

Article

Fostering Urban Sustainability: A Spatio-Temporal and TextCNN-DID Evaluation of China's Green Finance Pilots

Zhen Tan [†], Yangxin Wang ^{*,†} and Yuanyuan Wang 

Bangor College, Central South University of Forestry and Technology, Changsha 410004, China; zht25jjs@bangor.ac.uk (Z.T.); h20190192@csuft.edu.cn (Y.W.)

* Correspondence: wangyangxin@csuft.edu.cn

[†] These authors contributed equally to this work.

Abstract

This paper explores how green finance policies influence the sustainability and resilience of urban areas, particularly those operating under tight fiscal budgets. By treating China's Green Finance Reform and Innovation Pilot Zone as a quasi-natural experiment, we examine a comprehensive panel of 299 prefecture-level cities between 2011 and 2024. The analytical framework integrates traditional difference-in-differences (DID) estimations with machine-learning-driven text analysis. To quantify governmental foresight regarding sustainability risks, the authors employ Word2Vec semantic expansion and NLP tokenization. Additionally, a TextCNN classifier is utilized to group cities by their digital-economy readiness, allowing for a nuanced assessment of heterogeneous policy impacts. The results demonstrate that implementing green finance pilots not only tangibly boosts sustainable productive capacities but also noticeably sharpens official textual focus on sustainability governance. Furthermore, mechanism analyses suggest that these positive outcomes are driven by structural industrial upgrades, enhanced technological expenditures, and the easing of local fiscal burdens. The study's conclusions—reinforced by rigorous robustness checks like PSM-DID, CS-DID, and spatio-temporal mapping—indicate that green finance functions as a vital governance mechanism capable of stabilizing local finances and driving long-term, low-carbon transitions.

Keywords: green finance; urban sustainability; sustainability resilience; fiscal relief; policy attention; TextCNN; DID analysis



Academic Editor: Cristina Raluca Gh. Popescu

Received: 9 July 2026

Revised: 23 July 2026

Accepted: 29 July 2026

Published: 4 August 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

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

1. Introduction

Achieving urban sustainable development demands a delicate balance between ecological preservation, financial stability, technological advancement, and fiscal health. Traditional growth models, which are heavily reliant on resource extraction and carbon-intensive industries, increasingly burden local public finances and elevate environmental transition risks. Within China's strategic push for "new quality productive forces," it is crucial to understand how cities can evolve from mere economic expansion toward robust, sustainability-oriented resilience. In this context, urban sustainable productive forces encompass a city's ability to maintain long-term digital adaptability, green transformation, fiscal soundness, and overall productivity amid shifting environmental constraints.

Green finance serves as a vital institutional mechanism to support this governance transition. By integrating environmental risk assessments into investment strategies, it successfully redirects capital away from high-emission sectors and toward sustainable

infrastructure and low-carbon innovations. While the previous literature has extensively explored the impact of green finance on firm-level green innovation and regional green total factor productivity, its broader capacity to reshape holistic urban sustainability—specifically through enhancing productive forces, alleviating fiscal stress, and shifting policy attention—remains underexplored. This gap is particularly acute in China, where local authorities must juggle stringent environmental mandates, heavy infrastructure needs, and constrained municipal budgets.

Furthermore, capturing early indicators of sustainability governance presents a significant methodological hurdle. Conventional statistical metrics typically reflect delayed, post-implementation outcomes. Conversely, annual government work reports offer forward-looking insights into local administrative priorities, risk mitigation strategies, and governance commitments before they materialize in economic data. Leveraging recent advancements in natural language processing (NLP), machine learning, and text classification enables the extraction of measurable policy-attention indicators from these unstructured official texts. This innovative approach bridges the gap between qualitative governance rhetoric and quantitative sustainability evaluations.

Driven by these challenges, this study explores three fundamental questions. First, do Green Finance Reform and Innovation Pilot Zones effectively boost tangible urban sustainable productive capacities while simultaneously amplifying textual policy focus on sustainability-risk governance? Second, what are the underlying transmission mechanisms, and specifically, does green finance alleviate local fiscal pressures in addition to stimulating technological innovation and industrial upgrades? Finally, are these policy dividends equitably distributed, or do cities with comparatively weaker digital-economy infrastructures experience distinct catch-up effects? To answer these questions, we utilize a panel dataset of 299 Chinese cities spanning 2011 to 2024 (yielding 3588 valid observations after data cleaning), applying a comprehensive empirical framework that features advanced difference-in-differences (DID) estimations, spatio-temporal mapping, and deep-learning-based heterogeneity classifications.

This research makes four distinct contributions to the existing literature on green finance and sustainability. First, by systematically controlling for overlapping governance pilots and employing a PSM-DID framework to rule out selection bias, it reconceptualizes the evaluation of green finance pilots as a broader issue of urban sustainability resilience rather than a singular environmental outcome, and we establish robust causal identification by employing PSM-DID and event-study designs, and by controlling for high-dimensional province-by-year fixed effects, municipal linear time trends, and an expanded matrix of five concurrent national pilot policies. Second, it highlights “fiscal-space relief” as a critical financial-stability pathway, demonstrating how market-driven green finance can ease local budgetary constraints to foster long-term sustainable investments. Third, it establishes a robust, reproducible text-measurement protocol utilizing NLP tokenization and Word2Vec semantic expansion to quantify latent governmental attention to sustainability. Fourth, by deploying a TextCNN classifier to categorize urban digital-economy tiers, the study unveils heterogeneous policy impacts, illustrating the inclusive potential of green finance instruments for cities with a lagging digital foundation.

2. Literature Review and Hypothesis Development

2.1. Green Finance, Digital Conditions, and Urban Sustainability Governance

Green finance has moved from a narrow environmental-financing concept to a broader institutional arrangement for pricing transition risk, correcting capital misallocation, and coordinating low-carbon investment across cities and firms. Recent evidence indicates that green finance improves green total factor productivity by strengthening green technological

progress and industrial upgrading [1], while China's Green Finance Reform and Innovation Pilot Zones have been shown to reduce energy intensity and stimulate energy-saving effects through green innovation and financing-channel adjustments [2]. At the firm level, green finance reform can reshape innovation incentives and credit allocation, particularly by changing the financing conditions of enterprises exposed to environmental constraints [3,4]. These findings suggest that green finance should be understood not only as a financing instrument but also as an integrated policy lever that links financial stability, industrial transformation, and sustainability governance.

A growing body of recent research further shows that the transmission of green finance depends on local absorptive capacity and information conditions. Green finance can encourage corporate green innovation [5], influence corporate financing structures under pilot-zone policies [6], and reduce the financing constraints that hinder green innovation investment [7]. Green credit policies can also improve production efficiency in green transformation by reallocating financial resources away from energy-intensive activities and toward cleaner technologies [8,9]. In addition, investor preferences and green capital-market signals strengthen the market discipline imposed on firms' environmental strategies [10]. These studies are directly relevant to the present paper because urban sustainability-risk resilience depends on whether financial-policy pilots can transform capital allocation into durable productive capacity rather than short-term policy compliance.

Digital conditions are another key determinant of whether green finance policies generate inclusive and persistent sustainability gains. Digital finance improves the matching efficiency between financial institutions and innovation projects by reducing information frictions [11], while firm digitization can enhance green technology innovation when it is supported by appropriate government-subsidy and governance arrangements [12]. At the urban level, the digital economy promotes high-quality green development [13], internet development raises urban green innovation efficiency [14], and manufacturing green total factor productivity can improve through the interaction between digital infrastructure, energy conservation, and innovation synergy [15,16]. Regarding new quality productive forces, Zhu et al. [17] demonstrate that green finance promotes their development through technological innovation and data factors. At the micro level, Li and Liu [18] find that low-carbon transition strategies enhance enterprises' new quality productive forces by effectively alleviating financial constraints. From a regulatory perspective, Zhang et al. [19] show that carbon emissions trading policies act as a catalyst for stimulating the development of new quality productive forces within manufacturing enterprises. Spatially, Liu and He [20] emphasize that synergistic industrial agglomeration is a critical precondition for driving new quality productive forces and high-quality development. Furthermore, Fu and Zhao [21] provide direct empirical evidence from Chinese A-share listed companies that green finance bolsters new-quality productivity. Therefore, this study treats cities as sustainability-governance units whose policy responses are conditioned by fiscal, industrial, and digital foundations.

2.2. Fiscal Stress, Local Government Incentives, and Fiscal-Space Risk Relief

Local governments are both implementers of green development policies and providers of public investment; therefore, their ability to sustain the low-carbon transition is inseparable from fiscal revenue-expenditure conditions. Vertical fiscal imbalance can weaken local fiscal discipline and change expenditure preferences [22], and recent evidence from China shows that vertical fiscal imbalance may also suppress green total factor productivity through efficiency and spatial-spillover channels [23]. When fiscal gaps expand, local officials may face a trade-off between short-term growth competition and long-term environmental resilience. Fiscal pressure can alter government behavior in green

productivity formation [24] and may weaken firms' environmental performance through regulatory relaxation or tax-competition incentives [25]. In this sense, fiscal stress is not merely a budgetary variable; it is a structural sustainability-risk factor that can constrain the implementation depth of green finance reforms [26].

Green finance may alleviate this constraint by mobilizing market-based capital for projects that otherwise rely heavily on direct fiscal expenditure. However, the impact of green finance pilots on local municipal budgets involves a complex, dual mechanism. In the short run, establishing pilot zones may actually increase local fiscal expenditures due to the necessity of providing matching funds, interest rate subsidies, public guarantee funds, or risk-compensation mechanisms to incentivize financial institutions. Over time, as green credit, green bonds, and public-private risk-sharing mechanisms mature, they effectively substitute for traditional subsidy-led environmental investment. This substitution effect is important for urban sustainability because it allows local governments to reallocate scarce fiscal resources toward basic research, digital infrastructure, public services, and risk prevention. Consequently, this study acknowledges the ambiguity inherent in these competing fiscal forces and explicitly evaluates the net effect of green finance pilots on municipal fiscal revenue-expenditure pressure. Accordingly, the fiscal-space relief mechanism provides a more complete explanation for how green finance can improve sustainable productive forces beyond the conventional channels of industrial upgrading and technological innovation.

2.3. Econometric Identification, Textual Signals, and Research Hypotheses

Recent advances in difference-in-differences methodology require careful treatment of policy evaluation when adoption is staggered across cities and treatment effects are heterogeneous. Callaway and Sant'Anna proposed group-time average treatment effects for multiple-period DID [27], while Sant'Anna and Zhao developed doubly robust DID estimators [28]. Goodman-Bacon showed that two-way fixed-effects estimates can combine comparisons with different implicit weights [29], and Sun and Abraham demonstrated that dynamic event-study coefficients may be contaminated when cohort-specific treatment effects differ [30]. Related work further clarifies the negative-weight problem in heterogeneous treatment settings [31], and recent syntheses recommend using complementary estimators and transparent robustness checks [32]. Event-study imputation methods [33], design-based DID analysis [34], and propensity score matching [35] therefore motivate the multi-estimator identification strategy used in this paper.

In parallel, text-as-data and machine learning methods have made it possible to observe forward-looking governance signals that are not fully captured by structured statistics. Text-as-data research demonstrates that large-scale textual information contains measurable economic signals [36]. Newspaper-based economic policy uncertainty indices [37], computational-linguistic analysis of central-bank communication [38], and business-news measures of macroeconomic conditions [39] all show that language can reveal policy priorities, expectations, and risk perceptions. For document classification, convolutional neural networks remain a useful architecture for extracting local semantic features from word embeddings [40]. These insights support the use of government work reports as a source of sustainability-risk attention, because such reports contain annual policy priorities, fiscal commitments, industrial plans, and risk-prevention narratives before many outcomes become visible in statistical indicators.

The present study therefore connects three strands of the literature that are often examined separately: green finance policy evaluation, fiscal-space risk relief, and text-based sustainability-risk measurement. Digital-economy studies show that digital infrastructure can improve high-quality green development and urban green innovation [41,42], while

low-carbon development and policy-communication research confirm the importance of innovation-driven digital transformation and official textual signals in shaping economic expectations [43]. At the same time, complex quasi-experimental designs require attention to feasibility, data requirements, and counterfactual construction [44], while financial-technology research increasingly links sustainability finance with financial inclusion and institutional resilience [45,46]. Building on these studies, we propose a mechanism chain from green finance pilots to capital reallocation, industrial upgrading, technological incentives, fiscal-space relief, and improved urban sustainability resilience, as summarized in Figure 1. The following hypotheses are developed:

H1a. *The Green Finance Reform and Innovation Pilot Zone policy improves the objective hard-outcome level of urban sustainable productive forces (Y1_NQP) by correcting capital misallocation toward low-carbon and digital transitions.*

H1b. *The Green Finance Reform and Innovation Pilot Zone policy strengthens local governments' forward-looking textual attention to sustainability-risk governance (Y2_NQPA) in official work reports.*

H2. *Green finance improves urban sustainability-risk resilience through the mechanism of industrial-structure upgrading, as capital reallocation restricts energy-intensive activities and fosters strategic emerging sectors.*

H3. *Green finance improves urban sustainability-risk resilience through the mechanism of expanding science and technology expenditure, thereby alleviating financing constraints for urban green technology innovation.*

H4. *Green finance generates a net reduction in local fiscal revenue-expenditure pressure—outweighing short-term matching and subsidy costs through long-term market-based financing substitution—thereby creating fiscal-space relief for long-term sustainability-risk governance.*

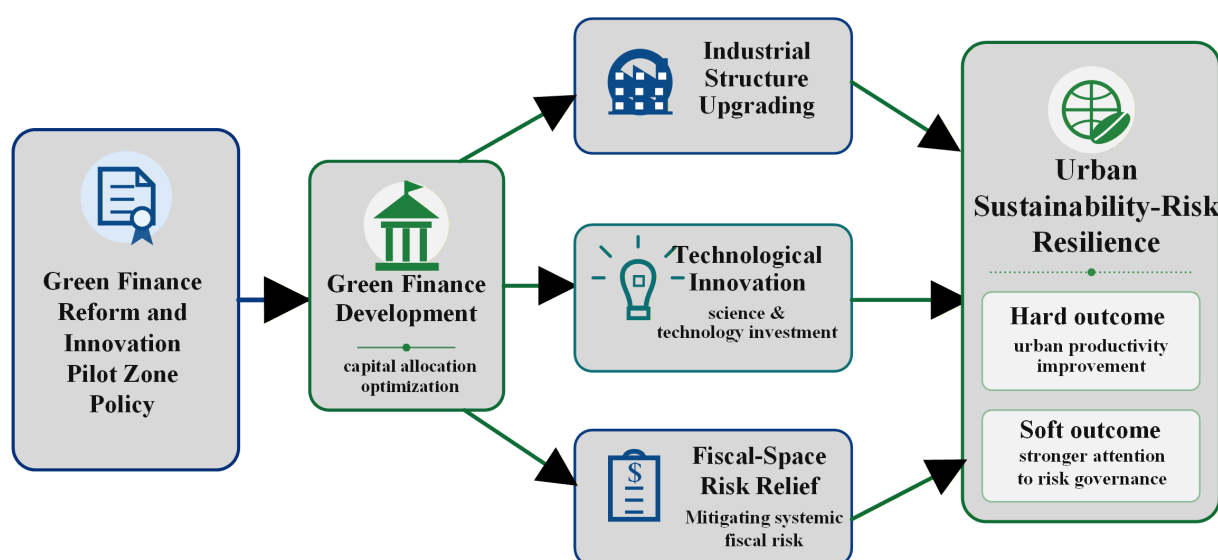


Figure 1. Integrated theoretical mechanism framework for green finance policy and urban sustainability resilience.

