The Interplay between Sustainability Initiatives and Capital Structure Dynamics using Machine Learning and Shap Analysis: Are we there yet?

V. M. Maluleke[1*] & J.H. Bührmann[2]

ARTICLE INFO

Article details
Submitted by authors 3 Nov 2024
Accepted for publication 26 Jun 2025
Available online 29 Aug 2025

Contact details
* Corresponding author
vukosiera@hotmail.com

Author affiliations
1 School of Mechanical, Aeronautical and Industrial Engineering, University of the Witwatersrand, Johannesburg, South Africa

2 School of Industrial Engineering, North-West University, Potchefstroom, South Africa

ORCID® identifiers
V. M. Maluleke
https://orcid.org/0009-0008-0923-1645

J.H. Bührmann
https://orcid.org/0000-0003-0657-9933

DOI
http://dx.doi.org//10.7166/36-2-3131

ABSTRACT

This research explores the relationship between sustainability performance and capital structure in industrial and resources firms listed on the Johannesburg and London Stock Exchanges. The study leverages environmental, social, and governance factors to assess their influence on capital structure decisions using machine learning techniques, including Shapley additive explanations analysis. The findings reveal that traditional financial metrics, such as operating profit margin and total assets, are more significant predictors of capital structure than sustainability factors. While environmental, social, and governance factors play a role, their impact on capital structure is limited. The study highlights the importance of integrating sustainable practices with financial performance.

OPSOMMING

Hierdie navorsing ondersoek die verband tussen volhoubaarheidsprestasie en kapitaalstruktuur in nywerheids- en hulpbronmaatskappye wat op die Johannesburgse en Londense Effektebeurse genoteer is. Die studie gebruik omgewings, maatskaplike en korporatiewe bestuursfaktore om hul invloed op kapitaalstruktuurbesluite te evalueer met behulp van masjienleer tegnieke, insluitend die Shapley additiewe verduidelikingsanalise. Die bevindings toon dat tradisionele finansiële maatstawwe, soos bedryfswinsmarge en totale bates, meer beduidende voorspellers van kapitaalstruktuur is as volhoubaarheidsfaktore. Hoewel omgewings, maatskaplike en korporatiewe bestuur faktore 'n rol speel, is hul impak op kapitaalstruktuur beperk. Die studie beklemtoon die belangrikheid van die integrasie van volhoubare praktyke met finansiële prestasie.

1. INTRODUCTION

The debate around the optimal mix of debt and equity financing has remained a central theme for decades in both the academic literature and corporate practice [1]. Despite significant research, no universal agreement exists about the ideal capital structure or a standardised method for determining the appropriate debt ratio [2]. Some researchers suggest a linear link between financial indicators and capital structure, while others propose non-linear relationships [3].

Decisions about the proportion of debt and equity and efforts to maintain a target leverage ratio continue to be key elements of corporate finance strategy [4]. Determining the right balance between debt and equity remains a focus of intense study, as achieving this balance is essential for firms aiming for long-term financial health. Overleveraged companies may resort to drastic cost-cutting – often at employees' expense – when facing financial distress [5].

Poor capital structure choices can lead to severe financial consequences, including distress and bankruptcy, especially if optimisation is overlooked [6]. The lack of a clear consensus on optimal capital structure complicates recommendations on ideal debt levels, which often vary by firm depending on their specific leverage targets and financial position [7].

Empirical studies on capital structure have largely relied on statistical models to investigate the factors influencing variations in corporate debt ratios. However, limited research explores non-linear models as alternative approaches to capturing the complexity of capital structure dynamics [8].

Numerous theories attempt to explain the determinants of capital structure, yet the significant variation between firms over time remains difficult to account for fully [9]. The mix of debt and equity financing is central to maximising a firm's value [10], and the academic debate typically focuses on three major theories: the trade-off theory, the pecking order theory, and the market timing theory.

Although these frameworks offer important conceptual insights, empirical research has struggled to differentiate clearly between the applicability and limitations of each, leaving the field with unresolved questions about the key drivers of capital structure [11]. The absence of a definitive method for determining target leverage ratios further fuels the ongoing search for optimal capital structure strategies [12]. In response, this study applies causal inference techniques to reveal the underlying causal links between capital structure decisions and other financial variables.

The seminal work of [13] pioneered examining capital structure's impact on firm value, laying the groundwork for subsequent theoretical developments. Their theorem, which emphasises the financing choices between debt and equity, underscores the allure of debt financing for firms that seek to capitalise on debt tax shields [6]. Building on this foundation, the trade-off theory proposes that capital structure reflects a delicate balance between the tax benefits of debt and the costs of bankruptcy [9].

However, this theory faces criticism of its assumptions about debt discipline and agency problems. In contrast, the pecking order theory offers insights into firms' financing preferences, highlighting information asymmetries between firms and investors as a key determinant of capital structure [6, 10]. This theory suggests that firms prioritise internal financing, such as retained earnings, over external debt financing, reflecting a cautious approach driven by asymmetrical information.

Overall, the pursuit of understanding capital structure dynamics remains an ongoing endeavour, with scholars continuing to grapple with the complexities of corporate financing decisions and their implications for firm value. This study aims to contribute to the scientific discourse on and research into the impact of sustainability performance, as measured by environmental, social, and governance (ESG) ratings, on a firm's capital structure. By exploring this relationship, it seeks to enrich the discussions of relevant theoretical frameworks and their potential influence on the interplay between key variables.

Specifically, the study intends to shed light on how ESG ratings influence various capital structure theories, elucidating firms' financing preferences based on internal funds and debt, and how these preferences may shift given varying ESG ratings. In addition, the research endeavours to deepen the understanding of how capital structure theories are affected by ESG considerations. Through empirical investigation and analysis, this study aims to advance knowledge in corporate finance and sustainability, offering valuable insights for academics and practitioners.

2. REVIEW OF THE LITERATURE

2.1. The theories of capital structure

Capital structure, a fundamental aspect of corporate finance, elucidates how companies procure funds to fuel their growth and operations [14]. Senior managers and shareholders continually grapple with questions about the optimal debt to equity financing ratio and its impact on firm value. Numerous research endeavours have investigated how to understand a firm's capital structure, recognising its pivotal role in maximising financial value [6].

The three prominent theories mentioned earlier – the trade-off theory, the pecking order theory, and the market timing theory – delineate the determinants of a firm's capital structure, despite extensive exploration, empirical evidence remains inconclusive across various global markets [15]. Modigliani and Miller [13] pioneered the examination of capital structure's effect on firm value, positing the trade-off theory. Initially, they contended that, in perfect capital markets, capital structure holds no sway over a firm's value. However, subsequent iterations of their theory acknowledged the tax benefits of debt financing [7].

The pecking order theory, proposed by [10], proposes that firms prioritise financing sources based on information asymmetries, preferring internal over external financing. Market timing theory, introduced by [16], suggests that firms adjust their financing choices based on prevailing market conditions, opting for debt when equity is overvalued and vice versa. Harris and Raviv [17] synthesised the capital structure literature, highlighting the factors influencing leverage ratios. They concluded that, while various theories offer insights, empirical nuances remain largely unaddressed.

Key variables influencing capital structure decisions in listed firms in the United States have been identified as using advanced model selection techniques [18]. Empirical investigations into capital structure theories have been extended through factor-analytic methods to address measurement difficulties [19]. The impact of leverage on stakeholder perceptions – particularly among employees and suppliers – has also been explored [5]. Recent innovations in capital structure analysis include machine learning techniques, which offer promising avenues for understanding complex financial dynamics. The potential of machine learning to overcome the limitations of traditional statistical methods has been emphasised [20].

This paper contributes to the discourse on capital structure determination, leveraging machine learning to unearth nuanced insights. By incorporating new variables and sophisticated analytical techniques, we aim to deepen understanding and refine existing theories.

