Proceedings of the 8th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2026.36189

Machine Learning Identification of Capital Structure Resilience in New Energy Firms under Energy Market Volatility Shocks

Yujia Liu[1], Xinyi Wang[2*]

[1]School of Earth Sciences and Engineering, Nanjing University, Nanjing, China
[2]School of Social Sciences, Nanyang Technological University, Singapore, Singapore
**Corresponding Author. Email: rara481846778@gmail.com*

Abstract. Global energy markets continue to fluctuate under geopolitical conflicts, supply demand mismatch, and the pressure of low carbon transition. These conditions place stronger external pressure on the financing stability and capital structure adjustment capacity of new energy firms. Photovoltaic and lithium battery firms usually have high R&D expenditure, long capacity expansion cycles, and strong demand for debt financing. High energy price volatility may affect their capital structure resilience through market expectations, financing costs, and cash flow pressure. Unbalanced panel data are constructed from A share listed photovoltaic and lithium battery firms from 2014 to 2024. High volatility energy price shocks are defined when the quarterly return volatility of Brent, WTI, or natural gas prices exceeds the 75th percentile of the sample period. Capital structure resilience labels are then constructed according to post shock leverage, interest coverage ratio, and cash to short term debt ratio. Logistic regression, random forest, XGBoost, and LightGBM are used for classification. SHAP is used to explain key feature contributions. The results show that XGBoost performs well in AUC, F1 score, and recall of low resilience firms. Short term debt ratio, cash to short term debt ratio, operating cash flow volatility, and Brent price volatility are important variables for distinguishing high resilience firms from low resilience firms. The analysis provides an explainable quantitative basis for financial risk identification, debt structure optimization, and investment screening under energy market shocks.

Keywords: Machine Learning, Capital Structure Resilience, New Energy Firms, Energy Price Volatility, Explainable AI

1. Introduction

Global energy markets have become increasingly volatile under geopolitical tension, supply-demand mismatch, and the pressure of low-carbon transition. Changes in international oil and natural gas prices may influence the capital structure stability of new energy firms through investment expectations, financing costs, and risk premiums [1]. Photovoltaic and lithium battery firms are located in capital-intensive segments of the new energy industry chain and are usually characterized

DOI: 10.54254/2755-2721/2026.36189

by high R&D expenditure, long capacity expansion cycles, and strong dependence on debt financing. Existing studies have begun to show that machine learning can capture capital structure dynamics and provide a methodological basis for identifying corporate financing conditions [2]. Based on unbalanced panel data of A-share photovoltaic and lithium battery listed firms from 2014 to 2024, high volatility in Brent, WTI, and natural gas prices is defined as the external shock context. Capital structure resilience labels are then constructed according to post-shock leverage, interest coverage, and cash-to-short-term-debt ratio, followed by explainable machine learning identification of high-resilience and low-resilience firms.

2. Literature review

2.1. Capital structure and financial stability of new energy firms

The capital structure of new energy firms has clear industrial specificity, as financing demand mainly comes from technological R&D, capacity expansion, and supply-chain construction. Debt structure directly affects the stability of corporate operations [3]. Photovoltaic and lithium battery firms often rely on external financing to support scale expansion during high-growth periods, while excessive short-term debt may narrow their cash flow safety margin and increase repayment pressure after market shocks. Oil price uncertainty may affect the investment behavior of renewable energy firms, which means that energy market changes influence not only traditional energy industries but also the financing and investment rhythm of new energy firms [4]. Capital structure resilience should therefore be observed through leverage changes, repayment capacity, and liquidity buffers. A single leverage ratio is not sufficient to reflect the financial pressure-bearing capacity of firms after external shocks.

2.2. Energy market volatility as an external shock context

Energy market volatility is suitable as an external shock context for identifying the capital structure resilience of new energy firms. International crude oil and energy price data released by public databases can support the measurement of high-frequency energy price changes and provide a reproducible basis for constructing energy price volatility [5]. Brent crude oil price series can also be used to calculate quarterly return volatility and then be matched with annual or quarterly firm-level financial data [6]. In the methodological design, energy price volatility should not be directly treated as the resilience label. It should be used to define the timing of external shocks. The actual resilience judgment should come from post-shock financial performance, including whether leverage deteriorates, whether interest coverage remains positive, and whether the cash-to-short-term-debt ratio falls below the industry median. This design creates a clear logical connection among energy shocks, financial performance, and machine learning identification.

2.3. Machine learning for financial resilience identification

Machine learning methods are capable of handling nonlinear relations and multivariable interactions in capital structure resilience identification. Existing studies have applied artificial intelligence models to capital structure prediction, indicating that financial indicators, firm size, and operating variables can be algorithmically combined to form strong classification capacity [7]. XGBoost uses a gradient boosting mechanism to process complex feature relationships and is suitable for identifying nonlinear patterns among high-dimensional financial variables [8]. Random forest

DOI: 10.54254/2755-2721/2026.36189

reduces the instability of a single model through multiple decision trees and can improve robustness in corporate classification tasks [9].

3. Experimental methods

3.1. Sample selection and data construction

Data include firm financial and market information and international energy prices. The sample covers A-share listed photovoltaic and lithium battery firms from 2014 to 2024 and forms an unbalanced panel, with each firm entering from its first complete fiscal year after listing. Firms are identified using Wind, Shenwan, and CSMAR classifications. ST firms, financial firms, firms with insufficient observations, and records with serious missing values are excluded, leaving about 80–100 firms. Firm-level variables include leverage, short-term debt, liquidity, interest coverage, ROA, cash flow, R&D intensity, firm size, and stock return volatility. Brent, WTI, and natural gas prices from EIA, FRED, and the World Bank are used to construct high-volatility energy shock windows.

3.2. Resilience label construction after energy price shocks

Capital structure resilience is classified through energy shock identification, post-shock financial observation, and resilience assessment. High-volatility energy shocks are defined as quarters when the return volatility of Brent crude oil, WTI crude oil, or natural gas exceeds the 75th percentile. Firm capital structure is then observed over the following four quarters. Firms with stable leverage, positive interest coverage, and cash-to-short-term-debt ratios at or above the industry-year median are classified as highly resilient. Firms showing rising leverage, weaker interest coverage, or below-median liquidity are classified as low-resilience firms. Energy price volatility therefore serves only as the external shock context, while resilience is determined by post-shock leverage stability, debt-servicing capacity, and liquidity buffers.

3.3. Model training, class imbalance processing and evaluation

Model training uses capital structure resilience as the classification target, with high resilience coded as 0 and low resilience as 1. Financial, market, and energy-price variables are used as inputs. To avoid information leakage, 2014–2020 data are used for training, 2021–2022 for validation, and 2023–2024 for testing. Logistic regression serves as the baseline, while random forest, XGBoost, and LightGBM capture nonlinear relationships [10]. Class imbalance is addressed using class weights, SMOTE, or scale_pos_weight. Performance is evaluated using AUC, F1, recall, and PR-AUC, with particular emphasis on identifying low-resilience firms. It also uses regularization terms to control model complexity. Its optimization objective is shown in Formula 1.

$$\mathrm{Obj}=\sum_{i=1}^{n} l\left(y_i,\hat{y}_i\right)+\sum_{k=1}^{K}\Omega\left(f_k\right) \tag{1}$$

Here, $l\left(y_i,\hat{y}_i\right))$ denotes the loss function between the true label and the predicted result. $\Omega(f_k)$ denotes the regularization term of the kth tree, which is used to reduce the risk of overfitting. After model training, SHAP is used to explain feature contribution. It analyzes the influence direction of short term debt ratio, cash to short term debt ratio, operating cash flow, R&D intensity, energy price volatility, and other variables on high resilience or low resilience classification [11]. The logic of SHAP value calculation is shown in Formula 2.

DOI: 10.54254/2755-2721/2026.36189

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{j\}) - f(S)] \tag{2}$$

Here, ϕ_j denotes the marginal contribution of the jth variable to the model prediction result. S denotes a feature subset that does not include this variable, and F denotes the full feature set. Through this method, the variables that best distinguish high resilience firms from low resilience firms can be further identified.

3.4. Error analysis and model boundary identification

Error analysis evaluates the identification stability of the machine learning model across different firm types. After training, test-set predictions are compared with the true resilience labels to identify false positive and false negative cases. Since low resilience is coded as 1, a false positive refers to a high-resilience firm classified as low resilience, while a false negative refers to a low-resilience firm classified as high resilience. False negatives deserve particular attention because financially fragile firms may be overlooked. Misclassified samples are further grouped by firm size, leverage, short-term debt ratio, cash-to-short-term-debt ratio, and operating cash flow volatility to identify systematic classification bias. This analysis does not alter model training but clarifies the model's identification boundary and supports further interpretation based on firm growth, refinancing capacity, and industrial-chain position.

4. Results

4.1. Capital structure resilience identification and model validation

Since the resilience label is derived from post shock financial performance after high volatility energy price shocks, model evaluation does not focus only on overall accuracy. It pays more attention to whether the model can identify low resilience firms in a stable way. After class weight and scale_pos_weight are used, ensemble learning models perform better than logistic regression. As shown in Table 1, XGBoost reaches an AUC of 0.842, an F1 score of 0.781, and a recall of 0.806 for low resilience firms. This indicates that the model can effectively capture the characteristics of firms whose capital structure deteriorates after shocks. LightGBM reaches an AUC of 0.831, which is close to the performance of XGBoost. This shows that gradient boosting models are suitable for handling nonlinear relationships between financial variables and energy volatility variables. The AUC of logistic regression is 0.721. It can provide a baseline judgment, but its ability to identify interactions among short term debt ratio, cash buffer, and operating cash flow is relatively limited. The performance of random forest lies between the two types of models. This suggests that multi tree ensemble methods can improve identification stability, while their interpretation accuracy on complex sample boundaries remains weaker than that of XGBoost.

Table 1. Model validation results for capital structure resilience identification

Model	AUC	F1 score	Recall of Low Resilience Firms	PR AUC
Logistic Regression	0.721	0.684	0.672	0.693
Random Forest	0.804	0.742	0.758	0.765
XGBoost	0.842	0.781	0.806	0.798
LightGBM	0.831	0.769	0.791	0.786

DOI: 10.54254/2755-2721/2026.36189

4.2. Feature contribution, shock response patterns and error analysis

As shown in Figure 1, the short-term debt ratio has the highest mean SHAP value of 0.184, with a misclassification rate of 19.2%, indicating that greater short-term repayment pressure is associated with low resilience. The cash-to-short-term-debt ratio has a mean SHAP value of 0.162 and a 15.8% misclassification rate, confirming the importance of liquidity buffers. Operating cash flow volatility shows a mean SHAP value of 0.137 but a higher misclassification rate of 20.7%, suggesting greater uncertainty in identifying resilience boundaries for firms with unstable cash flows. Brent price volatility has a mean SHAP value of 0.119 and mainly captures external shock intensity rather than directly determining classification. Misclassifications are concentrated among small, highly leveraged firms with high short-term debt and volatile cash flows. Because low resilience is coded as the positive class, false negatives are particularly important for risk identification. Expanding lithium battery firms may generate false positives because of strong refinancing capacity despite short-term repayment pressure, whereas small firms with weak disclosure, unstable cash flow, and limited refinancing capacity are more prone to false negatives. Model boundaries should therefore be interpreted in relation to firm growth, refinancing capacity, and industrial-chain position.

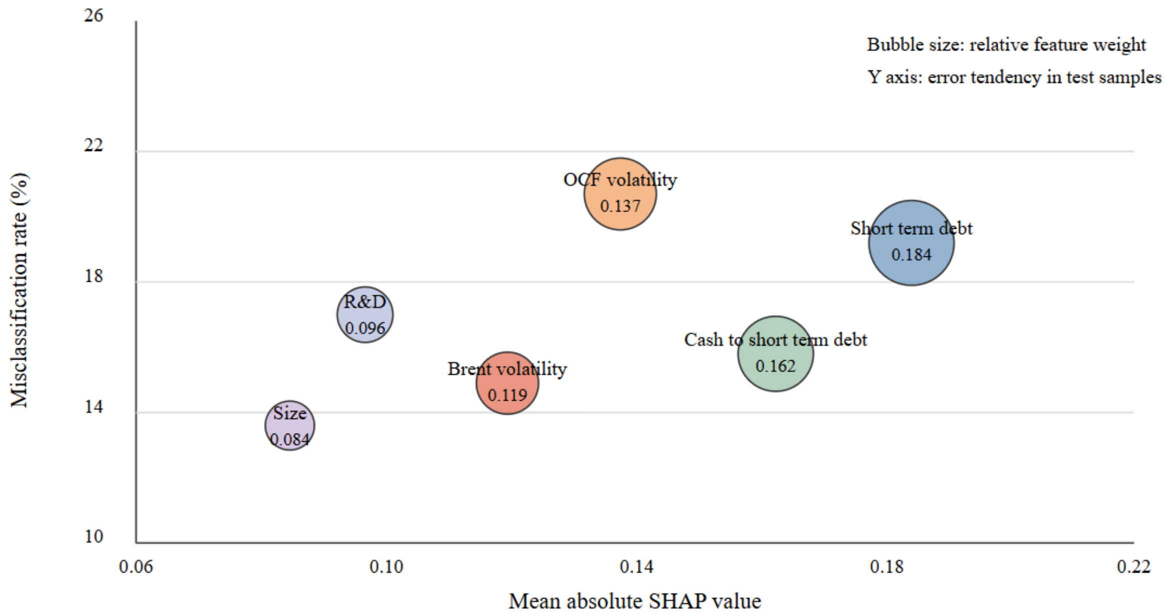

Figure 1. Bubble statistical chart of feature contribution and misclassification rate distribution

5. Discussion

The results show that energy price volatility affects capital structure resilience mainly by amplifying internal financial differences. Short-term debt has the highest contribution, while liquidity buffers and cash flow stability are also important. Brent volatility mainly captures external shock intensity. The results suggest that financing risk premiums, investor expectations, and liquidity buffers may be key mechanisms, although further panel regression, DID, or event-study tests are needed. Misclassification is more common among small, highly leveraged firms with high short-term debt and volatile cash flows, indicating the need to consider firm growth, refinancing capacity, and industrial-chain position.

DOI: 10.54254/2755-2721/2026.36189

6. Research innovation

The innovation lies in the research context, resilience label construction, and methodological application. International energy price volatility is used as the external shock context, linking energy market changes with capital structure risk. Resilience is measured through post-shock leverage changes, interest coverage, and cash-to-short-term-debt ratios, capturing repayment capacity, liquidity buffers, and structural stability. XGBoost, LightGBM, and SHAP are further applied to classify resilience and explain the joint effects of debt structure, liquidity, cash flow, and energy price volatility.

7. Conclusion

The analysis identifies capital structure resilience of new energy firms under energy price shocks through a framework of shock identification, resilience classification, model training, and error analysis. The sample covers A-share photovoltaic and lithium battery firms from 2014 to 2024. Results show that ensemble learning, particularly XGBoost, performs better than linear models in identifying complex resilience patterns. Short-term debt, liquidity, operating cash flow volatility, and Brent price volatility are key predictors. Explainable machine learning therefore provides an effective tool for assessing capital structure risk and financial stability under energy market shocks.

Author contribution

Yujia Liu and Xinyi Wang contributed equally to this paper.

References

[1] Xie, L., & Su, Z. (2025). Energy transition and corporate debt: Evidence from Chinese listed companies. *Energy Economics*, 108944.

[2] Bistuer-Talavera, C., et al. (2024). Capital structure decisions in the energy transition: Insights from Spain. *Utilities Policy, 91*, 101851.

[3] Huang, X., & Yang, L. (2025). Balancing the books: The role of energy-related uncertainty in corporate leverage. *Global Finance Journal, 64*, 101077.

[4] Bhattacherjee, P., & Mishra, S. (2025). Oil shocks and capital structure: Role of ESG across the globe. *International Review of Economics & Finance, 99*, 103982.

[5] Halder, A., & Kannadhasan, M. (2025). Energy uncertainty and corporate bankruptcy risk: International evidence. *Energy Economics*, 108908.

[6] Zhang, W., & Chiu, Y.-B. (2023). Country risks, government subsidies, and Chinese renewable energy firm performance: New evidence from a quantile regression. *Energy Economics, 119*, 106540.

[7] Lin, B., & Xie, Y. (2023). The impact of government subsidies on capacity utilization in the Chinese renewable energy industry: Does technological innovation matter? *Applied Energy, 352*, 121959.

[8] Lin, B., & Zhang, A. (2023). Government subsidies, market competition and the TFP of new energy enterprises. *Renewable Energy, 216*, 119090.

[9] Dang, T. H.-N., et al. (2023). Measuring the energy-related uncertainty index. *Energy Economics, 124*, 106817.

[10] Bai, L., et al. (2023). Diversification effects of China's carbon neutral bond on renewable energy stock markets: A minimum connectedness portfolio approach. *Energy Economics, 123*, 106727.

[11] Uddin, G. S., et al. (2023). Risk network of global energy markets. *Energy Economics, 125*, 106882.