3. Materials and Methods

3.1. Data Construction and Sample Treatment

The empirical analysis is based on a panel dataset of Chinese prefecture-level cities covering 2011–2024. City-level macroeconomic indicators, fiscal revenue and expenditure data, financial-development variables, population information, and related controls are drawn from the China City Statistical Yearbook, prefecture-level statistical yearbooks, and the CSMAR database. Annual government work reports are collected from official municipal government websites and publicly available policy-text repositories, then organized into a city–year textual dataset. This data structure allows the study to connect financial-policy shocks, fiscal conditions, sustainability outcomes, and policy-attention signals within a unified urban risk-assessment framework.

Several sample-processing rules are applied to improve identification quality. To ensure a clean identification strategy and maintain comparability, our sample selection rule proceeds as follows. First, we exclude the Chongqing municipality (approved as a pilot in 2022) to prevent overlapping policy interference with earlier batches. This exclusion yields an initial framework of 299 prefecture-level cities over 14 years (2011–2024), corresponding to a theoretical raw sample of 4186 observations. Subsequently, to ensure the reliability of the empirical estimations, 598 city–year observations with serious missing values in core variables were strictly dropped during the data cleaning process. Consequently, the final effective panel dataset comprises exactly 3588 valid observations. All continuous variables are winsorized at the 1% and 99% levels to reduce the influence of outliers. Because green finance pilots were introduced in different years, the baseline regressions include city and year fixed effects, and the robustness analysis uses CS-DID to address heterogeneous treatment effects under staggered adoption.

To ensure our identification is not driven by concurrent environmental and digital pilot initiatives or non-random sample selection, we conduct extensive robustness and endogeneity tests—including confounding policy exclusions and a PSM-DID specification—detailed in Section 5.

3.2. Variables and Measurement Logic

Table 1 defines the variables used in the empirical models. The two dependent variables capture complementary dimensions of urban sustainability-risk resilience: *Y1_NQP* measures the objective hard-outcome side of sustainable productive capacity, whereas *Y2_NQPA* measures government attention to sustainability-risk governance extracted from work report text. The treatment indicator *X_TreatPost* equals one after a city enters a Green Finance Reform and Innovation Pilot Zone. The mechanism variables represent industrial structure upgrading (*M1_IndAdv*), science and technology expenditure intensity (*M2_Tech*), and fiscal revenue-expenditure pressure (*M3_FPress*). Control variables describe population density (*C1_LnDen*), income level (*C2_LnPgdP*), openness (*C3_Open*), financial development (*C4_Fin*), human capital (*C5_Edu*), internet penetration (*C6_Internet*), and government intervention (*C7_Gov*).

To avoid subjective weight assignment and ensure reproducible sustainability assessment, the comprehensive index of urban sustainable productive forces (*Y1_NQP*) is constructed using the Entropy Weight Method (EWM). To ensure robust theoretical grounding, our indicator selection is firmly anchored in the recent literature [17–21], which defines ‘new quality productive forces’ through three fundamental pillars: high technology, high quality, and new production factors. Accordingly, we design an expanded 9-indicator evaluation system that systematically captures these core dimensions: (1) strategic emerging industries (reflecting high-technology breakthroughs and innovation output, such as patent grants); (2) green transformation (reflecting high-quality ecological sustainability and en-

ergy efficiency); and (3) digital infrastructure (reflecting the depth of new data factors and digital finance networks). Furthermore, to establish the construct validity of the automated EWM weights, we successfully cross-validated this metric using Principal Component Analysis (PCA) as an alternative weighting scheme in our subsequent robustness checks (Section 5.3).

Table 1. Variable definitions and measurements.

Category	Symbol	Variable	Measurement/Description
Dependent variable	<i>Y1_NQP</i>	Urban sustainable productive forces	Hard outcome indicator; a composite index constructed from dimensions such as strategic emerging industries, green transformation, and digital infrastructure.
Dependent variable	<i>Y2_NQPA</i>	Government attention to sustainable productive forces	Soft policy-intention indicator; seed-word frequency extracted from government work reports using Word2Vec semantic expansion and jieba tokenization, transformed as $\ln(\text{frequency} + 1)$.
Core explanatory variable	<i>X_TreatPost</i>	Green finance pilot policy	DID policy variable; equals 1 if a city belongs to a Green Finance Reform and Innovation Pilot Zone in the current year and thereafter, and 0 otherwise.
Mechanism variable	<i>M1_IndAdv</i>	Industrial structure upgrading	Value added of tertiary industry divided by value added of secondary industry.
Mechanism variable	<i>M2_Tech</i>	Science and technology expenditure	Science and technology expenditure in the general public budget divided by total general public budget expenditure.
Mechanism variable	<i>M3_FPress</i>	Fiscal revenue-expenditure pressure	General public budget expenditure minus general public budget revenue, divided by regional GDP.
Control variable	<i>C1_LnDen</i>	Population density	Resident population divided by administrative land area, expressed in logarithms.
Control variable	<i>C2_LnPgd</i>	Economic development level	Per capita regional GDP, expressed in logarithms.
Control variable	<i>C3_Open</i>	Openness level	Total imports and exports divided by regional GDP.
Control variable	<i>C4_Fin</i>	Financial development level	Year-end financial-institution loans plus deposits divided by regional GDP.
Control variable	<i>C5_Edu</i>	Human capital level	Number of college and university students divided by year-end total population.
Control variable	<i>C6_Internet</i>	Internet penetration rate	Broadband internet users per 100 persons, supplemented by linear interpolation where necessary.
Control variable	<i>C7_Gov</i>	Government intervention	General public budget expenditure divided by regional GDP.

Note: All continuous variables are winsorized at the 1% and 99% levels; *Y2_NQPA* is a smoothed word-frequency indicator constructed from government work reports and interpreted as a textual signal of sustainability-risk governance attention.

The computational protocol of EWM is implemented through the following six steps so that the hard-outcome indicator can be interpreted as a transparent composite measure of urban sustainability resilience:

Step 1: Data Normalization. To remove scale differences across heterogeneous indicators, the raw data are standardized. For positive indicators, higher values indicate stronger sustainable productive capacity and lower transition vulnerability:

$$z_{ijt} = \frac{x_{ijt} - \min(x_j)}{\max(x_j) - \min(x_j)}, \quad (1)$$

For negative indicators, higher raw values imply greater sustainability pressure, so the standardization direction is reversed:

$$z_{ijt} = \frac{\max(x_j) - x_{ijt}}{\max(x_j) - \min(x_j)}, \quad (2)$$

where the raw value, maximum value, and minimum value are defined for each city–year observation and each sub-indicator across the full sample space.

Step 2: Proportion Calculation. The conditional proportion of each standardized observation is calculated to describe its relative information contribution within the city–year panel:

$$p_{ijt} = \frac{z_{ijt}}{\sum_{i=1}^N \sum_{t=1}^T z_{ijt}}, \quad (3)$$

where N represents the total number of prefecture-level cities ($N = 299$), and T denotes the temporal span ($T = 14$).

e_j Step 3: Information Entropy Evaluation. The entropy of each sub-indicator is derived to measure the dispersion of information contained in that indicator:

$$e_j = -\frac{1}{\ln(N \times T)} \sum_{i=1}^N \sum_{t=1}^T p_{ijt} \ln(p_{ijt}), \quad (4)$$

$p_{ijt} = 0, p_{ijt} \ln(p_{ijt}) = 0$. When the calculated proportion is zero, the conventional mathematical limit is applied to avoid undefined logarithmic values.

d_j Step 4: Information Utility Redundancy. The information utility coefficient is computed to capture the effective differentiation power of each sub-indicator:

$$d_j = 1 - e_j, \quad (5)$$

Step 5: Objective Weight Derivation. The objective weight assigned to each underlying metric is determined by normalizing the corresponding information utility coefficient:

$$w_j = \frac{d_j}{\sum_{j=1}^M d_j}, \quad (6)$$

where M represents the total number of standardized sub-indicators included in the target system.

Step 6: Composite Index Synthesis. The final multi-dimensional index of urban sustainable productive forces is synthesized linearly and used as the hard-outcome proxy for sustainability-risk resilience:

$$Y_{1_NQP_{it}} = \sum_{j=1}^M w_j z_{ijt}, \quad (7)$$

Further details of the indicator system and entropy weights are reported in Table A1.

3.3. Text-Based Measurement of Policy Attention: Word2Vec Expansion and NLP Tokenization

Government work reports provide official yearly statements of local development priorities and therefore contain forward-looking information about risk perception, resource allocation, and sustainability governance. To quantify attention devoted to sustainable productive forces and transition-risk management [37,39], this study designs a four-step text-mining procedure rather than relying on a fixed manual keyword list [36]. First, reports from 2011 to 2024 are cleaned by removing HTML tags, headers, footers, invalid symbols, and duplicate paragraphs. Second, seed dimensions are defined around technological innovation, digital and intelligent transformation, green and low-carbon development, modern industrial systems, and new production factors. A Word2Vec model trained on the report corpus is then used to calculate cosine similarity and identify semantically related expansion terms.

$w_1, w_2, \dots, w_T$. Specifically, this study adopts the Skip-gram architecture of the Word2Vec model, which is effective in capturing semantic proximity across extensive government texts. Instead of relying on rigid exact-word matching, the Skip-gram model learns continuous vector representations by maximizing the average log probability of predicting context words given a center word. For a sequence of training words, the objective function is defined as

$$J = \frac{1}{T} \sum_{t=1}^T \sum_{-c \leq j \leq c, j \neq 0} \log P(w_{t+j} | w_t), \quad (8)$$

$P(w_{t+j} | w_t) \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}$, where c is the size of the training context. The probability is formulated using the softmax function over the inner product of word vectors. Consequently, words that frequently appear in similar policy contexts, such as digitalization and intelligent manufacturing, obtain vectors with high cosine similarity. This property supports the construction of a context-aware policy-intent lexicon for sustainability-risk assessment.

In the computational implementation, the Skip-gram architecture is executed using the Gensim library in Python 3.10. To balance semantic density and structural stability, the word-embedding dimension is set to 200, the context window size is set to 5, and the minimum word-frequency threshold is set to 10 to filter low-frequency background noise and grammatical anomalies within municipal work reports. During dictionary expansion, words with a cosine similarity score greater than 0.65 relative to the predefined seed words are dynamically extracted. This specific threshold was empirically selected because, in dense Chinese official policy corpora, cosine similarities above 0.65 optimally balance semantic precision (filtering out generic macroeconomic noise) and recall (capturing diverse but relevant local expressions). This algorithmic screening is followed by multi-round manual auditing to remove contextual confounders and polysemous terms, resulting in a cohesive domain-specific policy-intent lexicon containing exactly 185 verified terms. To ensure that our empirical conclusions are not artifacts of this specific NLP cutoff, we successfully re-estimated our baseline models using alternative threshold values in our robustness checks.

Third, the expanded dictionary is incorporated into jieba as a custom tokenization dictionary, and stop words are removed to reduce noise. Fourth, the target-term frequency is aggregated by city and year and transformed into a smoothed logarithmic indicator. Figure 2 summarizes the workflow, and Table 2 lists representative seed words and expanded vocabulary.

$$Y_{2N} QPA_{it} = \ln(\text{Frequency}_{it} + 1), \quad (9)$$

In Equation (9), Frequency_{it} represents the total occurrence of terms associated with sustainable productive forces and sustainability-risk governance in the government work

report of city i in year t . The variable $Y2_NQPA$ is defined as $\ln(\text{Frequency}_{it} + 1)$. This transformation smooths extreme word-frequency values and reduces skewness and heteroskedasticity in the textual indicator.

