

Article

Explainable AI in Corporate Finance: A SHAP-Based Approach to Enhancing Transparency

Aryan Amrollah Majdabadi ¹ and Hamid Mostofi ^{2,*} ¹ Department of Electrical Engineering and Computer Science, Technische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany; aryan_majdabadi@yahoo.com² Department of Work, Technology and Participation, Technische Universität Berlin, Marchstraße 23, 10587 Berlin, Germany

* Correspondence: mostofidarbani@tu-berlin.de

Abstract

The increasing adoption of machine learning models in corporate valuation has substantially improved predictive accuracy but at the cost of interpretability, a critical limitation in regulated financial environments. This study investigates whether SHAP (Shapley Additive Explanations) can systematically enhance the transparency of XGBoost based valuation models and examines whether the resulting insights extend those of classical linear regression. An empirical analysis was conducted on a cross-sectional dataset of U.S. publicly listed firms (2018), including more than 200 financial indicators. After systematic preprocessing and a hybrid feature selection procedure combining XGBoost importance, mutual information, and correlation based filtering, both an OLS regression and an XGBoost model were trained and validated. XGBoost achieved substantially higher predictive performance ($R^2 = 0.654$ vs. 0.364), while SHAP values provided transparent global and local explanations of model decisions. Both models consistently identified earnings before tax and EV to sales as primary value drivers; however, SHAP additionally showed nonlinear effects, threshold behaviors, and context dependent interactions, particularly for EBITDA margin, share buybacks, and asset based indicators, that remain undetectable in linear models. These findings confirm that SHAP significantly enhances model transparency and generates economically meaningful insights beyond classical regression, supporting its application in auditable, regulatory compliant financial modeling.

Keywords: explainable artificial intelligence; XAI; SHAP; corporate valuation; XGBoost; machine learning interpretability



Academic Editors: Jungho Baek and Darjan Karabasevic

Received: 6 June 2026

Revised: 29 August 2026

Accepted: 9 September 2026

Published: 14 September 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The application of machine learning has fundamentally transformed corporate valuation in recent years. In particular, tree-based ensemble models such as XGBoost enable the modeling of complex, non-linear relationships between financial indicators and firm value, often achieving substantially higher predictive accuracy than traditional regression approaches (Dwivedi et al., 2023).

However, this increase in predictive performance comes at the cost of interpretability. Unlike linear models, the contribution of individual variables in complex machine learning models cannot be directly interpreted in economic terms. In a financial context, this lack of transparency constitutes a critical limitation, as valuation models must not only be accurate but also interpretable and compliant with regulatory requirements (Cohen et al., 2021).

Against this backdrop, the field of Explainable Artificial Intelligence (XAI) has gained increasing relevance (Notovich et al., 2023). XAI methods aim to make complex model decisions more transparent and to reconcile predictive performance with interpretability (OECD, 2022). A prominent approach in this context is SHAP (Shapley Additive Explanations), which is grounded in cooperative game theory and enables a consistent decomposition of model predictions into individual feature contributions (Chan et al., 2025).

This study investigates the extent to which SHAP can systematically enhance the transparency of machine learning based valuation models and examines the additional insights that SHAP based explanations provide relative to classical linear regression approaches.

1.1. Corporate Valuation

Corporate valuation aims to determine the economic value of a firm as a basis for investment, financing, and transaction decisions. In practice, three principal valuation approaches have emerged: income-based methods such as the Discounted Cash Flow (DCF) model, market-based multiples, and regression-based approaches (Felden & Graffius, 2024).

The DCF method, which discounts expected future cash flows, is widely regarded as the theoretical benchmark in valuation practice (Steiger, 2010). Establishing a transparent and systematic link between operational corporate performance drivers and final value creation has long been recognized as a cornerstone of value-based governance and valuation theory (Copeland et al., 1992). In contrast, multiples-based approaches derive firm value from the market prices of comparable companies, providing a pragmatic and market-oriented perspective (Liu et al., 2002; Suozzo et al., 2001). Regression-based models extend this framework by empirically modeling the relationship between firm value and its underlying financial drivers (Götze et al., 2023; Wooldridge, 2016).

Despite their widespread use, these classical approaches exhibit several inherent limitations. DCF models rely heavily on assumptions regarding future cash flows and discount rates, making them sensitive to estimation uncertainty and potential bias (Damodaran, 2002; Fernandez, 2002). Multiples-based methods are constrained by the comparability of peer groups and are susceptible to market distortions (Andreas, 2007; Liu et al., 2002). While linear regression models offer a high degree of interpretability, they are inherently limited in capturing complex and non-linear relationships among financial variables (Majka, 2024; Wooldridge, 2016). Quantitative financial analysis has historically relied on these linear paradigms and structural asset-pricing assumptions championed by early capital market frameworks (Malkiel, 2018), which often fail to account for the volatile and asymmetric states of modern asset markets.

These structural limitations highlight the need for alternative approaches that can both capture complex data patterns and maintain economic interpretability (Steiger, 2010). In the context of increasing data availability, data-driven methods have therefore gained growing importance in corporate valuation.

1.2. Machine Learning in Finance

The increasing availability of large and heterogeneous financial datasets has significantly accelerated the adoption of machine learning methods in corporate valuation. In contrast to traditional linear models, machine learning approaches are capable of capturing complex and non-linear relationships between financial indicators and firm value, thereby enabling improved predictive performance (OECD, 2022; Wasserbacher & Spindler, 2022).

In the broader financial domain, these algorithmic advancements have driven a paradigm shift within Financial Planning and Analysis (FP&A), moving core corporate functions away from retrospective reporting toward forward-looking, analytically driven steering mechanisms. Beyond corporate valuation, the practical deployment of machine

learning spans credit risk modeling, fraud detection, algorithmic trading, and dynamic risk management (OECD, 2022), where models consistently outperform classical time-series approaches (Wasserbacher & Spindler, 2022).

In particular, tree-based ensemble methods such as gradient boosting have proven to be especially effective in this context. By sequentially combining multiple decision trees, these models are able to capture high-dimensional interactions among financial variables while simultaneously reducing bias and variance (Dixon et al., 2020). A prominent implementation of this approach is XGBoost, which incorporates regularization and robust training mechanisms to achieve strong predictive performance, particularly in structured financial datasets (Chen & Guestrin, 2016; Dixon et al., 2020).

Mechanistically, XGBoost extends classical boosting architectures by integrating explicit L_1 (Lasso) and L_2 (Ridge) regularization parameters directly into its objective function, preventing the excessive weighting of individual financial metrics and mitigating overfitting (Dixon et al., 2020). Additional hyperparameters, such as learning rate (shrinkage) (Dixon et al., 2020) and row/column sub-sampling (subsample and colsample) (Chen & Guestrin, 2016), further reduce variance when dealing with volatile and highly correlated accounting variables (Dixon et al., 2020). Crucially, the optimization process utilized by XGBoost leverages both first-order (gradient) and second-order (hessian) derivatives of the loss function (Chen & Guestrin, 2016). This curvature information accelerates convergence and stabilizes tree splits over heterogeneous corporate data structures (Chen & Guestrin, 2016). In contrast to linear regression methods that rely on a closed analytical solution (Wooldridge, 2016), the gradient boosting principle consists of building decision trees sequentially, where each new tree corrects the residuals of the previous models (Dixon et al., 2020). Furthermore, XGBoost provides a distinct operational advantage in empirical finance due to its capability to natively handle missing values. During training, the algorithm endogenously determines an optimal default splitting direction for missing values based on loss reduction, preserving model integrity without requiring subjective imputation frameworks during the tree-building phase (Chen & Guestrin, 2016).

However, this increase in model complexity introduces a fundamental limitation: the resulting models lack an explicitly interpretable functional form, making it difficult to attribute predictions to individual input variables (Arrieta et al., 2019). In the context of corporate valuation, this lack of interpretability constitutes a critical challenge, as valuation models must not only deliver accurate predictions but also provide economically meaningful explanations, which comply with regulatory requirements. This systemic conflict mirrors a classic epistemological divergence in modern statistics between data modeling cultures focused on simple parametric inference and algorithmic modeling cultures optimized for predictive mapping over complex patterns (Breiman, 2001).

This inherent tension between model complexity and interpretability motivates the integration of explainable methods and provides the foundation for the application of Explainable Artificial Intelligence (XAI) in data-driven valuation approaches.

1.3. Explainable Artificial Intelligence

Building on this tension between predictive power and interpretability, the field of Explainable Artificial Intelligence has emerged as a direct response to the opacity of complex models in regulated financial environments, where transparency and accountability are not optional but required (Von Eschenbach, 2021).

Explainable Artificial Intelligence (XAI) comprises methods designed to make the decision processes of complex models more transparent, thereby bridging the gap between predictive performance and interpretability. Beyond technical transparency, XAI

aims to provide explanations that are comprehensible, verifiable, and tailored to different stakeholder groups (Amarna et al., 2025; OECD, 2022).

In the financial domain, XAI serves several key functions. It reduces information asymmetries, supports regulatory compliance, and enables accountability in data-driven decision-making processes. Moreover, it facilitates the identification of potential biases and enhances the robustness of model-based decisions, thereby strengthening trust among investors, regulators, and other stakeholders (OECD, 2022).

Methodologically, XAI approaches can be categorized along several dimensions. A fundamental distinction is made between intrinsically interpretable models and post hoc explanation methods. While the former provide transparency by design, the latter generate explanations for complex black-box models without altering their internal structure (Arrieta et al., 2019; Dwivedi et al., 2023). In addition, a distinction can be drawn between global explanations, which describe overall model behavior, and local explanations, which focus on individual predictions. In financial applications, local explanations are particularly relevant, as they enable case-specific interpretability (Minh et al., 2022).

Despite these advances, many XAI methods exhibit limitations in terms of stability, consistency, and fidelity. In particular, approximation-based approaches may produce simplified or inconsistent explanations, which can limit their applicability in highly regulated environments (Rudin, 2018; Salih et al., 2025).

These limitations highlight the need for more robust and theoretically grounded explanation methods and motivate the use of approaches such as SHAP, which are discussed in the following section.

1.4. Shapley Additive Explanations

The theoretical foundation of SHAP is rooted in cooperative game theory, specifically in the Shapley values introduced by Shapley (1953), which were originally developed to ensure a fair allocation of collective gains among participants. In the context of machine learning, each feature is interpreted as a “player” contributing to the overall model prediction. The objective is to provide a consistent and fair attribution of the model output to individual features by considering all possible feature combinations (Lundberg & Lee, 2017).

A central property of SHAP is the additive decomposition of model predictions into individual feature contributions. Each prediction can be expressed as the sum of a baseline value and the marginal contributions of individual variables. This additive structure enables an intuitive yet theoretically grounded interpretation of complex models by representing predictions as a combination of distinct feature effects (Lundberg & Lee, 2017; Salih et al., 2025).

A key advantage of SHAP over alternative XAI methods lies in its axiomatic foundation. The resulting attributions satisfy essential properties such as consistency and local accuracy, ensuring that explanations are both stable and uniquely determined (Lundberg & Lee, 2017). In contrast to approximation-based approaches such as LIME, SHAP provides a consistent and comparable decomposition of model predictions.

For tree-based models such as XGBoost, the TreeSHAP algorithm enables an efficient and exact computation of Shapley values. By exploiting the structure of decision trees, feature contributions can be computed in polynomial time without relying on stochastic approximations (Chan et al., 2025). This results in deterministic and reproducible explanations, which is particularly important in financial contexts characterized by strict requirements for auditability and transparency.

Beyond its theoretical advantages, SHAP offers practical benefits. It supports both global and local interpretability and allows for the analysis of interaction effects between variables. This makes SHAP particularly well-suited for complex, non-linear financial

data, where interdependencies between features play a crucial role (Chan et al., 2025; Salih et al., 2025).

However, SHAP is not without limitations. In the presence of highly correlated features, attributions may be distributed across multiple variables, complicating interpretation. Additionally, the choice of baseline values can influence the resulting explanations, and out-of-distribution data may lead to unstable or misleading attributions (Salih et al., 2025; Slack et al., 2019).

Despite these limitations, SHAP represents a theoretically grounded and practically robust approach for systematically enhancing the transparency of complex machine learning models in financial applications.

1.5. Research Gap

Despite the growing application of machine learning models in corporate valuation, a fundamental tension between predictive performance and interpretability remains unresolved. Empirical studies across different financial domains show that tree-based machine learning models outperform traditional regression approaches in predictive accuracy. In particular, ensemble methods such as random forests have been shown to significantly outperform linear models in both asset pricing and corporate valuation contexts, particularly in forecasting key financial variables such as operating cash flows (Götze et al., 2023; Heidbrink, 2025).

Existing literature on explainable artificial intelligence has introduced a variety of post-hoc methods to address this limitation. However, many of these approaches rely on approximations of the underlying model behavior, which raises fundamental concerns regarding their fidelity and reliability. In particular, post-hoc explanation methods may fail to accurately reflect the true decision logic of complex models, thereby limiting their trustworthiness in high-stakes domains such as finance (Rudin, 2018). Furthermore, in the presence of highly correlated variables, a common characteristic of financial datasets, methods such as SHAP may yield misleading feature attributions that capture statistical associations rather than underlying causal relationships (Chan et al., 2025).

Although SHAP has emerged as a theoretically grounded method based on cooperative game theory that ensures consistent and locally accurate feature attributions, empirical evidence on its practical contribution to transparency in financial valuation contexts remains limited, particularly given known challenges such as feature correlation and instability in real-world financial data (Chan et al., 2025; Lundberg & Lee, 2017).

More importantly, current research lacks a systematic comparison between SHAP-based explanations of complex machine learning models and the interpretability of classical linear regression approaches. While traditional linear regression models are valued for their transparency and readily interpretable functional relationships, they are inherently limited in capturing complex, non-linear interactions. In contrast, advanced machine learning models can better represent such relationships but at the cost of interpretability (Götze et al., 2023). Conversely, SHAP provides local and global explanations for complex models, yet it remains unclear to what extent these explanations align with established financial value drivers or provide additional insights beyond traditional methods.

This gap highlights the need for a structured and empirical evaluation of whether explainable machine learning methods can not only improve transparency but also generate economically meaningful interpretations that are comparable to, or extend beyond, classical approaches.

1.6. Research Question and Hypotheses

Against this backdrop, the central research question of this study is formulated as follows:

To what extent can SHAP values enhance the transparency and interpretability of machine learning models, using XGBoost as a case study in corporate valuation, and how do the resulting insights compare to those obtained from linear regression models?

To address this question, the study uses an XGBoost model trained on a cross sectional dataset of U.S. listed firms, and applies SHAP to generate both global and local explanations, as well as systematically compares the resulting feature attributions against those derived from an OLS regression baseline.

Based on the theoretical considerations and existing literature, the following objective and hypotheses are derived:

Research Objective 1 (RO1). *To evaluate the extent to which SHAP values enhance model transparency in corporate valuation by assessing whether feature attributions yield economically consistent, non-linear threshold behaviors and interaction effects that align with established valuation theory.*

H2a. *The primary top-rank value drivers identified by SHAP align with those derived from classical OLS linear regression.*

H2b. *For secondary financial indicators, SHAP feature attributions diverge from OLS feature importance rankings by capturing non-linear threshold dynamics and context-dependent interactions that are flattened in linear models.*

To operationalize this comparative inquiry, financial variables are categorized ex-ante: Primary value drivers are defined as top-tier fundamental metrics that structurally anchor corporate valuation across corporate finance theory (specifically core operating earnings and direct market-pricing multiples, expected to dominate the top ranks), whereas secondary financial indicators encompass context-dependent metrics (such as operating margins, capital return policies, and asset buffers).

2. Methodology

2.1. Dataset

The empirical analysis is based on a financial dataset, which contains annual financial statement data and market information for publicly listed U.S. companies. The dataset comprises 225 financial and market-related variables for several thousand firms in the year 2018. These variables are derived from standardized balance sheets, cash flow statements, and income statements, thereby covering key dimensions of firm performance.

The primary dependent variable is market capitalization (Market Cap), which represents the market-based valuation of a firm. This measure is particularly suitable for addressing the research question, as it reflects both forward-looking market expectations and fundamental financial indicators (Andreas, 2007).

Following data import, non-numeric variables, text-based identifiers, and purely descriptive features were removed, as they do not contribute to quantitative modeling.

To ensure the robustness of the model estimation, a systematic analysis of missing values was conducted. Variables with more than 40% missing values were excluded, as their inclusion would lead to unstable and unreliable model estimates. Similarly, observations with more than 30% missing values were removed (Notovich et al., 2023).

Following primary missing value and target validation, the dataset retained $N = 3954$ observations across 216 features. Deduplication (removing exact and quasi-duplicate

attributes) reduced the feature space to 189 attributes while retaining the full sample of $N = 3954$ observations. Subsequent holdout partitioning yielded a training cohort of $N_{train} = 3163$ (80%) and an out-of-sample evaluation cohort of $N_{test} = 791$ (20%).

After filtering for valid observations, remaining missing values were imputed using median imputation. This method was chosen due to its robustness to outliers and its independence from distributional assumptions (Wooldridge, 2016).

Finally, all numerical variables were converted to float format to ensure numerical stability during subsequent data processing and model estimation.

2.2. Data Preprocessing and Preparation

Following the initial data cleaning, the dependent variable (Market Cap) was defined and the dataset was prepared for modeling. Observations with missing or negative values in the target variable were excluded, as they do not allow for meaningful economic interpretation.

Subsequently, a logarithmic transformation of the target variable was applied to reduce the pronounced right skewness typical of capital-intensive firms and to stabilize the regression estimates (Wooldridge, 2016).

To prevent data leakage, variables that directly contain or are derived from Market Cap were removed. In particular, Enterprise Value was excluded, as it incorporates Market Cap by definition. This step ensures that the model does not implicitly “learn” the target variable, which would otherwise result in artificially inflated predictive performance (Cohen et al., 2021).

While absolute Enterprise Value was excluded due to direct accounting identity with Market Cap, the relative ratio EV_to_Sales was retained as a standardized market multiple, as scaling by top-line revenue prevents direct mathematical target reconstruction while preserving essential valuation sentiment.

Next, potential redundancy among variables was addressed. A pairwise correlation analysis was conducted to identify features with near-perfect correlation ($|\rho| \geq 0.99999$) as well as exact duplicates. These variables were removed, as they do not provide additional information and may lead to numerical instability, particularly in regression models.

The near-perfect correlation threshold of $|\rho| \geq 0.99999$ was strictly implemented as an initial automated data-cleaning filter. The intentional choice of this highly conservative threshold served exclusively to eliminate exact database duplicates, identical descriptive features, and direct structural accounting identities without risking premature information loss or accidentally dropping valid proxy indicators. Stricter or lower thresholds were not applied at this preliminary stage to prevent the inadvertent omission of distinct economic variables. Instead, the rigorous filtering of actual multicollinearity among economically differentiated variables was subsequently handled during the iterative hybrid feature selection stage using a pre-specified heuristic threshold of 0.55 on the training data, which successfully balanced information retention and linear model stability. In addition, column names were standardized by removing special characters and harmonizing formats to facilitate interpretability.

For model training, the dataset was split into training and test sets using an 80/20 split (random_state = 42). All preprocessing steps were fitted exclusively on the training data and subsequently applied to the test data to avoid information leakage (Cohen et al., 2021).

Missing values in the independent variables were imputed using median imputation. In addition, outliers were treated using winsorization, where values below the 1st percentile and above the 99th percentile were capped at these respective thresholds. This approach reduces the influence of extreme observations without removing data points, thereby preserving information while improving model stability (Heidbrink, 2025).

The before–after analysis (Figure 1) indicates that winsorization effectively corrects extreme values in both the upper and lower tails of the EV_to_Sales distribution. In particular, the 99th percentile is moderately reduced, reflecting the attenuation of excessively high valuation levels, which are commonly observed in high-growth but unprofitable firms.

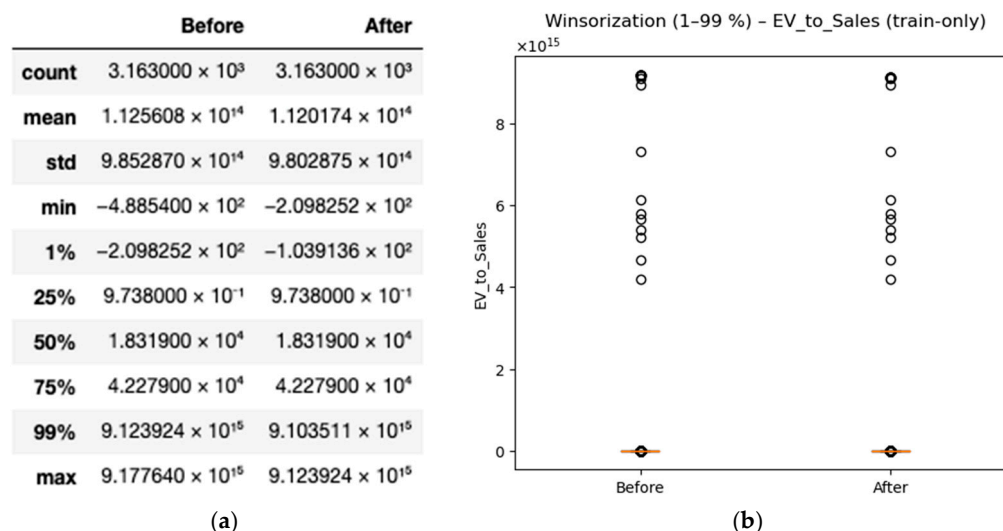


Figure 1. Before and after comparison of winsorization for the variable EV_to_Sales (1st–99th percentile, applied on training data only): (a) summary statistics before and after winsorization; (b) distributional comparison showing the reduction in extreme values through percentile-based capping.

At the same time, the median and interquartile range remain largely unchanged, confirming that the underlying economic structure of the distribution is preserved. This suggests that the transformation mitigates the influence of outliers without distorting the core distributional properties of the data.

As a result, the dataset provides a numerically stable, internally consistent, and economically interpretable basis for subsequent modeling and feature selection.

2.3. Feature Selection

The objective of feature selection is to derive a statistically and economically meaningful subset from the original 225 variables that provides both strong predictive power and high interpretability. To achieve this, a hybrid selection approach was developed, combining data-driven model indicators (XGBoost feature importance and mutual information) with an economic plausibility assessment. This ensures that variables are selected not only based on statistical relevance but also on their economic significance for corporate valuation.

In a first step, two independent methods were applied to the training data to generate a ranking of potentially relevant features:

1. XGBoost feature importance, which quantifies the contribution of each variable to the reduction in the model loss and captures non-linear effects as well as interactions between variables (Rane et al., 2023).
2. Mutual information (MI), which measures the dependency between an independent variable and the target variable without assuming a linear relationship (Notovich et al., 2023).

From each method, the top 25 variables were selected and subsequently combined. This approach reduces the risk of model-specific bias and results in a more stable and robust pool of candidate features.

To address multicollinearity, which can distort parameter estimates particularly in linear models (Heidbrink, 2025), an additional correlation-based filtering step was implemented. For each candidate feature, the pairwise correlation (train-only) with all previously selected variables was computed. If the correlation coefficient exceeded the threshold $|p| \geq 0.55$, only the feature with the higher absolute correlation with the target variable $|\text{corr}(\text{feature}, y)|$ was retained. This procedure was applied iteratively until a final set of 12 features was obtained.

The retention of exactly 12 features (Table 1) was designed as a pre-specified target choice to achieve an optimal balance between model parsimony, predictive capacity, and practical interpretability. In high-dimensional financial modeling, restricting the feature space prevents overfitting, eliminates critical pairwise collinearity for the linear benchmark, and ensures that post-hoc XAI attributions remain focused on the most economically meaningful value drivers.

Table 1. Comprehensive Overview and Economic Classification of the Selected Features.

Nr.	Technical Feature Name (Code)	Alias (Text)	Valuation Dimension	Economic Interpretation
1	10Y_Dividend_per_Share_Growth_per_Share	Div_Growth_10Y	Distribution/Governance	Long-term dividend increases as an indicator of sustainable earning power and capital discipline.
2	10Y_Operating_CF_Growth_per_Share	10Y Operating CF Growth	Cash Flow/Internal Financing	Long-term growth of operating cash flows as a robust indicator of corporate value creation.
3	5Y_Net_Income_Growth_per_Share	5Y Net Income Growth	Earnings Dynamics	Medium-term net income trends as a metric for earnings quality and economies of scale.
4	5Y_Operating_CF_Growth_per_Share	5Y Operating CF Growth	Liquidity/Efficiency	Recent cash flow dynamics, highly relevant for capturing short-term trend- or turnaround effects.
5	5Y_Revenue_Growth_per_Share	5Y Revenue Growth	Revenue/Expansion	Top-line growth acting as a primary driver of future market valuation multiples.
6	Acquisitions_and_disposals	Acquisitions & Disposals	Investment Strategy	M&A activity serving as a direct proxy for strategic corporate expansion or consolidation.
7	Earnings_before_Tax	Earnings before Tax	Profitability	Pre-tax income functioning as the core metric of structural operational earning capacity.
8	Dividend_Yield	Dividend Yield	Valuation/Capital Market	Dividend yield as an expression of direct cash generation and investor expectations.
9	EBITDA_Margin	EBITDA Margin	Efficiency/Profitability	Operating margin acting as a key indicator of internal cost efficiency and scalability.

Table 1. Cont.

Nr.	Technical Feature Name (Code)	Alias (Text)	Valuation Dimension	Economic Interpretation
10	EV_to_Sales	EV to Sales	Market Valuation	Capital-structure-adjusted revenue valuation, particularly critical for growth-oriented firms.
11	Graham_Net_Net	Graham Net-Net	Asset Structure/Risk Buffer	Net-Net working capital according to Graham, reflecting a fundamental asset-based margin of safety.
12	Issuance_buybacks_of_shares	Share Issuance & Buybacks	Capital Market Policy	Equity interventions functioning as a signal of management confidence and capital structure control.

A complementary sensitivity analysis across multiple correlation thresholds ($\rho = 0.50$ – 0.80 , evaluated strictly on the training partitions across the standardized 12-feature subset) confirmed that the pre-specified threshold of 0.55 represents a balanced, conservative compromise. While higher thresholds (e.g., $\rho = 0.60$ or $\rho = 0.80$) achieve comparable or marginally higher linear fits (reaching Holdout $R^2 = 0.450$ at $\rho = 0.80$), they do so by allowing stronger pairwise linear dependencies into the candidate selection before pruning. In contrast, the pre-specified cutoff of $\rho = 0.55$ strictly enforces pairwise correlations below $|\rho| = 0.50$ in the final matrix (observed $|\rho|_{\max} = 0.47$, as shown in Figure 2), inherently guaranteeing low variance inflation and stable OLS coefficient estimation without discarding informative feature interactions.

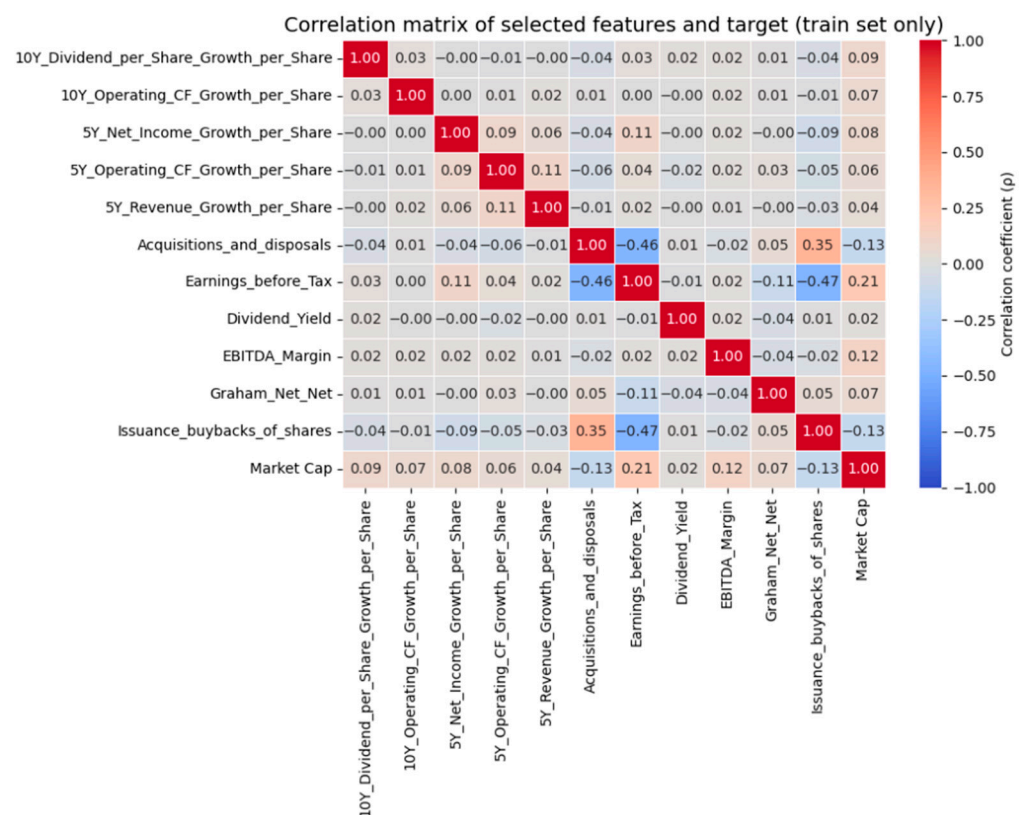


Figure 2. Correlation matrix of the selected features and the target variable (Market Cap), based on the training data. The figure illustrates the absence of strong multicollinearity and the overall correlation structure among variables.

The results (Table 2) confirm a clear trade-off between redundancy reduction and predictive performance across different correlation thresholds. While stricter thresholds improve model stability, they slightly reduce the explanatory power of OLS. In contrast, more lenient thresholds increase apparent OLS performance but introduce multicollinearity and reduce robustness. Overall, a threshold of $p = 0.55$ provides the best balance between stability and information retention.

Table 2. Sensitivity analysis of correlation thresholds and model performance (OLS vs. XGBoost).

Threshold (ρ)	Holdout R^2 (OLS)	Holdout R^2 (XGB)	Δ % (Improvement)	CV R^2 (OLS)	CV R^2 (XGB)
0.50	0.362	0.661	+82.5%	0.332	0.620
0.55	0.364	0.654	+79.6%	0.343	0.611
0.60	0.376	0.654	+73.9%	0.350	0.609
0.65	0.370	0.652	+76.5%	0.353	0.593
0.80	0.450	0.619	+37.7%	0.401	0.566

Following the data-driven selection, a qualitative validation of the final feature set was conducted. Each variable was assessed in terms of its economic relevance and assigned to one of several key valuation dimensions, including growth, profitability, valuation, financial structure, and capital market policy. This step ensures that the selected features collectively provide a coherent and economically meaningful representation of firm value.

To further validate the feature set, a full correlation matrix of the 12 selected variables, including the target variable, was computed. The maximum pairwise correlation is $|p|_{\max} = 0.47$, indicating the absence of critical multicollinearity. As expected, growth and profitability variables exhibit predominantly positive relationships with market capitalization, while variables such as *issuance_buybacks_of_shares* show a negative association. This confirms both the statistical independence and the economic directional consistency of the selected features.

Through the combination of data-driven relevance measures, correlation-based filtering, and economic validation, a compact, stable, and interpretable feature subset was obtained. The selected variables capture all key value dimensions and provide a robust foundation for the subsequent modeling and evaluation.

2.4. Modeling and Validation

To evaluate the predictive performance of the selected feature set, a methodological comparison between a linear baseline model and a non-linear ensemble approach was conducted. The objective was to assess the performance and robustness of both models under consistent conditions, thereby providing a solid basis for the subsequent explainability analysis using SHAP values.

Both models were implemented within a unified preprocessing pipeline to ensure comparability and prevent data leakage. The pipeline includes the following steps:

1. Median imputation for robust handling of remaining missing values
2. Quantile transformation to approximate a normal distribution and reduce skewness (Notovich et al., 2023)
3. Standardization (applied to OLS only)
4. Model estimation

Both the baseline Ordinary Least Squares (OLS) regression and the eXtreme Gradient Boosting (XGBoost) algorithm were integrated into identical, unified preprocessing

pipelines to prevent data leakage and guarantee structural comparability. Preprocessing comprised three sequential stages: missing value imputation using median values (`SimpleImputer(strategy = 'median')`), non-linear quantile mapping to a Gaussian distribution (`QuantileTransformer(output_distribution = 'normal')`), and feature standardization (`StandardScaler()`).

To ensure optimal performance and mitigate overfitting, the XGBoost model was parameterized with 600 decision trees (`n_estimators = 600`), a shrinkage rate of $\eta = 0.05$ (`learning_rate = 0.05`), a maximum tree depth of 4 (`max_depth = 4`), and subsampling ratios of 0.9 for both observations (`subsample = 0.9`) and features (`colsample_bytree = 0.9`), anchored with a deterministic seed (`random_state = 42`).

Model evaluation was executed on identical data partitions, employing both a fixed 80/20 train-test split (`Ntrain = 3163`; `Ntest = 791`, `random_state = 42`) and a stratified 10-fold cross-validation scheme (`KFold(n_splits = 10, shuffle = True, random_state = 42)`).

The linear regression model serves as an interpretable benchmark (Heidbrink, 2025). Its estimated coefficients allow for a direct economic interpretation of the relationships between variables and market capitalization. Positive coefficients indicate a positive association with firm value (e.g., Earnings before Tax, `EV_to_sales`), whereas negative coefficients reflect value-reducing effects (e.g., acquisitions and disposals, `Graham Net-Net`).

The occurrence of negative coefficients is economically plausible, as they may reflect capital outflows, acquisition effects, or accounting adjustments that temporarily reduce firm value. The coefficient analysis therefore serves as a validation of economically meaningful relationships rather than as a basis for variable elimination.

Model performance was first evaluated on a holdout test set (20% of the data). Three complementary evaluation metrics were used: R^2 , mean absolute error (MAE), and root mean squared error (RMSE).

The results indicate a substantial performance improvement, with the XGBoost model achieving an out-of-sample $R^2 = 0.654$ compared to $R^2 = 0.364$ for linear regression (an absolute gain of $\Delta R^2 = 0.290$), which can be attributed to the ability of gradient boosting methods to capture non-linear relationships, interactions, and heterogeneity among variables.

In addition, the lower error metrics (MAE = 0.77 vs. 1.35; RMSE = 1.48 vs. 2.00) demonstrate a more accurate and robust predictive performance of the XGBoost model. The reduction in RMSE by approximately 26% suggests that XGBoost is particularly effective in modeling highly valued or extreme observations. Note that error metrics (MAE and RMSE) were evaluated on the log-transformed target variable $\ln(y)$. Consequently, the ~26% error reduction reflects improvements on the log scale, corresponding to a reduction in relative percentage errors and multiplicative deviations in raw values. This effectively suppresses the influence of extreme financial outliers. Since the absolute value of RMSE depends on the scale of the data, its interpretation is most meaningful in relative terms. While linear regression exhibits larger average and extreme prediction errors, reflecting its simplifying assumptions, XGBoost is able to better adjust to complex patterns and reduce the influence of outliers.

To validate the robustness of these results, a 10-fold cross-validation was conducted. This approach evaluates model performance across multiple data splits and provides insight into its stability (Janiesch et al., 2021). The linear model shows moderate variability ($CV R^2 = 0.343 \pm 0.061$), whereas XGBoost achieves a higher average performance ($CV R^2 = 0.611 \pm 0.104$), albeit with slightly increased variance. These findings confirm that the observed performance differences are not driven by a specific data split but reflect systematic model characteristics.

Descriptively, the standard deviations across the 10 cross-validation folds (0.061 for OLS and 0.104 for XGBoost) indicate that XGBoost achieves a consistently higher baseline performance across iterations. While this descriptive comparison illustrates the non-linear model's stability across folds, it serves as an empirical benchmark rather than a formal inferential hypothesis test.

To address potential concerns regarding algebraic target dependency arising from the construction of the Enterprise Value metric, a formal ablation analysis was performed. Specifically, both the OLS linear baseline and the XGBoost model were re-estimated on the identical data partitions after completely excluding the EV_to_Sales ratio from the selected feature set (yielding an 11-feature baseline).

Table 3 presents the empirical performance comparison between the baseline 12-feature specifications and the re-estimated 11-feature ablation models.

Table 3. Comparative performance evaluation and ablation analysis with and without EV_to_Sales.

Model	R ² (Holdout)	MAE (Holdout)	RMSE (Holdout)	CV R ² (Mean)	CV R ² (Std. Dev.)
Baseline Models (with EV_to_Sales)					
OLS (Linear Regression)	0.364	1.353	2.001	0.343	0.061
XGBoost Regressor	0.654	0.765	1.475	0.611	0.104
Ablation Models (without EV_to_Sales)					
OLS (Linear Regression)	0.341	1.390	2.037	0.321	0.060
XGBoost Regressor	0.628	0.847	1.530	0.591	0.091

The empirical results demonstrate that excluding EV_to_Sales leads only to a marginal reduction in explanatory power across both architectures (Holdout R² decreases by 0.023 for OLS and by 0.026 for XGBoost). Crucially, the structural performance advantage of XGBoost over the linear baseline remains virtually identical ($\Delta R^2 = 0.287$ without EV_to_Sales vs. $\Delta R^2 = 0.290$ in the baseline model). Furthermore, the cross-validated stability of the non-linear ensemble is preserved (CV R² = 0.591 ± 0.091).

In addition to predictive metrics, the feature attribution hierarchy and non-linear patterns exhibit strong fundamental consistency under ablation:

- **Invariance of the Primary Driver:** Earnings before Tax remains the undisputed primary value driver across both models (OLS standardized coefficient $\beta = 0.9886$, Rank 1; mean |SHAP| = 1.0007, Rank 1).
- **Reconfiguration and Persistence of Secondary Drivers:** In the absence of EV_to_Sales, capital market policy and balance sheet safety metrics move to the immediate top tier in SHAP importance, led by Share Issuance & Buybacks (Rank 2; mean |SHAP| = 0.3305) and Graham Net-Net (Rank 3; mean |SHAP| = 0.3170).
- **Persistence of Non-Linear Differentiation:** While EBITDA Margin slightly shifts from Rank 3 to Rank 4 (mean |SHAP| = 0.2535), it continues to demonstrate a substantial rank divergence against the linear baseline (Rank 4 in SHAP vs. Rank 10 in OLS | β |), confirming that the model captures operational efficiency via non-linear threshold dynamics rather than simple linear slopes.

This confirms that the predictive superiority of the tree-based ensemble and the core SHAP attribution dynamics are driven by genuine fundamental corporate accounting relationships rather than an algebraic artifact or mechanical target leakage from EV_to_Sales.

3. Results

3.1. Global SHAP Analysis

Following the evaluation of model performance in Section 2.4, the explainable component of the XGBoost model is examined in detail. The objective is to make the relationships learned by the model between the selected financial indicators and market capitalization transparent, thereby empirically testing Objective RO1. For this purpose, SHAP values are employed, which assign each feature and observation a marginal contribution to the model prediction.

A positive SHAP value indicates that a given feature increases the predicted market capitalization, whereas a negative value decreases it (Rane et al., 2023). The global analysis is based on SHAP values computed on the test dataset, which was not used during model training. This ensures that the interpretation reflects genuine out-of-sample behavior rather than overfitting effects.

The global importance of features is first assessed using the mean absolute SHAP values ($\text{mean}(|\text{SHAP}|)$), which measure the average impact of each variable on the model output.

The results (Figure 3) show that profitability, market-based valuation, and capital structure or transaction-related variables represent the dominant value drivers identified by the model. This structure is economically plausible and aligns with established valuation frameworks. In both discounted cash flow (DCF) models and multiples-based approaches, expected earnings and their capitalization in financial markets are considered key determinants of firm value (Heidbrink, 2025).

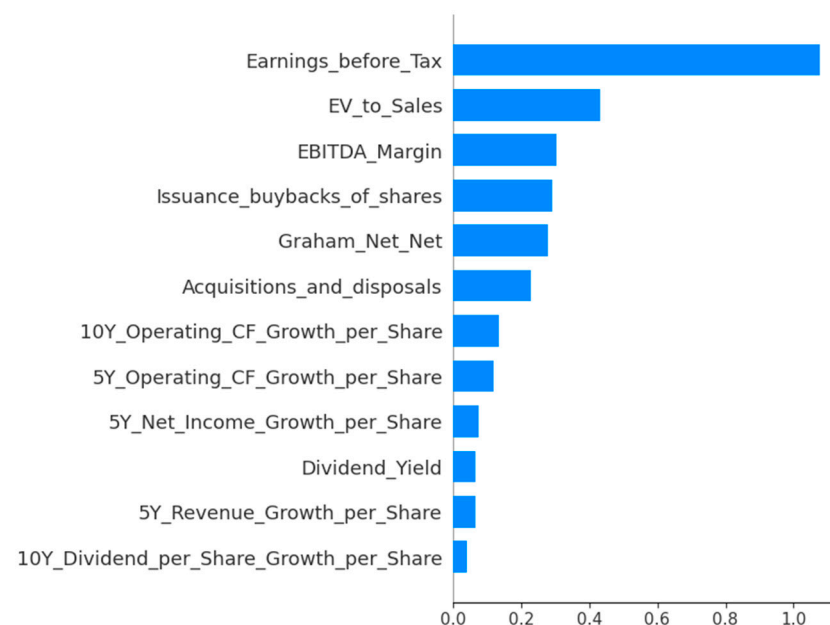


Figure 3. Global SHAP feature importance based on mean absolute SHAP values (test dataset).

The SHAP summary plot provides a more detailed view of the direction and heterogeneity of feature effects. For each variable, SHAP values are displayed conditional on the feature magnitude.

A detailed inspection of the SHAP distributions (Figure 4) reveals that the model captures economically meaningful patterns. Earnings before Tax exhibits the strongest effect: higher pre-tax earnings are associated with higher predicted market capitalizations, while negative values reduce firm value (Damodaran, 2002).

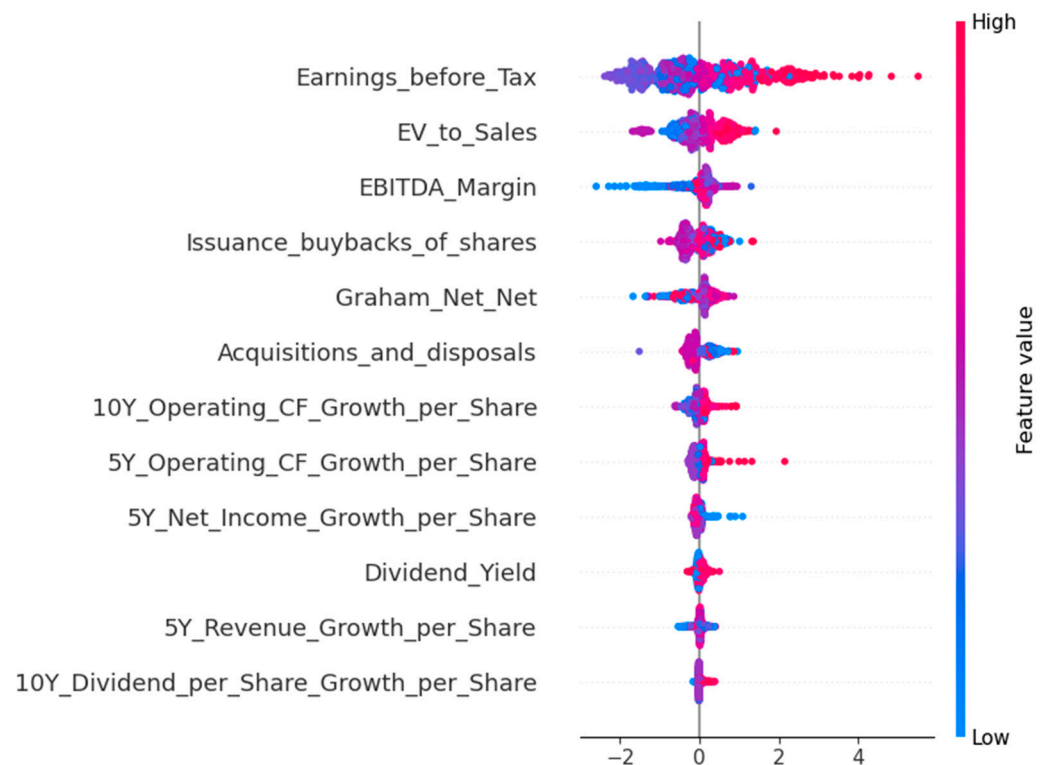


Figure 4. SHAP summary plot showing both magnitude and direction of feature contributions across all observations.

An inverse pattern is observed for EV_to_Sales: lower multiple values are associated with positive SHAP contributions, whereas higher multiples generate negative contributions within the model (Liu et al., 2002). In a contemporaneous cross-sectional setting, this pattern reflects a model-dependent conditional association: within the multi-tree ensemble, high EV-to-sales ratios frequently characterize capital-intensive or distressed firms where top-line revenues fail to translate proportionally into market capitalization, leading the algorithm to apply a downward adjustment relative to the baseline expectation. Crucially, as demonstrated in our ablation analysis in Section 2.4, this conditional relationship represents an auxiliary market-sentiment proxy rather than a distorting mechanical driver, as the model's core predictive capacity and primary earnings attributions remain fully robust when the metric is excluded.

EBITDA Margin shows a clear positive relationship with market capitalization. High margins (represented by higher feature values) are associated with positive SHAP values, while low margins exert a negative effect. Compared to Earnings before Tax, the dispersion of SHAP values is smaller, indicating that EBITDA margin acts as a stabilizing factor that contributes consistently, but with lower sensitivity to extreme observations.

Share Issuance & Buybacks reveals a non-linear pattern. Negative values (share buybacks) are associated with positive contributions, whereas positive values (equity issuance) tend to reduce predicted firm value. This reflects the common market interpretation of buybacks as signals of financial strength, while equity issuance may imply dilution effects (Damodaran, 2002).

Graham Net-Net, a measure rooted in asset-based valuation, exhibits strong SHAP variations, particularly at extreme values. Firms with higher balance sheet safety margins tend to be valued more positively, while weaker asset backing leads to negative contributions (Fernandez, 2002; Liu et al., 2002). This indicates that the model incorporates classical valuation principles beyond purely earnings-based metrics.

For acquisitions and disposals, both positive and negative contributions are observed. Moderate acquisition activity appears to increase firm value, likely reflecting growth expectations, whereas excessive or frequent transactions are associated with negative effects, potentially indicating increased risk or integration challenges (Heidbrink, 2025).

Growth-related variables show a lower but still systematic influence. Positive growth rates are associated with slightly positive SHAP values, while negative growth reduces predicted firm value. This suggests that the model treats growth as a supporting factor rather than a primary value driver, with profitability and stability playing a more dominant role in the dataset.

Finally, Dividend Yield and Div_Growth_10Y play a secondary but consistent role. Moderate dividend yields contribute positively, whereas very high yields, often associated with declining stock prices, have a negative effect. Similarly, sustainable but moderate dividend growth is associated with slightly positive contributions. This indicates that dividend policy primarily acts as a signal of financial stability rather than a core value driver.

The dispersion of SHAP values across individual observations provides descriptive visual evidence of non-linear distribution patterns and heterogeneity in feature contributions. While global summary plots visually suggest context-dependent variations, particularly for Earnings before Tax and EV_to_Sales, they serve as qualitative indications of underlying data complexity rather than formal mathematical tests of multi-variable interaction terms.

Overall, the results of the global SHAP analysis support Objective RO1 in several ways:

1. Transparent identification of key value drivers: Global SHAP values clearly reveal which variables dominate the model output. The ranking is both statistically consistent and economically plausible, making the model's decision structure interpretable.
2. Interpretability of effect direction: The summary plots indicate whether variables increase or decrease firm value. High profitability and margins increase predicted market capitalization, whereas high valuation multiples or weak growth reduce it.
3. Explanation of heterogeneous effects: The variation in SHAP values demonstrates that the model differentiates between firm-specific contexts, capturing interactions between variables that influence valuation outcomes.
4. Integration of economic logic: The alignment between SHAP results and established financial value drivers (profitability, valuation, growth, dividend policy) indicates that explainable machine learning methods complement rather than contradict traditional valuation approaches.

These findings provide strong empirical support for RO1: SHAP significantly enhances the transparency of the XGBoost model by making the importance, direction, and magnitude of individual financial indicators observable and interpretable.

3.2. Comparison of OLS and SHAP

Building on the insights from the global SHAP analysis in Section 3.1, this section examines the extent to which the value drivers identified by SHAP align with those derived from the linear OLS regression model. The objective is to empirically test Hypotheses H2a and H2b by comparing the relative importance and ranking of features across both modeling approaches.

For each model, feature importance was quantified as follows:

- For the OLS regression, by the absolute magnitude of standardized coefficients ($|\beta|$)

- For the XGBoost model, by the mean absolute SHAP values (mean $|\text{SHAP}|$)

Subsequently, features were ranked (1 = most important), and the Spearman rank correlation coefficient was computed to assess the similarity between both rankings.

When interpreting the Spearman rank correlation ($r_s = 0.224, p = 0.48$), appropriate statistical caution is required. Given the small feature dimension ($N = 12$), the test possesses low statistical power, meaning that the lack of statistical significance does not formally prove divergence. Rather, this metric should be viewed as descriptive evidence of weak linear-rank agreement, stemming from fundamental conceptual differences: standardized OLS coefficients estimate constant global slopes, whereas mean absolute SHAP values quantify average local predictive contributions across non-linear feature distributions. For visualization, the rankings were compared in a two-dimensional plot, complemented by a tabular overview of absolute rank positions.

The results (Figure 5, Table 4) reveal both convergence and divergence between the OLS and XGBoost + SHAP models:

1. Consistent primary drivers: Both models identify Earnings before Tax and EV_to_Sales as the dominant determinants of firm value. These variables rank first and second in both approaches, reflecting the central importance of profitability and market-based valuation. This consistency highlights their robustness as fundamental value drivers.
2. Differences in non-linear and interaction effects: Variables such as EBITDA Margin, Share Issuance & Buybacks, and Graham Net-Net receive substantially higher importance in the SHAP ranking than in the OLS model. This indicates that these features exhibit non-linear or context-dependent effects that are not fully captured by linear regression. SHAP captures interaction effects, such as share buybacks being value-enhancing primarily for profitable firms, or margins acting as stabilizing rather than amplifying factors under certain conditions.
3. Smoothing in OLS vs. differentiation in SHAP: The OLS coefficients are more evenly distributed across features, reflecting the additive and linear structure of the model. In contrast, SHAP concentrates importance on a smaller set of dominant variables while assigning more precise and differentiated contributions to others. This suggests that XGBoost is more sensitive to economically relevant extreme scenarios and better captures complex market dynamics.
4. Low rank correlation: The low correlation coefficient (0.224) indicates only partial agreement between the rankings. While the key drivers are consistent, the relative importance of secondary variables differs substantially. The non-significant p -value further suggests that no systematic correspondence between the two ranking structures can be established, emphasizing the added methodological value of SHAP.

From an economic perspective, these findings imply that linear valuation approaches such as OLS provide stable and interpretable baseline relationships but fail to capture non-linearities and interaction effects, such as threshold effects, diminishing returns, or valuation distortions. In contrast, XGBoost combined with SHAP acts as a bridge between structural interpretability and empirical market dynamics, revealing how the relevance of individual features varies across different firm contexts.

Overall, the empirical findings provide strong support for Hypothesis 2a (H2a) and partial, descriptive support for Hypothesis 2b (H2b). The identical top-ranking of Earnings before Tax across both models confirms H2a, demonstrating that primary core earnings drivers remain invariant across linear and non-linear paradigms. Simultaneously, the descriptive rank shifts observed in secondary indicators—such as EBITDA Margin, Share Issuance & Buybacks, and Graham Net-Net—provide qualitative support for H2b, illustrating that SHAP captures context-sensitive nuances that linear models inherently smooth.

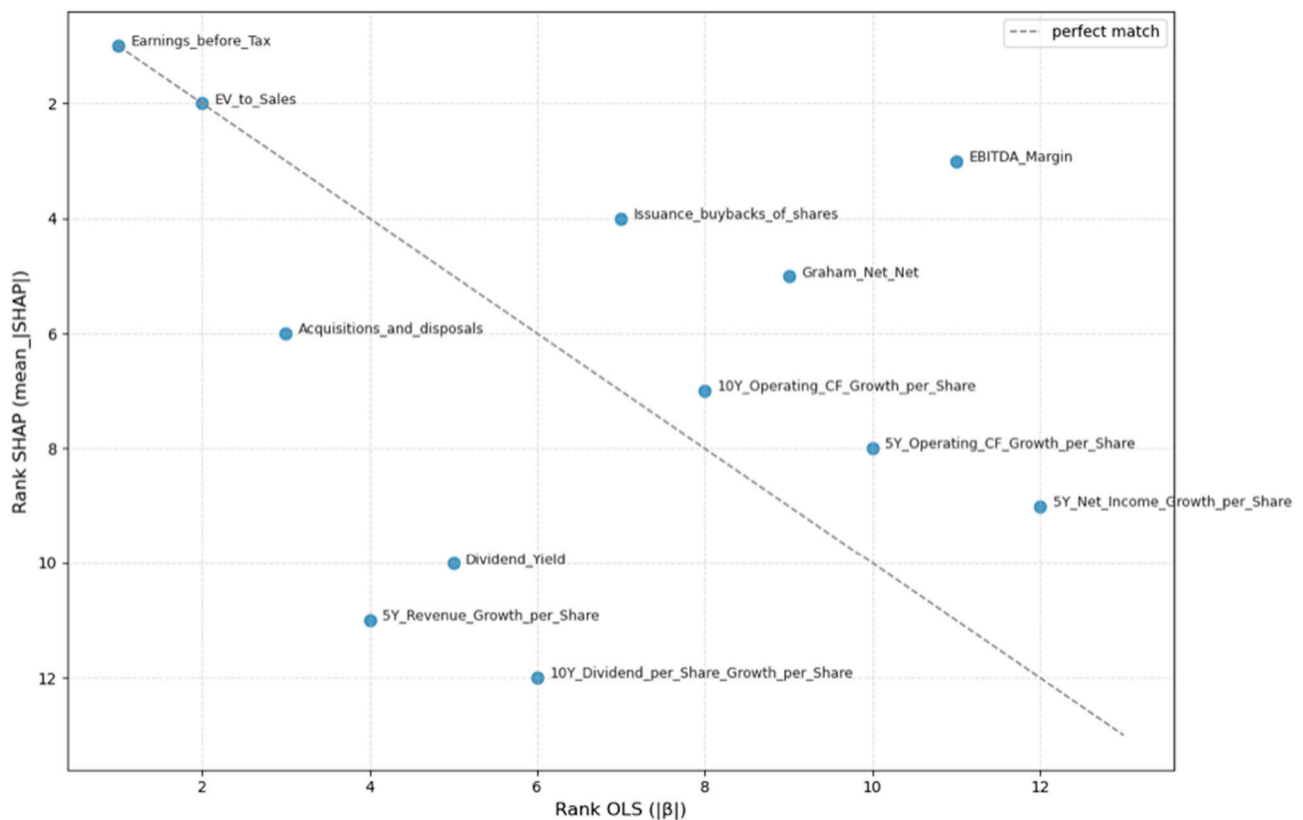


Figure 5. Scatter plot comparing feature importance rankings derived from OLS and SHAP, highlighting deviations from perfect rank alignment.

Table 4. Tabular comparison of feature rankings across linear and non-linear modeling approaches.

Feature	Rank OLS (Standardized)	Rank XGBoost + SHAP
Earnings before Tax	1	1
EV to Sales	2	2
EBITDA Margin	11	3
Share Issuance & Buybacks	7	4
Graham Net-Net	9	5
Acquisitions & Disposals	3	6
10Y Operating CF Growth	8	7
5Y Operating CF Growth	10	8
5Y Net Income Growth	12	9
Dividend Yield	5	10
5Y Revenue Growth	4	1
Div_Growth_10Y	6	12

3.3. Local SHAP Analysis

While the previous analyses focused on global relationships between financial indicators and market capitalization, this section examines model behavior at the level of individual firms. Using three representative cases (high, medium, and low predicted market capitalization), the analysis demonstrates how SHAP enables transparent, instance-level interpretation of model predictions.

Case A: High Predicted Market Capitalization

The first case (index 534) represents a firm with a very high predicted market capitalization, significantly above the sample mean. The local SHAP decomposition shows that the model's baseline expectation $E[f(x)]$ is substantially increased by a series of strong positive feature contributions.

As shown in Figure 6, the dominant driver is Earnings before Tax: exceptionally high pre-tax earnings generate the largest positive SHAP contribution and shift the prediction upward. This is fully consistent with the global analysis, where profitability was identified as the most important value driver. Additional positive contributions stem from EV_to_Sales and 10Y Operating CF Growth. A moderate EV/Sales multiple indicates a relatively attractive valuation, while strong long-term operating cash flow growth reflects the sustainability of the firm's earnings capacity.

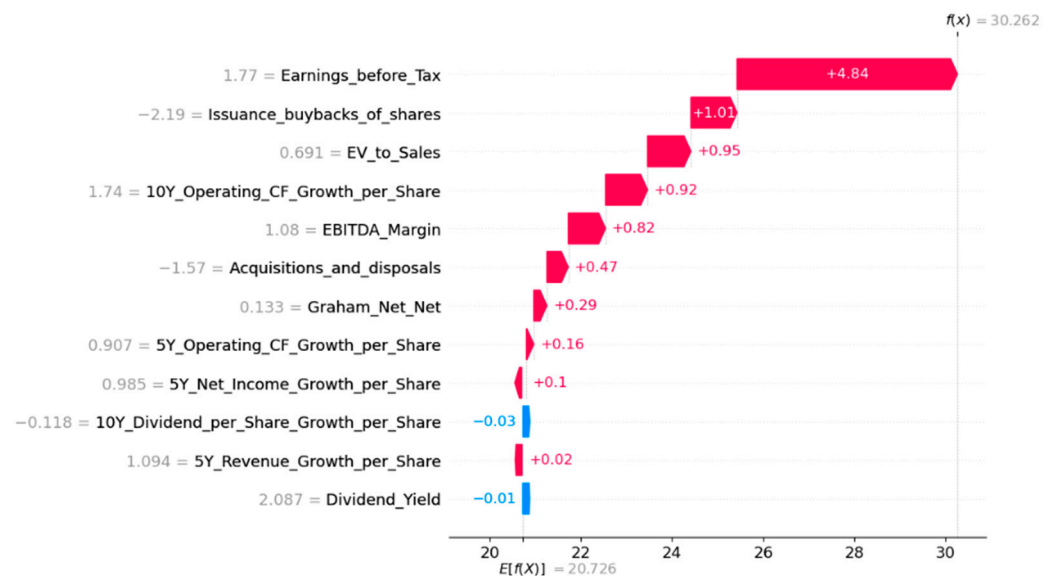


Figure 6. Local SHAP waterfall plot for a firm with high predicted market capitalization (Note: Feature values displayed on the left vertical axis represent normalized, quantile-transformed values reflecting relative sample percentiles rather than raw financial statement amounts).

Particularly noteworthy are the capital structure and transaction-related variables Share Issuance & Buybacks and acquisitions and disposals. In this case, substantial share buybacks generate a strong positive SHAP contribution, suggesting that the model interprets these actions as signals of financial strength. This aligns with established financial theory, where buybacks are often viewed as mechanisms to return excess capital to shareholders and enhance firm value (Fernandez, 2002; Steiger, 2010). Moderate divestments also contribute positively, indicating a perceived strategic focus.

Negative contributions are comparatively minor. Overall, the firm can be characterized as highly profitable, growth-oriented, and attractively positioned from both a valuation and capital market perspective. The high predicted market capitalization is therefore coherently explained by the local SHAP contributions.

Case B: Medium Predicted Market Capitalization

The second case (index 577) represents a firm with a predicted market capitalization close to the sample median. In this instance, the local SHAP values exhibit a more balanced structure, with positive and negative contributions largely offsetting each other.

On the positive side in Figure 7, EBITDA Margin plays a central role, complemented by moderate contributions from Graham Net-Net and acquisitions and disposals. These

variables indicate a firm with stable profitability, a certain degree of asset backing, and targeted investment activity.

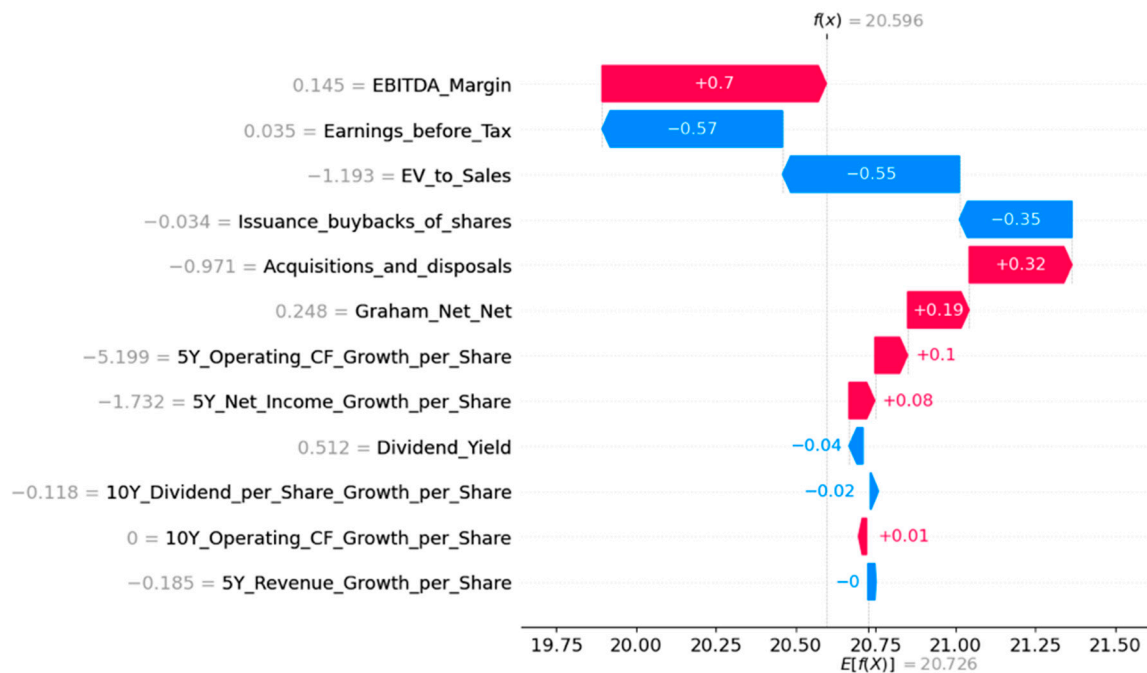


Figure 7. Local SHAP waterfall plot for a firm with medium predicted market capitalization (Note: Feature values displayed on the left vertical axis represent normalized, quantile-transformed values reflecting relative sample percentiles rather than raw financial statement amounts).

Conversely, a relatively high EV_to_Sales multiple exerts downward pressure on the prediction, suggesting that the firm may already be priced at a premium. The absence of significant share buybacks and comparatively moderate growth rates further dampens the valuation.

This case illustrates that the model does not simply distinguish between “high” and “low” value firms but captures intermediate positions, where strong fundamentals are partially offset by valuation constraints or weaker growth dynamics.

Case C: Low Predicted Market Capitalization

The third case (index 78) represents a firm with a very low predicted market capitalization, significantly below the sample mean. The waterfall plot shows that the baseline expectation is substantially reduced by multiple strong negative SHAP contributions.

Again in Figure 8, profitability emerges as the central factor: Earnings before Tax is strongly negative, leading to a substantial downward adjustment. This effect is reinforced by a significantly negative EBITDA Margin, indicating weak operational performance.

Additional negative contributions arise from asset structure and capital market signals. Although a higher Graham Net-Net value definitionally indicates substantial asset backing and balance-sheet safety, its negative local SHAP contribution in this distressed scenario illustrates an important non-linear, context-dependent pattern: when accompanied by severe operational losses (heavily negative Earnings before Tax and EBITDA Margin), high asset backing fails to support valuation, reflecting capital market skepticism toward unprofitable firms with idle assets (a classic ‘value trap’). Positive values in Share Issuance & Buybacks indicate equity issuance, which is interpreted as value-dilutive. Weak revenue and earnings growth, combined with unremarkable dividend indicators, further reinforce the negative outlook.

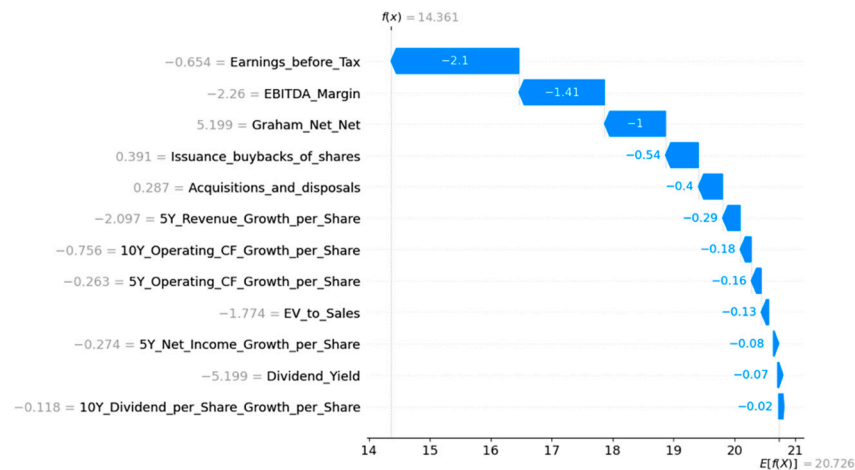


Figure 8. Local SHAP waterfall plot for a firm with low predicted market capitalization (Note: Feature values displayed on the left vertical axis represent normalized, quantile-transformed values reflecting relative sample percentiles rather than raw financial statement amounts).

Overall, the firm can be characterized as a combination of weak profitability, limited asset backing, and an unattractive capital market profile. The low predicted market capitalization is therefore not a “black-box” outcome but can be clearly explained through the SHAP decomposition.

4. Discussion

4.1. Positioning of Results Within the Existing Literature

The empirical findings can be linked to central assumptions of the financial economics literature and simultaneously demonstrate that modern explainable AI methods are capable of capturing established theoretical concepts in a more differentiated manner (Chan et al., 2025). The dominant value drivers—profitability, operating performance, and valuation multiples—correspond to the fundamental factors identified (Fama & French, 2006) as key determinants of long-term firm value. Variables such as Earnings before Tax, EBITDA Margin, and EV_to_Sales reflect these relationships (Fernandez, 2002) and highlight the importance of stable earnings as a core indicator of future firm performance.

The SHAP analysis extends this classical perspective by revealing non-linear mechanisms that cannot be captured by linear models (Salih et al., 2025). This is particularly evident for valuation metrics: while EV_to_Sales exhibits a positive relationship in the linear model, SHAP values indicate that high multiples exert a negative effect beyond a certain threshold (Chan et al., 2025; Damodaran, 2002). This observation can be related to the concept of asymmetric market regimes described by (R. I. Hamilton & Papadopoulos, 2024), according to which valuation processes are not proportional but vary structurally across different states of the market. SHAP makes such regime shifts empirically observable and thereby provides a more nuanced perspective on market reactions (Lundberg & Lee, 2017).

The role of profitability is also refined by the SHAP analysis. The lower dispersion of SHAP values for EBITDA Margin compared to Earnings before Tax suggests that the model interprets margin-based measures as more stable indicators. This effect can be linked to the logic of the DuPont Analysis framework (Soliman, 2007), which conceptualizes operating margins as indicators of sustainable efficiency and resilience. Accordingly, markets appear to reward consistent profitability more strongly than volatile earnings components (Soliman, 2007).

Another result that closely aligns with existing theory concerns Share Issuance & Buybacks. Negative values (share buybacks) contribute positively to firm valuation, whereas

positive values (equity issuance) have a negative impact. This pattern is consistent with capital structure signaling theory (Chan et al., 2025), according to which buybacks signal financial strength, while equity issuance is often interpreted as value-dilutive. SHAP enables a quantitative assessment of these effects at the observation level, thereby making context-specific market reactions directly observable (Dwivedi et al., 2023; R. I. Hamilton & Papadopoulos, 2024).

Overall, the findings suggest that SHAP not only confirms theoretically expected relationships but also provides additional insights into the structure of economic mechanisms. The method reveals interdependencies, threshold effects, and asymmetric response patterns that remain hidden in classical linear approaches (Chan et al., 2025; Rane et al., 2023). As a result, a more differentiated view of corporate valuation emerges, one that is consistent with the existing literature while simultaneously extending it through empirical evidence.

4.2. Methodological Positioning of the Models

The comparison between OLS and SHAP highlights that both approaches provide distinct yet complementary perspectives on corporate valuation. Linear regression, as a transparent baseline model, captures the average marginal effect of individual variables (Götze et al., 2023), but relies on the assumption of additive and constant effect directions (Lundberg & Lee, 2017). This structural simplification limits its ability to represent interactions and non-linear relationships among highly interdependent financial variables (Salih et al., 2025).

SHAP extends this linear paradigm through a flexible, observation-specific attribution framework (Dwivedi et al., 2023). Instead of estimating global average effects, SHAP quantifies the relative contribution of each feature for every individual observation (Dwivedi et al., 2023). As a result, variable contributions become visible not only at the global model level but also within the specific context of their interaction. This capability corresponds to the concept of “local approximation of causality” described in the XAI literature (Murdoch), where explanations are derived from data-driven structural properties rather than fixed model parameters (Murdoch et al., 2019).

From a methodological perspective, this distinction is reflected in the observed rank correlation coefficient of 0.224 between the OLS coefficients and the mean absolute SHAP values. Rather than indicating instability, this result points to fundamentally different levels of explanation: OLS captures linear, global effects, whereas SHAP reflects a more granular and context-sensitive weighting logic (Arrieta et al., 2019). These differences should therefore not be interpreted as contradictions but as an epistemological extension of model interpretation. In this sense, the underlying model logic is not replaced but refined (Salih et al., 2025).

The observed rank shifts are not random model deviations but instead reveal a fundamental limitation of linear regression. The OLS model imposes a globally constant and additive effect for each variable. This assumption implies that financial indicators operate independently of firm-specific context, interactions with other variables, and valuation levels. For key efficiency and profitability metrics such as EBITDA Margin, this represents a strong simplification, as their economic relevance typically emerges only in combination with earnings levels, growth expectations, and capital structure.

The low importance of EBITDA Margin in the OLS ranking is particularly informative: while it ranks only 11th in the linear model, it is among the top three features in the SHAP-based ranking (rank 3). This pronounced shift indicates that the linear model systematically underestimates the role of operating margins, as it captures their effect only as an isolated average contribution. In contrast, the SHAP analysis demonstrates that their relevance

primarily arises from context-dependent interactions with valuation and growth variables, which cannot be adequately represented in a linear framework.

A similar pattern is observed for other financially relevant variables. The ranking of Share Issuance & Buybacks improves significantly in the SHAP-based model (rank 4) compared to OLS (rank 7), while asset-based indicators such as Graham Net-Net are also assigned higher importance in the non-linear model (rank 5 vs. rank 9). These shifts suggest that the economic significance of capital structure decisions and asset-based valuation metrics is highly state-dependent. The linear model fails to capture such context-sensitive effects, leading to a smoothing of important signals related to capital structure and firm fundamentals. As a result, the derived importance profiles may appear statistically stable but can distort key economic valuation dimensions.

Furthermore, SHAP attributions provide descriptive illustrations consistent with state-dependent valuation dynamics (Chan et al., 2025). In tree-based ensembles, feature splits implicitly account for high-dimensional feature dependencies, which SHAP visualizes at the observation level (R. I. Hamilton & Papadopoulos, 2024; Murdoch et al., 2019). While these visual attributions reflect context-dependent patterns that OLS flattens into constant additive slopes (Arrieta et al., 2019), they represent descriptive post-hoc mappings rather than formal second-order interaction tests. This aligns with the established assumption in financial economics that market mechanisms are context-dependent (Dwivedi et al., 2023).

At the same time, the comparison underscores the role of linear regression as a necessary reference framework (Rudin, 2018). OLS provides a clear structural baseline, delivers expectation-based effects (Wooldridge, 2016), and serves as an interpretable point of departure (Arrieta et al., 2019). SHAP builds upon this foundation by identifying where and why the linear structure reaches its limits (Dwivedi et al., 2023). This complementary relationship is also emphasized in the current XAI literature, where explainable AI is not viewed as a replacement for classical models but as an integrative layer bridging statistical simplicity and data-driven complexity (Salih et al., 2025).

Overall, the methodological comparison provides clear empirical support for both Hypothesis 2a (H2a) and Hypothesis 2b (H2b). Both approaches identify identical core value drivers at the macroeconomic top tier, confirming H2a, but differ substantially in their assessment of secondary feature rankings and relative importance, fully supporting H2b. While OLS provides robust, globally valid directional relationships, it remains limited to constant additive average effects and thus systematically flattens interaction-driven and context-dependent mechanisms. In contrast, XGBoost combined with SHAP successfully uncovers these underlying non-linear threshold dynamics and regime-dependent influence patterns without sacrificing theoretical alignment.

4.3. Theoretical and Practical Implications

The findings of this study demonstrate that the application of explainable AI methods, particularly SHAP, has important theoretical and practical implications for financial modeling. From a theoretical perspective, SHAP enables an empirically grounded extension of classical valuation frameworks by decomposing complex model structures into additive and interpretable components (Chan et al., 2025). Through its foundation in cooperative game theory, SHAP provides a level of transparency that is not only interpretable but also scientifically verifiable (Chan et al., 2025). In this sense, it fulfills key requirements of a “rigorous science of interpretable AI,” as emphasized in the recent XAI literature (Doshi-Velez & Kim, 2017).

A central theoretical implication concerns the empirical validation of established financial economic assumptions. The SHAP values not only confirm fundamental relationships, such as the importance of profitability and stable earnings in line with the earnings per-

sistence hypothesis associated with Eugene Fama and Kenneth French (Fama & French, 2006), but also quantify threshold effects and non-linear regime shifts that classical models can only capture qualitatively (Dixon et al., 2020). This opens new avenues for empirical theory testing: interaction effects, asymmetric market responses, and context-dependent valuation patterns are no longer merely assumed but can be directly observed in the data (Chan et al., 2025). SHAP thus functions as a methodological bridge between theoretical model assumptions and empirical structure.

The practical relevance of these findings is particularly evident in financial applications. SHAP provides both global and local explanations, enabling complex valuation models to become auditable, documentable, and communicable across different stakeholder groups. In areas such as credit risk assessment, corporate analysis, and investment decision-making, the transparent representation of model logic can enhance trust among analysts, investors, and regulatory authorities (Salih et al., 2025). The availability of local explanations, for example through waterfall plots, allows for case-specific interpretability and supports accountable decision-making at the individual firm level (Rane et al., 2023).

An additional implication relates to regulatory requirements. The EU Artificial Intelligence Act defines transparency, traceability, and accountability as core criteria for the use of AI in high-risk domains, including finance (Burton et al., 2024). The results of this study suggest that SHAP is well suited to meet these requirements from a methodological standpoint: feature attributions are reproducible, quantitatively precise, and based on clearly defined axioms. Consequently, SHAP can support both model validation and auditability, thereby operationalizing regulatory guidelines in practice (Minh et al., 2022). In conjunction with established regulatory frameworks such as Basel III and the General Data Protection Regulation (GDPR), SHAP can serve as an integrative tool to strengthen transparency and governance in data-driven financial models (Chan et al., 2025).

Furthermore, the findings have implications for responsible AI and governance. The transparency generated by SHAP reduces information asymmetries between data analysts, management, and external stakeholders (Notovich et al., 2023), thereby facilitating accountability in algorithmic decision-making processes (Heidbrink, 2025; Von Eschenbach, 2021). In this sense, SHAP can be understood as a tool that connects scientific validity, regulatory requirements, and managerial practice (Rane et al., 2023).

At the same time, the results highlight important limitations. SHAP captures statistical relationships rather than causal effects (R. I. Hamilton & Papadopoulos, 2024), and its correct economic interpretation therefore requires domain expertise to contextualize feature attributions within underlying financial mechanisms (Salih et al., 2025). Moreover, practical implementation requires careful consideration of challenges such as feature correlation and the choice of baseline values (Lundberg & Lee, 2017; Salih et al., 2025). Overall, the integration of explainable AI extends beyond technical implementation and necessitates analytical as well as organizational capabilities, an issue that is further discussed in the following section on limitations.

4.4. Methodological and Practical Limitations

Despite the high methodological rigor and empirical relevance of this study, several limitations must be acknowledged. These arise from the underlying dataset, model assumptions, and the inherent properties of explainable methods such as SHAP. Rather than diminishing the validity of the findings, these limitations define the boundaries of generalizability and provide valuable directions for future research.

Although XGBoost combined with SHAP enables a powerful and transparent modeling framework, the methodology entails structural constraints. A central issue concerns multicollinearity (Wooldridge, 2016): in the presence of highly correlated input

variables, SHAP distributes attribution values proportionally across multiple features (Salih et al., 2025), which complicates the clear interpretation of economically related indicators. This is particularly relevant in financial datasets, where profitability and margin-based variables are often strongly interrelated, potentially leading to overlapping attributions (Chan et al., 2025).

In addition, SHAP is sensitive to the choice of baseline values, meaning that the reference point influences the magnitude of the attributed effects (Lundberg & Lee, 2017). In heterogeneous datasets, this can limit the comparability of explanations across observations or model specifications.

Furthermore, SHAP values capture statistical associations rather than causal mechanisms (Murdoch et al., 2019). They indicate which variables the model considers relevant but do not establish whether these relationships are causally valid in an economic sense (R. I. Hamilton & Papadopoulos, 2024). This underscores a fundamental distinction between predictive inference and causal inference: while the former optimizes the minimization of expected loss to forecast market capitalization based on observational patterns, the latter strictly requires isolating invariant structural parameters under counterfactual interventions. Because tree-based models and post-hoc attribution methods like SHAP operate purely on conditional probabilities, they map systemic associations rather than the physical mechanisms of the underlying financial market. Consequently, high feature importance in our model indicates strong predictive utility within the 2018 regime, but does not imply that manipulating these corporate metrics would directly generate a corresponding change in market value. To bridge this epistemological gap, future research should integrate formal causal graphs and explore structural causal inference frameworks, such as Causal SHAP, to explicitly model the directed dependencies and address the confounding structures inherent in financial data. Consequently, theoretical validation through domain expertise remains essential for meaningful interpretation.

The stability of the results is also context-dependent. Although TreeSHAP provides deterministic outputs (Dwivedi et al., 2023), variations in hyperparameters, feature selection, or training data may lead to slight differences in attribution values (Janiesch et al., 2021). This sensitivity highlights the importance of robust validation strategies and critical model assessment. Specifically, while TreeSHAP delivers exact and deterministic mathematical solutions for a fixed ensemble, full second-order interaction matrices (SHAP interaction values) and empirical stability assessments across extensive resamples were beyond the intended scope of this comparative baseline study, representing important avenues for future computational research.

The theoretical interpretation of the empirical findings is grounded in established financial frameworks, including the models of (Fama & French, 2006), the DuPont Analysis (Soliman, 2007), and capital structure signaling theory (Chan et al., 2025). While these frameworks provide valuable heuristic guidance, they capture only specific aspects of real-world market dynamics. The relationships identified through SHAP must therefore be interpreted within the context of the model and data: they reflect patterns present in the training data rather than the full complexity of economic causality (Salih et al., 2025; Wasserbacher & Spindler, 2022).

Moreover, this study focuses on interpretation within a single model framework. A cross-model comparison, such as between SHAP and alternative explanation methods like LIME or Integrated Gradients, was beyond the scope of this analysis. As a result, it remains unclear to what extent the observed patterns generalize across different model architectures. In addition, the economic interpretation of certain interaction effects, particularly those related to growth variables or dividend metrics, remains partly qualitative and requires further empirical validation.

From a practical and regulatory perspective, SHAP contributes significantly to the transparency of data-driven valuation models but does not address all relevant dimensions. While it ensures quantitative interpretability at the level of individual model decisions, it does not replace comprehensive governance structures across the model lifecycle. Aspects such as data provenance, bias detection, and ethical evaluation remain outside its methodological scope (Burton et al., 2024).

Another limitation concerns the practical implementation of interpretable AI in finance. The benefits of SHAP can only be fully realized when accompanied by sufficient domain expertise. Misinterpretation by users lacking financial or methodological knowledge may lead to so-called “illusory interpretability,” where explanations are formally correct but economically misleading (Notovich et al., 2023). This challenge is particularly relevant in regulatory contexts, where explainable methods may be incorrectly perceived as direct evidence of model validity (Salih et al., 2025).

Finally, computational complexity represents a further constraint. As the number of features and dataset size increase, the computational burden of SHAP grows significantly (Dwivedi et al., 2023), limiting its scalability in real-time applications or more complex deep learning environments (R. I. Hamilton & Papadopoulos, 2024). This limitation highlights the need for future research on more efficient approximation techniques and model-agnostic explanation frameworks that balance performance and interpretability.

Taken together, these limitations underline that explainable methods such as SHAP, despite their substantial analytical value, do not replace theoretical modeling or economic judgment. Instead, their strength lies in complementing and validating classical valuation approaches (Salih et al., 2025).

In conjunction with the theoretical and practical implications discussed in Section 4.3, SHAP contributes to increased trust through its transparency, reproducibility, and compatibility with regulatory requirements. At the same time, the identified limitations demonstrate that this trust does not arise from the method itself but from its responsible application and critical interpretation.

SHAP can therefore be understood as a bridge between data-driven performance and scientific rigor: an approach that enables transparency and trust while remaining dependent on methodological discipline, theoretical grounding, and domain expertise.

4.5. Study-Specific Limitations and Directions for Future Research

In addition to the general methodological and practical limitations discussed above, this study is subject to several study-specific constraints arising from the chosen dataset, the observation period, and the empirical design.

The dataset underlying this analysis is based on a cross-sectional sample of U.S. firms for the year 2018. This temporal and geographical restriction limits the generalizability of the findings to other market phases or countries characterized by different regulatory and accounting frameworks (Andreas, 2007). Crucially, the choice of a single cross-sectional year from 2018 reflects a relatively stable, pre-COVID macroeconomic environment characterized by steady economic growth and standard baseline volatility. Because the data precedes the severe systemic shocks, supply chain disruptions, and shifting monetary policies associated with the pandemic and subsequent recovery periods, this specific temporal window may limit the immediate generalizability of our findings to periods of intense market stress or structural macroeconomic crises. As no panel structure was available, time-dependent dynamics such as business cycles, monetary policy shifts, or structural breaks could not be explicitly modeled. As no panel structure was available, time-dependent dynamics such as business cycles, monetary policy shifts, or structural breaks could not be explicitly modeled (J. D. Hamilton, 1989).

The data preprocessing procedures applied, such as winsorization and median imputation, enhance robustness but may also smooth extreme values, thereby partially attenuating real market variability (Cohen et al., 2021). In addition, the dataset exhibits sectoral heterogeneity, which may introduce minor weighting effects. Overall, the findings should therefore be interpreted primarily as representative of developed capital markets with high information efficiency.

Due to the quantile transformation applied during preprocessing, SHAP values reflect feature impact based on relative ranks (percentiles) within the sample distribution rather than raw absolute values (e.g., a higher relative rank of EBITDA margin rather than an absolute percentage threshold).

Despite these limitations, the study opens several promising avenues for future research, which directly arise from the constraints related to data, design, and model choice.

A key research direction concerns the integration of causal inference methods into explainable AI frameworks. While SHAP provides transparent representations of statistical relationships, it does not allow for causal interpretation. Future research could therefore build on approaches such as “Causal SHAP,” which integrate causal discovery algorithms like the Peter-Clark Algorithm or IDA. Key challenges include the treatment of latent variables, the modeling of causal uncertainty, and scalability to high-dimensional datasets (Ng et al., 2025).

Another important research avenue involves the extension to time-series and panel data. The present cross-sectional analysis does not capture dynamic effects, which are highly relevant in financial applications. Future studies could apply XAI methods such as SHAP to time-series models, for example through approaches like TSHAP, to analyze time-varying value drivers, business cycles, or regime shifts. The development of context-sensitive explanations for real-time applications, such as high-frequency trading, also represents a largely unexplored but practically relevant field (Nguyen & Ifrim, 2026).

Finally, a systematic comparison of different XAI methods represents a further research opportunity. The exclusive use of SHAP in this study leaves open the question of whether alternative explanation techniques may yield more robust or actionable insights. Future research should therefore compare attribution-based methods such as SHAP with local and counterfactual approaches, including LIME and Counterfactual Explanations. In particular, counterfactual explanations offer significant potential for auditability, fairness analysis, and “actionable recourse,” which are of particular importance in regulated domains such as finance (Bensafi & Usun, 2023).

The reflection of these limitations concludes the discussion and leads to the final section, in which the research question is addressed and the main findings are summarized.

5. Conclusions

The central research question of this study, whether explainable approaches, specifically the combination of XGBoost and SHAP, enable a transparent and economically interpretable valuation of firms, can be clearly answered based on the empirical results.

The analysis was guided by a central research objective (RO1) to evaluate whether explainable methods enhance model transparency by revealing economically consistent, non-linear threshold behaviors, alongside two core sub-hypotheses: that primary value drivers align across both linear and non-linear paradigms (H2a), while secondary financial indicators diverge in their rankings due to context-dependent interactions (H2b).

The findings demonstrate that modern machine learning models can achieve high predictive performance while remaining interpretable through explainable methods. Core financial value drivers, including earnings, profitability, and valuation multiples, are consistently identified across both linear and non-linear approaches. However, whereas linear

regression is constrained by its assumption of additive and constant marginal effects, which leads to systematic underestimation of context-dependent variables related to efficiency and capital structure, SHAP further reveals non-linear effects, threshold behaviors, and interaction patterns that classical regression cannot capture. The combination of XGBoost and SHAP therefore proves superior not only in predictive accuracy but also in the richness and nuance of the economic interpretation it affords.

The study therefore demonstrates that explainable AI methods provide substantial added value beyond purely predictive models. They enable a nuanced analysis of economic relationships without sacrificing model performance. In the context of corporate valuation, SHAP emerges as a particularly suitable tool for explaining complex model decisions in a transparent and economically meaningful way.

The empirical findings clearly confirm Objective RO1, as SHAP significantly enhances the transparency and interpretability of machine learning models. Hypothesis H2a is fully supported, confirming that primary fundamental drivers align across paradigms, while Hypothesis H2b receives partial, descriptive support due to observed rank shifts in secondary indicators that reflect underlying data heterogeneity.

The contribution of this study can be summarized along three dimensions:

- **Methodological:** Demonstrating the complementarity of high-performance machine learning models and explainable methods in simultaneously improving predictive accuracy and interpretability.
- **Theoretical:** Providing an empirical refinement of classical financial valuation frameworks through the identification of non-linear effects, threshold dynamics, and context-dependent relationships.
- **Practical:** Contributing to governance, auditability, and regulatory compliance in data-driven valuation models within financial contexts.

Ultimately, these findings provide important conceptual and practical implications for financial modeling. For financial analysts, the integration of SHAP values with tree-based machine learning offers a viable methodological framework to bridge predictive capacity with transparent equity analysis, illustrating non-linear patterns that standard linear models omit. From a governance and regulatory perspective (e.g., within the context of emerging guidelines such as the EU Artificial Intelligence Act), this framework serves as an empirical case study illustrating how post-hoc explainability techniques can operationalize traceability and transparency in algorithmic valuation pipelines. While the empirical validation is situated within a cross-sectional 2018 baseline and does not substitute for empirical user-decision studies or longitudinal panel testing, it outlines how explainable machine learning can support structured, auditable corporate financial workflows without defaulting to black-box opacity.

Author Contributions: Conceptualization, A.A.M. and H.M.; methodology, A.A.M. and H.M.; formal analysis, A.A.M.; data curation, A.A.M.; writing—original draft preparation, A.A.M.; writing—review and editing, A.A.M. and H.M.; visualization, A.A.M.; supervision, H.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: We acknowledge support by the Open Access Publication Fund of TU Berlin. AI-assisted tools were used for grammar correction, punctuation, and language editing. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
ML	Machine Learning
XAI	Explainable Artificial Intelligence
SHAP	Shapley Additive Explanations
XGBoost	Extreme Gradient Boosting
OLS	Ordinary Least Squares
R ²	Coefficient of Determination
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
DCF	Discounted Cash Flow
EV	Enterprise Value
EBITDA	Earnings Before Interests, Taxes, Depreciation and Amortization
P/E	Price-to-Earnings Ratio
M&A	Mergers and Acquisitions
EU	European Union

References

- Amarna, A., Hani, L. B., Mohsin, H., Al-daeef, H., Al-alawneh, N. A., & Samara, H. (2025). The impact of financial transparency on the quality of financial and banking services. *An-Najah University Journal for Law and Economics*, 1(2), 151–168. [\[CrossRef\]](#)
- Andreas, S. (2007). Comprehensive multiples valuation. In *Equity Valuation Using Multiples* (pp. 48–83). Gabler. [\[CrossRef\]](#)
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2019). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *arXiv*. [\[CrossRef\]](#)
- Bensafi, A.-E.-H., & Usun, S. (2023). Explainable AI models in financial risk prediction: Bridging accuracy and interpretability in modern finance. *Electronics Communications, and Computing Summit*, 1(1), 67–75.
- Breiman, L. (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 16(3), 199–231. [\[CrossRef\]](#)
- Burton, C., De Boel, L., Padova, Y., & Theodorakis, N. (2024). *10 Things you should know about the EU artificial intelligence act*. Wilson Sonsini Goodrich & Rosati. Available online: https://www.wsgr.com/a/web/qrkz1SnNzWw6nk7B3oAyDa/10-things-you-should-know-about-the-eu-artificial-intelligence-act_v2.pdf (accessed on 28 August 2026).
- Chan, C.-P., Tsai, C.-H., Tang, F.-K., & Yang, J.-H. (2025). A SHAP-based comparative analysis of machine learning model interpretability in financial classification tasks. *Journal of Applied Economic Sciences*, 20(3), 385. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *arXiv*. [\[CrossRef\]](#)
- Cohen, S. N., Snow, D., & Szpruch, L. (2021). Black-box model risk in finance. *arXiv*. [\[CrossRef\]](#)
- Copeland, T. E., Koller, T., & Murrin, J. M. (1992). Valuation: Measuring and managing the value of companies. *Long Range Planning*, 25(1), 125. [\[CrossRef\]](#)
- Damodaran, A. (2002). *The dark side of valuation: Firms with no earnings, no history and no comparables*. New York University.
- Dixon, M. F., Halperin, I., & Bilokon, P. (2020). *Machine learning in finance: From theory to practice*. Springer International Publishing. [\[CrossRef\]](#)
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. *arXiv*. [\[CrossRef\]](#)
- Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., & Ranjan, R. (2023). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, 55(9), 194. [\[CrossRef\]](#)
- Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, 82(3), 491–518. [\[CrossRef\]](#)
- Felden, B., & Graffius, M. (2024). Grundlagen der unternehmensbewertung. In *Unternehmensbewertung von KMU* (pp. 11–37). Springer Fachmedien Wiesbaden. [\[CrossRef\]](#)

- Fernandez, P. (2002). *Valuation methods and shareholder value creation* (SSRN 621985). Social Science Research Network (SSRN). Available online: <https://ssrn.com/abstract=621985> (accessed on 28 August 2026).
- Götze, T., Gürtler, M., & Witowski, E. (2023). Forecasting accuracy of machine learning and linear regression: Evidence from the secondary CAT bond market. *Journal of Business Economics*, 93(9), 1629–1660. [CrossRef]
- Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2), 357. [CrossRef]
- Hamilton, R. I., & Papadopoulos, P. N. (2024). Using SHAP values and machine learning to understand trends in the transient stability limit. *IEEE Transactions on Power Systems*, 39(1), 1384–1397. [CrossRef]
- Heidbrink, L. B. (2025). Methodik und datenbasis. In *Quantitative prognosen in der externen unternehmensrechnung* (pp. 125–180). Springer Fachmedien Wiesbaden. [CrossRef]
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. [CrossRef]
- Liu, J., Nissim, D., & Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research*, 40(1), 135–172. [CrossRef]
- Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *arXiv*. [CrossRef]
- Majka, M. (2024). *The role of regression analysis in financial modeling*. Available online: https://www.researchgate.net/publication/384208120_The_Role_of_Regression_Analysis_in_Financial_Modeling (accessed on 28 August 2026).
- Malkiel, B. G. (2018). *Burton Malkiel's a random walk down wall street*. Macat Library. [CrossRef]
- Minh, D., Wang, H. X., Li, Y. F., & Nguyen, T. N. (2022). Explainable artificial intelligence: A comprehensive review. *Artificial Intelligence Review*, 55(5), 3503–3568. [CrossRef]
- Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(44), 22071–22080. [CrossRef] [PubMed]
- Ng, W. Y., Wang, L. R., Liu, S., & Fan, X. (2025). Causal SHAP: Feature attribution with dependency awareness through causal discovery. *arXiv*. [CrossRef]
- Nguyen, T. L., & Ifrim, G. (2026). TSHAP: Fast and exact SHAP for explaining time series classification and regression. In R. P. Ribeiro, B. Pfahringer, N. Japkowicz, P. Larrañaga, A. M. Jorge, C. Soares, P. H. Abreu, & J. Gama (Eds.), *Machine learning and knowledge discovery in databases. Research track* (Vol. 16016, pp. 60–77). Springer Nature. [CrossRef]
- Notovich, A., Chalutz-Ben Gal, H., & Ben-Gal, I. (2023). Explainable artificial intelligence (XAI): Motivation, terminology, and taxonomy. In L. Rokach, O. Maimon, & E. Shmueli (Eds.), *Machine learning for data science handbook* (pp. 971–985). Springer International Publishing. [CrossRef]
- OECD. (2022). *OECD framework for the classification of AI systems: Vol. 323. OECD digital economy papers*. OECD. [CrossRef]
- Rane, N., Choudhary, S., & Rane, J. (2023). Explainable artificial intelligence (XAI) approaches for transparency and accountability in financial decision-making. *SSRN Electronic Journal*. [CrossRef]
- Rudin, C. (2018). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *arXiv*. [CrossRef]
- Salih, A. M., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Lekadir, K., & Menegaz, G. (2025). A perspective on explainable artificial intelligence methods: SHAP and LIME. *Advanced Intelligent Systems*, 7(1), 2400304. [CrossRef]
- Slack, D., Hilgard, S., Jia, E., Singh, S., & Lakkaraju, H. (2019). Fooling LIME and SHAP: Adversarial attacks on post hoc explanation methods. *arXiv*. [CrossRef]
- Soliman, M. T. (2007). The use of dupont analysis by market participants. *SSRN Electronic Journal. SSRN Paper 1101981*. [CrossRef]
- Steiger, F. (2010). The validity of company valuation using discounted cash flow methods. *arXiv*. [CrossRef]
- Suozzo, P., Cooper, S., Sutherland, G., & Deng, Z. (2001). *Valuation multiples: A primer*. UBS Warburg. Available online: <http://rgquintero.com/media/91f679831d8e9521ffff80e8ffffe906.pdf> (accessed on 28 August 2026).
- Von Eschenbach, W. J. (2021). Transparency and the black box problem: Why we do not trust AI. *Philosophy & Technology*, 34(4), 1607–1622. [CrossRef]
- Wasserbacher, H., & Spindler, M. (2022). Machine learning for financial forecasting, planning and analysis: Recent developments and pitfalls. *Digital Finance*, 4(1), 63–88. [CrossRef]
- Wooldridge, J. M. (2016). *Introductory econometrics: A modern approach* (6th ed.). Cengage Learning.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.