2.2. The need to incorporate environmental, social and governance factors

Incorporating ESG factors into corporate strategies is increasingly recognised as not just a moral imperative but also a source of competitive advantage and positive returns. Research by [21] supports this notion, demonstrating a positive correlation between strong ESG performance and financial performance. Furthermore, Orlitzky *et al.* [22] conducted a meta-analysis that indicated that firms with superior corporate social performance tend to outperform their peers financially over the long term.

A major driver of firms' adoption of ESG practices is the potential to enhance their financial outcomes. Research suggests that strong corporate governance structures are linked to better financial performance, with firms that demonstrate sound ESG practices often experiencing superior returns [23]. This is supported by findings from Dimson *et al.* [24], who observed that companies that effectively address ESG issues tend to generate positive abnormal returns. Beyond financial metrics, ESG integration also contributes to building corporate reputation – an increasingly valuable asset in a market that values social responsibility. According to Lozano [25], adopting a comprehensive approach to sustainability not only strengthens internal practices but also appeals to ethically minded consumers, ultimately reinforcing brand loyalty and competitive positioning.

Moreover, integrating ESG principles into business operations can result in significant efficiencies and cost reductions. As noted by Lozano [25], firms that emphasise sustainability often innovate in ways that lower resource use, streamline supply chains, and cut waste, thus ultimately reducing their operational costs. In parallel, ESG performance is gaining importance among institutional investors. Flammer [26] shows that companies that actively engage with ESG issues are more likely to attract capital from socially responsible investors and institutional funds, thereby improving their financial resilience and funding opportunities.

Overall, embedding ESG considerations in corporate strategy can yield a range of advantages, from stronger financial performance and improved reputation to operational efficiency and enhanced access to capital. By committing to sustainability, companies not only support broader societal goals but also secure a competitive edge in the increasingly ESG-conscious global economy.

2.3. Environmental, social and governance ratings

Environmental (E), social (S) and governance (G) ratings serve as crucial indicators of a firm's ability to navigate ESG risks and opportunities, having a direct impact on financial performance. These ratings offer valuable insights for investors, guiding their investment decisions. Comprising three distinct components – E-rating, S-rating, and G-rating – ESG ratings provide a holistic view of a company's sustainability efforts. Even if a firm has a low E-rating, robust S- and G-ratings can compensate, leading to a favourable combined ESG rating. Each rating encompasses specific criteria and benchmarks for assessment, reflecting the company's performance in environmental stewardship, social responsibility, and corporate governance [27].

The E-rating focuses on environmental factors, including resource use, emissions reduction, and innovation. Companies are evaluated based on their efficiency in resource management and emission reduction strategies, as well as their innovation in environmental technologies and processes. The S-rating assesses social sustainability under four key themes: workforce practices, human rights adherence, community engagement, and product responsibility. Last, the G-rating evaluates governance practices related to management integrity, shareholder rights, and integrating financial, social, and environmental considerations into corporate decision-making processes. Collectively, these ratings provide investors with a comprehensive understanding of a company's sustainability performance and its potential for long-term financial success [27].

2.4. Causality

Causality remains crucial in our quest to gain intellectual understanding of the universe and of its contents. It examines the establishment of cause-and-effect relationships. In the realm of model interpretation, it is imperative for a model to offer explanations that resonate with decision-makers and that accurately portray the underlying reasons behind its decisions, including a coherent interpretation of cause-and-effect analyses [28]. Presently, explainable artificial intelligence (XAI) models, particularly those designed to decipher pre-trained black box models (i.e., post hoc model-agnostic models), focus on constructing models on local interpretations, offering approximations to the predictive black box [28].

Finding causal relationships between features and predictions in observational data poses a formidable challenge, and is a crucial step in elucidating predictions [28]. In observational data analysis, the counterfactual framework is rooted in the assumption that well-defined causal states exist that can transparently explain all members of the population of interest [29].

Causal effects are elucidated through comparisons of outcomes related to exposures to alternative causal states. Let Y represent the target variable, namely the firm's capital structure. We propose that there are clearly defined causal situations that kickstart the process of understanding causality within the counterfactual framework. We do this by identifying potential outcomes for random variables that cover all individuals in the population we are studying [13]. For a binary cause, we denote the potential outcome random variables as Y^0 and Y^1.

Furthermore, let Y_i^0 denote the potential outcome in the control state for individual i, and Y_i^1 denote the potential outcome in the treatment state for individual i. This notation signifies the individual-level causal effect of the treatment. To assess counterfactuals, a causal model encompassing all pertinent information about changes in the antecedent must be established. The individual-level causal effect (Δ_i) is defined as the difference between the potential outcomes:

$$\Delta_i = Y_i^1 - Y_i^0, \tag{1}$$

In addition, the individual-level causal effect can be defined as the difference between the expectations of individual-specific random variables:

$$E[Y_i^1] - E[Y_i^0], \tag{2}$$

Here, E[.] denotes the expectation operator from probability theory, capturing the disparity in the expectation of capital structure between treatment and control states. For a binary cause with two causal states and associated potential outcome variables Y^0 and Y^1, and a corresponding causal exposure variable Z specified to assume two values (0 for control, 1 for treatment), the outcome variable Y is defined as follows:

$$Y = Y^0 \text{ if } Z = 0,$$

$$Y = Y^1 \text{ if } Z = 1. \tag{3}$$

2.5. Structural causal models

Structural causal models (SCMs) provide a mathematical framework for representing and reasoning about causal relationships among variables. SCMs use graphical models to depict causal dependencies, offering an intuitive visualisation in which variables are depicted as nodes, and the relationships between variables are illustrated as edges in a graph. This graphical representation provides a clear and insightful way to understand the causal connections among various elements in the model.

2.5.1. Mathematical framework

Let X be a set of variables $\{X_1, X_2, \ldots, X_n\}$, each representing a specific aspect of the system under consideration. Associate each variable with a structural equation that expresses its value as a function of its direct causes. For X_i, this can be denoted as:

$$X_i = f_i(\mathrm{Pa}_i, U_i), \tag{4}$$

where Pa_i represents the set of direct causes (parents) of X_i, and U_i is an exogenous variable representing unobserved influences. The exogenous variables, denoted as $\{U_1, U_2, \ldots, U_n\}$, represent external influences that are not influenced by other variables in the model.

2.6. Shapley additive explanations

The Shapley additive explanations (SHAP) method, introduced by [30], uses game theory principles to explain individual predictions. Drawing from the concept of Shapley values in economics and game theory, SHAP offers a method to allocate a game's payoff among a group of players equitably. This concept can be directly applied to XAI approaches.

In SHAP, a prediction task can be compared to a game in which each feature value of the instance serves as a player collaborating to achieve a gain. This gain represents the disparity between the Shapley value of the prediction and the average of the Shapley values across feature values of the instance being explained.

One of the distinguishing features of SHAP is its assignment of importance values to each feature for a specific prediction. Notably, SHAP introduces a novel class of additive feature importance measures, and provides theoretical insights that show the existence of a unique solution in this class that possesses desirable properties [28].

In SHAP, an explanation model $g(z_0)$ is given by a linear combination of Shapley values ϕ_j of a feature j with a coalitional vector, $\mathbf{z}'$, of maximum size M.

$$g(z') = \varphi_0 + \sum_{j=1}^{M} \varphi_j z'_j \tag{5}$$

To distribute the payoff among a finite group of participants in a collaborative game equitably, SHAP relies on four fundamental fairness principles. The first principle is additivity, which dictates that the total contributions must sum up to the overall game outcome. The second principle, symmetry, ensures that a player who contributes more to the game cannot receive less reward than those who contribute less. The third principle, efficiency, mandates that predictions should be accurately attributed to the corresponding feature values. Finally, the fourth principle, dummy, asserts that a feature with no influence over the outcome should be assigned a Shapley value of zero [28].