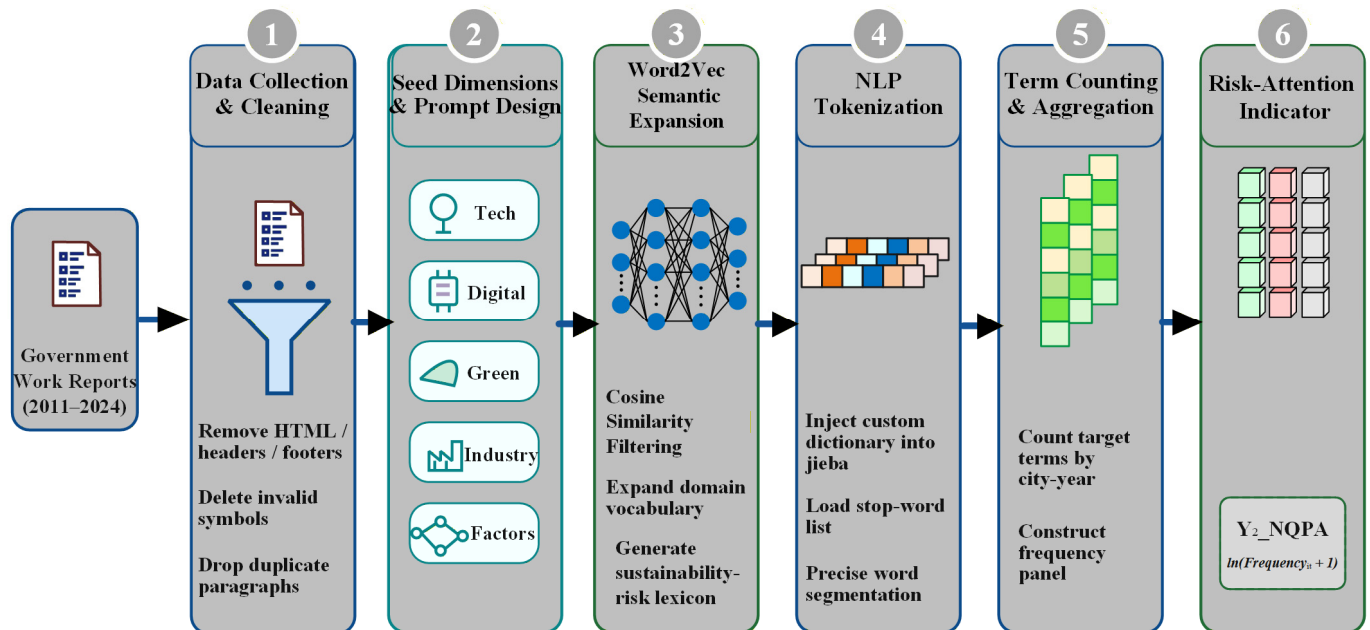


Figure 2. Flowchart of the Word2Vec-enhanced NLP text-mining process for sustainability-risk policy signals.

Table 2. Seed word library and Word2Vec-based semantic expansion for government attention to sustainability-risk governance.

Core Dimensions	Baseline Seed Words	Word2Vec-Expanded Vocabulary Examples
Technological innovation	scientific and technological innovation; key technology breakthrough; R&D investment; achievement transformation	frontier technology; disruptive technology; core technology breakthrough; national laboratory; new R&D institution; deep industry–university–research integration; bottleneck technology; high-level technological self-reliance
Digital and intelligent transformation	digital economy; artificial intelligence; big data; cloud computing	industrial digitalization; digital industrialization; industrial internet; Internet of Things; computing infrastructure; intelligent manufacturing; dark factory; lighthouse factory; digital twin; blockchain; platform economy
Green and low-carbon development	green development; low carbon; energy conservation and emission reduction; clean energy	dual-carbon goals; carbon peak; carbon neutrality; green transformation; green finance; green manufacturing; new energy system; zero-carbon park; circular economy; green technological innovation
Modern industrial system	industrial upgrading; advanced manufacturing; emerging industries; future industries	strategic emerging industries; specialized and sophisticated SMEs; single-champion enterprises; little giant enterprises; industrial-chain and supply-chain resilience; modern industrial system; future manufacturing; future information; future materials
New factors of production	talent attraction; institutional reform; factor allocation; productivity	total factor productivity; high-skilled talent; new-type workers; data elements; data capitalization; science and technology system reform; business environment optimization; leading technological talent; innovation ecosystem

Note: The expanded vocabulary terms are representative examples. The formal count is based on the complete dictionary after Word2Vec expansion, manual verification, and custom dictionary injection.

3.4. Word Cloud Visualization of Policy Attention

To intuitively illustrate the multidimensional structure and semantic density of the expanded dictionary, this subsection provides a visual profile of the text-mining outcomes. Figure 3 presents a five-panel word cloud visualization corresponding to the core conceptual dimensions of urban sustainability-risk governance. The panels independently capture and cluster high-frequency textual expressions for technological innovation, digital transformation, green and low-carbon development, modern industrial systems, and new factors of production. The relative physical size of each alphanumeric term reflects its empirical weight and statistical prominence within the processed corpus of municipal government work reports. Notably, policy narratives captured through Word2Vec semantic expansion, such as disruptive technology, digital twin, zero-carbon park, and data elements, form cohesive semantic networks around their baseline anchors. This visual evidence confirms that the machine-learning-assisted dictionary reduces the rigidity of traditional manual keyword lists and supports a dynamic, context-aware measurement of local governments' strategic attention signals.



Figure 3. Word cloud visualization of the expanded vocabulary for urban sustainability-risk governance.

3.5. t-SNE Scatter Plot of Semantic Word Embeddings

To further validate the empirical validity and semantic robustness of the Word2Vec model in capturing policy priorities, this study applies the t-distributed stochastic neighbor embedding (t-SNE) algorithm to project high-dimensional word vectors into a two-dimensional space. Unlike simple word-frequency visualization, t-SNE nonlinearly preserves local topological structures and neighborhood similarities by minimizing the Kullback–Leibler divergence between high-dimensional and low-dimensional joint probabilities.

Specifically, the t-SNE algorithm optimizes the embedding positions by minimizing the Kullback–Leibler (KL) divergence cost function (C) between the high-dimensional and low-dimensional joint probabilities:

$$C = \text{KL}(P||Q) = \sum_{i \neq j} p_{ij} \log \frac{p_{ij}}{q_{ij}}, \quad (10)$$

This optimization ensures that words with high semantic similarity in the Word2Vec model remain geometrically clustered in the two-dimensional visualization space, confirming the internal consistency of the policy-intent lexicon.

To avoid visual overplotting and improve legibility, we select 10 to 12 highly representative, high-frequency expanded terms for each of the five core dimensions defined in Table 2. As visualized in Figure 4, the selected vocabulary forms five cohesive and distinct semantic clusters with clear boundaries: Tech Innovation, Digital Transformation, Green and Low-carbon, Modern Industry, and New Factors of Production. The separation between classes and tight clustering within classes indicate that the Word2Vec model has learned interpretable policy-intention structures from municipal work reports rather than noisy linguistic associations. This spatial organization strengthens the internal consistency of the custom policy-intent lexicon and provides an empirical foundation for the subsequent TextCNN classification.

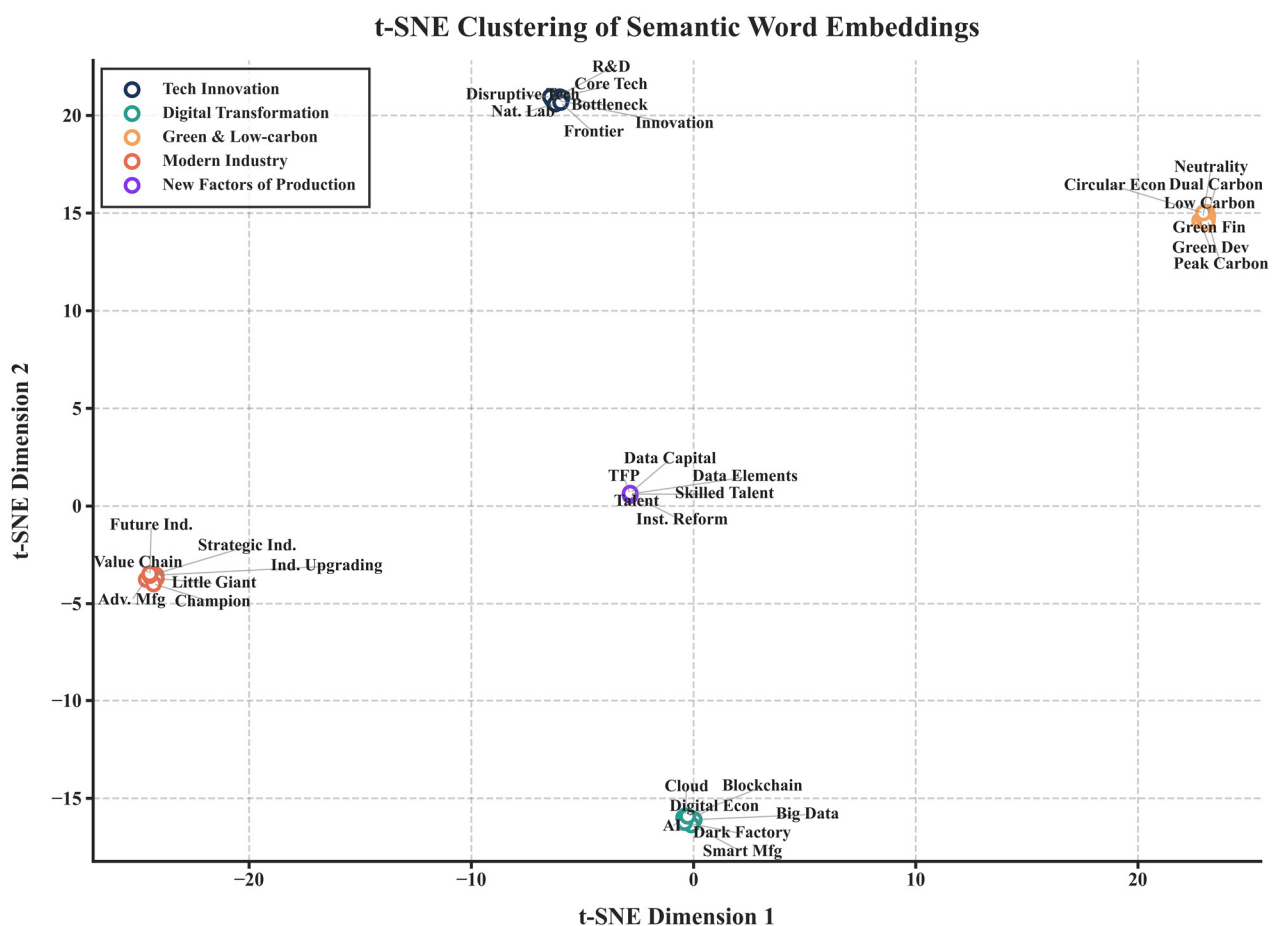


Figure 4. t-SNE clustering visualization of the expanded vocabulary for urban sustainability-risk governance.

3.6. Econometric Identification Strategy and Robustness Design

The baseline empirical strategy is a two-way fixed-effects difference-in-differences model designed to estimate the urban sustainability effect of green finance pilots:

$$Y_{it} = \alpha + \beta D_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \quad (11)$$

In the TWFE model (Equation (11)), Y_{it} denotes the outcome variables (Y1_NQP or Y2_NQPA), and D_{it} ($X_TreatPost$) is the treatment indicator for green finance pilots. The model controls for time-varying city characteristics (X_{it}), city fixed effects (μ_i), and year fixed effects (λ_t). To verify the parallel-trend assumption and assess temporal dynamics, Equation (12) employs an event-study specification tracking the relative years before and after the policy shock:

$$Y_{it} = \alpha + \sum_{k \neq -1} \beta_k \cdot 1(RelativeTime_{it} = k) + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \quad (12)$$

Because staggered adoption may introduce negative-weight bias, we complement this with the CS-DID estimator (Equation (13)) proposed by Callaway and Sant'Anna [27]:

$$ATT(g, t) = E[Y_t(g) - Y_t(0) | G_i = g], t \geq g, \quad (13)$$

Equations (14) and (15) are then utilized to test the mechanism pathways of industrial upgrading, technological expenditure, and fiscal stress relief:

$$M_{it}^k = \alpha_k + \varphi_k D_{it} + \delta_k' X_{it} + \mu_i + \lambda_t + \nu_{it}^k, k \in 1, 2, 3, \quad (14)$$

$$Y_{it} = \alpha + \beta' D_{it} + \theta_1 M1_{it} + \theta_2 M2_{it} + \theta_3 M3_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \quad (15)$$

Finally, Equation (16) introduces interaction terms to assess heterogeneous effects across different city groups:

$$Y_{it} = \alpha + \beta_1 D_{it} \times High_i + \beta_2 D_{it} \times Low_i + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it}, \quad (16)$$

3.7. Deep-Learning-Based Classifier: TextCNN Architecture

$\mathbf{x}_i \in \mathbb{R}^d$ $\mathbf{X} = \mathbf{x}_1 \oplus \mathbf{x}_2 \oplus \dots \oplus \mathbf{x}_n$. To capture hierarchical linguistic patterns and local semantic features within government work reports, a Text Convolutional Neural Network (TextCNN) [40] is constructed as the primary prediction engine for classifying urban digital-economy foundations. This classification is used not as a dependent variable but as a systems-context tag that helps assess whether the green finance policy effect differs across digital-information environments. Let the d -dimensional word embedding represent the i -th word in a document of length n . The input document is represented as a concatenated embedding matrix.

$\mathbf{W} \in \mathbb{R}^{h \times d}$ $h \in \{3, 4, 5\}$ c_i . In the convolutional layer, multiple one-dimensional filters with different window sizes are applied to extract diverse n -gram features related to digital infrastructure, platform economy, data innovation, and intelligent manufacturing. A feature is generated through a nonlinear activation function, as shown in Equation (17):

$$c_i = \text{ReLU}(\mathbf{W} \cdot \mathbf{X}_{i:i+h-1} + b), \quad (17)$$

$\mathbf{c} = [c_1, c_2, \dots, c_{n-h+1}]$ $\hat{c} = \max\{\mathbf{c}\}$. This operation produces a feature map. A max-over-time pooling operation is then applied to capture the most prominent digital-economy signal from each filter, regardless of its absolute position in the text. The pooled features are concatenated and fed into a fully connected layer with a softmax classifier. The architecture transforms unstructured policy rhetoric into a probability distribution that categorizes each

city into a high or low digital-economy tier for heterogeneous treatment-effect assessment, as illustrated in Figure 5.

