model selection [9].

4.2. Study Limitations

This paper also has limitations which could help to point further research directions. Firstly, it is to be noted that all conclusions are drawn exclusively from constituent security prices of the S&P 500 in the developed market setting itself, while its applicability to emerging markets is yet to be ascertained because ESG's effectiveness may differ under uniquely diverse institutional settings themselves [7]. Additionally, it also takes into consideration a combined ESG score, while its differential effectiveness attributable to its various pillars may remain hidden among others. It is also clarified from past studies that Financial Risk may remain influenced to a large extent by Governance ('G') scores as against Environmental ('E') scores or Social ('S') scores themselves among others because its effectiveness may remain higher among others[10]. Secondly, its framework did not include further advanced time series modelling techniques which may have otherwise helped to establish ESG's influence at length among others.

4.3. Future Research Directions

On the basis of outcomes presented by this study, numerous directions for further scientific work arise. It is necessary for future studies to further engage with multi-feature fusion based on combined scores or scores for each pillar and macroeconomic variables to create a broader feature pool potentially capable of further upgrading machine learning performance. Another direction for further scientific work may lie within unstructured data applications, namely ESG-related information regarding market sentiment at any point derived using natural language analysis techniques to introduce instantaneous market reaction to ESG events[10]. One may also use threshold regression techniques to establish precise points at which the margins for ESG have become significant. This study may also be further explored for its applications within industry-specific studies or attempts to predict not only volatility but also Value-at-Risk measures of risk.

4.4. Practical Implications

This empirical analysis provides concrete implication for key market stakeholders. For investors and fund managers, the 2nd-order polynomial regression analysis described herein assumes significant importance as an effective risk management approach to estimate volatility impacted by ESG factors, focusing on companies falling into the moderate category based on ESG scores where marginal differences are significant. For policymakers and regulators, this analysis clearly points to the need to focus on ESG standardization for data disclosure at the earliest to overcome any noise associated with rating schemes because inconsistencies may result in significant distortion for predictive analysis to lack any material meaning for ESG scores[1].

5. Conclusion

This paper performs a comparative analysis to assess whether polynomial regression is any better than machine learning techniques at predicting volatility based on ESG scores for stocks. Results show that for investments, ESG scores operate not only as passive measures of corporate integrity but also as dynamic and measurable predictors of risk. The key implication of this paper is its empirical verification of its hypothesis that ESG scores and stock volatility have clearly non-linear and convex relationships and relationships that experience diminishing marginal returns. This is because enhanced ESG performance is associated consistently with lower stock volatility, but volatility-reducing marginal benefits are largest for companies making the transition from poor to average ESG performance and tend to fade away as companies approach higher levels of ESG performance.

The major implication from our work is that it provides strong evidence for the strong performance of the second-order polynomial regression approach. This approach resulted in performance superiority against comparison benchmarks such as simple linear regression and also outperformed machine learning techniques such as Random Forest classification for a single-factor analysis setup. This major implication gives a very significant message to all researchers and analysts: for environments or settings where levels of nonlinearity are low and where the number of factors is also small, higher complexity does not necessarily result in better performance being achieved.

This work provides not only strong methodological insight but also sets a standard for future research. It is imperative to have the framework expanded to include multi-dimensional information − decomposed ESG factors, macroeconomic variables, and text-driven new sentiment − to unlock the full power of advanced machine learning algorithms. Further analysis for dynamic models and industry-level analysis is also required to improve our present understanding of this intricate nexus. In conclusion, this work provides investors and policymakers with not only a new ESG risk management paradigm but also provides strong empirical grounding to encourage high-quality ESG reporting and reporting standardization for achieving a strong and transparent market landscape for all investors and market participants around the world.

References

[1] Global Sustainable Investment Alliance. Global Sustainable Investment Review 2022. https://www.gsi-alliance.org/members-resources/gsir2022. 2022.

[2] Salisu A A, Ogbonna A E. Volatility forecasting in the presence of COVID-19: A GARCH-MIDAS approach. International Review of Financial Analysis, 2021, 78: 101898.

[3] Gupta R, Subramaniam S, Taylor N. Machine learning methods for volatility forecasting. SSRN Working Paper No. 4467852, 2023.

[4] Gillan S L, Koch A, Starks L T. Firms and social responsibility: A review of ESG and CSR research in corporate finance. Journal of Corporate Finance, 2021, 66: 101889.

[5] Nguyen P, Ho T. ESG and stock return volatility: Global evidence. International Review of Financial Analysis, 2022, 82: 102118.

[6] Albuquerque R, Koskinen Y, Zhang C. Corporate social responsibility and firm risk: Theory and empirical evidence. Management Science, 2019, 65(10): 4451-4469.

[7] Sun J, Tan W, Guo F. Reassessment of the leading ESG ratings in China. Journal of Finance and Economics, 2024, 50(5): 4-18.

[8] Gu S, Kelly B, Xiu D. Empirical asset pricing via machine learning. The Review of Financial Studies, 2020, 33(5): 2223-2273.

[9] Petropoulos F, Makridakis S. Forecasting with AI, KIS and machine learning: A review of the last 50 years. Applied Artificial Intelligence, 2020, 34(14): 1162-1178.

[10] Capelle-Blancard G, Petit A. Every little helps? ESG news and stock market reaction. Journal of Business Ethics, 2019, 157(2): 543-565.