To assess the pre-existing model f when only a subset S of features are part of the model, a conditional expected value formulation was used to remove the other features. The formulation is shown in Eq(6) [28]:

$$E[f\left(\boldsymbol{X} \middle| X_{(s=x_s)}\right]. \tag{6}$$

3. DATA AND METHODOLOGY

3.1. Data sampling

The data sample consisted of non-financial firms listed on the Johannesburg Stock Exchange (JSE) and the London Stock Exchange (LSE), including their E, S and G factors. The JSE data had 164 records, and the LSE data had 490 records. The data also included the Global Industry Classification Standard (GICS) sector and industry names of the firms. Financial firms were excluded owing to differences in accounting standards from those used in resources, telecommunications, and industrial firms; those sectors adhere to distinct reporting and accounting standards, leading to differences in capital structure.

To address the issue of missing data, imputation was used, adopting the mean of observed values for variables with data gaps. This approach was selected for its simplicity and effectiveness in preserving the overall statistical properties of the data.

3.1.1. Rationale for mean imputation

1. **Preservation of data distribution:**
   The mean is a measure of central tendency that accounts for all observed values, providing a balanced estimate representative of the overall dataset. By imputing missing values with the mean, the imputation minimises the distortion of the data distribution, especially in cases where the data is symmetrically distributed.
2. **Simplicity and ease of interpretation:**
   Mean imputation is computationally straightforward, making it a practical choice in many research scenarios. The mean is a commonly understood statistical measure, which aids in the transparency and interpretability of the data-processing steps.

Below is the list of variables that formed part of the dataset.

Table 1: List of Variable Names and Their Types

Variable Names	Type
Company common name	Independent
GICS Sector name	Independent
GICS Industry name	Independent
Country of incorporation	Independent
Leverage ratio	**Dependent**
ESG score	Independent
Environmental pillar score	Independent
Social pillar score	Independent
Governance pillar score	Independent
Company market cap	Independent
S&P rating	Independent
Tangible assets	Independent
Price to book value per share	Independent
Net income after minority interest	Independent
Dividend payout ratio	Independent

3.2. Data splitting

The dataset was divided into a training set (X_{train}, y_{train}) and a test set (X_{test}, y_{test}), with X_{train} containing $r \times m$ samples and X_{test} containing $(1 - r) \times m$ samples, where r is the chosen ratio.

3.3. Training model - random forest model

The prediction model was trained using the random forest algorithm. The random forest model is initialised with hyperparameters, such as the number of trees $n_{\text{estimators}}$ and the maximum depth of each tree $\max_{\text{depth}}$.

3.4. Construction of decision trees

For $k = 1$ to $n_{\text{estimators}}$:

1. **Bootstrap sampling**: For each iteration k, a bootstrap sample $X^k_{bootstrap}$ of size m' is drawn from the original training dataset X_{train}. This sampling method ensures that each tree in the ensemble is trained on a slightly different subset of the data, which promotes model diversity. Mathematically, the bootstrap sample can be expressed as:

$$X^k_{bootstrap} = \{ x_i, x_{i_j} . . . , x_{jm}' \}, \qquad \mathbf{x_{ij}} \sim X_{\text{train}}, \tag{7}$$

   where x_{i_j} represents data points sampled with replacement from the training set X_{train}.

2. **Decision tree growth:** A decision tree T_k is then grown using the bootstrap sample $X^k_{bootstrap}$. The tree is constructed by recursively partitioning the data based on feature values that minimise a chosen impurity measure. For each node, the optimal split is determined by evaluating potential splits θ [31]:

$$Split(X^k_{bootstrap}) = \arg\min[\text{Impurity}(X_{left}, X_{right})], \tag{8}$$

   where impurity is a function that measures the impurity (e.g., Gini impurity or entropy) of the resulting subsets X_{left} and X_{right} after the split. For example, the Gini impurity for a split is given by [31]:

$$Gini(Xleft, Xright) = \frac{|X_{left}|}{|X^k_{bootstrap}|}\text{Gini}(X_{left}) + \frac{|X_{right}|}{|X^k_{bootstrap}|}\text{Gini}(X_{right}), \tag{9}$$

   where $Gini(X) = 1 - \sum_c {p_c}^2$ and p_c is the proportion of class c in the node. The growth process continues until the tree reaches a predefined maximum depth ($\max_{\text{depth}}$) or other stopping criteria, such as a minimum number of samples per leaf node.

In this context, the term “estimators” refers to the individual decision trees that are constructed as part of the ensemble method. The ensemble consists of multiple decision trees, each trained on a different bootstrap sample of the training data. The number of estimators, denoted by $n_{\text{estimators}}$, represents the total number of decision trees in the ensemble. Each tree T_k contributes to the final model by providing a vote or prediction; these are then aggregated to produce the final prediction.

3.5. Aggregation of predictions

The predictions from each decision tree in the ensemble are aggregated to form the final prediction for each input x_j. This aggregation can be done in different ways, depending on whether the task is classification or regression. For a regression task, the final prediction $\widehat{y_J}$ for a given input x_j is obtained by averaging the predictions from all trees [28]:

$$\widehat{y_J} = \frac{1}{n_{estimators}} T_k(x_j), \tag{10}$$

where $\widehat{y_J}$ is the predicted value for sample j, and $T_k(x_j)$ is the prediction of the k^{th} tree for sample j. For a classification task, the final class label $\widehat{y_J}$ is determined by taking the majority vote among the predictions of all trees [28]:

$$\widehat{y_j} = \mathrm{mode}\left\{T_k(x_j) \mid k = 1, \ldots, n_{\text{estimators}}\right\}. \quad (11)$$

The process of aggregating predictions leverages the diversity of the individual trees to reduce variance and to improve overall model performance. The final model benefits from the ensemble learning approach by achieving a more robust and accurate prediction than any single decision tree.

3.6. Feature importance analysis

Feature importance was assessed using the random forest regressor, an ensemble learning method that aggregates multiple decision trees to enhance predictive performance. The process involved the following steps:

1. **Model training**: The random forest regressor was trained on the dataset, with encoded features and target values. Each tree in the ensemble contributed to predictions by averaging outputs from individual trees.

2. **Visualisation**: The computed importance scores were visualised using horizontal bar plots, facilitating the identification of key predictors. These plots displayed features that were sorted by their impact on model performance.

3.7. Feature importance thresholding

A threshold-based approach was used to optimise feature selection and enhance model interpretability for feature importance. Initially, feature importance scores were computed to quantify each feature's impact on the model's accuracy. A threshold value of 0.01 was established to filter out less significant features, retaining only those with importance scores above this cutoff. Applying this threshold isolated the most influential features, streamlining the model and focusing the analysis on key predictors. This approach simplified the model, improving its clarity and performance. This methodology provided a robust method for evaluating the significance of the predictors, aiding in identifying the most influential factors and enhancing the interpretability of the model.

4. RESULTS

4.1.1. Johannesburg stock exchange results

Figure 1 depicts the ranking of the predictors used to elucidate capital structure, in which the latter served as the dependent variable, while E, S, G, and other financial variables served as independent variables. The determination of feature importance was conducted through the random forest regressor model. The figure illustrates how these predictors had an impact on the predictability of capital structure.

The feature importance analysis revealed that, while sustainability factors such as environmental and social considerations held importance, their predictive value in capital structure decisions was comparatively lower than other financial metrics. Operating profit margin (OPM), total assets, and company name emerged as the top predictors, overshadowing sustainability metrics in their influence. Although sustainability variables such as the environmental pillar score, ESG score, and social pillar score held notable significance, their lower feature importance scores suggested that they did not play as dominant a role in shaping capital structure decisions as traditional financial metrics.

Consequently, while these sustainability factors indicated a growing awareness of ESG considerations in financial decision-making, their lower predictive value suggested that they might not yet be the primary drivers of capital structure choices. This nuanced understanding underscores the ongoing integration of sustainability initiatives into financial strategies, albeit at a pace that may not fully reflect their purported importance in the current landscape of financial decision-making.