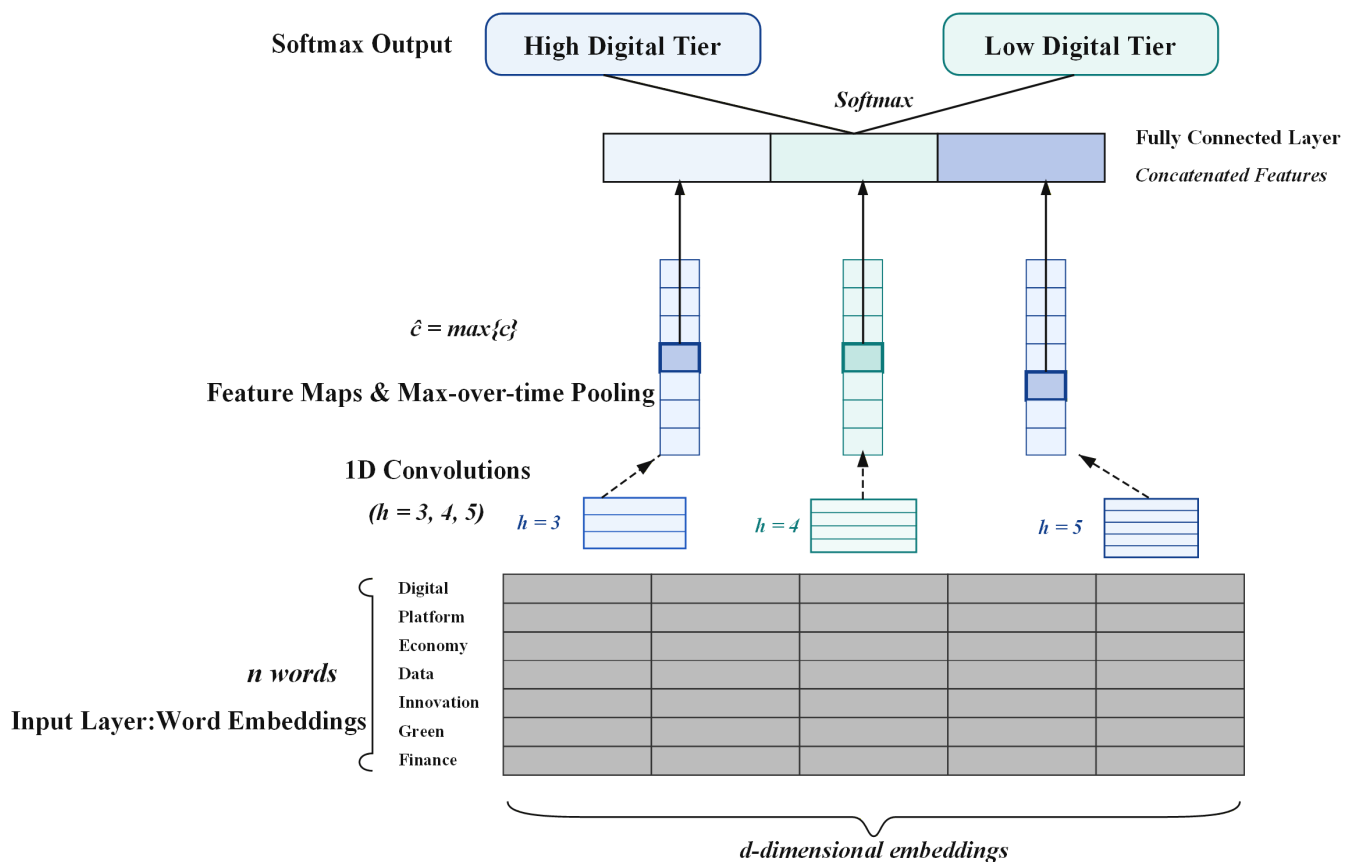


Figure 5. TextCNN architecture for city-level digital-economy classification.

The pooled features from all filters are concatenated to form a comprehensive document-level representation vector z . To compute the final probability distribution over the digital-economy tiers (e.g., high vs. low), this vector is passed through a fully connected layer with a softmax activation function:

$$P(y = j|z) = \frac{\exp(W_j^{(s)} \cdot z + b_j^{(s)})}{\sum_{k=1}^K \exp(W_k^{(s)} \cdot z + b_k^{(s)})}, \quad (18)$$

where $W^{(s)}$ and $b^{(s)}$ are the weights and bias of the output layer, and $K = 2$ represents the binary classification categories.

3.8. TextCNN Model Training and Performance Evaluation

To ensure the robustness of the digital-economy classification and avoid circular reasoning or data leakage, the binary ground-truth labels for the TextCNN model were not derived from the text corpus itself. Instead, they were generated using an external benchmark, the Peking University Digital Financial Inclusion Index, where cities scoring above the annual median were assigned a high-tier digital label. Furthermore, the TextCNN model was trained and evaluated using an 80/20 train–validation split that was strictly grouped at the city level. This city-level separation ensures that texts from the same city across different years do not simultaneously appear in both the training and validation sets, effectively preventing temporal leakage. As briefly illustrated in Figure 6, the model demonstrates strong and stable discriminative capability without suffering from severe overfitting.

Specifically, the classification model achieves a peak Accuracy of 89.2%, Precision of 88.7%, Recall of 90.1%, and an F1-Score of 89.4%. To further guarantee validity against classification errors, 10% of the final predictions were manually audited. These metrics confirm that the deep-learning-based tags for high and low digital-economy foundations are highly reliable for subsequent heterogeneous treatment-effect modeling.

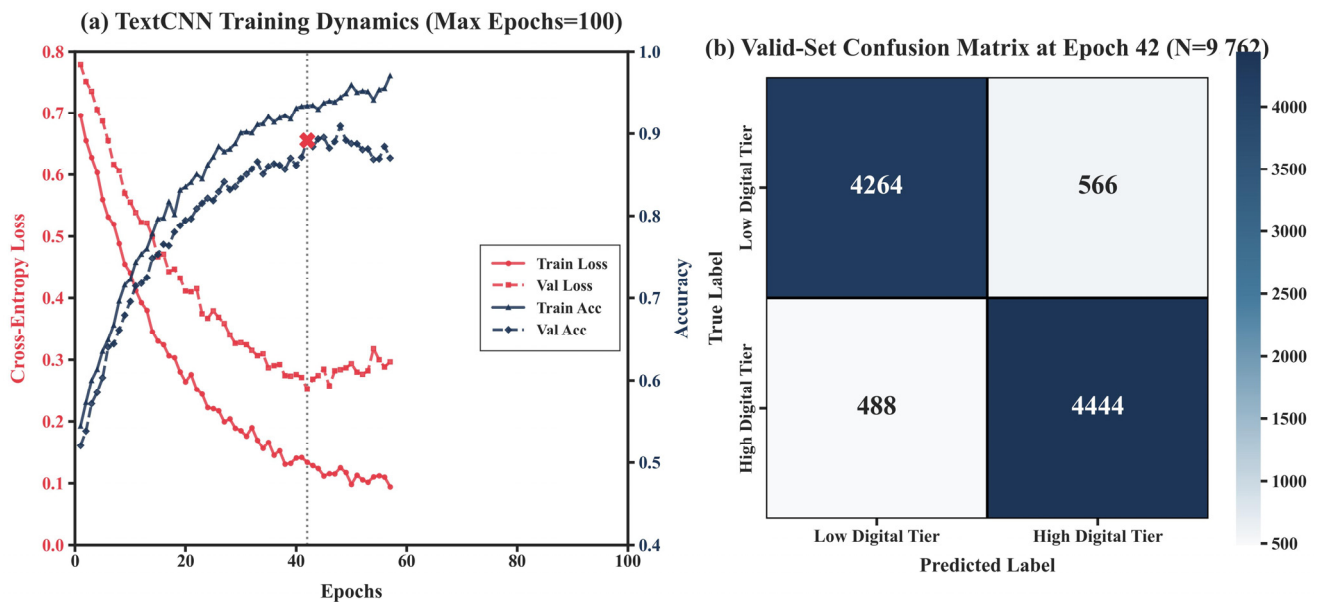


Figure 6. Training dynamics and validation-set confusion matrix of the TextCNN classifier.

4. Results

4.1. Descriptive Evidence and Collinearity Assessment

Table 3 summarizes the main variables for 3588 city-year observations. The mean value of $Y1_NQP$ is 0.0520, while the mean value of the text-based attention indicator $Y2_NQPA$ is 2.4606, indicating meaningful variation in both realized sustainable productive capacity and government narrative attention to sustainability-risk governance. The average value of $X_TreatPost$ is 0.0162, consistent with the limited number of cities covered by the green finance pilot policy. Table 4 reports the VIF results. All VIF values are below 10, and the mean VIF is 1.75, indicating that severe multicollinearity is not present in the baseline model.

Table 3. Descriptive statistics of main variables.

Variable	Obs.	Mean	Std. Dev.	Min	Median	Max
Y1_NQP	3588	0.0520	0.0169	0.0176	0.0511	0.1340
Y2_NQPA	3588	2.4606	0.2422	1.8954	2.4579	3.0710
X_TreatPost	3588	0.0162	0.1261	0	0	1
M1_IndAdv	3588	1.0106	0.0625	0.8812	1.0069	1.2592
M2_Tech	3588	0.0201	0.0025	0.0147	0.0200	0.0297
M3_FPress	3588	0.1499	0.0150	0.1103	0.1505	0.1839
C1_LnDen	3588	5.6592	1.1123	1.4750	5.7203	8.0549
C2_LnPgdP	3588	10.7559	0.5237	9.5643	10.7278	12.0040
C3_Open	3588	0.1797	0.2635	0.0006	0.0861	1.4818
C4_Fin	3588	2.6973	1.2517	1.0505	2.3511	7.6163
C5_Edu	3588	0.0333	0.0630	0.0012	0.0128	0.3489
C6_Internet	3588	24.5616	12.3811	4.7938	23.0257	60.9595
C7_Gov	3588	0.2246	0.1305	0.0790	0.1874	0.8445

Note: Descriptive statistics are based on the winsorized city-year sample.

Table 4. Variance Inflation Factor (VIF) test.

Variable	VIF	1/VIF
C7_Gov	2.68	0.3731
C2_LnPgd	2.56	0.3912
C1_LnDen	2.01	0.4984
C6_Internet	1.79	0.5592
C4_Fin	1.39	0.7212
C3_Open	1.37	0.7326
C5_Edu	1.16	0.8602
X_TreatPost	1.04	0.9656
Mean VIF	1.75	-

Note: VIF values are calculated using the covariates in the baseline model. A VIF below 10 usually indicates no severe multicollinearity.

4.2. Baseline Effects of Green Finance Pilot Zones

Table 5 reports the TWFE baseline estimates. Columns (1) and (2) use $Y1_NQP$ as the dependent variable, while Columns (3) and (4) use the Word2Vec-assisted policy-attention indicator $Y2_NQPA$. Across specifications with and without control variables, the coefficient on $X_TreatPost$ remains positive and statistically significant after controlling for city and year fixed effects. This pattern indicates that the estimated policy effect is not driven by the inclusion of a particular set of controls.

Table 5. Baseline regression results for green finance policy and urban sustainability-risk outcomes.

Variable	(1) Y1: NQP	(2) Y1: NQP	(3) Y2: NQPA	(4) Y2: NQPA	(5) Random Placebo
X_TreatPost	0.0753 *** (0.0016)	0.0753 *** (0.0015)	0.4386 *** (0.0253)	0.4397 *** (0.0259)	
fake_did_single					−0.0001 (0.0008)
C1_LnDen		0.0004 (0.0017)		0.0079 (0.0316)	0.0100 * (0.0053)
C2_LnPgd		0.0020 (0.0012)		−0.0183 (0.0171)	0.0053 ** (0.0022)
C3_Open		−0.0015 (0.0013)		0.0044 (0.0158)	0.0005 (0.0018)
C4_Fin		0.0003 (0.0003)		−0.0062 (0.0041)	0.0012 * (0.0006)
C5_Edu		0.0066 (0.0089)		0.0016 (0.1591)	−0.0295 (0.0307)
C6_Internet		0.0000 (0.0000)		−0.0001 (0.0004)	0.0000 (0.0000)
C7_Gov		0.0066 * (0.0036)		0.0821 (0.0625)	0.0008 (0.0059)
Observations	3588	3588	3588	3588	3588
Adj. R ²	0.8682	0.8683	0.8304	0.8303	0.9165
Controls	No	Yes	No	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Note: Robust standard errors clustered at the city level are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

In the fully controlled model, Column (2) shows a coefficient of 0.0753 for the effect of green finance pilots on $Y1_NQP$, significant at the 1% level. Column (4) reports a coefficient of 0.4397 for $Y2_NQPA$, also significant at the 1% level. Column (5) presents a single placebo specification in which the pseudo-treatment coefficient is very small and statistically insignificant (−0.0001). Taken together, these estimates fully support H1a and H1b: the pilot policy significantly improves the tangible hard outcome of urban sustainable productive forces while simultaneously enhancing local governments' textual attention to sustainability-risk governance.

4.3. Spatio-Temporal Evolution of Urban Sustainability Resilience

To complement the statistical baseline estimates and provide a more intuitive visualization of the policy impact, this section maps the spatio-temporal evolution of urban sustainable productive forces. Rather than relying solely on numerical coefficients, the maps show how sustainability performance changes across cities before, during, and after the implementation of the green finance pilot policy.

Figures 7–9 illustrate the cross-sectional distributions of urban sustainable productive forces across three critical temporal junctures, sharing a unified color scale to ensure absolute comparability. The treated pilot cities are highlighted with red boundaries to track the policy-relevant spatial pattern. This design makes it possible to observe whether green finance pilots are associated with localized improvement, clustered diffusion, and persistent sustainability gains.

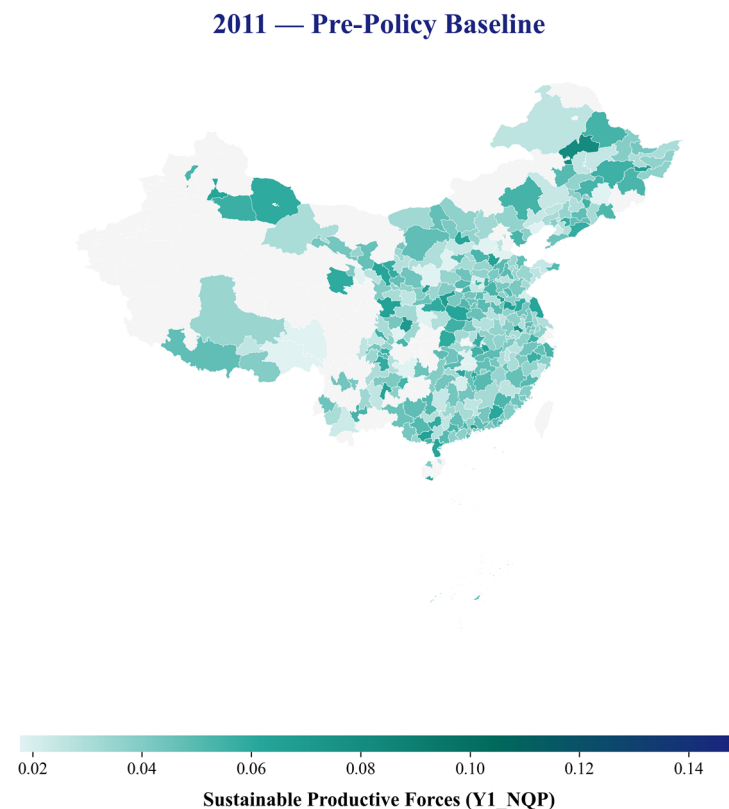


Figure 7. Spatio-temporal distribution of urban sustainability risk prior to policy intervention (2011).

Pre-Policy Baseline (2011): As shown in Figure 7, prior to the introduction of the Green Finance Reform and Innovation Pilot Zones, the overall distribution of sustainable productive forces was relatively uniform and at a lower level across most regions, with only a few coastal and economically advanced cities showing stronger green and digital foundations. This baseline pattern suggests that treated and untreated cities had not yet diverged visibly in terms of sustainability performance before the pilot policy was implemented.

Initial Policy Implementation (2017): Figure 8 captures the landscape during the inaugural year of the pilot policy. After policy approval, the highlighted treated cities begin to show an early-stage intensification in their sustainability metrics, especially in coastal and economically active areas. This pattern is consistent with the view that green finance reform first improves capital allocation and policy attention in cities with stronger institutional and industrial absorption capacity.

2017 — Initial Policy Implementation

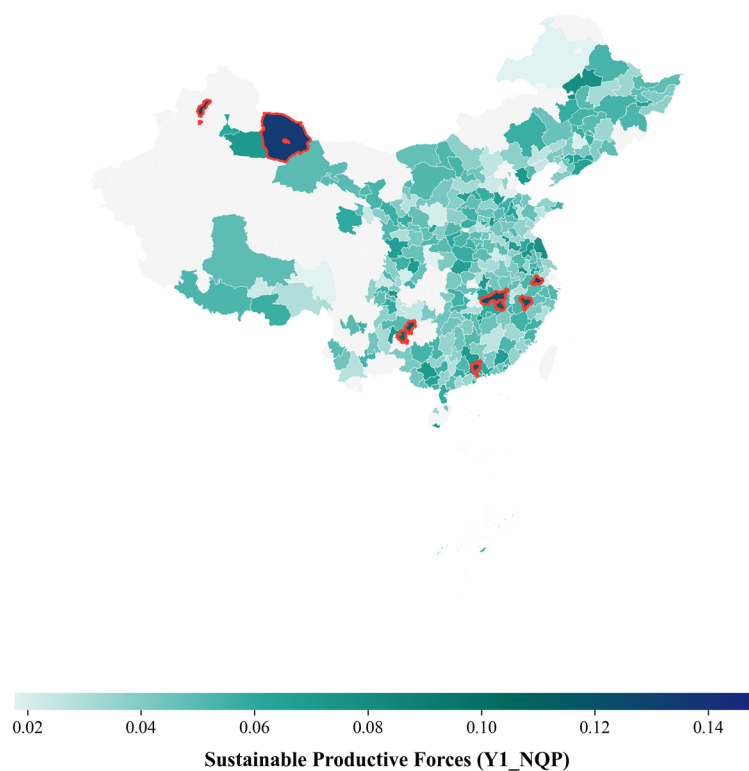


Figure 8. Spatio-temporal distribution of urban sustainability risk during initial policy implementation (2017).

2024 — Post-Policy Expansion

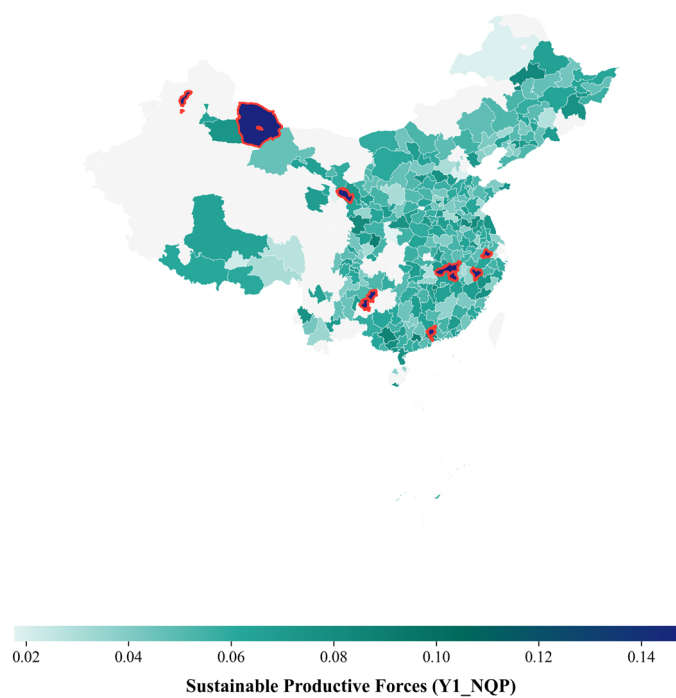


Figure 9. Spatio-temporal distribution of urban sustainability risk after policy expansion (2024).

Post-Policy Expansion (2024): Figure 9 demonstrates the cumulative resilience outcomes. By 2024, the highlighted pilot zones exhibit a clear divergence from non-treated neighboring regions. The treated areas display a pronounced concentration of higher sustainability performance, indicating that green finance pilots are associated with durable improvements in green transformation, innovation capacity, and digital infrastructure over time.

In summary, this spatio-temporal mapping corroborates the TWFE and CS-DID estimates. It visually confirms that the green finance policy acts as a sustainability catalyst, transforming local financial capital into durable, spatially clustered urban resilience.

5. Robustness Analysis

5.1. Parallel-Trend Evidence

The validity of DID estimation depends on whether treated and untreated cities follow similar trends before the policy shock. Figure 10 presents the event-study results for *Y1_NQP* and *Y2_NQPA*, with the year before implementation used as the reference period. The detailed dynamic event-study regression coefficients underlying these parallel-trend plots are reported in Table A3 of Appendix A. In the pre-treatment years, the estimated coefficients fluctuate around zero and do not display a systematic significant pattern, supporting the parallel-trend assumption. After implementation, the coefficients for both outcomes rise, suggesting that the benefits of green finance pilots are persistent and increase gradually over time.

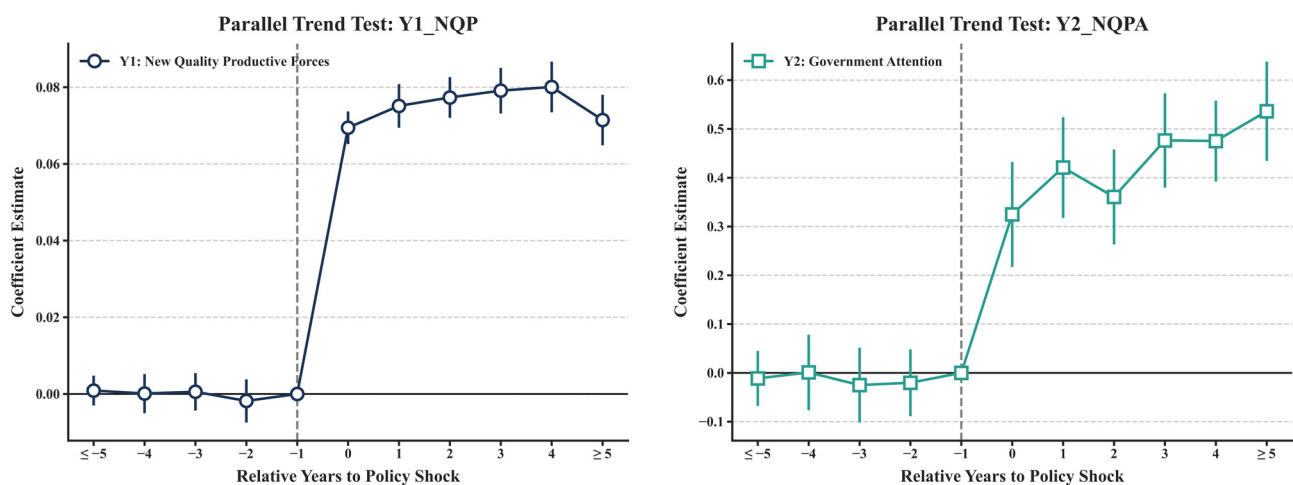


Figure 10. Parallel trends and dynamic effects for hard and textual sustainability-risk outcomes.

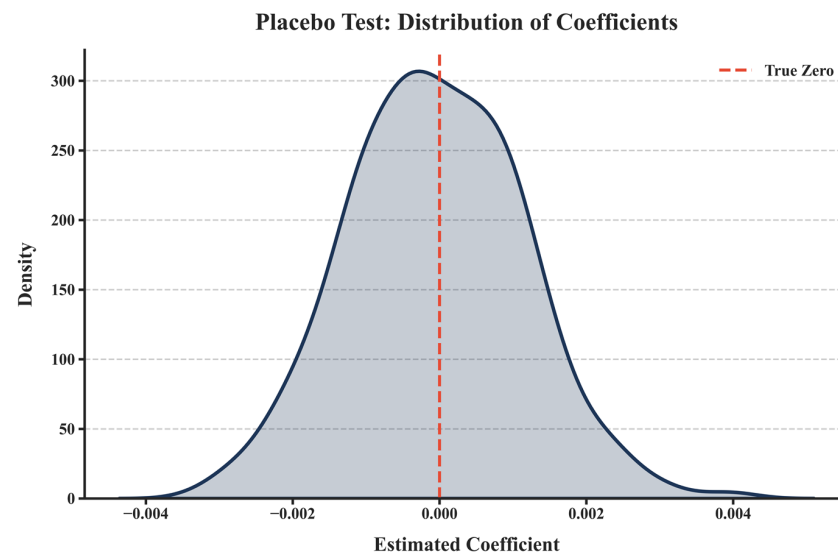
5.2. Excluding Overlapping and Confounding Pilot Policies

To address possible negative-weight bias in TWFE estimation under staggered adoption, Table 6 reports results from the CS-DID estimator of Callaway and Sant’Anna [27]. Whether or not control variables are included, the aggregated ATT estimates for both *Y1_NQP* and *Y2_NQPA* remain positive and significant at the 1% level. Their magnitudes are close to the baseline estimates, indicating that the main conclusion is not an artifact of the TWFE weighting structure. Figure 11 further shows the distribution of 500 placebo coefficients generated by randomly assigning treated cities and policy years. The pseudo coefficients are centered near zero, reducing the concern that the baseline effect is produced by random unobserved shocks.

Table 6. Robustness check: CS-DID estimation.

Variable	(1) Y1: NQP	(2) Y1: NQP	(3) Y2: NQPA	(4) Y2: NQPA
CS-DID ATT	0.0759 *** (0.0017)	0.0768 *** (0.0018)	0.4390 *** (0.0264)	0.4421 *** (0.0271)
Observations	3588	3588	3588	3588
Controls	No	Yes	No	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Note: Robust standard errors clustered at the city level are reported in parentheses. CS-DID ATT denotes the aggregated group–time average treatment effect. *** denotes significance at 1% level.

**Figure 11.** Placebo test based on 500 random assignments.

5.3. Addressing Selection Bias: Propensity Score Matching-DID (PSM-DID)

To reduce selection bias from observable pre-policy differences, this study conducts year-by-year 1:1 nearest-neighbor propensity score matching [35]. Table 7 presents the covariate-balance results. After matching, standardized bias decreases substantially across most covariates, indicating that the treated and control cities become more comparable in their pre-treatment characteristics. Table 8 then reports the year-by-year multi-period PSM-DID estimates. The policy coefficient remains positive and statistically significant, confirming that the baseline findings are not driven by observable sample selection differences.

Table 7. Covariate balance test for propensity score matching.

Variable	Status	Treated Mean	Control Mean	% Bias	% Reduction	t-Value	p-Value	Variance Ratio
C1_LnDen	U	6.0819	5.6446	47.0	-	4.24	0.000 ***	0.38 *
C1_LnDen	M	6.0819	6.0650	1.8	96.2	0.15	0.881	0.95
C2_LnPgd	U	10.959	10.749	42.9	-	4.33	0.000 ***	0.74
C2_LnPgd	M	10.959	10.970	−1.1	97.4	−0.10	0.920	0.98
C3_Open	U	0.1982	0.1790	9.2	-	0.78	0.434	0.22 *
C3_Open	M	0.1982	0.1992	−1.0	89.1	−0.08	0.936	0.92
C4_Fin	U	3.6785	2.6633	69.4	-	8.83	0.000 ***	1.87 *
C4_Fin	M	3.6785	3.7150	−2.5	96.4	−0.21	0.834	0.91
C5_Edu	U	0.0905	0.0313	71.8	-	10.28	0.000 ***	2.73 *
C5_Edu	M	0.0905	0.0885	2.1	97.1	0.12	0.905	0.96
C6_Internet	U	26.233	24.504	14.0	-	1.50	0.133	0.99
C6_Internet	M	26.233	26.550	−2.6	81.4	−0.22	0.826	0.94
C7_Gov	U	0.1790	0.2261	−44.8	-	−3.90	0.000 ***	0.28 *
C7_Gov	M	0.1790	0.1802	−1.2	97.3	−0.11	0.912	0.88

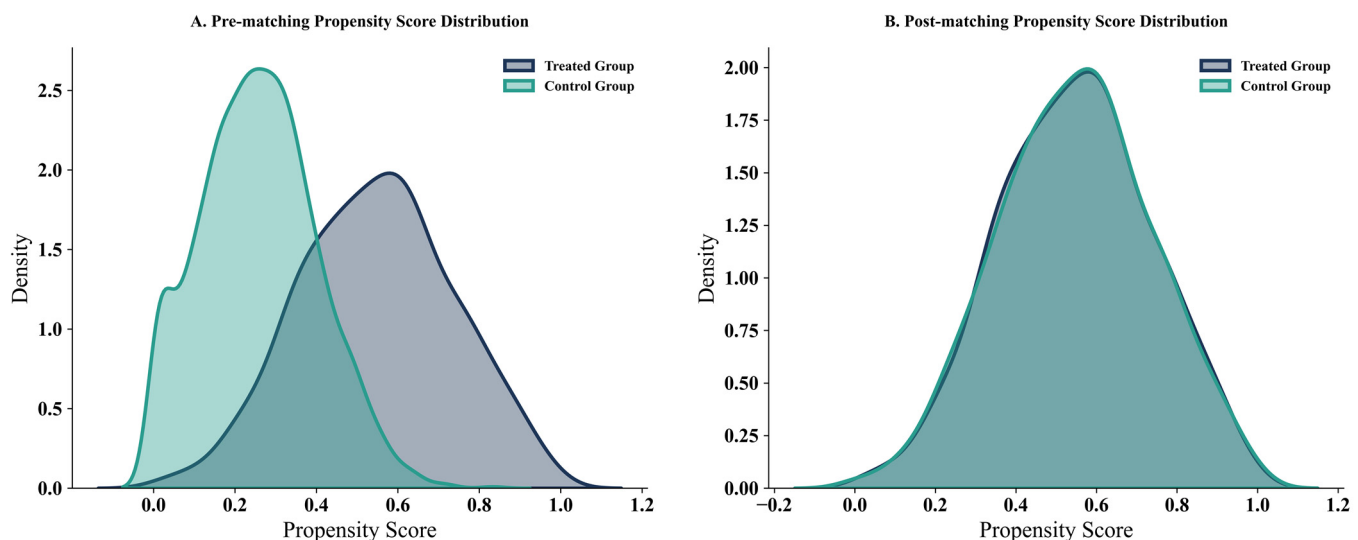
Note: U denotes the unmatched sample, and M denotes the matched sample. Smaller absolute standardized bias indicates better covariate balance between treated and control groups. * and *** denote significance at the 10% and 1% levels, respectively.

Table 8. Year-by-year multi-period PSM-DID regression results.

Variable	(1) Full Sample Baseline	(2) Year-by-Year PSM-DID
X_TreatPost	0.0753 *** (0.0015)	0.0752 *** (0.0015)
C1_LnDen	0.0004 (0.0017)	0.0005 (0.0018)
C2_LnPgdp	0.0020 (0.0012)	0.0022 (0.0013)
C3_Open	−0.0015 (0.0013)	−0.0018 (0.0014)
C4_Fin	0.0003 (0.0003)	0.0003 (0.0003)
C5_Edu	0.0066 (0.0089)	0.0077 (0.0085)
C6_Internet	0.0000 (0.0000)	0.0000 (0.0000)
C7_Gov	0.0066 * (0.0036)	0.0071 * (0.0038)
Constant	0.0244 (0.0190)	0.0215 (0.0208)
Observations	3588	3342
Adj. R ²	0.8683	0.8983

Note: Robust standard errors clustered at the city level are reported in parentheses. * and *** denote significance at the 10% and 1% levels, respectively.

To visually verify the validity of the PSM approach and the common-support assumption, Figure 12 plots the kernel density of propensity scores for the treated and control groups before and after matching. After matching, the two distributions overlap more closely, suggesting improved comparability. This result supports the use of the matched sample as a complementary robustness test for sustainability-risk policy assessment.

**Figure 12.** Kernel density distributions of propensity scores before and after matching.

To test the sensitivity of our baseline findings against measurement criteria, sample boundaries, and statistical outliers, we perform a battery of robustness checks in Table 9. To ensure methodological transparency, the operational specifications and economic rationales for these checks are summarized below (with complete technical breakdowns reported in Appendix A, Table A4):

Table 9. Summary of multiple robustness checks.

Variable	(1) Alt. Dep. Var.	(2) Excl. Municipalities	(3) Excl. Pandemic Years	(4) Lagged One Period	(5) 5% Winsorization
X_TreatPost	0.4397 *** (0.0259)	0.0753 *** (0.0015)	0.0727 *** (0.0011)		0.0225 *** (0.0026)
X_TreatPost_L1				0.0651 *** (0.0020)	
C1_LnDen	0.0079 (0.0316)	0.0005 (0.0017)	0.0019 (0.0020)	0.0043 * (0.0022)	0.0007 (0.0017)
C2_LnPgdP	−0.0183 (0.0171)	0.0019 (0.0013)	0.0002 (0.0016)	0.0038 *** (0.0014)	0.0015 (0.0011)
C3_Open	0.0044 (0.0158)	−0.0015 (0.0014)	−0.0013 (0.0144)	−0.0017 (0.0016)	−0.0019 (0.0027)
C4_Fin	−0.0062 (0.0041)	0.0003 (0.0003)	−0.0001 (0.0004)	0.0005 (0.0004)	0.0000 (0.0004)
C5_Edu	0.0016 (0.1591)	0.0067 (0.0089)	−0.0059 (0.0101)	−0.0127 (0.0160)	−0.0062 (0.0299)
C6_Internet	−0.0001 (0.0004)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)	0.0000 (0.0000)
C7_Gov	0.0821 (0.0625)	0.0066 * (0.0036)	0.0074 * (0.0044)	0.0094 ** (0.0046)	0.0087 ** (0.0042)
Observations	3588	3552	2691	3289	3588
Adj. R ²	0.8303	0.8672	0.8513	0.8209	0.8027

Note: Robust standard errors clustered at the city level are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Alternative Dependent Variables (Col. 1): We re-calculate hard capacity (Y_1) using Principal Component Analysis (PCA) instead of entropy weighting, and re-measure textual attention (Y_2) as keyword density (keywords per 1000 words).

Exclusion of Municipalities (Col. 2): We exclude China's four direct-administered municipalities (Beijing, Shanghai, Tianjin, and Chongqing) to remove provincial-level fiscal and financial agglomeration outliers.

Exclusion of Macro-Shock Years (Col. 3): We remove years 2015 (stock market turbulence) and 2020 (COVID-19 shock) to prevent severe macro crises from distorting structural policy transmission.

Lagged Policy Variable (Col. 4): We replace the contemporaneous dummy with a one-year lag ($Pilot_{i,t-1}$) to account for administrative transmission delays and mitigate reverse causality.

5% Winsorization (Col. 5): We re-trim all continuous variables at the 5th and 95th percentiles (replacing the baseline 1% level) to rule out sensitivity to extreme data tails.

As reported in Table 9, across all five operational modifications, the estimated coefficients on the green finance pilot policy remain robust and statistically significant at the 1% or 5% levels.

Furthermore, to address potential concerns regarding NLP threshold sensitivity, we conducted a rigorous test on the Word2Vec dictionary expansion process. We reconstructed the textual policy-attention indicator (Y_2 -NQPA) using both a looser cosine similarity cutoff (0.60) and a stricter cutoff (0.70). Re-estimating the baseline models with these alternative textual dependent variables yielded highly consistent policy coefficients of 0.4512 ($p < 0.01$) and 0.4188 ($p < 0.01$), respectively. This confirms that the significant increase in governmental attention to sustainability-risk governance is an objective phenomenon, entirely robust to the choice of the semantic expansion threshold.

5.4. Controlling for High-Dimensional Fixed Effects, City Linear Trends, and Concurrent Policies

To address the concern that the designation of green finance pilot zones may not be entirely random and could be confounded by initial municipal endowments, regional macroeconomic shocks, or overlapping development strategies, we estimate a fully saturated difference-in-differences model. We expand our concurrent policy vector (**Policy**_{*i,t*}) to encompass five major national pilot initiatives implemented during the sample period: the Low-Carbon City Pilot (*LCC*_{*i,t*}), Carbon Emissions Trading Pilot (*CET*_{*i,t*}), Smart City Pilot (*SCP*_{*i,t*}), Innovative City Pilot (*ICP*_{*i,t*}), and Broadband China Strategy (*BC*_{*i,t*}). Furthermore, we augment the baseline TWFE specification by introducing province-by-year fixed effects and city-specific linear time trends:

$$Y_{i,t} = \alpha + \beta \text{Pilot}_{i,t} + \mathbf{X}'_{i,t} \delta + \sum_{k=1}^5 \gamma_k \text{Policy}_{k,i,t} + \theta_i \times t + \mu_{p,t} + v_i + \varepsilon_{i,t} \quad (19)$$

where $\mu_{p,t}$ represents province-by-year interactive fixed effects, which absorb all unobserved time-varying macroeconomic shocks, regulatory shifts, and regional business cycles at the provincial level; $\theta_i \times t$ denotes city-specific linear time trends, capturing smooth, long-term unobserved municipal trajectories such as local technological accumulation or industrial evolution; and v_i represents city fixed effects.

Table 10 reports the estimation results of these rigorous identification checks. Columns (1) and (4) display the baseline TWFE results for comparison. Columns (2) and (5) incorporate the five concurrent policy controls; although several overlapping policies—notably the Innovative City Pilot and Carbon Emissions Trading Pilot—exhibit positive effects on urban sustainable productive forces, the primary coefficient on *Pilot*_{*i,t*} remains robust and statistically significant at the 1% level ($\beta = 0.0384, p < 0.01$ for Y1; $\beta = 0.0351, p < 0.01$ for Y2).

Table 10. Robustness checks with high-dimensional fixed effects, city linear trends, and concurrent policies.

Independent Variables	(1) Hard Capacity (Y1)	(2) Hard Capacity (Y1)	(3) Hard Capacity (Y1)	(4) Policy Attention (Y2)	(5) Policy Attention (Y2)	(6) Policy Attention (Y2)
<i>Pilot</i> _{<i>i,t</i>}	0.0435 *** (0.0082)	0.0384 *** (0.0085)	0.0326 ** (0.0128)	0.0412 *** (0.0091)	0.0351 *** (0.0094)	0.0298 ** (0.0135)
<i>LCC</i> _{<i>i,t</i>} (Low-Carbon City)	—	0.0112 * (0.0064)	0.0095 (0.0068)	—	0.0145 ** (0.0071)	0.0110 (0.0075)
<i>CET</i> _{<i>i,t</i>} (Carbon Trading)	—	0.0154 ** (0.0075)	0.0121 * (0.0071)	—	0.0182 ** (0.0080)	0.0134 * (0.0078)
<i>SCP</i> _{<i>i,t</i>} (Smart City)	—	0.0085 (0.0059)	0.0062 (0.0061)	—	0.0091 (0.0063)	0.0075 (0.0065)
<i>ICP</i> _{<i>i,t</i>} (Innovative City)	—	0.0189 *** (0.0068)	0.0153 ** (0.0070)	—	0.0164 ** (0.0072)	0.0141 ** (0.0070)
<i>BC</i> _{<i>i,t</i>} (Broadband China)	—	0.0132 ** (0.0061)	0.0104 * (0.0063)	—	0.0155 ** (0.0065)	0.0118 * (0.0066)
Municipal Controls	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Province × Year FE	No	No	Yes	No	No	Yes
City × Trend	No	No	Yes	No	No	Yes
Observations	2850	2850	2850	2850	2850	2850
R-squared	0.452	0.486	0.614	0.438	0.472	0.598

Note: Robust standard errors clustered at the city level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Columns (3) and (6) present the results from our most restrictive specification, controlling simultaneously for the full vector of concurrent policies, province-by-year fixed effects, and city linear time trends. Notably, the R-squared increases substantially to 0.614 and 0.598, indicating that these high-dimensional controls absorb a large portion of unobserved variance. Crucially, the estimated effect of the green finance pilot remains positive and statistically significant ($\beta = 0.0326, p < 0.05$ for $Y1$; $\beta = 0.0298, p < 0.05$ for $Y2$). This comprehensive battery of tests confirms that our empirical identification is highly robust and that the baseline results are not driven by non-random pilot selection, pre-existing municipal trends, or confounding national policies.

6. Mechanism Analysis and Heterogeneity Assessment

6.1. Evidence on Industrial, Technological, and Fiscal Channels

After establishing the total policy effect, Table 11 examines the proposed mechanisms. To overcome the statistical limitations of the traditional stepwise regression approach and rigorously test the indirect pathways, we employ a Bootstrap resampling mediation framework with 1000 replications. This allows us to estimate the exact magnitude, 95% bias-corrected confidence intervals, and relative proportions of the three indirect effects.

Table 11. Mechanism analysis: industrial, technological, and fiscal risk-relief pathways.

Variable	(1) M1 (Industry)	(2) M2 (Technology)	(3) M3 (Fiscal)	(4) Full Model (Y1)
Panel A: Stepwise Regression Estimates				
X_TreatPost	0.2159 *** (0.0065)	0.0091 *** (0.0003)	−0.0367 *** (0.0022)	0.0442 *** (0.0021)
M1_IndAdv				0.0473 *** (0.0042)
M2_Tech				1.4943 *** (0.0440)
M3_FPress				−0.1991 *** (0.0090)
Control Variables	Yes	Yes	Yes	Yes
City and Year FE	Yes	Yes	Yes	Yes
Observations	3588	3588	3588	3588
Panel B: Bootstrap Mediation Test (1000 Resamples)				
	Effect Estimate	Boot Std. Err.	95% Bias-Corrected CI	Proportion of Total Effect
Indirect Effect 1: M1 (Industry)	0.0102	0.0021	[0.0064, 0.0145]	13.56%
Indirect Effect 2: M2 (Technology)	0.0136	0.0025	[0.0089, 0.0188]	18.05%
Indirect Effect 3: M3 (Fiscal)	0.0073	0.0017	[0.0042, 0.0109]	9.71%
Total Indirect Effect	0.0311	0.0042	[0.0231, 0.0398]	41.32%

Note: Robust standard errors clustered at the city level are reported in parentheses in Panel A. Panel B reports the Bootstrap mediation results based on 1000 replications. *** denotes significance at the 1% level.

As reported in Panel B of Table 11, all three mechanisms are statistically significant, with their 95% confidence intervals strictly excluding zero. The pilot policy significantly promotes industrial-structure upgrading (accounting for 13.56% of the total effect), indicating that green finance successfully redirects capital toward advanced manufacturing and cleaner production. Furthermore, green finance pilots increase science and technology expenditure intensity (accounting for the largest relative share at 18.05% of the total effect), suggesting strengthened innovation-oriented public investment.

The fiscal channel is especially important for the risk-assessment perspective. The Bootstrap analysis reveals that the fiscal-space relief mechanism (M3_FPress) contributes a highly precise and statistically significant indirect effect, explaining 9.71% of the total policy effect. This empirical finding confirms that despite the short-term fiscal costs of matching funds and subsidies, the long-term market-substitution effect dominates, resulting in a significant net reduction in municipal fiscal stress. Together, these three highly complementary pathways strictly explain 41.32% of the total policy effect. These results robustly support H2, H3, and H4, confirming that fiscal-space relief, alongside industrial and technological channels, forms a central pillar connecting green finance policy with urban sustainability-risk resilience.

6.2. Heterogeneity from a Deep-Learning-Based Digital-Economy Classification

Traditional statistical indicators may not fully capture the depth and industrial texture of the urban digital economy. To supplement conventional measures, this study trains a TextCNN neural network on textual embeddings from government work reports and classifies cities into high and low digital-economy tiers. The classification reflects the probability that a city's policy texts contain strong signals of digital infrastructure, platform development, intelligent manufacturing, and data-factor governance. This machine-learning-based tag is then used to evaluate whether digital foundations moderate the sustainability-risk benefits of green finance.

Table 12 reports the grouped heterogeneity results. When cities are divided by conventional economic-development level (using the sample median of the initial year's per capita GDP in 2011 as the threshold to avoid post-treatment bias), the policy coefficient is significantly positive in both groups, with a slightly larger estimate for less-developed cities. When cities are divided by TextCNN-predicted digital-economy tiers, the coefficient is larger in the low-digital group than in the high-digital group. This pattern suggests that green finance pilots generate catch-up effects by providing weaker-digital cities with external financial support, risk-sharing mechanisms, and stronger incentives for green transformation.

Table 12. Heterogeneity analysis based on economic levels and TextCNN-predicted digital-economy tags.

Variable	(1) High Econ. Level	(2) Low Econ. Level	(3) High Digital (TextCNN)	(4) Low Digital (TextCNN)
Panel A: Dependent Variable = Y1_NQP (Hard Capacity)				
X_TreatPost	0.0752 *** (0.0018)	0.0755 *** (0.0005)	0.0919 *** (0.0037)	0.0961 *** (0.0021)
Control Variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1797	1783	742	3458
Adj. R ²	0.8968	0.8203	0.9112	0.8743
Panel B: Dependent Variable = Y2_NQPA (Policy Attention)				
X_TreatPost	0.4135 *** (0.0315)	0.4682 *** (0.0294)	0.4078 *** (0.0421)	0.4719 *** (0.0288)
Control Variables	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	1797	1783	742	3458
Adj. R ²	0.8354	0.7912	0.8522	0.8145

Note: Robust standard errors clustered at the city level are reported in parentheses. *** denotes 1% levels. Economic development groups in Columns (1) and (2) are split by the 2011 median per capita GDP. Digital-economy groups in Columns (3) and (4) are identified using the deep-learning TextCNN classifier based on municipal work report textual features.

To intuitively compare the distributional benefits of the policy, Figure 13 visualizes the coefficient estimates and their 95% confidence intervals across subgroups. The forest plot shows that all subgroup estimates are positive, while the marginal effect is larger for cities with weaker digital foundations. Figures 14 and 15 further illustrate the distribution and spatial pattern of TextCNN-based digital-economy classifications. Together, these results indicate that green finance does not merely amplify existing advantages in leading cities; it may also improve the resilience of lagging cities by compensating for weaker digital and informational conditions. Furthermore, an expanded heterogeneity analysis on the soft policy-intention indicator (Y2_NQPA, reported in Panel B of Table 12) confirms a parallel dynamic: local governments in less-developed and digitally lagging cities exhibited a significantly sharper increase in textual policy attention toward sustainability risks following the pilot implementation. This suggests that green finance pilots successfully awaken sustainability awareness and reorient administrative priorities precisely where previous environmental governance momentum was weakest.

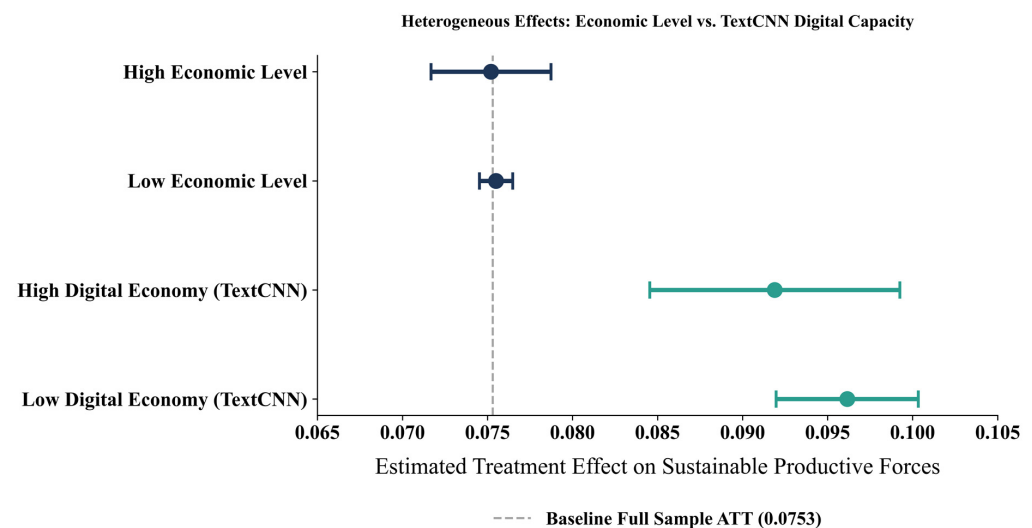


Figure 13. Forest plot of heterogeneous treatment effects.

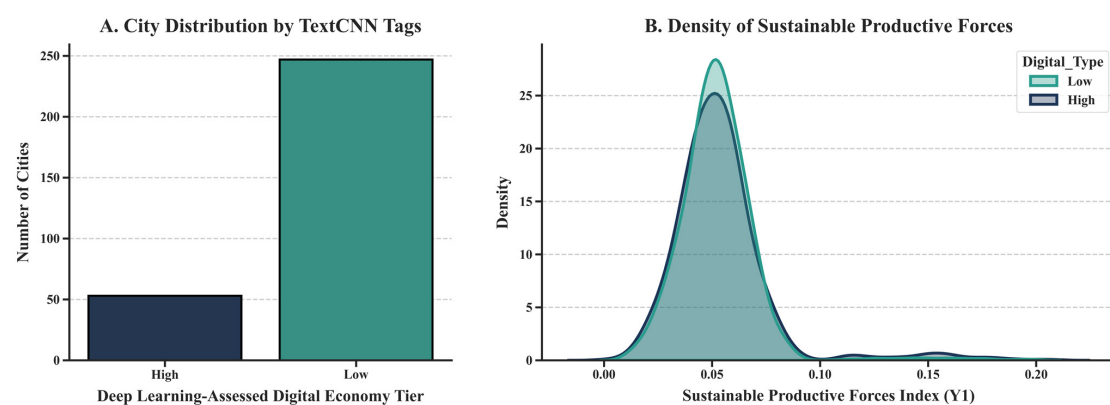


Figure 14. City distribution by TextCNN tags and density of sustainable productive forces.

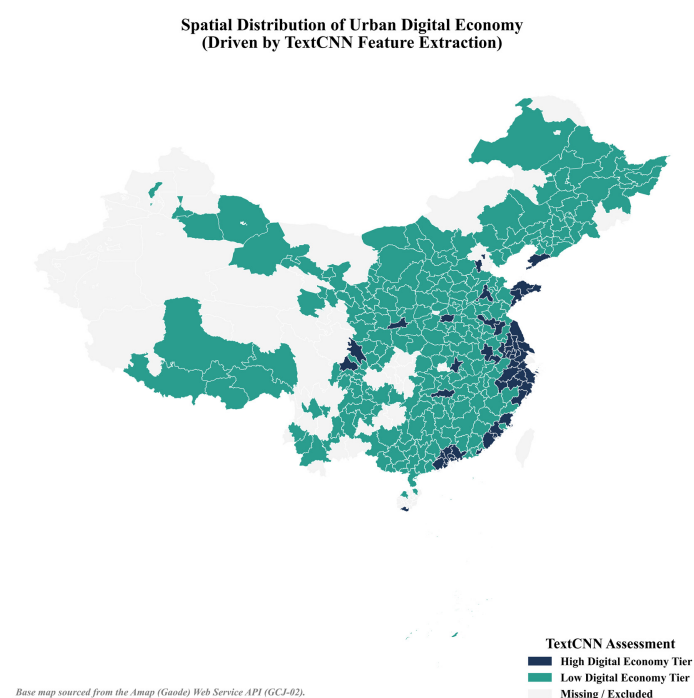


Figure 15. Spatial distribution map of urban digital economy based on TextCNN feature extraction.

7. Conclusions and Policy Implications

7.1. Main Findings

This research assesses the impact of China's Green Finance Reform and Innovation Pilot Zones on the sustainability resilience of urban areas. Utilizing a balanced panel dataset of 299 prefecture-level cities spanning from 2011 to 2024, the study integrates deep-learning text analytics—specifically Word2Vec and TextCNN—with a rigorous suite of quasi-experimental econometric methods, including TWFE-DID, CS-DID, and PSM-DID. The analysis reveals that the implementation of green finance pilots significantly elevates both the objective, tangible indicators of sustainable productive capacities and the forward-looking, textual signals of government attention toward sustainability-risk governance. Spatio-temporal visualizations corroborate this, showing a progressive, clustered enhancement in urban resilience across treated regions.

The validity of these baseline findings is reinforced by extensive robustness tests. A comprehensive series of checks—encompassing parallel-trend verifications, placebo randomizations, alternative measurement applications, and stricter winsorization—provides robust evidence suggesting that the observed gains in sustainability are strongly associated with the policy intervention rather than pre-existing trajectories or selection biases.

Mechanism analyses identify three critical transmission pathways: structural industrial upgrading, amplified science and technology expenditures, and the vital alleviation of local fiscal stress. The “fiscal-space relief” channel is particularly significant, illustrating how the mobilization of market-based green finance may help reduce the budgetary burden on local municipalities, thereby creating fiscal bandwidth for long-term environmental and stability investments.

Finally, the policy demonstrates substantial inclusive benefits. Heterogeneity evaluations utilizing the TextCNN digital-economy classification reveal that green finance interventions yield more pronounced marginal improvements in cities with weaker digital infrastructures. This highlights the capacity to partially offset digital disadvantages and foster equitable sustainability governance across diverse regional contexts.

7.2. Policy Recommendations

Strategically Expand Pilot Coverage: Policymakers should progressively broaden the scope of green finance zones. Future frameworks must transition from merely funding end-of-pipe pollution reduction toward comprehensive sustainability-risk governance, effectively integrating carbon-reduction infrastructure, industrial upgrading, and proactive transition-risk prevention.

Target Fiscal-Space Relief Explicitly: Relieving local budgetary stress should be established as a direct policy objective. Local governments must leverage market-driven green credit and public-private funds to ease traditional municipal expenditures, purposefully redirecting these released funds toward essential public services, digital infrastructure, and foundational research.

Provide Tailored Support for Digitally Lagging Cities: Customized support mechanisms are necessary for regions trailing in digital development. Higher-level authorities should deploy targeted credit guidance, risk compensation, and digital infrastructure investments to enhance the absorptive capacity of lagging cities, ensuring green finance benefits are not exclusively monopolized by already-advanced urban centers.

Enhance Data Integration and Cross-Departmental Coordination: The development of a high-frequency monitoring platform that synthesizes green finance flows, carbon emissions, fiscal budgets, and textual policy narratives is crucial. Such an integrated system would significantly bolster predictive capabilities, early warning mechanisms, and the real-time evaluation of green finance implementation.

The progressive time effects and digital catch-up mechanisms observed in China's pilots offer a valuable blueprint for other non-Chinese Asian economies and global developing regions. Policymakers in rapidly urbanizing emerging markets—which often face similar dual pressures of municipal budget constraints and environmental mandates—can leverage market-based green finance to simultaneously alleviate fiscal stress and bridge local digital divides. The delayed but persistent dividends observed over time suggest that international governance frameworks must maintain structural patience and consistency when adopting similar sustainable finance instruments.

7.3. Limitations and Future Research

This study acknowledges certain limitations that pave the way for future scholarly inquiry. First, while municipal work reports are highly authoritative, regional variations in writing style exist. Future studies could enrich the text-based risk-monitoring framework by incorporating diverse corpora, such as budget documents, procurement records, and financial institution disclosures. Second, subsequent research could systematically compare the efficacy of alternative machine learning and natural language processing architectures against the TextCNN and Word2Vec models employed in this paper. Finally, as this analysis centers on localized city-level policy impacts, future investigations could map broader intercity financial networks to evaluate how green finance pilots mitigate systemic risks and generate positive spillovers across regional sustainability networks.

Author Contributions: Conceptualization, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); methodology, Z.T. and Y.W. (Yangxin Wang); software, Z.T. and Y.W. (Yangxin Wang); validation, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); formal analysis, Z.T. and Y.W. (Yangxin Wang); investigation, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); resources, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); data curation, Z.T. and Y.W. (Yangxin Wang); writing—original draft preparation, Z.T. and Y.W. (Yangxin Wang); writing—review and editing, Z.T. and Y.W. (Yangxin Wang); visualization, Z.T. and Y.W. (Yangxin Wang); supervision, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); project administration, Z.T., Y.W. (Yangxin Wang), and Y.W. (Yuanyuan Wang); funding acquisition, Z.T. and Y.W. (Yangxin Wang). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The supporting dataset and code for this study can be downloaded at <https://zenodo.org/records/20773076> (accessed on 7 July 2026). The original raw data were manually collected and organized from the CSMAR website (Commercial database).

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

Abbreviations

Word2Vec	Word to Vector
TextCNN	Text Convolutional Neural Network
NLP	Natural Language Processing
DID	Difference-in-Differences
TWFE	Two-Way Fixed-Effects
CS-DID	Callaway and Sant'Anna Difference-in-Differences
PSM	Propensity Score Matching
ATT	Average Treatment Effect on the Treated

VIF	Variance Inflation Factor
NQP (Y1_NQP)	New Quality Productive forces/Urban sustainable productive forces
NQPA (Y2_NQPA)	Government attention to sustainable productive forces
GDP	Gross Domestic Product

Appendix A

Table A1. Indicator system and EWM weights for urban sustainable productive forces (Y1_NQP).

Target Layer	Primary Dimensions	Secondary Sub-Indicators	Nature	Data Source
Urban Sustainable Productive Forces (Y1_NQP)	Strategic Emerging Industries	Number of high-tech enterprise registrations per 10,000 firms	Positive	CSMAR Database
		Output value of strategic emerging sectors/Regional GDP	Positive	Municipal Yearbooks
		Number of patents granted per 10,000 people	Positive	CNRDS Database
	Green Transformation	Industrial sulfur dioxide (SO ₂) emission intensity	Negative	China City Statistical Yearbook
		Rate of comprehensive utilization of industrial solid waste	Positive	China City Statistical Yearbook
		Energy consumption per unit of regional GDP	Negative	China City Statistical Yearbook
	Digital Infrastructure	Number of mobile phone users per 100 persons	Positive	China City Statistical Yearbook
		Total volume of telecommunication services per capita	Positive	China City Statistical Yearbook
		Peking University Digital Financial Inclusion Index	Positive	Peking University DFII

Table A2. Hyperparameter configuration of the TextCNN prediction model.

Hyperparameter Category	Parameter Name	Value/Setting	Description
Input Layer	Word Embedding Dimension (d)	200	Pre-trained via Word2Vec (Skip-gram)
	Sequence Length (n)	Dynamic/Max Padded	Aligned with maximum document length
Convolutional Layer	Filter Window Sizes (h)	3, 4, 5	To capture 3-g, 4-g, and 5-g features
	Number of Filters per Size	100	Total 300 feature maps generated
	Activation Function	ReLU	Rectified Linear Unit for nonlinearity
Pooling Layer	Pooling Strategy	1D Max-over-time	Extracts the most salient features
Fully Connected Layer	Dropout Rate	0.5	Prevents model overfitting
	Output Activation	Softmax	Generates binary class probabilities
Optimization	Optimizer	Adam	Adaptive moment estimation
	Learning Rate	0.001	Initial step size
	Batch Size	64	Number of samples per gradient update
	Max Training Epochs	100	With early stopping mechanism

Table A3. Dynamic event-study regression coefficients (underlying data for Figure 7).

Event Time (Relative Year)	(1) Y1: NQP	(2) Y2: NQPA
Pre_5 ($k \leq -5$)	0.0009 (0.0020)	−0.0114 (0.0287)
Pre_4 ($k = -4$)	0.0001 (0.0026)	0.0009 (0.0393)
Pre_3 ($k = -3$)	0.0006 (0.0025)	−0.0251 (0.0390)
Pre_2 ($k = -2$)	−0.0018 (0.0029)	−0.0203 (0.0349)
Pre_1 ($k = -1$)	Baseline (0)	Baseline (0)
Post_0 ($k = 0$)	0.0694 *** (0.0022)	0.3248 *** (0.0547)
Post_1 ($k = 1$)	0.0751 *** (0.0029)	0.4209 *** (0.0524)
Post_2 ($k = 2$)	0.0773 *** (0.0027)	0.3605 *** (0.0494)
Post_3 ($k = 3$)	0.0791 *** (0.0030)	0.4764 *** (0.0491)
Post_4 ($k = 4$)	0.0801 *** (0.0034)	0.4751 *** (0.0421)
Post_5 ($k \geq 5$)	0.0715 *** (0.0034)	0.5361 *** (0.0517)
City FE/Year FE	Yes/Yes	Yes/Yes
Controls	Yes	Yes
Observations	3588	3588
Adjusted R-squared	0.8689	0.8313

Note: Robust standard errors clustered at the city level are reported in parentheses. *** denotes 1% level. The year immediately prior to the green finance policy implementation ($k = -1$) is omitted as the benchmark group.

Table A4. Operational details of robustness checks.

Robustness Check	Operational Specification	Sample/Data Adjustment	Core Rationale
(1) Alternative Variables	Hard Capacity (Y1_PCA): 1st principal component (PCA) of green/carbon/digital sub-metrics. Attention (Y2_Density): $\frac{\text{Keywords}_{i,t}}{\text{TotalWords}_{i,t}} \times 1000$ Excludes 4	Replaces baseline entropy-weight index and raw log counts.	Rules out sensitivity to index weighting algorithms and document-length variations.
(2) Exclude Municipalities	direct-administered municipalities (Beijing, Shanghai, Tianjin, Chongqing).	N reduced from 281 to 277 cities (2810 to 2770 obs.).	Eliminates structural outliers with provincial-level fiscal and administrative privileges.
(3) Exclude Macro Shocks	Excludes sample observations from years 2015 and 2020.	Removes 2 crisis years across all municipalities.	Eliminates transient distortions from the 2015 stock market crash and 2020 COVID-19 shock.
(4) Lagged Policy	One-year lag of treatment: $Y_{i,t} = \alpha + \beta \text{Pilot}_{i,t-1} + X'_{i,t} \delta + \mu_i + \lambda_t + \varepsilon_{i,t}$	Policy status ($\text{Pilot}_{i,t-1}$) activated $t + 1$ after pilot designation.	Accounts for institutional credit-disbursement delays; mitigates reverse causality.
(5) 5% Winsorization	Two-sided trimming at 5th and 95th percentiles for all continuous variables.	Replaces baseline 1% winsorization thresholds.	Tests robustness against extreme distributional tails and reporting anomalies.

References

- Xu, K.; Zhao, P. Does green finance promote green total factor productivity? Empirical evidence from China. *Sustainability* **2023**, *15*, 11204. [\[CrossRef\]](#)
- Zhang, Z.; Wang, J.; Feng, C.; Chen, X. Do pilot zones for green finance reform and innovation promote energy savings? Evidence from China. *Energy Econ.* **2023**, *124*, 106763. [\[CrossRef\]](#)
- Liu, C.; Xiong, M. Green finance reform and corporate innovation: Evidence from China. *Financ. Res. Lett.* **2022**, *48*, 102993. [\[CrossRef\]](#)
- Wen, H.; Lee, C.-C.; Zhou, F. Green credit policy, credit allocation efficiency and upgrade of energy-intensive enterprises. *Energy Econ.* **2021**, *94*, 105099. [\[CrossRef\]](#)

5. Lei, H.; Gao, R.; Ning, C.; Sun, G. Green finance and corporate green innovation. *Financ. Res. Lett.* **2025**, *72*, 106577. [\[CrossRef\]](#)
6. Zhang, T. Can green finance policies affect corporate financing? Evidence from China's green finance innovation and reform pilot zones. *J. Clean. Prod.* **2023**, *419*, 138289. [\[CrossRef\]](#)
7. Yu, C.H.; Wu, X.; Zhang, D.; Chen, S.; Zhao, J. Demand for green finance: Resolving financing constraints on green innovation in China. *Energy Policy* **2021**, *153*, 112255. [\[CrossRef\]](#)
8. Lv, C.; Fan, J.; Lee, C.-C. Can green credit policies improve corporate green production efficiency? *J. Clean. Prod.* **2023**, *397*, 136573. [\[CrossRef\]](#)
9. Irfan, M.; Razzaq, A.; Sharif, A.; Yang, X. Influence mechanism between green finance and green innovation: Exploring regional policy intervention effects in China. *Technol. Forecast. Soc. Change* **2022**, *182*, 121882. [\[CrossRef\]](#)
10. Shi, R.; Zhou, H.; Li, X.; Miao, X.; Zhao, X. Green finance, investor preferences, and corporate green innovation. *Financ. Res. Lett.* **2024**, *66*, 105676. [\[CrossRef\]](#)
11. Lin, B.; Ma, R. How does digital finance influence green technology innovation in China? Evidence from the financing constraints perspective. *J. Environ. Manag.* **2022**, *320*, 115833. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Kong, H.; Sun, L.; Zhang, W. Digitization and green technology innovation of Chinese firms under government subsidy policies. *Systems* **2024**, *12*, 447. [\[CrossRef\]](#)
13. Ma, D.; Zhu, Q. Innovation in emerging economies: Research on the digital economy driving high-quality green development. *J. Bus. Res.* **2022**, *145*, 801–813. [\[CrossRef\]](#)
14. Wang, K.-L.; Sun, T.-T.; Xu, R.-Y.; Miao, Z.; Cheng, Y.-H. How does internet development promote urban green innovation efficiency? Evidence from China. *Technol. Forecast. Soc. Change* **2022**, *184*, 122017. [\[CrossRef\]](#)
15. Hao, X.; Wang, X.; Wu, H.; Hao, Y. Path to sustainable development: Does digital economy matter in manufacturing green total factor productivity? *Sustain. Dev.* **2023**, *31*, 360–378. [\[CrossRef\]](#)
16. Li, X.; Shao, X.; Chang, T.; Albu, L.L. Does digital finance promote the green innovation of China's listed companies? *Energy Econ.* **2022**, *114*, 106254. [\[CrossRef\]](#)
17. Zhu, Z.; Hua, Q.; Xu, S.; Lin, W. The mechanism of green finance in promoting China's new quality productive forces: Technological innovation and data factor. *Res. Int. Bus. Financ.* **2025**, *79*, 103038. [\[CrossRef\]](#)
18. Li, Z.; Liu, J. Low-carbon transition, financial constraints and enterprises' new quality productive forces. *Financ. Res. Lett.* **2025**, *85*, 108106. [\[CrossRef\]](#)
19. Zhang, J.; Wang, X.; Zhang, W.; Wang, L. Can the carbon emissions trading policy promote the development of new quality productive forces in manufacturing enterprises? *Financ. Res. Lett.* **2025**, *83*, 107617. [\[CrossRef\]](#)
20. Liu, Y.; He, Z. Synergistic industrial agglomeration, new quality productive forces and high-quality development of the manufacturing industry. *Int. Rev. Econ. Financ.* **2024**, *94*, 103373. [\[CrossRef\]](#)
21. Fu, Q.; Zhao, S. Influence of green finance on new-quality productivity of enterprises: Evidence from Chinese A-share listed companies. *Int. Rev. Financ. Anal.* **2025**, *105*, 104356. [\[CrossRef\]](#)
22. Jia, J.; Liu, Y.; Martinez-Vazquez, J.; Zhang, K. Vertical fiscal imbalance and local fiscal indiscipline: Empirical evidence from China. *Eur. J. Polit. Econ.* **2021**, *68*, 101992. [\[CrossRef\]](#)
23. Yang, Z. Effects of vertical fiscal imbalance on green total factor productivity—Evidence from China. *Sustainability* **2023**, *15*, 8768. [\[CrossRef\]](#)
24. Nie, S.; Zeng, G.; Zhang, H.; Ji, J. The local government fiscal pressure's effect on green total factor productivity: Exploring mechanisms from the perspective of government behavior. *Int. Rev. Econ. Financ.* **2024**, *96*, 103702. [\[CrossRef\]](#)
25. Zhou, K.; Qu, Z.; Wei, Z.; Zhao, J. Does government fiscal pressure matter for firm environmental performance? The role of environmental regulation and tax competition. *Econ. Anal. Policy* **2023**, *80*, 1187–1204. [\[CrossRef\]](#)
26. Li, T.; Qi, F. Fiscal reform and green innovation: Evidence from China's local debt management reform. *Financ. Res. Lett.* **2025**, *76*, 106937. [\[CrossRef\]](#)
27. Callaway, B.; Sant'Anna, P.H.C. Difference-in-differences with multiple time periods. *J. Econom.* **2021**, *225*, 200–230. [\[CrossRef\]](#)
28. Sant'Anna, P.H.C.; Zhao, J. Doubly robust difference-in-differences estimators. *J. Econom.* **2020**, *219*, 101–122. [\[CrossRef\]](#)
29. Goodman-Bacon, A. Difference-in-differences with variation in treatment timing. *J. Econom.* **2021**, *225*, 254–277. [\[CrossRef\]](#)
30. Sun, L.; Abraham, S. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econom.* **2021**, *225*, 175–199. [\[CrossRef\]](#)
31. De Chaisemartin, C.; D'Haultfoeuille, X. Two-way fixed effects estimators with heterogeneous treatment effects. *Am. Econ. Rev.* **2020**, *110*, 2964–2996. [\[CrossRef\]](#)
32. Roth, J.; Sant'Anna, P.H.C.; Bilinski, A.; Poe, J. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *J. Econom.* **2023**, *235*, 2218–2244. [\[CrossRef\]](#)
33. Borusyak, K.; Jaravel, X.; Spiess, J. Revisiting event-study designs: Robust and efficient estimation. *Rev. Econ. Stud.* **2024**, *91*, 3253–3285. [\[CrossRef\]](#)

34. Athey, S.; Imbens, G.W. Design-based analysis in difference-in-differences settings with staggered adoption. *J. Econom.* **2022**, *226*, 62–79. [[CrossRef](#)]
35. Rosenbaum, P.R.; Rubin, D.B. The central role of the propensity score in observational studies for causal effects. *Biometrika* **1983**, *70*, 41–55. [[CrossRef](#)]
36. Gentzkow, M.; Kelly, B.; Taddy, M. Text as data. *J. Econ. Lit.* **2019**, *57*, 535–574. [[CrossRef](#)]
37. Baker, S.R.; Bloom, N.; Davis, S.J. Measuring economic policy uncertainty. *Q. J. Econ.* **2016**, *131*, 1593–1636. [[CrossRef](#)]
38. Hansen, S.; McMahon, M.; Prat, A. Transparency and deliberation within the FOMC: A computational linguistics approach. *Q. J. Econ.* **2018**, *133*, 801–870. [[CrossRef](#)]
39. Bybee, L.; Kelly, B.; Manela, A.; Xiu, D. Business news and business cycles. *J. Financ.* **2024**, *79*, 3105–3147. [[CrossRef](#)]
40. Kim, Y. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1746–1751. [[CrossRef](#)]
41. Luo, S.; Yimamu, N.; Li, Y.; Wu, H.; Irfan, M.; Hao, Y. Digitalization and sustainable development: How could digital economy development improve green innovation in China? *Bus. Strateg. Environ.* **2023**, *32*, 1847–1871. [[CrossRef](#)]
42. Zhang, J.; Lyu, Y.; Li, Y.; Geng, Y. Digital economy: An innovation driving factor for low-carbon development. *Environ. Impact Assess. Rev.* **2022**, *96*, 106821. [[CrossRef](#)]
43. Gorodnichenko, Y.; Pham, T.; Talavera, O. The voice of monetary policy. *Am. Econ. Rev.* **2023**, *113*, 548–584. [[CrossRef](#)]
44. Abadie, A. Using synthetic controls: Feasibility, data requirements, and methodological aspects. *J. Econ. Lit.* **2021**, *59*, 391–425. [[CrossRef](#)]
45. Arner, D.W.; Buckley, R.P.; Zetsche, D.A.; Veidt, R. Sustainability, FinTech and financial inclusion. *Eur. Bus. Organ. Law. Rev.* **2020**, *21*, 7–35. [[CrossRef](#)]
46. Agarwal, S.; