# Green Finance Policy Impact Analysis Based on Integrated Knowledge Graph

**Lin Lin, Xingcai Yao, Jingli Liang**
School of Economics, Guangzhou City University of Technology, Guangzhou 510800, Guangdong, China
Corresponding author: XingcaiYao (e-mail: lin88104@163.com)

**ABSTRACT** Due to the inherent fragmentation of policy, market, ESG and carbon data, tracking genuine sustainable impact remains challenging. This paper addresses prevailing issues in green finance policy impact assessments—specifically, unaccounted data heterogeneity, omitted-variable bias, and fragmented multi-source information—by applying an analytical approach based on an integrated Knowledge Graph (KG). The method utilizes a multidimensional KG constructed from policy, market, ESG (Environmental, Social, and Governance), and carbon data to mitigate fragmentation in green finance assessment. Natural Language Processing (NLP) is employed to structure policy knowledge, and a graph embedding algorithm is designed to quantify dynamic policy impacts on the scale of green credit, corporate emission reduction, and regional carbon intensity. Validation via a Difference-in-Differences (DID) model using data from 2016 to 2023 in China's green finance pilot zones demonstrates a significant positive policy effect: the green credit share increases by an average of 2.15 percentage points, and reductions in carbon intensity are confirmed. The KG method achieves 89.7% accuracy, representing a 22.5-percentage-point improvement over traditional regression methods. The study concludes that integrating KG technology substantially enhances visualization and causal inference capabilities in green finance policy impact analysis, enabling policymakers to trace cause-and-effect relationships—from the implementation of a new green credit scheme to measurable sustainability improvements—thus providing robust decision support for optimizing the policy tool mix.

**INDEX TERMS** green finance, knowledge graph, policy impact, difference-in-differences, low-carbon manufacturing

## I. INTRODUCTION

THE global climate governance process is accelerating, driving the advancement of "dual carbon" goals and exerting immediate, intense pressure on high-impact sectors. For the industry—a major contributor to global emissions—this acceleration necessitates a rapid and comprehensive transition toward low-carbon and circular practices [1,2]. Evaluating the effectiveness of green finance policies has thus become a central topic in academic research [3,4]. Policymakers and regulators urgently need to assess policy effectiveness and optimize policy tools. Given the imperative for rapid sectoral transition—especially in industries, where green policies must influence everything from sourcing to factory wastewater treatment—precise and dynamic measurement of policy impacts is required, spanning macroeconomic effects to micro-level transmission mechanisms.

Existing research investigates green finance at macro, meso, and micro levels. At the macro level, numerous studies have explored the overall effects of green finance on sustainable development [5], energy transformation [6], and carbon emission reduction [7]. For instance, some scholars have highlighted the role of green bonds in promoting energy efficiency investments and economic growth [8], while others have analyzed the mechanisms by which green finance, financial technology, and inclusive finance affect energy efficiency [9]. At the meso level, research focuses on the impact of policies on industrial transformation and regional development [10]. At the micro level, scholars employ quasi-natural experimental methods to evaluate the impact of green finance policies on enterprise development [11] and green credit behavior in banks [12]. Additional studies address the economic consequences of green finance information disclosure [13] and the structural impact of international green capital allocation [14]. Despite these advances, current evaluation methods have notable limitations. First, research often relies on traditional econometric models (such as panel regression [15] and Difference-in-Differences (DID) [16]). However, systematic application of Knowledge Graph (KG) to public policy analysis—especially within green finance—is still nascent. While some studies have begun to explore KG for financial risk control, little research has focused on structured modeling of policy text and the integration of policy clauses, implementation entities, and market responses into a unified graph framework for dynamic impact path analysis.

Traditional dimensionality-reduction tools, such as principal component analysis (PCA) or factor analysis, can only summarize covariation structures in tabular data and cannot encode multi-step causal chains or heterogeneous policy–market interactions. In contrast, KG integrates structured and unstructured data, models the relational topology among policies, institutions, and projects, and enables path-based causal interpretation—essential for tracing complex green finance transmission mechanisms.

## II. BUILDING A GREEN FINANCE POLICY KG

### A. MULTI-SOURCE HETEROGENEOUS DATA COLLECTION AND PREPROCESSING

Data collection comprehensively covers four dimensions: policy text, market entity behavior, environmental performance, and macroeconomic context. Specific sources include full policy texts from official release platforms such as the China Government Network and the People's Bank of China; ESG (Environmental, Social, and Governance) ratings and annual financial reports of listed companies from the Wind and CSMAR databases; daily trading data from local carbon emissions exchanges in Shanghai and Hubei; and the National Bureau of Statistics' Regional Industrial Structure and Energy Consumption Statistical Yearbook. All data are uniformly collected from 2016 to 2023, encompassing the critical period from the establishment to the deepening development of China's green finance reform pilot zones, thereby ensuring the integrity and comparability of time series analysis. In this paper, "Integrated KG" denotes a KG constructed from heterogeneous sources (policy, market, ESG, carbon), whereas "KG" refers to the general representation after integration. For clarity, "Integrated KG" is used only when emphasizing multi-source fusion; otherwise, "KG" is used. For unstructured policy text data, Python-based format parsing and text extraction are utilized. Regular expressions and custom rule dictionaries are applied to clean the text, removing headers, footers, irrelevant symbols, and formatting tags to ensure the purity of input for the subsequent Natural Language Processing (NLP) module. For structured corporate ESG and financial data, original tables are obtained via database interfaces and then processed using the Pandas library for data fusion and alignment. As a high-frequency time series, carbon trading data preprocessing focuses on resolving inconsistent data granularity. By calculating average transaction price and cumulative trading volume on a weekly or monthly basis, data is aggregated to match the analysis frequency of policy and corporate data, with missing data on non-trading days forwardfilled.

Effective entity connection requires identification and normalization. A specialized entity alias dictionary is constructed, and a fuzzy matching algorithm based on edit distance and Jaccard similarity is employed to align and merge entities from different sources that refer to the same real-world object. Regional economic and industrial data codes are standardized according to the "National Economic Industry Classification" standard published by the National Bureau of Statistics, ensuring consistency for "region" and "industry" entities. All cleaned, converted, and aligned data are stored in a structured MySQL database, with each entry assigned a unique identifier and timestamp.

