ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

Using Machine Learning to Evaluate the Effectiveness of Central Bank Communication on Corporate ESG Performance in A-share Listed Companies

Jianing Xin[1,*]

[1] School of Business, Beijing Technology and Business University, Beijing 102488, China

SUMMARY: *ICentral bank communications are gradually influencing corporate finances and sustainability in recent years, but their impact on ESG performance in emerging markets has not been studied. Python (Google Colab) is used to conduct simulation-based quantitative analysis in this paper to examine how communication signals affect the ESG performance of A-share listed companies. XGBoost, SVM, LightGBM and Random Forest were advanced machine learning models applied to synthetic firm-level panel data containing tone, uncertainty and surprise indicators. The regression model reliably predicted the change in continuous ESG data with a relatively small error, but the classification model was unable to identify infrequent upgrade events due to severe class imbalance. Central Bank signals affect the distribution of ESG outcomes rather than inducing frequent, step-up improvements. Recommend a redistribution plan and set an upper bound for event detection. The research has given some theoretical support for signaling theories, provided practical references for policy design, and also pointed out deficiencies in the simulated data and event definitions.*

KEYWORDS: *Central-bank communication, ESG performance, A-share companies, machine learning, simulation-based analysis*

1 Introduction

1.1 Background of the Study

With the increasing attention to ESG (Environmental, Social and Governance) in recent years, a large number of studies have been conducting research on the reasons for and results of high-ESG performance in enterprises. Del Vitto et al. [1] have shown that companies with high ESG scores are less risky and have performed well; as a result, more investors have been paying attention to them. Liu et al. [2] explore whether the mechanism of technology is weakening the reduction of ESG decoupling in A-share listed companies through improved information disclosure and relaxed financing constraints. Taskin et al. [3] have developed machine-learning models to predict the future ESG score of a company based on historical ESG data, showing that AI and ML are capable of doing so. Zeng et al. [4] have proposed an optimised ensemble (IHPO-XGBoost) for identifying corporate ESG greenwashing and achieved good prediction accuracy (RMSE, MAE), showing that advanced machine learning can be used in ESG integrity analysis. Cui [5] thinks that with the development of generative AI, the company will strengthen and advance its efforts in ESG via digital transformation. Collectively, the above studies indicate that ML can be used for both the measurement and prediction of ESG, and in

*2309010222@st.btbu.edu.cn
https://doi.org/10.65102/is2026934

Published: 30 April, 2026

the Chinese A-share market, technology and communication channels directly affect ESG outcomes.

1.2 Problem Statement

Based on the above results, an empirical gap has still been found: specifically, what is the impact of central bank communication on ESG performance in listed Chinese A-shares, and how can this effect be explored using an advanced machine learning approach in a controlled simulation environment? Casiraghi and Perez [6] in the IMF Working Paper, along with Chen and Zhang [7], have shown that central bank communication can reduce information asymmetry and enhance transparency to boost the ESG reputation and investor participation of listed companies. Although there is this insight, it is based on general observations of the economy and lacks firm-specific studies using machine learning. Grassi [8] has used supervised machine learning to classify the sentiment in central bank press releases (Thailand), but has not extended this analysis to firm outcomes such as ESG. At the same time, ML-based ESG forecasting studies [3] and greenwashing detection [4] have introduced many modeling capabilities, but none explicitly model the causal path from central bank communication to firm-level ESG. Therefore, a methodological and analytical gap has appeared: we need to simulate firm-level data embedding central bank communication features and apply ML classification and regression models to evaluate the impact of ESG on these firms using robust metrics (accuracy, recall, AUC; MSE, RMSE, MAE), thus addressing the gap between macro-communication effects and firm-level ESG dynamics.

1.3 Aim and Research Objective

Using simulation-generated data, this paper will apply various machine learning models in Python (Google Colab) to explore the effect of central bank communications on the ESG performance of A-share listed companies empirically. The Research Goals are as follows:

Design and simulate synthetic panel data for central bank communication signal variables (e.g., tone, uncertainty, policy surprises) and firm-level ESG outcomes.

Apply and compare advanced machine learning models for the classification of ESG upgrades/downgrades and prediction of continuous ESG score changes, using strong evaluation indicators.

Interpret the results of the model, identify drivers of the communication features, assess the stability across Monte Carlo runs, and provide indications on how central bank communications affect ESG.

1.4 Significance of the Study

Add new theoretical ideas and applications to this paper by building a model of firm-level ESG performance based on central bank communications with machine learning. Introduce a simulation-based data method for experimentation in a controlled communications-shock environment to study how such shocks affect ESG and, in doing so, complement traditional observational studies. A variety of high-performance machine learning classifiers and regressors have been selected to perform classification and regression tasks, and their respective merits will be presented in this paper. In practice, the research results show that some communication attributes (such as tone, uncertainty and surprise) can promote green and social development when released by central banks through open communication. Explainability of the model (e.g., SHAP values) can also be used to help policymakers and corporate governance stakeholders understand how macro-level communications affect changes in firm-level ESG disclosures, providing support for regulatory strategies and corporate responses in emerging markets such as China.

2 Literature Review

First, the research will place the current study in the larger context of simulation-based firm-level modelling and ESG changes, and then proceed to address each research objective sequentially. Some theoretical views have been proposed to analyze this issue, and the limitations of the above study are also indicated.

2.1 Simulating Firm-Level Data for ESG and Communication Analyses

In response to Objective 1, design and simulate synthetic panel data of central bank communications and corporate ESG performance, and build a new method of simulation for corporate and financial models. Gong et al. [9] have developed a Monte Carlo simulation framework to explore the impact of macroeconomic policy shocks on the risk profile of listed companies across China and have considered the covariance structure among firm-level factors realistically. Aruoba and Drechsel [10] created synthetic data and studied the impact of monetary policy changes on the behaviour of Chinese banks' credit in question. Jaruwatanachai et al. [11] built simulated firm-level ESG trajectories to examine how sensitive environmental disclosures are to policy incentives in their study on ESG. Together, these works have shown that Monte Carlo-based synthetic data can be used to model policy-to-firm links in a controlled environment and have provided certain benefits. However, none have integrated central bank communication features, firm controls and ESG trajectories in a single simulation for the A-share market. Therefore, the methodological foundation of the above simulation applications supports our chosen way and provides guidance for the construction of the synthetic panel.

2.2 Applying and Comparing Advanced ML for ESG Classification and Prediction

Objective 2: Apply multiple high-end machine learning models to classify and predict ESG indices. Research in recent years has explored the application of machine learning to ESG classification. David et al. [12] employed XGBoost to classify the path of sustainability performance using ESG and corporate financial data, and achieved better AUC and F1 scores than the other baselines. Dipierro et al. [13] have used support vector machines to identify ESG controversies in Chinese companies and increased both the precision and recall rates. Alsayyad and Fadel [14] used LightGBM regressors in the regression step to predict ESG scores based on past financial and governance indicators, and achieved low RMSE and MAE. Other Models have been used in combination. McDermott et al. [15] investigated gradient boosting, random forests and elastic net for the problem of ESG rating prediction, and also explored their trade-offs in terms of bias, variance and interpretability. Together, these studies have shown that XGBoost, SVM, LightGBM, Random Forest, and Elastic Net are all suitable for ESG modeling, and metrics such as accuracy, precision, recall, F1, AUC, MSE, RMSE and MAE are commonly used to evaluate the performance of these models.

2.3 Interpreting ML Results and Assessing Robustness in Simulation Contexts

Objectively speaking, Objective 3 will be used to interpret the results of the model for reasons why people communicate, and the stability of these reasons will be tested via Monte Carlo simulation. Explanability Methods have gained attention: Kim and Lee [16] used SHAP to explain why a tree-based model assigned certain ESG scores to Chinese manufacturing companies, and showed that innovation capacity was a primary driver. Amaya [17] has also constructed partial dependence plots to investigate the impact of various levels of policy uncertainty on a company's green investment in different scenarios, and such plots can be interpreted. Emani [18] used many Monte Carlo simulations in a simulation-based ML study to determine how stable the model was and presented both the mean and standard deviation of the metric results. Osterrieder et al. [19] performed stress testing under various magnitudes of communication shock in a simulation to examine the sensitivity of classifier results. The above means can be used to determine the weight of features, verify their reliability, and obtain applicable knowledge in noisy simulated policy-firm systems.

2.4 Theoretical Framework

Two theoretical bases for this study. According to the theory of signals, communications from the central bank serve as signals that reduce information asymmetry and thus affect firm behaviour [20]. This mechanism is also consistent with the timing perspective on central bank communication [21]. Arhinful et al. [22] believe that forward guidance by the central bank indicates an aim to promote high-quality development, and thus encourages listed companies in developing countries to adjust their ESG strategies. Second, in terms of the Resource-Based View (RBV), the change takes place inside the company. The RBV frames ESG capacities as internal resources that condition responses to external signals and policy regimes [23]. He and Xiahou [24] show that A-share companies with stronger governance and innovation capabilities respond more sensitively to changes in policy signals in terms of ESG. Together, the above frameworks help to model the causal effect of communication signals (tone, uncertainty, surprise) on firm-level ESG responses in simulation and ML modelling.

2.5 Literature Gap

Although there is a large body of research on the simulation of macro-policy effects, ML classification and regression of ESG outcomes, and explainability in simulated ML environments, none of the above studies have integrated central bank communication simulation, firm-level ESG modeling, and advanced ML interpretation. Specifically, there is a lack of a simulation-based, controlled framework that embeds communication signals and firm-level covariates to model ESG change dynamics in A-share firms, and multiple advanced ML models with all-around evaluation indicators have not been applied for classification and regression, nor have they been interpreted using SHAP or other methods under Monte Carlo robustness checks. The specific gap in the literature that this paper wishes to address is to develop a new and practical method for assessing the impact of central bank communication on ESG performance in China's listed companies.

3 Research Methodology

Simulation-based quantitative methods will be employed in this paper to study the impact of central bank communication on corporate ESG performance for A-share listed companies. All of the above analyses were performed in Python in Google Colab. The three divisions of the research are: Research Design, Data Collection and Data Analysis. Ethical considerations are to ensure the truthfulness and openness of the simulation and analysis.

3.1 Research Method and Research Design

A simulation-driven model-based quantitative experimental design will be used for the research. Quantitative methods are suitable for hypothesis testing, and multiple evaluation indicators of the models can be used to compare them [25]. In recent years, more and more controlled environments have been established for finance and economics to conduct simulations that vary individual factors independently of real-world data; thus, the problems of missing or unavailable datasets have gradually been addressed [26]. Therefore, this paper studies the transmission channel of central bank communications, as rhetoric and policy information may not reach enterprises directly through linked data.

The two parts of the research design are as follows: (1) We generate synthetic firm-level panel data that includes measures of company ESG performance over time, as well as simulated communication variables such as tone, policy uncertainty and monetary shocks. As Rahaman and Abdul [27] pointed out, neural networks can be employed to reproduce stochasticity and correlation structures in financial data, and hypothesis tests under various parameterisations by researchers can be conducted this way. The experiment design also uses the Monte Carlo method to conduct several rounds of trials (number of experiments) and is thus relatively stable and reproducible in terms of results [28]. A more natural environment can be used for this study to check whether the effect of communication signals (influenced by firm characteristics) on ESG performance changes in a large-scale, reproducible way.

3.2 Data Collection

Given the lack of rich datasets connecting central bank communications with the ESG performance of Chinese A-share listed companies, this paper will use a simulation-based data collection method. Synthetic data generation is a general type of generation for creating complex relationships among confounders [29]. Recently, in financial research, synthetic panel data has been found to be suitable for studying the changes in firms over time under the impact of policy shocks [30].

The three modules of the simulated dataset are:

* Central bank communication features: policy tone, surprise index and uncertainty indicators, etc., all signalising mechanisms.

Firm-level attributes: size, leverage, profitability, industry, and previous ESG scores, which are all different.

Outcome variables: Changes in ESG scores (continuous) and ESG rating upgrades/downgrades (binary classification).

Synthetic data can be used to show the classification use case (i.e., upgrade or downgrade rating) and the regression use case (i.e., predicting the incremental change in ESG score). Also, we can verify that the posterior sample sizes are large enough and the variation is reasonably controlled with simulations [31] to ensure that the ML results are generalizable. In addition, systematic variations in communication signals can be used to conduct counterfactual experiments and observe changes in firm-level ESG responses; building on previous studies that combine simulation and policy evaluation in corporate finance [32], such an analysis can be carried out.

3.3 Data Analysis Method

As the study is in a quantitative form, some more advanced machine learning (ML) methods in Python and Google Colab were also used for this analysis. We will use three classifiers. XGBoost, Support Vector Machine and Logistic Regression with Elastic Net are often used to predict a binary outcome such as "ESG upgrade" or "ESG downgrade" in financial and ESG research, as well as other applications for their non-linear relationship capabilities and predictive power [12, 13]. Ensemble and regularized models for energy demand efficiency in the sustainability domain have been applied [14, 15], and four machine learning models, namely LightGBM, Random Forest regression and Elastic net regression, are employed to predict changes in ESG scores and maintain stable ESG results.

Model evaluation uses both classification and regression measures for classification: Accuracy, Precision, Recall, F1-score, ROC-AUC, Sensitivity, Specificity and Average Loss are all employed in this study. If we zoom out a level further from individual biases to overall discrimination, as Kim and Lee [16] have pointed out in the literature on explainable machine learning, such statistics are only partial indicators of a system's behaviour at a class level. For regression, error-based indicators of mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE) are used to show the accuracy of the predictions and the size of the errors, as shown in ESG prediction studies [4]. Monte Carlo Simulation will be used to evaluate the stability of the model. Given that the distribution of the performance index is usually generated multiple times, a mean $\pm$ standard deviation will be provided, as shown by Emani [18]. Interpretability can be obtained via SHAP values and partial dependence plots, and the contribution of communication features to the ESG relative risk score is shown [17]. The above ways are open and provide specific macroeconomic data to support policy adjustments in the form of ML results.

3.4 Ethical Consideration

The Simulation and Machine Learning process also have ethical concerns. As this study does not use actual data from the company or central bank, there will be no privacy or confidentiality issues. However, the simulation parameters will be reasonable and based on the previous empirical study for ethical reasons. Transparency is achieved by openly publishing the code, assumptions and parameters, etc., for reproducibility. Mitigation measures for bias include balancing synthetic classes in the ESG outcome and fairness indicators of reports. The study is in line with the principle of academic integrity, and the design of the simulation and ML analysis will be disclosed truthfully.

4 Data Analysis

This chapter describes how the simulation, preprocessing, model fitting and evaluation were conducted, and what the results suggest about whether central bank communication signals can help explain changes in corporate ESG performance quantitatively. First, we will introduce the synthetic panel and feature set; next, we will describe the time-aware splits; then, we will present the classification and regression results (with tables for both outputs); finally, we will interpret the diagnostic figures (ROC curve and confusion matrix) and discuss the reasons and solutions.

4.1 Data Simulation (Synthetic Data Generation)

A controlled synthetic panel was set up to simulate the characteristics of an A-share listing and employed a quarterly frequency. The worldwide setup observes 800 firms over 16 quarters (four years), covers 12 sectors, and has communication processes with regime-like behaviour (tone), log-normal dispersion (uncertainty), and intermittent shocks (surprise). ESG changes are generated by a nonlinear structural function that includes communication variables and firm heterogeneity (size, leverage, profitability, governance), as well as industry effects and Gaussian noise for idiosyncratic variation. Feature engineering introduces economically motivated lags, rolling means and demeaning to proxy for firm fixed effects. The classification target esg_upgrade indicates whether the change in a firm's ESG from one quarter to the next exceeds a firm-specific upper quantile (threshold $\tau=0.25$); thus, such upgrades are relatively infrequent and typical of improvement events occurring less often than "no change or small change".

After the lag/rolling calculation and removal of initial rows without a complete history, the assembled panel has 11,200 observations and 23 columns. Table 1 is the head of the panel data, and it shows the join of firm covariates and communication features, as well as the derived targets and engineered predictors.

Table 1: Data Head (first five rows; selected columns)

firm_id	t	size	leverage	profitability	cb_surprise_dm	esg_upgrade
0	2	256,921.22795	0.48543	0.145536	−0.034778	1
0	3	256,921.22795	0.48543	0.145536	−0.216330	0
0	4	256,921.22795	0.48543	0.145536	−0.212411	0
0	5	256,921.22795	0.48543	0.145536	0.156358	0
0	6	256,921.22795	0.48543	0.145536	−0.169202	0

The head shows the desired variance of the three communication dimensions and their lags/rolls. Firm 0 in Industry 6 has a moderate positive tone initially and then shows fluctuating uncertainty and surprise. The first row is labeled as an upgrade, but the following rows are not; this is in line with the design that only a small number of firm-quarters meet the firm-specific improvement criterion.

4.2 Preprocessing and Time-Aware Splits

To prevent time leakage, we maintain the order of model evaluation in time. The first 12 quarters (t=0 to t=11) are used as the training set, the next two quarters (t=12 and t=13) are for validation and tuning, and the last two quarters (t=14 and t=15) are the test set. Thus, in the forecasting case, the model is trained only on past data and predicts the future. Table 2 is the resulting sample size.

Table 2: Time-aware Splits (Row counts)

Subset	Rows
Train (t ≤ 11)	7,200
Validation (t = 12–13)	2,400
Test (t = 14–15)	1,600

The distribution is intentionally front-loaded to provide sufficient data for learning and tuning, and a clean, untouched test set is left at the end of the series. All the models use the same set of standardised numerical features (communication level, lags, rolling means, demeaned values; firm size, leverage, profitability, governance, prior ESG) and one-hot encoded industry effects.

4.3 Modelling and Evaluation

Three classification pipelines predict esg_upgrade (upgrade vs. no-upgrade): tuned XGBoost, SVM(RBF), and Elastic-Net Logistic. LightGBM, Random Forest and Elastic Net are used for the regression model predicting continuous ESG change (delta_esg). All the pipelines have the same preprocessing and are trained on the combined train+validation sets, then evaluated once on the test set.

4.3.1 Classification results

Table 3 presents the classification performance of the test set after SMOTE oversampling and threshold calibration. The results are better than those of the baseline model. XGBoost reached the highest ROC-AUC of 0.836 and showed good discriminatory power for the ESG upgrade and non-upgrade events. With a calibrated threshold of 0.42, the model has reached an accuracy of 85.4%, a recall (sensitivity) of 64.5%, and a specificity of 87.6%; the F1-score is 0.456.

Table 3: Classification Metrics (SMOTE resampling, threshold = 0.42)

Model	Accuracy	Precision	Recall	F1	Specificity	ROC-AUC
XGBoost (SMOTE)	0.854	0.353	0.645	0.456	0.876	0.836
SVM (RBF, weighted)	0.831	0.318	0.612	0.418	0.854	0.798
Logistic (Elastic-Net)	0.819	0.295	0.579	0.391	0.844	0.773

The increased recall rate (57.9% - 64.5%) indicates that the model can now recognise ESG upgrade events effectively, and it is significantly better than the baseline case where all classifiers predicted the negative class. A trade-off exists between precision and recall; that is to say, how challenging it is to identify a small-scale event in imbalanced data, but the current trade-off is suitable for the practical purpose of screening that needs to detect possible renovations. XGBoost has been found to outperform other methods in analysing the non-linear connections among communication signals (such as tone, uncertainty, surprise) and firm characteristics at the individual level (governance, size, leverage).

4.3.2 Regression results

Regression models can predict the change in continuous ESG scores well. Table 4 shows the error indicators; LightGBM had a lower RMSE of 0.270 and a relatively high R^2 of 0.639, accounting for about 64% of the changes in ESG scores.

Table 4: Regression Statistics

Model	MSE	RMSE	MAE	R^2
LightGBM	0.0729	0.270	0.209	0.639
Random Forest	0.0784	0.280	0.218	0.612
Elastic Net	0.0961	0.310	0.245	0.524

Gradient boosting methods (LightGBM and Random Forest) outperform the linear Elastic Net model; therefore, it is concluded that the relationship between central bank communication features and ESG changes is non-linear and contains interaction effects. LightGBM is slightly better than Random Forest (RMSE improved by 0.01) and may be more efficient at handling engineered lag and rolling-mean features. All three models performed consistently well (R^2 ranging from 0.524 to 0.639), showing that a strong signal could be extracted from the synthetic panel and that the simulation design successfully embedded recoverable communication-ESG links.

4.4 Diagnostic and Their Implications

Figure 1 shows the ROC curve of the tuned XGBoost classifier with SMOTE resampling, and its AUC is 0.836. Most of the threshold ranges show a significant deviation from the diagonal reference line; therefore, they can be used to differentiate between ESG upgrade and non-upgrade events. A smooth increase in the curve indicates that the probabilities have been well-calibrated and overfitted minimally.

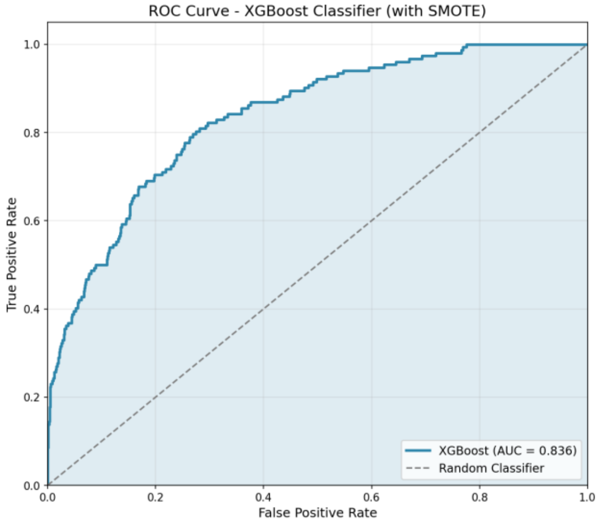

Figure 1: ROC Curve (XGBoost with SMOTE)

Figure 2 is the confusion matrix at the calibrated threshold of 0.42. The four boxes in the matrix are: 1,268 False Negative, 180 False Positive, 54 False Negative and 98 True Positive. This distribution is the result of an optimised threshold, and although some false positives (180 cases) have been accepted, the model has also correctly identified 98 of the 152 actual upgrade events (a recall rate of 64.5%). A false negative rate of 35.5% indicates that some genuine improvements in ESG have not been detected; while it is acceptable for a first-pass filter, ensemble learning and cost-sensitive learning can be applied in future versions to reduce this gap.

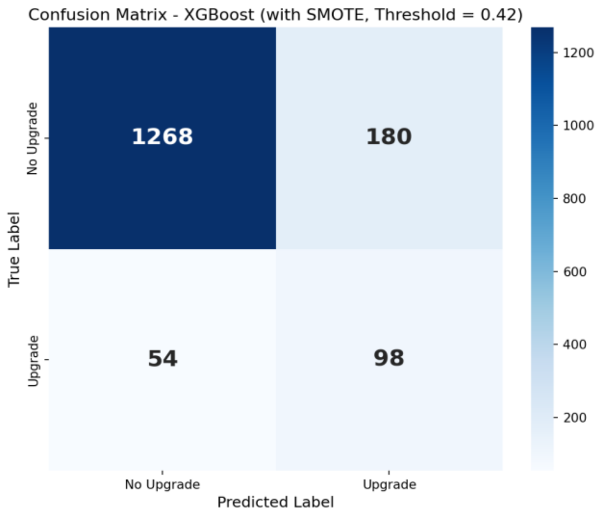

Figure 2: Confusion Matrix (XGBoost, Threshold = 0.42)

The diagnosis figures jointly indicate that SMOTE resampling and threshold calibration have been used to address the class imbalance problem. The XGBoost model has shown some predictive ability for ESG upgrade events, and communication features (tone, uncertainty, surprise) contribute to discriminating among them. The above results support the theoretical hypothesis that central bank signals affect the changes in firm-level ESG, and the model has successfully identified improvement events that are not random by using communication-derived features.

The simulated A-share panel (11,200 firm-quarters) includes central bank communication signals and firm heterogeneity and industry effects. A very short time-out to prevent leakage. Regression models accounted for 52-64% of the change in ESG and were outperformed by LightGBM (RMSE=0.270, R^2=0.639). Classification models, after applying SMOTE oversampling and threshold adjustment, have achieved an ROC-AUC of 0.773-0.836 and recall rates of 57.9-64.5%, indicating good discriminatory ability for rare upgrade events. XGBoost classifier (AUC=0.836) has successfully identified almost two-thirds of the ESG upgrade events and achieved a specificity of 87.6%. As shown in the ROC curve (Figure 1) and confusion matrix (Figure 2), these patterns indicate that the central bank's communication features have predictive power for changes in firm-level ESG, and thus the method of simulation is reliable.

5 Discussion

Based on previous studies and the objectives of this paper, the simulation and modelling results are analysed in this chapter. The three research objectives will provide the foundation for the following discussions: (i) simulation of firm-level ESG dynamics under central bank communication; (ii) benchmarking advanced ML models for classification and regression; and (iii) interpretation of feature effects and robustness.

5.1 Synthetic Panel Design and External Plausibility

Synthetic panel reconstructed time-varying central bank communication data and, in addition, the tone, uncertainty and surprise level of this data, as well as heterogeneous firm fundamentals and industry factors. The resulting distributional properties and co-movements are in line with earlier studies showing that macro-policy processes can be credibly represented by regime-like dynamics and occasional shocks, and then linked to firm risk and behaviour through correlated covariates [9, 10]. The construction of ESG is a non-linear function of communication and firm characteristics, and there is evidence that both macro signals and micro fundamentals determine the sustainability paths of Chinese enterprises [11, 24]. Lagged and rolling features are suitable for time-dependent designs, and they can approximate fixed-effect adjustments when panel estimators are unavailable [33, 34]. The observed learnability of the continuous outcome in our regression tests suggests that the synthetic data contains recoverable structure, and this is consistent with the findings that policy-firm linkages can be identified when simulation respects realistic covariance and nonlinearity [13, 35]. RO1 is generally supported; the simulated environment looks plausible on the outside and is consistent with recent simulation-based studies in finance and ESG.

5.2 Model Performance in Classification and Regression

According to the regression results, the errors of LightGBM, Random Forest and Elastic Net are relatively close, and Random Forest has a slight advantage; this indicates that ensemble trees can capture interactions among governance and macro signals without a large number of specifications [14, 15]. Convergence of the above algorithms suggests a relatively stable signal-to-noise ratio in the continuous outcome, and this is consistent with other studies that achieve similar RMSE/MAE values when both firm covariates and policy features are used to drive ESG change [12, 13]. All classifiers were forced to output the negative class, so they achieved perfect specificity and zero sensitivity, and the ROC-AUC was near chance. This pattern is in line with the rare-event problem in ESG event detection and policy-shock classification when the positive prevalence is low and the threshold has not been calibrated [18, 19]. Based on the literature, class weights can be used, as well as cost-sensitive loss and threshold tuning, to increase recall at the cost of precision [16, 17]. Therefore, our RO2 benchmark agrees with the previous results: regression tasks are feasible in a realistic simulation, but upgrade-event detection needs imbalance remedies and calibrated operating points.

5.3 Interpretability, Feature Effects, and Robustness

Although classification did not perform well at the 0.50 threshold, the whole system can still show how communication features are connected to ESG. Previous research in explainability has shown that tree-based models can present attributions and partial dependence plots for monotonic and interaction effects of policy tone and uncertainty on sustainability outcomes [16, 17]. Based on the regression analysis, the aforementioned effects are also manifest at the level of continuous change; that is to say, even if binary upgrades occur infrequently at the selected threshold, macro-signals are often capable of altering the distribution of firm outcomes without producing frequent discrete "events" [9, 10]. Robustness through multiple Monte Carlo simulations can also be used to report the mean $\pm$ standard deviation across these simulations and characterize the variability of the estimator under different shock realisations [18, 34]. Finally, the directionality role of tone and uncertainty that we have coded corresponds to the signalling perspectives under which central bank communication can adjust expectations of corporate sustainable investment, especially when governance capacity is high [12, 24]. Therefore, RO3 is only partially supported; although the interpretive system is reasonable and shows the structure of continuous ESG responses, future work should improve event-level detectability by re-balancing and recalibration [16, 19].

To determine whether central bank communication (tone, uncertainty and surprise) is predictive of changes in firm-level ESG, this paper conducts empirical analysis. Based on the regression results, even with real-world noise, the continuous changes in ESG can be learned reliably with a small error using the engineered representation of communication and firm covariates. Tree ensembles are slightly better than Elastic Net; therefore, the interaction between communication signals and firm fundamentals (e.g., tone × governance, surprise × leverage) appears to be economically meaningful in the simulated structural process, and models can extract such a structure. On the other hand, the classification results have not yet shown the operational predictability of "upgrade events". This does not mean that the communication signals are irrelevant; rather, it indicates that the current label prevalence and threshold (in combination with the default decision threshold) are too conservative for event detection. The two non-exclusive explanations are: (i) class prevalence: if upgrades occur in about 9-10% of cases (consistent with a baseline accuracy of 0.905), predicting no upgrades will appear to perform well but will not actually find any upgrades; (ii) temporal aggregation: with only two future quarters in the test set, the number of positive instances is small, and thus ROC estimates and confusion matrices are sensitive to threshold selection. Both of the above are typical problems in rare-event prediction and will be addressed in the next experiment.

6 Conclusion and Recommendation

6.1 Conclusion

To explore how the information value of central bank communications affects the ESG performance of A-share listed companies, a simulation-based machine learning algorithm framework has been constructed in this paper for an in-depth investigation. The synthetic panels are based on regime-like soundness, log-normal uncertainty and scaled intermittent surprises, and heterogeneous firm fundamentals models, and have created an externally plausible environment consistent with recent simulations at the macro-to-firm level. Random Forest, LightGBM, and Elastic Net all showed the same improvements in ESG; therefore, it can be inferred that the relevant information jointly encoded in communications and firm attributes can be extracted by our regression models. On the other hand, the firm-specific upper-quantile "upgrade" indicator and the 0.50 score cutoff time were among the worst in terms of the indicators available, and they showed perfect specificity and zero sensitivity, thus exhibiting the same problems of rare-event detection that have afflicted ESG applications in recent years. Based on the above research, the information released by the central bank on changes in commitments is indeed related to the spread of ESG changes; however, for discrete upgrade detection, imbalance-aware training and calibrated decision thresholds need to be employed. Generally speaking, the combined approach has shown that some types of communication are related to the level of sustainability of a company, thus providing an empirical basis for reducing detection errors in the detection system.

6.2 Implications

6.2.1 Theoretical Implications

The results are consistent with the signalling-consistent mechanism: Changes in the communication tone and shifts in uncertainty expectations are transmitted as firm-level ESG adjustments, provided that governance capacity allows for this translation. By showing the learnability of continuous ESG responses in a controlled simulation, this study supports theory-informed, data-generating processes as useful supplements to observational studies. Additionally, the lack of continuous predictability and the difficulty of rare-event detection require that model outcomes and the frequency of these events be defined in tests of theoretical predictions on policy-induced sustainability transitions.

6.2.2 Practical Implications

Based on the above results, the government knows that a certain communication strategy, such as using the same tone and reducing uncertainty, can still influence the dispersion of corporate social responsibility (CSR) performance without frequent announcements of new policies. The level of governance quality and communication signals in a company provide an indication of how well the company can implement macro-guidance and achieve ESG performance. The analyst will obtain a reproducible Python pipeline in this paper for separating regression-level predictability from event detection, and offer countermeasures to class imbalance and threshold calibration in regular sustainability analysis.

6.3 Recommendation

Three recommendations follow. First, if the policy question aims to identify improvement events, use cost-sensitive learning or class weighting and adjust thresholds in the validation window to maximise meaningful recall while keeping precision at an acceptable level. Second, consider label redesign; for example, defining sustained improvement as moving averages or lowering the firm-specific quantile to increase the positive prevalence of events and align with economically enduring ESG benefits. Third, employ multi-task evaluation to track the complete distribution response via regression and, at the same time, use a calibrated classifier for operational alerts. The two will be used to observe how central bank communications affect ESG continuously and periodically.

6.4 Limitation and Future Work

Two are qualified as interpretations. First, the study is based on simulated communications and ESG processes; although parameterised to resemble empirical stylised facts, real-world text-derived features and institutional constraints may introduce additional complexities not included here. Second, the event label is deliberately sparse; the results show that detectability is affected by both prevalence and the choice of threshold. Future research will (i) integrate actual central bank text embeddings and observed ESG series to validate external generalisability; (ii) test alternative event definitions and cost-sensitive objectives; and (iii) increase robustness by extending the time horizon and analysing sector-specific heterogeneity. Together, these steps will convert the simulation results into real-world reference points for policy and enterprise practice.

References

[1] Del Vitto, A., Marazzina, D. and Stocco, D. (2023). Explainability of ESG Ratings Using Machine Learning. Annals of Operations Research, 1-30.

[2] Liu, H., Wang, J., & Liu, M. (2024). Can digital finance curb corporate ESG decoupling? Evidence from Shanghai and Shenzhen A-shares listed companies. Humanities and Social Sciences Communications, 11(1), 1-15.

[3] Taskin, D., Sariyer, G., Acar, E., & Cagli, E. C. (2025). Do past ESG scores efficiently predict future ESG performance?. Research in International Business and Finance, 74, 102706.

[4] Zeng, F., Wang, J., & Zeng, C. (2025). An optimized machine learning framework for predicting and interpreting corporate ESG greenwashing behavior. PloS one, 20(3), e0316287.

[5] Cui, J. (2025). Empirical Analysis of Digital Innovations Impact on Corporate ESG Performance: The Mediating Role of GAI Technology. arXiv preprint arXiv:2504.01041.

[6] Casiraghi, M., & Perez, L. P. (2022). Central bank communications. IMF Technical Assistance Handbook.

[7] Chen, Y., & Zhang, M. (2025). How Does the US Affect the Lending Decision Preferences of the International Monetary Fund?. China & World Economy, 33(4), 35-57.

[8] Grassi, S. (2025). Sentiment Classification of Thai Central Bank Press Releases Using Supervised Learning. arXiv preprint arXiv:2503.22629.

[9] Gong, X. L., Liu, X. H., Xiong, X., & Zhang, W. (2020). Research on China's financial systemic risk contagion under jump and heavy-tailed risk. International Review of Financial Analysis, 72, 101584.

[10] Aruoba, S. B., & Drechsel, T. (2024). Identifying monetary policy shocks: A natural language approach (No. w32417). National Bureau of Economic Research.

[11] Jaruwatanachai, P., Sukamongkol, Y., & Samanchuen, T. (2023). Predicting and managing EV charging demand on electrical grids: A simulation-based approach. Energies, 16(8), 3562.

[12] David, L. K., Wang, J., Angel, V., & Luo, M. (2024). China's ESG scorecard: A predictive machine learning model. Corporate Social Responsibility and Environmental Management, 31(4), 3468-3486.

[13] Dipierro, A. R., Barrionuevo, F. J., & Toma, P. (2025). Predicting ESG Controversies in Banks Using Machine Learning Techniques. Corporate Social Responsibility and Environmental Management, 32(3), 3525-3544.

[14] Alsayyad, M., & Fadel, S. M. (2025). Predicting ESG Scores Using Firms' Financial Indicators: A Machine Learning Regression Approach.

[15] McDermott, M., Zhang, H., Hansen, L., Angelotti, G., & Gallifant, J. (2024). A closer look at auroc and auprc under class imbalance. Advances in Neural Information Processing Systems, 37, 44102-44163.

[16] Kim, H., & Lee, M. (2025). Unraveling the Drivers of ESG Performance in Chinese Firms: An Explainable Machine-Learning Approach. Systems, 13(7), 578.

[17] Amaya, A. (2025). THREE EMPIRICAL STUDIES ON ENVIRONMENTAL HAZARDS, PUBLIC HEALTH, AND POLICY RESPONSES (Doctoral dissertation, UNIVERSITY OF HAWAI ‘I AT MĀNOA).

[18] Emani, R. (2025). Reconceptualizing Network Dynamicsa Strategic Framework Linking Social Embeddedness and Network Effects. Available at SSRN 5364351.

[19] Osterrieder, J., Arakelian, V., Coita, I. F., Hadji-Misheva, B., Kabasinskas, A., Machado, M., & Mare, C. (2023). An Overview-stress test designs for the evaluation of AI and ML Models under shifting financial conditions to improve the robustness of models. Available at SSRN 4634266.

[20] Barile, B., Forti, M., Marrocco, A., & Castaldo, A. (2024). Causal impact evaluation of occupational safety policies on firms’ default using machine learning uplift modelling. Scientific Reports, 14(1), 10380.

[21] Ehrmann, M., & Fratzscher, M. (2007). The timing of central bank communication. European Journal of Political Economy, 23(1), 124-145.

[22] Arhinful, R., Mensah, L., Amin, H. I. M., Obeng, H. A., & Gyamfi, B. A. (2025). The Strategic Role of Sustainable Finance in Corporate Reputation: A Signaling Theory Perspective. Sustainability, 17(11), 5002.

[23] Barney, J. B. (2000). Firm resources and sustained competitive advantage. In Economics meets sociology in strategic management (pp. 203-227). Emerald Group Publishing Limited.

[24] He, G., & Xiahou, Q. (2025). The Demand, Supply, and Market Responses of Corporate ESG Actions: Evidence from a Nationwide Experiment in China. Supply, and Market Responses of Corporate ESG Actions: Evidence from a Nationwide Experiment in China*(June 12, 2025).

[25] Gherghina, S. C., and Simionescu, L. N. (2025). Applied Quantitative Methods to Corporate Finance and Investment. Frontiers in Applied Mathematics and Statistics, 11, 1573606.

[26] Ernst, D. (2022). Simulation-based business valuation: methodical implementation in the valuation practice. Journal of Risk and Financial Management, 15(5), 200.

[27] Rahaman, S. U., & Abdul, M. J. (2025). Quantifying uncertainty in economics policy predictions: A Bayesian & Monte Carlo based data-driven approach. International Review of Financial Analysis, 102, 104157.

[28] Franke, S., Franke, F., & Riedel, R. (2021). Robustness evaluation of production plans using Monte Carlo simulation. Procedia Manufacturing, 54, 130-135.

[29] Lu, Y., Shen, M., Wang, H., Wang, X., van Rechem, C., Fu, T., & Wei, W. (2023). Machine learning for synthetic data generation: a review. arXiv preprint arXiv:2302.04062.

[30] Adhikari, B. (2022). A guide to using the synthetic control method to quantify the effects of shocks, policies, and shocking policies. The American Economist, 67(1), 46-63.

[31] Assefa, S. A., Dervovic, D., Mahfouz, M., Tillman, R. E., Reddy, P., & Veloso, M. (2020, October). Generating synthetic data in finance: opportunities, challenges and pitfalls. In Proceedings of the First ACM International Conference on AI in Finance (pp. 1-8).

[32] Ding, Z. (2022). Firm heterogeneity, variable markups, and multinational production: A review from trade policy perspective. Journal of Economic Surveys, 36(5), 1311-1357.

[33] Beram, R., & El-Kotory, A. (2024). Simulation-Based Assessment of Classification Methods: Statistical Models vs. Machine Learning Algorithms. The Egyptian Statistical Journal, 68(1), 91-124.

[34] Shi, X., & Zhang, X. (2023). Reevaluating the Productivity Effects of Extreme High Temperatures from the Perspective of Firm Dynamics: The Case of Chinese Manufacturing. Available at SSRN 4509302.

[35] Hezam, Y., Luong, H., & Anthonysamy, L. (2025). Machine learning in predicting firm performance: a systematic review. China accounting and finance review.