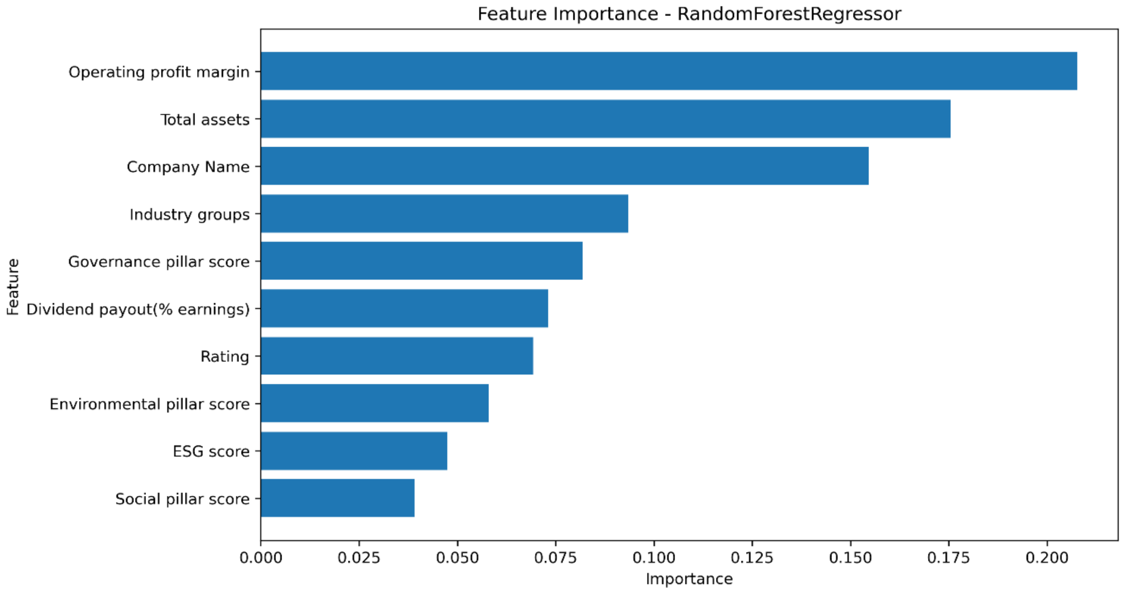

Figure 1: A feature importance plot for capital structure predictors - JSE.

4.1.2. Interpretation of Johannesburg Stock Exchange results using Shapley additive explanations

To interpret the model's predictions, the SHAP model was applied to interpret the predictions of the trained random forest regressor model. The SHAP values quantified the impact of financial variables, including each E, S, and G feature, on the model's predictions. Figures 2 – 5 depict various plots that illustrate the results from the SHAP values for predictors in the JSE dataset. OPM, total assets, industry groups and company name emerged as pivotal factors influencing capital structure decisions, as highlighted by their dominant presence in the force plot and their consistently high SHAP values in instances in the SHAP mean, beeswarm, violin and waterfall plots.

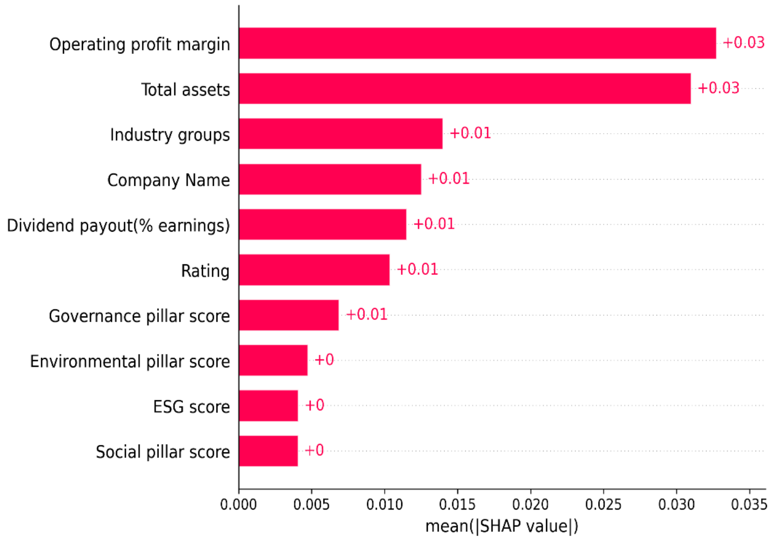

Figure 2: A mean SHAP value plot of the financial predictors - JSE.

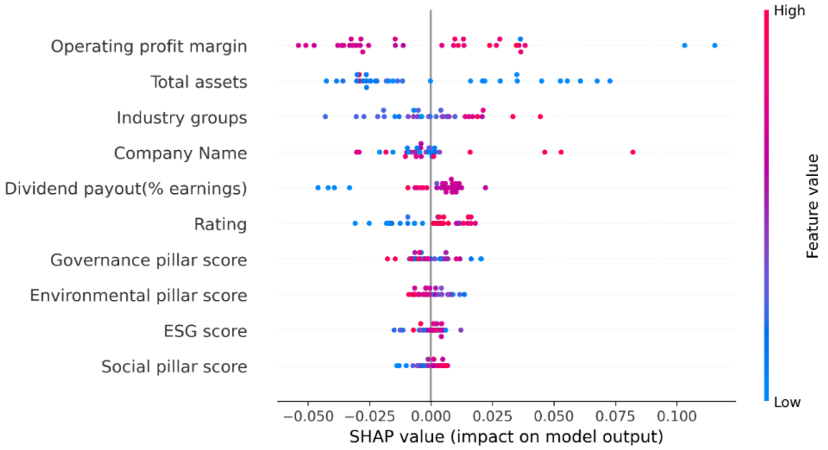

Figure 3: A SHAP value view of the financial predictors using a beeswarm plot - JSE.

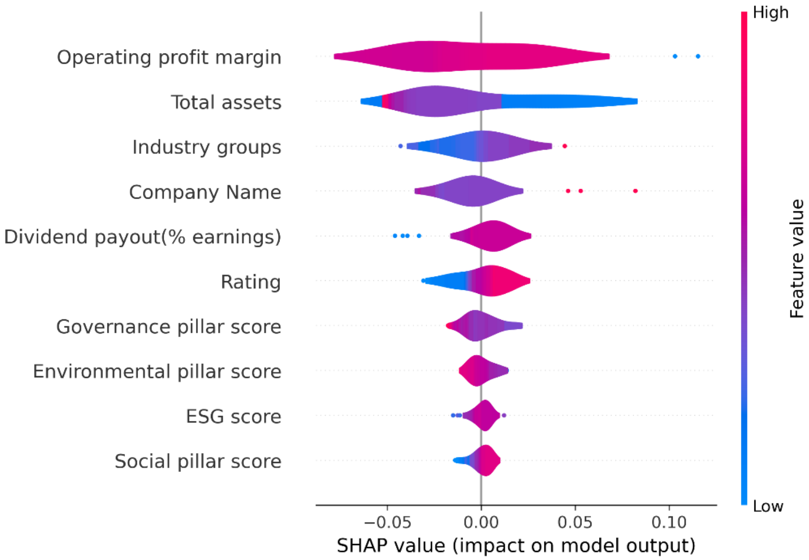

Figure 4: A SHAP value view of the financial predictors using a violin plot - JSE.

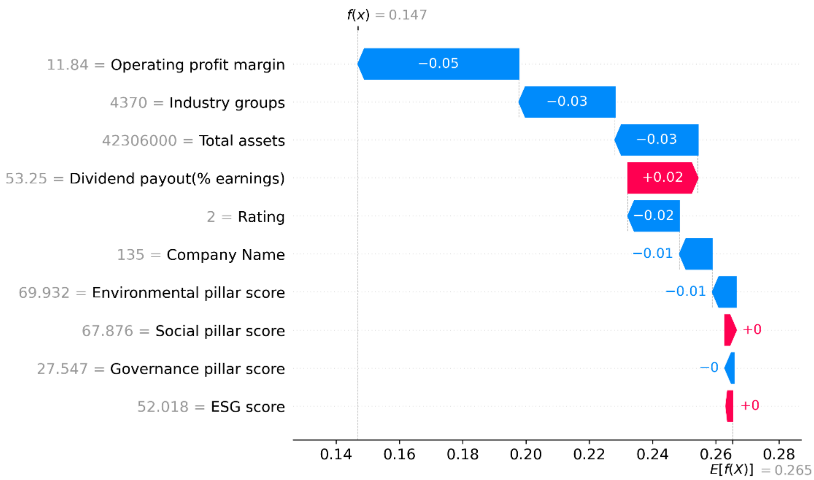

Figure 5: A SHAP value view of the financial predictors using a waterfall plot - JSE.

4.1.3. Interpretation of London Stock Exchange results using Shapley additive explanations

To elucidate the model's predictions, the SHAP framework was used to interpret the outcomes of the trained random forest regressor model. A similar analytical approach, akin to the one conducted for the JSE, was undertaken, focusing on the LSE's industrial, resources, and technological sectors. The objective was to compare the findings and to scrutinise potential associations between ESG factors and capital structure. The investigation concentrated on the financing behaviours of companies listed on the LSE. Figure 6 provides the relative importance of the different features in influencing the model's output.

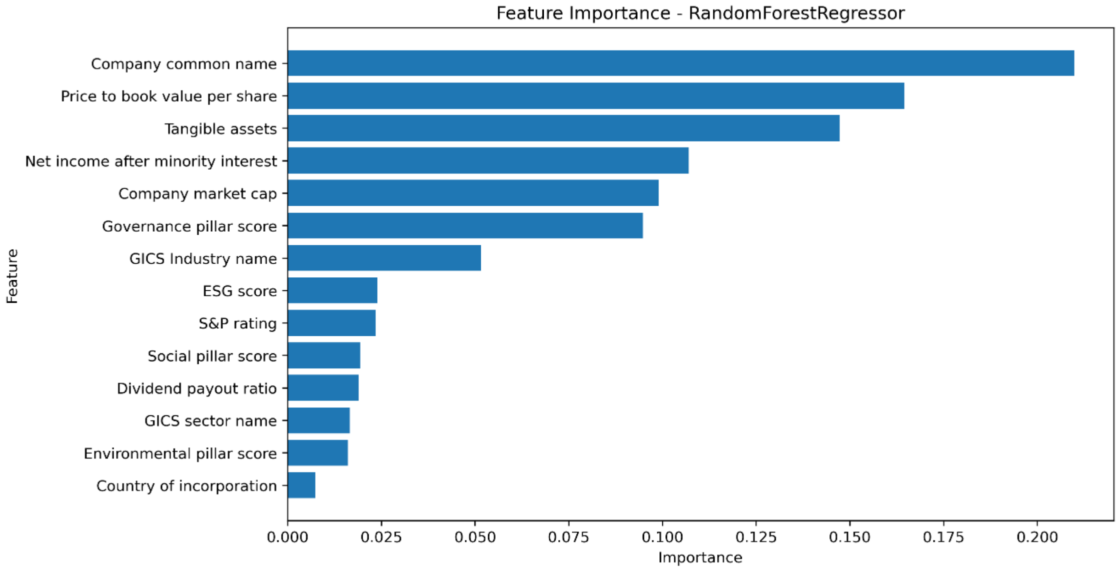

Figure 6: The feature importance plot of the capital structure predictors – LSE.

Figures 7 – 10 depict various plots that illustrate the results from the SHAP values for predictors in the LSE dataset. Tangible assets, company market cap, net income after minority interest and company common name were the most influential factors in capital structure decisions based on the SHAP mean, beeswarm, violin and waterfall plots. The results revealed that, consistent with those of the JSE, the sustainability initiatives did not significantly influence financing practices in advanced economies.

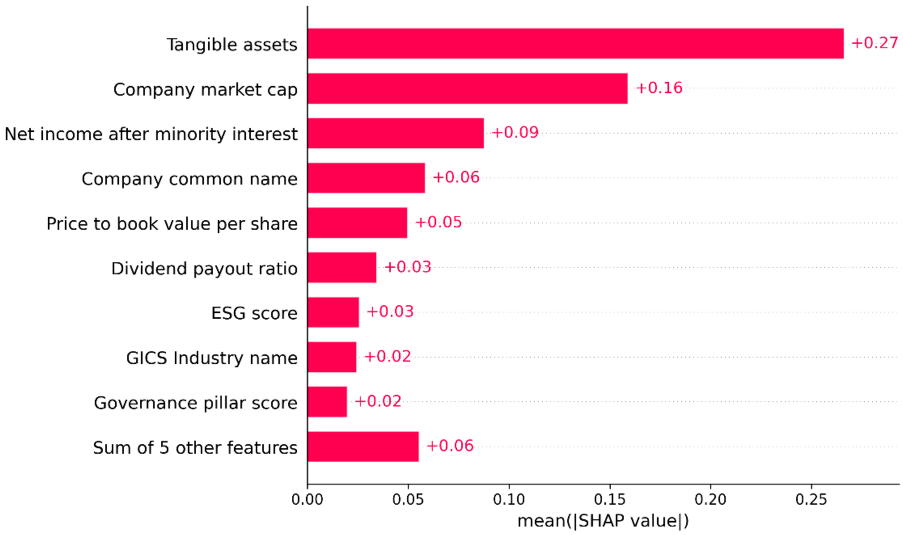

Figure 7: A mean SHAP value plot of the financial predictors – LSE.

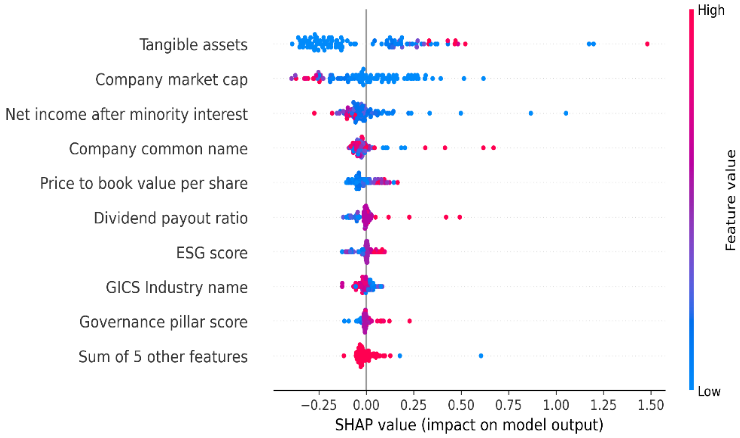

Figure 8: A SHAP value view of the financial predictors using a beeswarm plot - LSE.

Figure 9: A SHAP value view of the financial predictors using a violin plot - LSE.

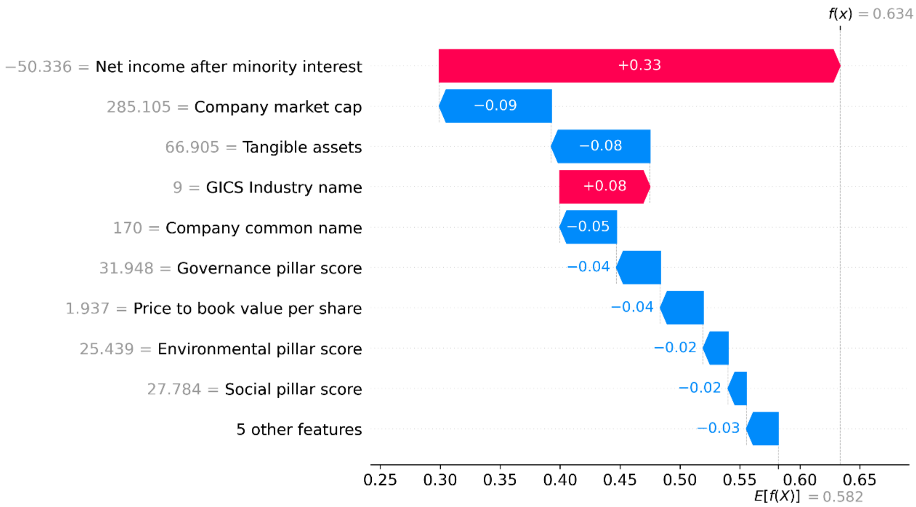

Figure 10: A SHAP value view of the financial predictors using a waterfall plot - LSE.

Conversely, while present, sustainability factors such as the environmental pillar score, the ESG score, and the social pillar score demonstrated comparatively lower contributions and exhibited greater variability in their impact, suggesting a less pronounced role in shaping capital structure decisions. These findings underscore the continued significance of conventional financial metrics in guiding capital structure dynamics, with sustainability considerations playing a crucial and varied role in the decision-making process.

4.1.4. Statistical analysis of environmental, social, and governance factors for Johannesburg Stock Exchange

Based on the findings of this study, it was observed that there was no discernible strong correlation between the ESG factors and the capital structure of JSE-listed industrial, resources, and technological firms. This is clear from a plot of the actual vs predicted values, shown in figure 11. As a result, it could be concluded that sustainability initiatives do not notably sway financing practices in these sectors. However, it is crucial to emphasise the broader benefits of sustainability efforts, as they contribute positively to objectives aligned with the sustainable development goals of the United Nations (UN).

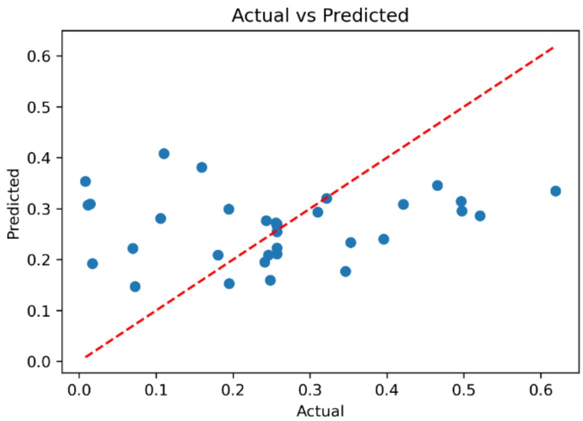

Figure 11: A plot of the actual vs predicted values - JSE.

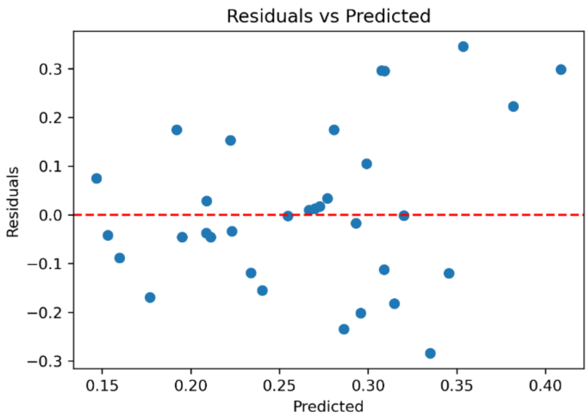

Figure 12: A plot of the residuals vs predicted values - JSE.

In addition, an analysis of the residuals vs features plot – a method used to visualise potential relationships between residuals and predictors – did not reveal any significant associations among the environmental (E), social (S), and governance (G) factors (see figure 12). This corroborates the earlier conclusion that these factors do not substantially influence the predictability of capital structure among JSE-listed firms.

A residual plot of each E, S, and G predictive factor was conducted, see figures 13 - 15. The plots were used as a diagnostic tool to assess the relationships between the E, S, and G independent variables and the residuals. The independence of the residuals ensured that the model's predictions were accurate and reliable for the range of the independent variable. If the residuals were dependent on the predictors, it would have meant that the model's errors differed from the score, leading to inconsistent and potentially misleading predictions. Figure 16 provides a plot of the overall residuals vs ESG score for the JSE dataset. Similar to figures 13 – 15, this figure also indicated no significant relationship between the residuals and the ESG score.

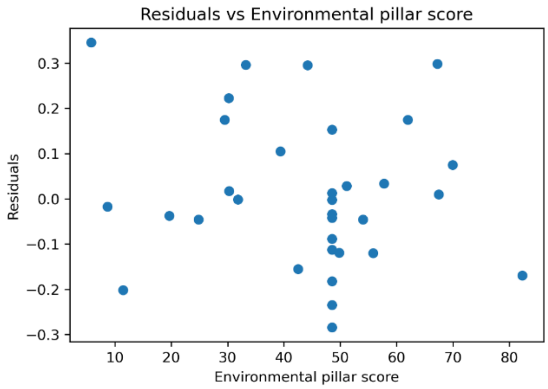

Figure 13: A plot summarising the residuals vs environmental score - JSE.

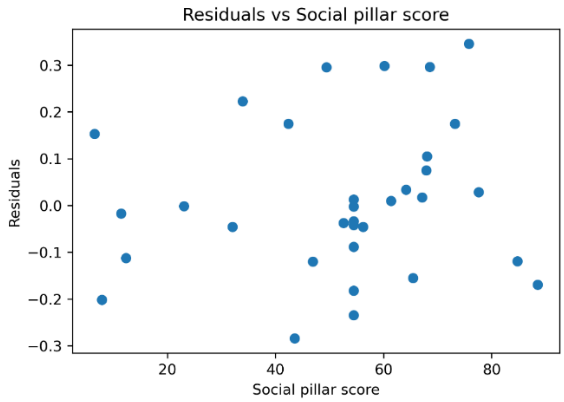

Figure 14: A plot summarising the residuals vs social score - JSE.

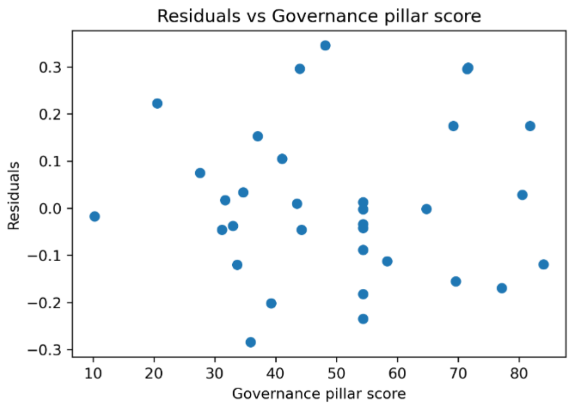

Figure 15: A plot summarising the residuals vs governance score - JSE.

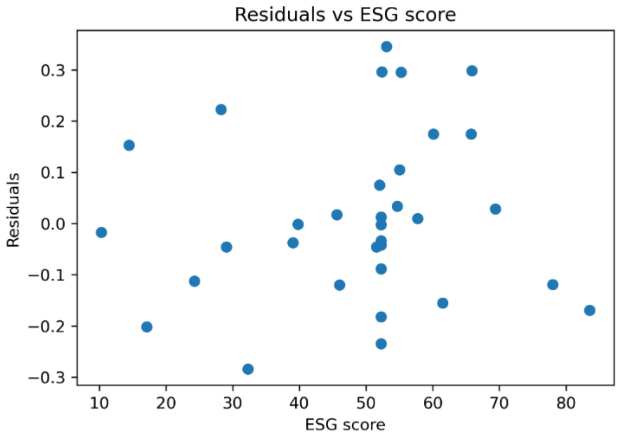

Figure 16: A plot summarising the overall residuals vs ESG score – JSE.

4.1.5. Statistical analysis of environmental, social, and governance factors for the London Stock Exchange

Interestingly, while sustainability factors may lack predictive power, financial variables contribute more substantially to the predictability of capital structure. Notably, an increase in predictability performance was observed in the LSE dataset, compared to that of the JSE dataset. There were also fewer outliers than observed in the JSE data; the actual vs predicted plot for the LSE (figure 17) shows a better prediction than that for the JSE market (figure 11).

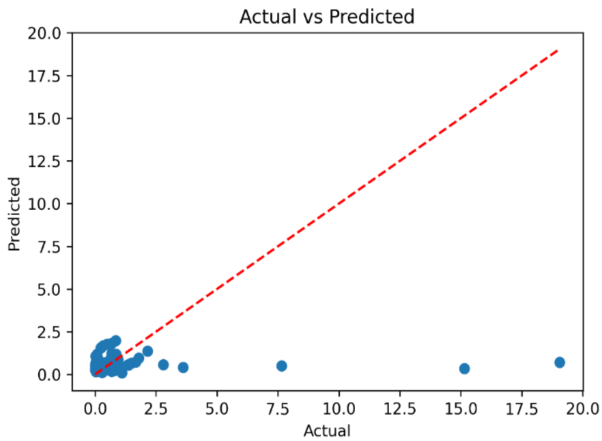

Figure 17: A plot of the actual vs predicted values - LSE.

Similar to the JSE dataset, the residuals vs features plot for the LSE dataset, illustrated in figure 18, did not reveal any significant associations among the E, S, and G factors. However, the picture does not completely rule out the existence of dependency. Nonetheless, this corroborates the earlier conclusion that these factors do not substantially influence the predictability of capital structure among LSE-listed firms.

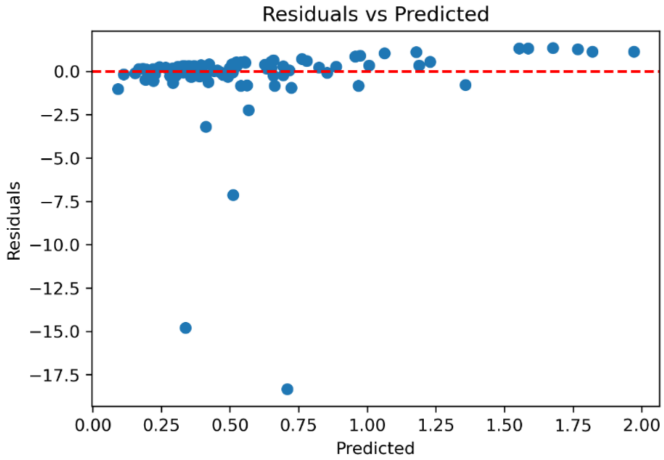

Figure 18: A plot of the residuals vs predicted values - LSE.

The residual plots using E, S, and G, as well as the overall ESG score as a predictive factor for the LSE dataset are illustrated in figures 19 – 22. The plots indicate that majority of the time the residual values are close to zero, suggesting that sustainability factors had low predictive power for the LSE dataset.

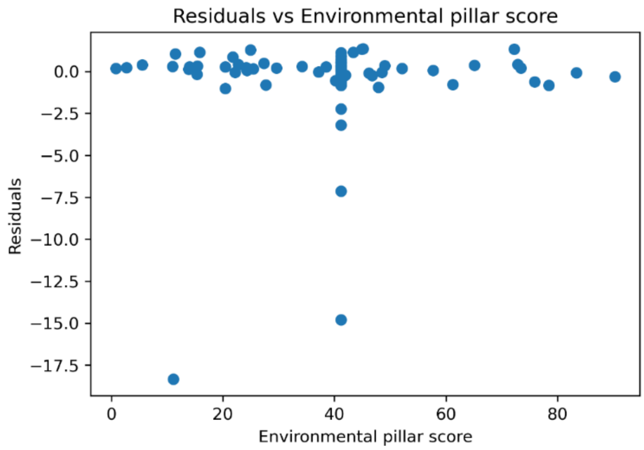

Figure 19: A plot summarising the residuals vs environmental score – LSE.

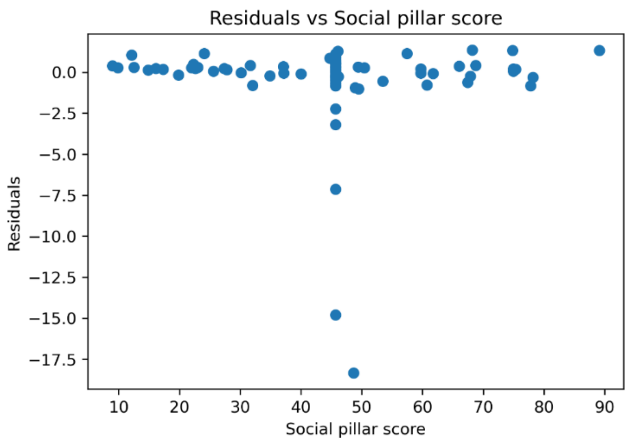

Figure 20: A plot summarising the residuals vs social score – LSE.

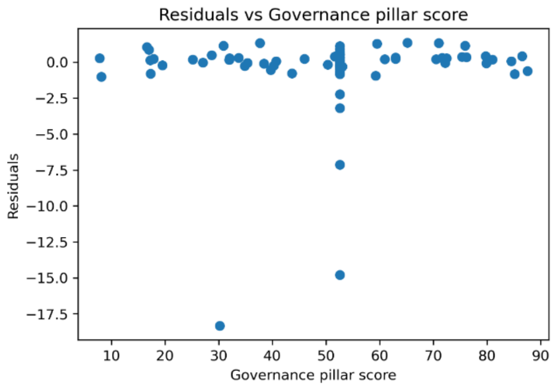

Figure 21: A plot summarising the residuals vs governance score - LSE.

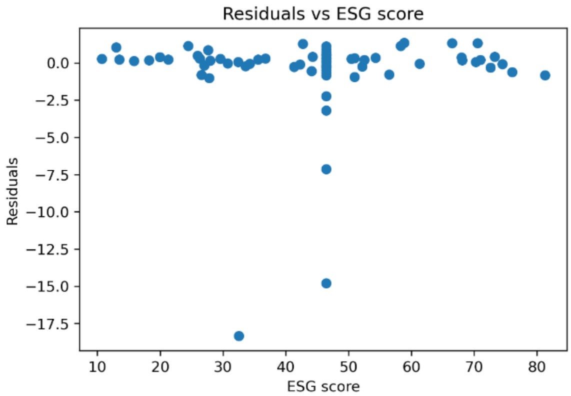

Figure 22: A plot summarising the overall residuals vs ESG score - LSE.

5. SUMMARY OF RESEARCH FINDINGS

This paper investigated the relationship between sustainability performance and capital structure, using ESG ratings and leverage ratios. The study focused on publicly listed companies in the JSE and the LSE, specifically targeting the resources, industrial, and technological sectors. This focus was because of their shared legislative framework on sustainable finance, which aligns with the Paris Agreement and the UN's climate change requirements.

5.1. Key findings and conclusions

1. **Lack of clear relationship**: The analysis did not find a clear relationship between sustainability initiatives and optimal capital structure. It concluded that sustainability performance does not significantly influence a firm's ability to raise capital or to change its capital structure.
2. **Encouragement of sustainability initiatives**: Despite the lack of influence on capital structure, sustainability initiatives should still be encouraged for their broader beneficial effects and alignment with global climate action goals.
3. **Theoretical alignment**: The research has contributed to the field by analysing results through the lenses of trade-off theory, pecking order theory, and agency theory. The findings align somewhat with traditional financial theories, particularly the trade-off and pecking order theories.
4. **Feature importance analysis**:
    - **JSE**: Financial metrics such as OPM, total assets, industry groups and company name emerged as the top predictors of capital structure for the JSE dataset. While significant, sustainability factors such as the environmental pillar score, the ESG score, and the social pillar score had lower predictive value than the traditional financial metrics.

- **LSE**: Tangible assets, company market cap, net income after minority interest and company common name were the most influential factors in capital structure decisions based on the SHAP analysis. Similar to the pattern observed in the JSE dataset, financial variables were observed to play a more substantial role in predicting capital structure than sustainability factors.

5.2. Implications

1. **Current financial sector priorities**: The results reflected the current priorities in the financial sector, highlighting that sustainability factors are not yet primary drivers of capital structure decisions. This underscores the need for a greater integration of sustainable objectives with financial profitability to make sustainability a more critical consideration in investment decisions.
2. **Ongoing transition**: The findings indicated a growing awareness of ESG considerations in financial decision-making. However, the lower predictive value of sustainability factors suggests that the transition towards fully integrating these initiatives into financial strategies is ongoing, and may not yet reflect their full importance.

6. CONCLUSION

This study has used advanced machine learning techniques and a rigorous feature importance analysis to investigate the interplay between sustainability performance and capital structure in publicly listed firms. The approach integrated SHAP values to interpret the influence of ESG factors on the leverage ratio. The analysis focused on firms listed on the JSE and the LSE, paying particular attention to the resources, industrial, and technological sectors.

The methodology involved several key steps. First, the data was preprocessed, including handling missing values through mean imputation and encoding categorical variables to prepare the dataset for machine learning models. A random forest regressor model was trained to predict the leverage ratio based on the ESG scores and other financial metrics. The SHAP values were then calculated to quantify the impact of each predictor on the leverage ratio, providing a nuanced understanding of feature importance.

The key findings from this study reveal that, while ESG factors are increasingly acknowledged in financial decision-making, their direct impact on capital structure decisions remains limited. Specifically, the results indicate that traditional financial metrics such as OPM, tangible or total assets, industry groups, company name and company market cap are significantly more predictive of capital structure than ESG factors. This suggests that, despite growing attention to sustainability, these traditional metrics continue to play a dominant role in shaping capital structure decisions.

The analysis has also highlighted that sustainability initiatives, although not a primary driver of capital structure, align with broader corporate social responsibility and regulatory expectations. Therefore, while ESG factors do not yet substantially influence capital structure, they contribute to the overall strategic alignment of firms with global sustainability goals.

For future research, several directions are proposed. First, further investigation into the long-term effects of integrating ESG factors into financial strategies could provide deeper insights into their potential impact on capital structure over time. This would include studying how evolving regulatory frameworks and increasing investor pressure might influence the role of sustainability in financial decision-making.

In addition, expanding the research to encompass a wider range of industries and geographic regions could yield a more comprehensive understanding of the global implications of ESG factors for capital structure. Finally, exploring alternative methodologies and advanced machine learning techniques could enhance the precision and depth of feature importance analysis, offering more refined insights into the interaction between sustainability performance and financial outcomes.

Overall, this research underscores the need for continued exploration and adaptation of financial strategies to incorporate sustainability considerations, with the ultimate goal of achieving a balance between financial performance and responsible investment practices.

REFERENCES

[1] **K. Karlsen and N. Mathisen**. Capital structure and machine learning techniques in Scandinavia. Master's thesis, Norwegian University of Science and Technology, Faculty of Economics and Management, NTNU Business School, 2021.

[2] **W. Tarantin, Jr and M. Ribeiro do Valle**. *Capital structure: The role of the funding sources on which Brazilian listed companies are based*. Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Departamento de Contabilidade, 2015.

[3] **S. C. Myers**. The capital structure puzzle. *The Journal of Finance*, 39(3):575–592, 1984.

[4] **S. Amini, R. Elmore, Ö. Öztekin, and J. Strauss**. Can machines learn capital structure dynamics? *Journal of Corporate Finance*, 70:102073, 2021.

[5] **V. Maksimovic and S. Titman**. Financial policy and reputation for product quality. *The Review of Financial Studies*, 4(1):175–200, 1991.

[6] **L. J. Chen and S. Y. Chen**. How the pecking-order theory explains capital structure. *Journal of International Management Studies*, 6(3):92–100, 2011.

[7] **J. H. Binsbergen and J. R. Graham**. An empirical model of optimal capital structure. *Journal of Applied Corporate Finance*, 23(4):34-59, 2011.

[8] **H. T. Pao**. A comparison of neural network and multiple regression analysis in modeling capital structure. *Expert Systems with Applications*, 34(3):1607–1615, 2008.

[9] **H. Ai, M. Z. Frank, and A. Sanati**. The trade-off theory of corporate capital structure. *Oxford research encyclopedia of economics and finance*. Oxford University Press, UK, 2021.

[10] **S. C. Myers and N. S. Majluf**. Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2): 187-221, 1984

[11] **T. C. Wilson**. *Value and capital management: A handbook for the finance and risk functions of financial institutions*. Wiley, Hoboken NJ, 2015.

[12] **F. Giglio**. The capital structure through the Modigliani and Miller model. *International Business Research*, 15(11):11-16, 2022.

[13] **F. Modigliani and M. H. Miller**. The cost of capital, corporation finance, and the theory of investment. *The American Economic Review*, 48(3):261–297, 1958.

[14] **T. T. Gaytan**. AI-based prediction of capital structure: Performance comparison of ANN, SVM, and LR models. PhD thesis, Tecnológico de Monterrey, Monterrey, Mexico, 2022.

[15] **Y. Zhu**. Capital structure: The case of firms issuing debt. *Australian Journal of Management*, 37(2):283–295, 2012.

[16] **M. Baker and J. Wurgler**. Market timing and capital structure. *Journal of Finance*, 57(1):1–32, 2002.

[17] **M. Harris and A. Raviv**. The theory of capital structure. *The Journal of Finance*, 46(1):297–355, 1991.

[18] **M. Z. Frank and V. K. Goyal**. Capital structure decisions: Which factors are reliably important? *Financial Management*, 38(1):1–37, 2009.

[19] **S. Titman, R. Wessels, J. Franks, D. Mayers, R. Masulis, and W. Torous**. The determinants of capital structure choice. *The Journal of Finance*, 43(1), 1988.

[20] **Y. Goeletsis, C. Papaloukas, T. Exharos, and C. Katsis**. Bankruptcy prediction through artificial intelligence, in Information Resources Management Association (ed.), *Machine learning: Concepts, methodologies, tools and applications.* IGI Global Scientific Publishing, 2012, pp. 684-693.

[21] **R. G. Eccles and G. Serafeim**. The impact of corporate sustainability on organizational processes and performance. *Management Science*, 59(5):1045–1061, 2013.

[22] **M. Orlitzky, F. L. Schmidt, and S. L. Rynes**. Corporate social and financial performance: A meta-analysis. *Organization Studies*, 24(3):403–441, 2003.

[23] **J. D. Margolis and J. P. Walsh**. Misery loves companies: Rethinking social initiatives by business. *Administrative Science Quarterly*, 48(2):268–305, 2003.

[24] **E. Dimson, O. Karakas, and X. Li**. Active ownership. *The Review of Financial Studies*, 28(12):3225–3268, 2015.

[25] **R. Lozano**. A holistic perspective on corporate sustainability drivers. *Journal of Cleaner Production*, 81:166–177, 2015.

[26] **C. Flammer**. Does corporate social responsibility lead to superior financial performance? A regression discontinuity approach. *Management Science*, 61(11):2549–2568, 2015.

[27] **L. Lindkvist and O. Saric**. Sustainability performance and capital structure: An analysis of the relationship between ESG rating and debt ratio. Master's thesis, Umeå University, 2020.

[28] **Y. L. Chou, C. Moreira, P. Bruza, C. Ouyang, and J. Jorge**. Counterfactuals and causability in explainable artificial intelligence: Theory, algorithms and applications. *Information Fusion*, 81:59–83, 2022.

[29] **C. Winship and S. L. Morgan**. The estimation of causal effects from observational data. *Annual Review of Sociology*, 25(1):659–706, 1999.

[30] **S.M. Lundberg and S.I. Lee**. A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30: 4765–4774, 2017.
[31] **L. Breiman, J. Friedman, R.A. Olshen and C.J. Stone**. Classification and regression trees. Chapman and Hall/CRC, 2017.