To avoid treating the Integrated KG as an implicit black box, the schema is explicitly defined. The KG comprises six entity types—Policy Document, Policy Instrument, Financial Institution, Enterprise, Green Project, and Region—each with standardized attributes detailed in Section 2.2. Relationship schema includes Contains Instrument, Incentive/Constraint Object, Provide Credit, Implement Emission Reduction, Located in Region, and Synergy Effect. All relationships are extracted using a hybrid pipeline: (1) Bidirectional Encoder Representations from Transformers (BERT)–Bidirectional Long Short-Term Memory (BiLSTM)–Conditional Random Field (CRF) (BERT–BiLSTM–CRF) sequence labeling for entity recognition; (2) dependency-based rule templates and a fine-tuned BERT classifier for relation extraction; and (3) manual validation for ambiguous cases. This explicit schema ensures full transparency in KG construction and enables reproducible embedding learning.

### B. KG MODEL LAYER DESIGN

The model layer design strictly follows domain-driven design principles. Through theoretical analysis of green finance policy transmission mechanisms and data exploration, six core entity types and their key attributes are distilled, and relationship types between them are systematically defined.

Core entities are identified according to their functional roles in the policy transmission chain. The policy document entity serves as the starting point for analysis, with attributes including document number, title, issuing authority, issuance date, and policy effectiveness level, making policy text traceable and classifiable. The policy instrument entity is abstracted from the policy document and includes instrument name, instrument type (categorized as incentive, constraint, or guidance), and target area. This independent definition enables analysis of the combined effects of different instruments. The financial institution entity includes "institution name, type, green credit balance, and green credit share. The enterprise entity focuses on enterprise name, industry classification, carbon emission intensity, ESG score, and green project investment amount. High-carbon enterprises are identified if their industry's carbon emission intensity exceeds the national average. The green project entity is the ultimate vehicle for capital flows and environmental benefits, with attributes including project name, sector, total investment amount, annual emission reduction, and project status. Regional entities are used for spatial analysis at the macro level, with attributes including administrative division code, name, annual carbon intensity, and proportion of green industry added value. Relationship definitions aim to accurately represent potential transmission paths of policy impacts. Design encompasses static affiliation and spatial relationships, as well as dynamic behavioral and impact relationships. Policy documents and instruments are connected via the Includes Instrument relationship, representing specific measures deployed by macro policy. Instruments and market entities are linked through Constraint Object or Incentive Object relationships. The funding link between financial institutions and enterprises is represented by the Provide Credit relationship, which allows quantification of financial support intensity. The connection between enterprises and environmental performance is captured via Implement Emission Reduction, linked to their specific green projects. All market entities are connected to regional entities via Located in Region, aggregating micro-level behaviors to the macro-regional level. To analyze policy synergy, a Synergy Effect relationship connects policy instruments with complementary or reinforcing objectives or implementation.

### C. KNOWLEDGE EXTRACTION BASED ON NLP

Following model layer design and data preprocessing, a systematic NLP knowledge extraction process transforms unstructured policy text into graph-structured knowledge. The core task is to automatically identify and extract entities and relationships from policy text that meet model definitions, populating KG instance data. This process addresses challenges of semantic understanding and structured representation, employing sequence labeling and relationship extraction.

A deep learning-based sequence labeling method is used for entity identification in policy texts. To ensure domain adaptability, a corpus of 500 green finance-related policy documents is manually annotated using the BIO (Begin, Inside, Outside) framework for character-level labels. Entity types include policy instruments, target agencies, policy objectives, and constraints. Using this annotated corpus, a pretrained language model, BERT, is fine-tuned. Its contextual semantic representation capabilities address issues such as reference, ellipsis, and long-range dependencies in policy language. Pre-processed policy sentences are fed into the finetuned BERT model to obtain contextual embeddings, followed by a BiLSTM to capture sequence features, and CRF layer to output the optimal label sequence, accurately identifying entities in the text.

Inter-entity relationship extraction is modeled as a semantic relationship classification problem based on entity pairs. After entity recognition, syntactic analysis and rule templates preliminarily identify entity pairs within the same or adjacent sentences. A joint extraction strategy classifies relationships between entity pairs. Contextual information—including entity descriptions, text fragments, and sentence semantics—is combined to form an input sequence for another fine-tuned BERT classification model. A fully connected layer and softmax function classify relationships into predefined categories. This allows the model to learn that specific verb phrases, such as to make a request for something, typically indicate constrained object relationships.

Extraction performance is evaluated using a manually annotated test set of 5,200 sentences. The entity extraction module achieves a precision of 0.92 and recall of 0.88, while relation extraction achieves a precision of 0.89 and recall of 0.84. These metrics ensure the policy-related components ("P" dimension of Integrated KG) meet reliability standards for downstream quantitative modeling.

After quality verification, extracted entities and relationships are linked and fused with those from structured data sources and imported in batches into a Neo4j graph database.

### D. KNOWLEDGE FUSION AND STORAGE

Post-NLP extraction, entity and relationship instances may exhibit redundancy, ambiguity, and inconsistency, preventing formation of a high-quality KG. Thus, a systematic knowledge fusion process—comprising entity linking and disambiguation, attribute alignment, and conflict resolution—is implemented to integrate fragments from multiple sources into a unified, consistent, and complete knowledge base. Entity linking determines whether descriptions extracted from different sources refer to the same real-world object, enabling their merger into a unique KG entity. A standard name dictionary for financial institutions, including official names, abbreviations, and aliases, is used as a benchmark, with a hybrid matching strategy based on rules and embedding similarity. Rule-based methods rely on exact string matching, containment relationships, and a custom synonym list. For cases where rules are insufficient, cosine similarity between character-level ngram feature vectors and bag-of-words model vectors is calculated; entities are merged if combined similarity exceeds a preset threshold (0.85, based on manual evaluation). Attribute alignment and conflict resolution address attribute values from different sources, using authority priority rules to ensure each attribute retains the most authoritative and recent value.

To describe the scale and topological characteristics of the constructed KG, Table 1 summarizes the numbers of core entity and relationship instances.

**TABLE 1.** Statistics of Core Entity and Relationship Instances in the Constructed KG

| Entity Type | Quantity | Relationship Type | Quantity |
|---|---|---|---|
| Policy Documents | 1,248 | Contains Tools | 2,915 |
| Policy Tools | 867 | Constraint Objects | 4,203 |
| Financial Institutions | 425 | Incentive Objects | 1,876 |
| Enterprises | 8,762 | Provide Credit | 28,741 |
| Green Projects | 5,639 | Implement Emission Reductions | 6,501 |
| Regions | 31 | Located in Regions | 15,024 |
| Total | 16,972 | Synergy | 342 |
| | | Total | 59,602 |

Table 1 shows that a total of 16,972 entities and 59,602 relationships have been integrated, successfully transforming previously fragmented multi-source data into a coherent whole. This comprehensive network provides a robust foundation for quantitative analysis of policy variables' influence on green credit, corporate emission reductions, and regional carbon intensity through complex pathways.

Figure 1 illustrates the complete construction process and technology stack from raw data to the final queryable KG. The rich connections depicted provide a solid foundation for subsequent in-depth analysis of policy impact mechanisms.

## III. ANALYSIS OF POLICY IMPACT TRANSMISSION PATH BASED ON GRAPH EMBEDDING

### A. GRAPHICAL REPRESENTATION OF POLICY IMPACT TRANSMISSION

To quantitatively analyze the dynamic impact of green finance policies, the policy impact transmission process is formally defined as a graph computation problem on a KG. Specifically, policy impact is modeled as the set of all valid paths from a source node representing policy, traversing predefined relationship types, through the graph network, and ultimately reaching a target node representing market or environmental performance.

A depth-first search algorithm based on path pattern constraints is used to systematically discover and enumerate all potential impact paths. The algorithm's core input is a set of predefined "meta-path" templates, which specify allowable sequences of node and relationship types. These meta-paths are grounded in green finance theory and graph pattern analysis. Starting from a specific policy node, the algorithm executes Cypher queries constrained by these meta-paths in the Neo4j graph database. Traversal depth is typically set to three to five hops to balance computational complexity and transmission logic integrity.

### B. GRAPH EMBEDDING AND RELATION REPRESENTATION LEARNING

Upon defining the graphical representation of policy impact transmission paths, the next step is to quantify the strength of path associations by converting symbolic representations (entities and relationships) in the KG into low-dimensional, dense numerical vectors (embeddings). The objective is to capture the rich semantic and structural information within the graph, such that entities with similar semantics or close associations have similar vector representations. Two mainstream embedding algorithms—TransE (Translational Embedding) [17] and ComplEx (Complex Embeddings) [18]—are compared and adopted to address the complexity of different relationship types in the graph.

The TransE model treats relations as translation operations in the entity vector space. For any triple (head entity h, relation r, tail entity t), the TransE model aims for the learned vector representation to satisfy:

$$h + r \approx t \tag{1}$$

During training, a low-dimensional vector is randomly initialized for each entity and relationship (dimension set to 200 via grid search). These vectors are optimized by maximizing the score for correct triplets and minimizing the score for incorrect triplets (negative samples created by replacing h or t). The triplet scoring function is:

$$f_r(h, t) = -\|h + r - t\|_{L_1/L_2} \tag{2}$$

L1/L2 refer to the L1 or L2 norm, respectively.

Adam is used as the optimizer, with a learning rate of 0.001. Training employs a margin loss function and negative sampling to improve efficiency and embedding quality. The TransE model is simple and efficient, particularly suited for one-to-one relationships and hierarchical graph structures.

However, the Green Finance KG contains numerous symmetric and one-to-many/many-to-one relationships, which TransE handles less effectively. Therefore, the ComplEx model is used in parallel, embedding in complex vector space with a scoring function:

$$f_r(h, t) = \mathrm{Re}(\langle r, h, \bar{t}\rangle) \tag{3}$$

$\bar{t}$ is the conjugate of t. ComplEx leverages the multiplication property of complex numbers to naturally model symmetric, asymmetric, and inverse relationships. Its training process is similar to TransE, but parameters are initialized in the complex domain, requiring complex gradient calculations. Embedding quality is evaluated using link prediction tasks, specifically Mean Reciprocal Rank (MRR) and Hits@N metrics on the validation set.

### C. QUANTITATIVE ALGORITHM FOR POLICY IMPACT INTENSITY

After obtaining low-dimensional vector representations, a policy impact intensity quantification algorithm based on path embedding aggregation is proposed [19]. The goal is to convert symbolic transmission paths into computable numerical indicators, accurately measuring the association strength between policy and market nodes. Figure 2 presents the vector space calculation diagram for policy impact transmission paths.

Algorithm inputs in Figure 2 include the entity and relationship embedding matrices learned via the ComplEx model, the policy node p to be evaluated, the target market node m, and the set P of all paths connecting p and m. The output is a continuous impact strength score:

$$s(p, m) \in [0, 1] \tag{4}$$

A higher score indicates greater potential impact of policy node p on market node m.

First, each path $P_k$ in P is converted to a corresponding embedding sequence. Based on ComplEx, the path is mapped into a vector sequence. Next, a function calculates the influence strength of a single path; inspired by KG reasoning models, the vector representation of $P_k$ is defined

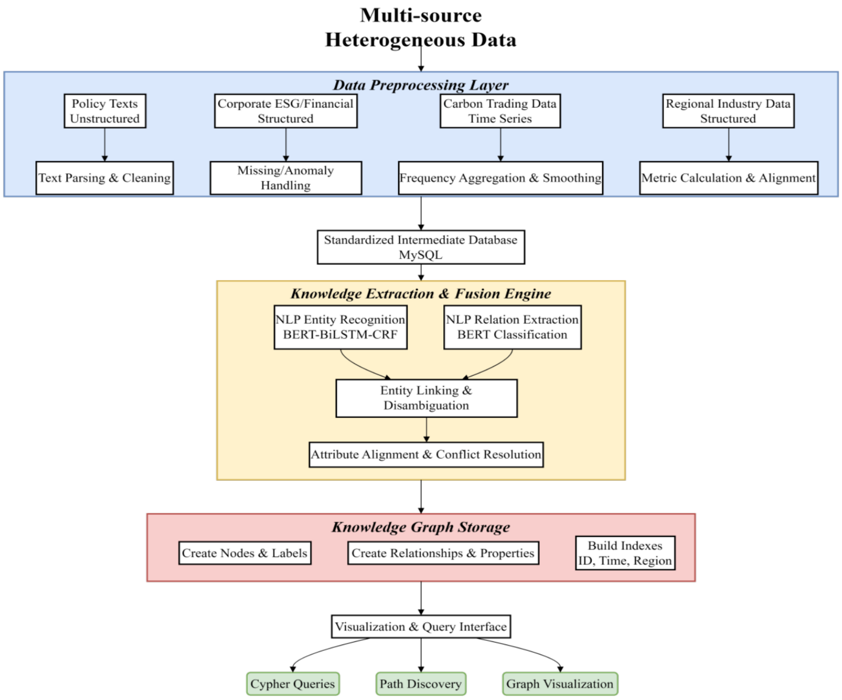


**FIGURE 1.** Architecture of the Green Finance KG Construction

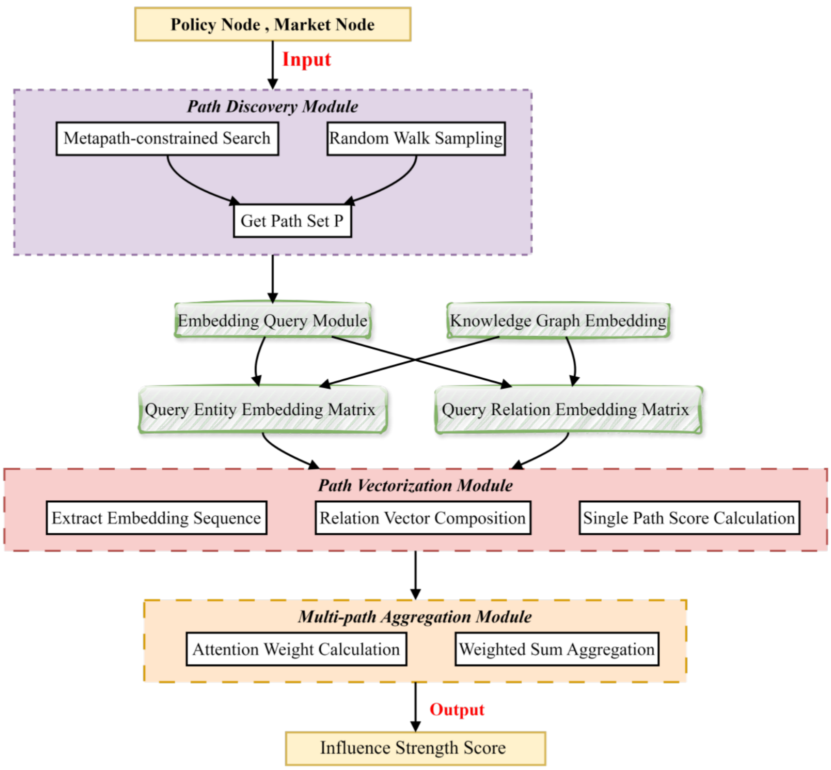


**FIGURE 2.** Vector space calculation framework for policy impact transmission paths

as a combination of all relationship vectors along the path [20]. The relationship combination operator captures path semantics, and the compatibility score between the path vector and the head and tail entity vectors is calculated:

$$s_k = \sigma\left(\mathrm{Re}(\langle v_{e_1}, v_{\mathrm{path}_k}, \bar{v}_{e_L}\rangle)\right) \tag{5}$$

Here, $\mathrm{Re}(\cdot)$ denotes the real part, and $\sigma$ denotes the sigmoid function, which normalizes the score to the range

[0, 1]. $v_{e_1}$, $v_{\text{path}_k}$, and $v_{e_L}$ represent different embedding vectors.

Finally, a multi-path aggregation mechanism is implemented. Since multiple transmission paths often exist from policy to market performance, strength scores are aggregated using an attention mechanism for weighted fusion, allowing the model to learn the importance of different paths. The attention score for each path is calculated as follows:

$$\alpha_k = \text{softmax}\left(w^T \cdot \tanh\left(W \cdot [v_{e_1} \| v_{\text{path}_k} \| v_{e_L}] + b\right)\right) \tag{6}$$

where softmax$(\cdot)$ denotes the normalized attention function, W and w are learnable parameters, and $\|$ represents vector concatenation. The final impact strength score is the weighted sum of all path scores:

$$s(p, m) = \sum_{k=1}^{|P|} \alpha_k \cdot s_k \tag{7}$$

This method accounts for path reliability and captures the relative importance of various transmission channels, yielding a comprehensive assessment of policy impact intensity.

The KG-based policy impact score is incorporated into econometric analysis as a continuous variable representing policy exposure intensity in the DID model. Specifically, the aggregated embedding score serves as a proxy for policy intensity, reflecting multi-hop policy–market linkages rather than a binary treatment indicator. This bridges KG and DID frameworks: DID identifies average treatment effects, while the embedding score provides heterogeneous treatment intensity across regions and industries.

## IV. EXPERIMENTAL VERIFICATION AND ANALYSIS

### A. EXPERIMENTAL DESIGN AND DATA DESCRIPTION

An empirical testing framework based on a quasi-natural experiment is used to compare key performance indicators between the experimental group (green finance reform pilot zones, EG) and the control group (non-pilot zones, CG) before and after policy implementation. The observation period is 2016–2023, with the first batch of pilot zones (Zhejiang, Jiangxi, Guangdong, Guizhou, Xinjiang) approved in 2017 as the treatment group. This period covers both the baseline (2016–2017) and policy effect observation (2018–2023) phases, effectively capturing dynamic policy effects. The CG is constructed for comparability, employing propensity score matching (PSM) to select provinces most similar to the treatment group in terms of economic development, industrial structure, financial deepening, and carbon emission intensity. This ensures the groups meet the parallel trend hypothesis before policy implementation. Entropy balance matching further enhances robustness in subsequent regression analysis.

Experimental data are derived from the constructed green finance KG, aggregated into panel data format. Core variables include explained, policy, and control variables. Explained variables focus on three policy impact dimensions: "green credit scale" (share of green credit in provincial loan balances), "corporate emission reduction" (carbon emissions per unit operating revenue), and "regional carbon intensity" (provincial CO2 emissions per unit gross domestic product (GDP)). The core policy variable is the interaction term DID (Treat × Post), where Treat is a group dummy (EG = 1, CG = 0) and Post is a time dummy (2017 and later = 1, before = 0). Control variables encompass macroeconomic and financial factors that may influence results. Descriptive statistics and tests of between-group differences for 2016 data (pre-policy) are presented in Table 2.

**TABLE 2.** Descriptive statistics and Between-Group Difference Tests of Main Variables (Pre-Policy, 2016)

| Variable | EG Mean | SD | CG Mean | SD | Mean Difference t-test p- value |
|---|---|---|---|---|---|
| GCR (%) | 5.82 | 1.95 | 4.13 | 1.68 | 0.024** |
| CCI (tons/10,000 CNY) | 1.57 | 0.43 | 1.62 | 0.51 | 0.341 |
| RCI (tons/10,000 CNY) | 0.89 | 0.29 | 0.94 | 0.33 | 0.287 |
| PcGDP (log) | 11.15 | 0.38 | 10.92 | 0.41 | 0.018** |
| TIP (%) | 48.65 | 7.12 | 45.83 | 8.04 | 0.102 |
| UR (%) | 61.34 | 8.56 | 56.89 | 9.12 | 0.033** |
| LDR (%) | 75.68 | 10.23 | 78.45 | 11.56 | 0.214 |

Note: *p<0.1, **p<0.05, ***p<0.01; GCR: Green Credit Ratio; CCI: Corporate Carbon Intensity; CNY: Chinese yuan; RCI: Regional Carbon Intensity; PcGDP: Per Capita Gross Domestic Product; TIP: Tertiary Industry Proportion; UR: Urbanization Rate; LDR: Loan-to-Deposit Ratio; SD: Standard Deviation.

Table 2 shows that the EG has a significantly stronger foundation for green finance development, with a mean green credit ratio of 5.82% versus 4.13% for the CG (p = 0.024). EG also exhibits higher economic development (log PcGDP: 11.15 vs. 10.92, p = 0.018) and urbanization (61.34% vs. 56.89%, p = 0.033). However, baseline environmental indicators show no significant differences, indicating comparable initial environmental performance and satisfying the parallel trend assumption required for the DID model.

### B. ASSESSMENT OF THE NET EFFECT OF POLICIES

The DID model quantifies the net effect of green finance reform and innovation pilot zone policies [21]. The baseline regression model is:

$$Y_{it} = \beta_0 + \beta_1 Treat_i \times Post_t + \gamma X_{it} + \mu_i + \lambda_t + \epsilon_{it} \tag{8}$$

$Y_{it}$ represents the outcome variable; $Treat_i$ is the treatment group dummy variable; $Post_t$ is the post-policy period dummy variable; $X_{it}$ is a vector of control variables; $\mu_i$ and $\lambda_t$ are individual fixed effects (FE) and time FE, respectively; $\epsilon_{it}$ is the random error term. The core coefficient $\beta_1$ of the model is the DID estimator, reflecting the mean difference in outcome changes between the EG and CG after policy implementation. Continuous variables are winsorized at the 1% level to mitigate outliers, and standard errors are clustered at the province level for serial correlation.

Before regression analysis, the parallel trends assumption must be verified. Figure 3 shows annual mean trends in green credit ratios for EG and CG from 2016 to 2023.

Figure 3, with green credit ratio (%) on the vertical axis, compares trends before and after policy implementation. Pre-policy (2016–2017), both groups exhibit parallel trends, satisfying the assumption. Post-policy (2017–2023), EG experiences rapid growth (5.1% to 10.3%), while CG increases modestly (4.3% to 5.9%), demonstrating the significant effect of pilot zone policies on green credit ratios.

An event study-based dynamic effect test confirms the parallel trend assumption and shows that policy effects strengthen over time.

Table 3 presents results from three baseline regressions.

**TABLE 3.** DID Regression Results for Policy Effects

| Variable | (1) GCR | (2) CCI | (3) RCI |
|---|---|---|---|
| DID (Treat × Post) | 2.15*** (0.48) | -0.23*** (0.07) | -0.11*** (0.03) |
| PcGDP (log) | 0.82* (0.42) | -0.05 (0.06) | -0.03 (0.02) |
| TIP | 0.11** (0.04) | -0.01 (0.01) | -0.005 (0.003) |
| UR | 0.09 (0.06) | -0.02* (0.01) | -0.008** (0.003) |
| LDR | -0.03 (0.02) | -0.004* (0.002) | -0.002** (0.001) |
| EIP | 0.45*** (0.12) | -0.06*** (0.02) | -0.03*** (0.01) |
| $\mu_i$ | Yes | Yes | Yes |
| $\lambda_t$ | Yes | Yes | Yes |
| Observations | 240 | 240 | 240 |
| R2 | 0.78 | 0.65 | 0.71 |

Note: Cluster-robust standard errors in parentheses; *p < 0.1, **p < 0.05, ***p < 0.01; EIP: Environmental Investment Proportion.

In Table 3, column (1) uses green credit ratio as the dependent variable, with a significantly positive DID coefficient at the 1% level ($\beta_1 = 2.15$, standard error (SE) = 0.48), indicating that pilot area policy increased green credit ratio by approximately 2.15 percentage points on average over the 2018–2023 post-policy period. Columns (2) and (3) show significantly negative DID coefficients for corporate and regional carbon intensity, confirming effective promotion of emission reductions and carbon intensity reduction.

### *C. COMPARATIVE ANALYSIS OF KG EVALUATION ACCURACY*

After assessing net policy effects, the accuracy advantage of the integrated KG method (M1) over traditional regression (M2) is examined. A policy effect prediction task is constructed: models are trained using 2016–2019 data to predict the direction of change in regional green credit ratios for 2020–2023, with predictions compared to observed values. The traditional method (M2) employs a FE logistic regression using macroeconomic indicators and regional characteristics, while the KG method incorporates a policy impact intensity score as a key predictor variable, quantifying the combined impact of policies via complex transmission pathways.

Here, “traditional regression” refers to an FE logistic regression that lacks the ability to model crosspolicy dependencies or multi-layered market interactions, resulting in omitted-variable bias. The KGenhanced model mitigates this bias by explicitly encoding policy–firm–region pathways and embedding their interactions into dense numerical representations.

Both methods use identical training and test set partitioning, and five-fold cross-validation assesses generalization. Predictive performance is evaluated via accuracy, precision, recall, and F1 score, as shown in Figure 4.

Figure 4 demonstrates the substantial advantages of the integrated KG method in assessing green finance policy impact. Accuracy increases from 67.2% to 89.7%, and precision from 65.8% to 88.3%. This improvement results from KG technology’s capacity to deconstruct complex policy transmission mechanisms. By constructing a multidimensional network with 59,602 relationships across policy clauses, corporate behaviors, financial data, and regional characteristics, and applying a graph embedding algorithm to quantify dynamic policy impacts via multi-hop paths, the method overcomes bias inherent in traditional approaches that ignore policy instrument correlations and synergies.

The accuracy metric evaluates the prediction task using DID covariates to classify whether a region’s green credit ratio rises or falls, with the KG score included as a predictive feature, enabling capture of complex policy transmission signals. Thus, the 89.7% figure reflects predictive classification accuracy, not causal-effect estimation accuracy.

To elucidate the KG method’s advantages, the Green Credit Special Guidance Policy is selected as a representative example, and graph algorithms and visualization tools implemented in Neo4j are used to generate the complete policy impact transmission network, as shown in Figure 5.

Figure 5 uses a node layout to visually illustrate the complex transmission mechanism of green finance policy impacts. The red policy tool node PT (impact value 800) radiates through three core pathways: the first influences Enterprise X via Bank A (node size 600, weight 0.8), connecting to Green Project 1 (weight 0.9) and ultimately impacting regional carbon intensity (RCI, intensity 0.9). The second pathway connects Enterprise Y via Bank B (node size 700, weight 0.9), branching into two green projects (weights 0.8, 0.7) before converging at the target node. The third pathway influences Enterprise Z via Bank C (weight 0.7). The multi-path transmission model reveals differences in policy effectiveness, with the second path as the main channel due to high-intensity correlation between banks and enterprises. Enterprise Y connects two green projects for a synergistic effect. The central position of the regional carbon indicator node (size 900) as the terminal convergence point verifies that policy effects ultimately improve regional environmental performance, illustrating how KG decomposes abstract policy–market causal relationships into quantifiable

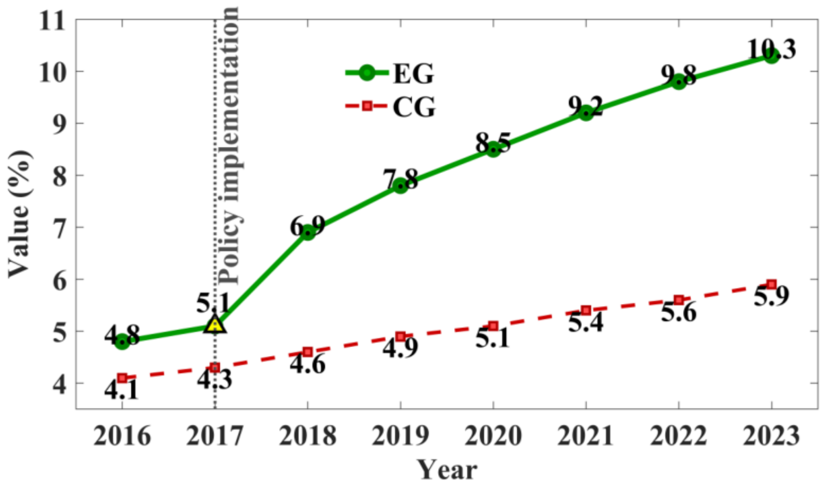


FIGURE 3. DID-Based Trend Analysis of Green Credit Ratios

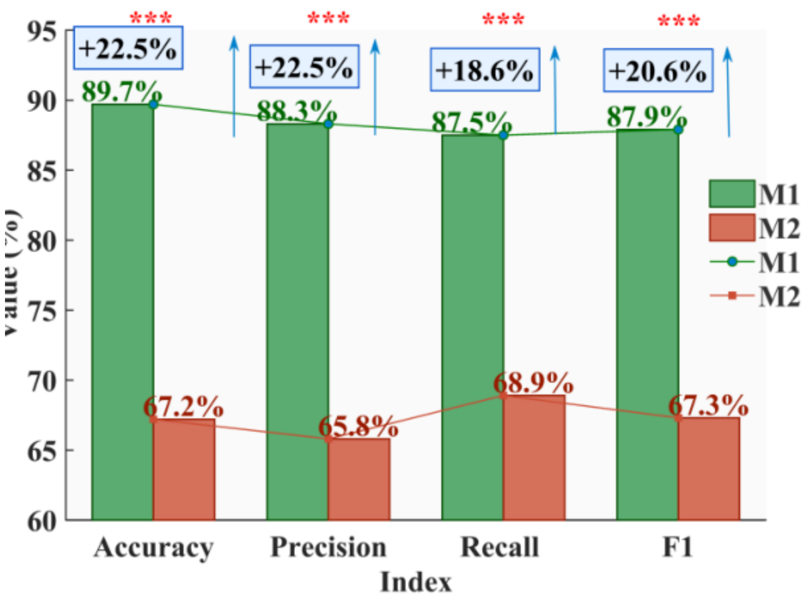


FIGURE 4. Predictive Performance Comparison of Green Finance Policy Impact Assessment Methods. *** indicates a p-value less than 0.01, meaning it is significant at the 1% statistical level. This signifies extremely high statistical significance, indicating that the improvement in prediction accuracy of the KG method is unlikely to be due to random factors

network paths.

### D. ANALYSIS OF POLICY SYNERGY EFFECT MECHANISM

Upon verifying the KG method's accuracy advantage, its association analysis capability is leveraged to explore the synergy mechanism among multiple policy tools. Policy synergy is quantified based on KG topological relationships by constructing a policy synergy network, where nodes represent different policy tools and edges represent synergistic relationships. Edge weights are computed using three dimensions: co-occurrence intensity (frequency of two policies appearing together in the same region and time window), association path overlap (degree of overlap in paths from different tools to the same market performance node), and semantic similarity (cosine similarity of policy tool embedding vectors). The final Policy Synergy Index (PSI) is defined as the logarithm of the weighted degree centrality of the node.

Panel data from 2016–2023 are used to calculate annual average PSI for each pilot region and examine its relationship with green credit growth rate. Figure 6 presents a scatter plot and fitted curve between PSI and green credit growth rate, with each point representing a pilot region-year observation.

Figure 6, based on 240 simulated PSI and green credit growth rate observations, uses color gradients to reflect PSI intensity and its impact on credit expansion. A clear positive relationship emerges: higher policy coordination correlates with faster green credit growth. As PSI rises from 0.2 to

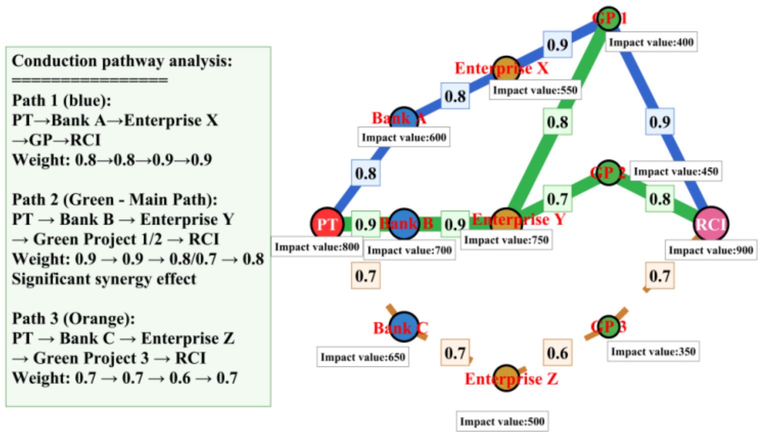


FIGURE 5. KG-Based Green Finance Policy Impact Transmission Network

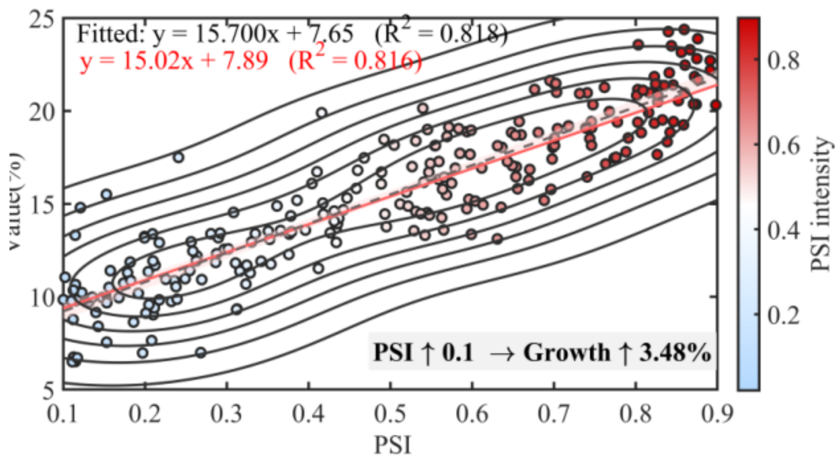


FIGURE 6. Relationship Analysis Between Policy Coordination and Green Credit Growth

0.8, green credit growth increases from 9% to 21%. Data points cluster around the regression equation y = 15.02x + 7.89; the solid red line shows theoretical fit, and the dashed black line shows sample fit, with high R2 correlation. The caption "PSI up 0.1 → Growth up 3.48%" highlights the marginal effect. The confidence interval narrows at higher PSI, indicating more accurate predictions. Density contours reveal most data are concentrated in PSI range 0.4– 0.7. These results indicate that policy tool synergy significantly amplifies green credit expansion, providing a quantitative basis for optimizing policy mixes.

A FE panel regression model is established [22] to quantify this relationship:

$$\text{Green Credit Growth}_{it} = \alpha + \beta PSI_{it} + \gamma X_{it} + \mu_i + \lambda_t + \epsilon_{it} \tag{9}$$

Regression results are presented in Table 4.

TABLE 4. Regression results of policy synergy effect mechanism

| Variable | (1) Basic Model | (2) Including Regional FE | (3) Full Model |
|---|---|---|---|
| PSI | 0.348*** (0.045) | 0.339*** (0.042) | 0.315*** (0.038) |
| PcGDP (log) | - | - | 0.892* (0.467) |
| TIP | - | - | 0.086** (0.035) |
| UR | - | - | 0.124* (0.065) |
| LDR | - | - | -0.028 (0.019) |
| EIP | - | - | 0.378*** (0.098) |
| $\mu_i$ | No | Yes | Yes |
| $\lambda_t$ | No | No | Yes |
| Constant | 8.215*** (0.823) | 8.943*** (0.785) | 6.327** (2.541) |
| Observations | 240 | 240 | 240 |
| R2 | 0.283 | 0.412 | 0.536 |
| Adjusted R2 | 0.280 | 0.395 | 0.508 |

Note: Cluster-robust standard errors in parentheses; *p < 0.1, **p < 0.05, ***p < 0.01.

Table 4 verifies the explanatory power of PSI constructed by KG for green credit growth. In the basic model (1), the coefficient is 0.348 (p < 0.01), indicating every 0.1 unit increase in PSI promotes green credit growth by 3.48

percentage points. Adding regional FE (2) maintains significance (0.339), increasing explanatory power (R2 from 0.283 to 0.412). In the full model (3) with all controls and double FE, PSI remains significant (0.315, $p < 0.01$), even after controlling for PcGDP (0.892), TIP (0.086), UR (0.124), and EIP (0.378). This confirms that policy tool coordination yields a "1 + 1 > 2" effect, providing empirical evidence for optimizing policy combinations.

### *E. MULTIDIMENSIONAL HETEROGENEITY ANALYSIS*

Following verification of policy synergy effects, KG's fine-grained analytical capabilities are used to examine heterogeneity in green finance policy impact across industries and regions. This analysis, based on policy impact intensity scores, constructs a two-dimensional industry-region matrix. Enterprise nodes are classified into eight key industries according to national standards: energy supply, heavy chemical, manufacturing, construction, transportation, agriculture/forestry/animal husbandry/fishery, services, and high-tech. Regional nodes are grouped into four sectors based on geography and economic development: eastern coastal, central, western, and northeastern regions. For each industry-region combination, average policy impact intensity is calculated, forming 64 analysis units (8 industries × 8 regions).

Policy impact intensity visually demonstrates differentiated patterns, as shown in Figure 7.

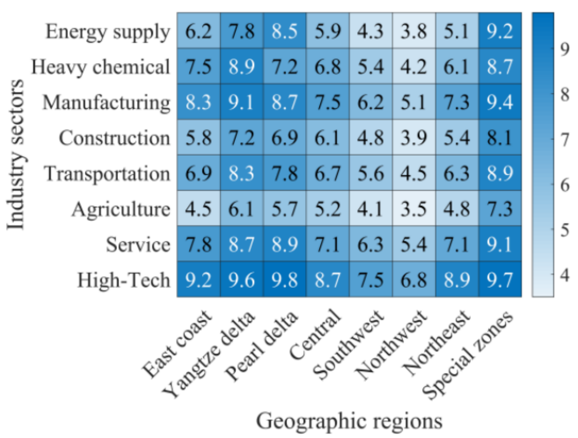


**FIGURE 7.** Distribution Analysis of Policy Impact Intensity Across Industries and Regions. Note: Tech denotes technology.

Figure 7 presents a cross-analysis of industry and region dimensions (eight geographic regions on the horizontal axis). Data show high-tech industries have an impact intensity of 9.7 in special economic zones, manufacturing excels in the Yangtze River Delta (9.1) and special economic zones (9.4), while agriculture is lowest in the northwest (3.5). Regionally, special economic zones, Yangtze River Delta, and Pearl River Delta rank highest, with the northwest lowest. This gradient reveals that policy effectiveness is shaped by industrial characteristics and regional infrastructure: technology-intensive industries (high-tech, manufacturing) are more responsive due to financing needs, while resource-based industries (agriculture, energy) lag. The eastern coastal region leverages financial infrastructure for policy amplification, while western regions are constrained by economic capacity. This heterogeneous analysis demonstrates KG's ability to identify policy targeting, providing a quantitative basis for designing differentiated policy tools.

## V. CONCLUSION

Sustainable investment plays a pivotal role in transforming high-impact sectors, such as the industry's transition to green manufacturing. This paper empirically demonstrates the significant advantages of an integrated KG approach for policy impact assessment, constructing a green finance KG with 16,972 entities and 59,602 relationships. Experiments reveal that this method improves assessment accuracy from 67.2% to 89.7%, accurately quantifying policy effects in green finance reform pilot zones—increasing green credit by 2.15 percentage points and reducing carbon intensity. PSI analysis further shows that a 0.1-unit increase in PSI leads to a 3.48% increase in green credit. These findings demonstrate that KG technology, by deconstructing complex policy impact pathways effectively addresses biases inherent in traditional methods and provides a reliable quantitative foundation for optimizing green finance policies.

## AUTHOR CONTRIBUTIONS

Conceptualization –Lin Lin, Xingcai Yao and Jingli Liang; methodology – Lin Lin, Xingcai Yao and Jingli Liang; investigation – Lin Lin, Xingcai Yao and Jingli Liang; writing-original draft preparation – Lin Lin and Xingcai Yao. All authors have read and agreed to the published version of the manuscript.

## CONFLICTS OF INTEREST

The authors declare no conflict of interest.

## FUNDING

This research received no external funding.

## ACKNOWLEDGEMENTS

Not applicable.

## REFERENCES

[1] X. C. Tan, Y. Wang, B. H. Gu, L. S. Kong, and A. Zeng, "Research on the national climate governance system toward carbon neutrality—A critical literature review," Fundamental Research, vol. 2, no. 3, pp. 384-391, 2022, doi: 10.1016/j.fmre.2022.03.010.

[2] R. Cai, H. ZHU, W. Li, Y. Xiao, and Z. Liu, "Development path of energy science and technology under "dual carbon" goals: Perspective of multi-energy system integration," Bulletin of Chinese Academy of Sciences (Chinese Version), vol. 37, no. 4, pp. 502-510, 2022, doi: 10.16418/j.issn.1000-3045.20220215001.

[3] P. K. Ozili, "Green finance research around the world: a review of literature," International Journal of Green Economics, vol. 16, no. 1, pp. 56-75, 2022, doi: 10.1504/IJGE.2022.125554.

[4] Y. Zhao, "Research on the Synergistic Path of Green Finance and High-Quality Development of the Real Economy Under the" Dual Carbon" Goal," Frontiers in Economics & Policy Modeling, vol. 1, no. 1, pp. 1-5, 2025, doi: 10.64229/3m3ws839.

[5] A. Afzal, E. Rasoulinezhad, and Z. Malik, "Green finance and sustainable development in Europe," Economic research-Ekonomska istraživanja, vol. 35, no. 1, pp. 5150-5163, 2022, doi: 10.1080/1331677X.2021.2024081.

[6] B. Zhang and Y. Wang, "The effect of green finance on energy sustainable development: a case study in China," Emerging Markets Finance and Trade, vol. 57, no. 12, pp. 3435-3454, 2021, doi: 10.1080/1540496X.2019.1695595.

[7] M. S. Meo and M. Z. Abd Karim, "The role of green finance in reducing CO2 emissions: An empirical analysis," Borsa Istanbul Review, vol. 22, no. 1, pp. 169-178, 2022, doi: 10.1016/j.bir.2021.03.002.

[8] Y. Ning, J. Cherian, M. S. Sial, S. Álvarez-Otero, U. Comite, and M. Zia-Ud-Din, "Green bond as a new determinant of sustainable green financing, energy efficiency investment, and economic growth: a global perspective," Environmental Science and Pollution Research, vol. 30, no. 22, pp. 61324-61339, 2023, doi: 10.1007/s11356-021-18454-7.

[9] H. Liu, P. Yao, S. Latif, S. Aslam, and N. Iqbal, "Impact of Green financing, FinTech, and financial inclusion on energy efficiency," Environmental science and pollution research, vol. 29, no. 13, pp. 18955-18966, 2022, doi: 10.1007/s11356-021-16949-x.

[10] B. Gu, F. Chen, and K. Zhang, "The policy effect of green finance in promoting industrial transformation and upgrading efficiency in China: analysis from the perspective of government regulation and public environmental demands," Environmental Science and Pollution Research, vol. 28, no. 34, pp. 47474-47491, 2021, doi: 10.1007/s11356-021-13944-0.

[11] Q. Cui, X. Ma, S. Zhang, and J. Liu, "Does the implementation of green finance regulation promote the highquality development of enterprises? Evidence from a quasi-natural experiment in China," Environmental Science and Pollution Research, vol. 30, no. 43, pp. 97786-97807, 2023, doi: 10.1007/s11356-023-29355-2.

[12] I. Akomea-Frimpong, D. Adeabah, D. Ofosu, and E. J. Tenakwah, "A review of studies on green finance of banks, research gaps and future directions," Journal of Sustainable Finance & Investment, vol. 12, no. 4, pp. 1241-1264, 2022, doi: 10.1080/20430795.2020.1870202.

[13] C. Liu and S. S. Wu, "Green finance, sustainability disclosure and economic implications," Fulbright Review of Economics and Policy, vol. 3, no. 1, pp. 1-24, 2023, doi: 10.1108/FREP-03-2022-0021.

[14] M. Mohsin, A. Dilanchiev, and M. Umair, "The impact of green climate fund portfolio structure on green finance: empirical evidence from EU countries," Ekonomika, vol. 102, no. 2, pp. 130-144, 2023, doi: 10.15388/Ekon.2023.102.2.7.

[15] S. Yadav I, D. Pahi, and R. Gangakhedkar, "The nexus between firm size, growth and profitability: new panel data evidence from Asia–Pacific markets," European Journal of Management and Business Economics, vol. 31, no. 1, pp. 115-140, 2022, doi: 10.1108/EJMBE-03-2021-0077.

[16] M. Wooldridge J, "Simple approaches to nonlinear difference-in-differences with panel data," The Econometrics Journal, vol. 26, no. 3, pp. C31-C66, 2023, doi: 10.2139/ssrn.4183726.

[17] D. Song, F. Zhang, M. Lu, S. Yang, and H. Huang, "DTransE: Distributed translating embedding for knowledge graph," IEEE Transactions on Parallel and Distributed Systems, vol. 32, no. 10, pp. 2509-2523, 2021, doi: 10.1109/TPDS.2021.3066442.

[18] Z. Li, H. Liu, Z. Zhang, T. Liu, and N. N. Xiong, "Learning knowledge graph embedding with heterogeneous relation attention networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 33, no. 8, pp. 3961-3973, 2021, doi: 10.1109/TNNLS.2021.3055147.

[19] G. Munda, "Qualitative reasoning or quantitative aggregation rules for impact assessment of policy options? A multiple criteria framework," Quality & Quantity, vol. 56, no. 5, pp. 3259-3277, 2022, doi: 10.1007/s11135-021-01267-8.

[20] Y. Zhao, X. Wang, J. Chen, Y. Wang, W. Tang, X. He, et al., "Time-aware path reasoning on knowledge graph for recommendation," ACM Transactions on Information Systems, vol. 41, no. 2, pp. 1-26, 2022, doi: 10.1145/3531267.

[21] Y. Ren, Z. Ding, and J. Liu, "How green finance boosts carbon efficiency in agriculture: a quasi-experiment from China," China Agricultural Economic Review, vol. 16, no. 2, pp. 267-289, 2024, doi: 10.1108/CAER-08-2023-0228.

[22] Z. Liang and E. Nasruddin, "Impact of Green Finance on High-Quality Economic Development: A Panel Data Regression," Prague Economic Papers, vol. 33, no. 5, pp. 543-564, 2024, doi: 10.18267/j.pep.876.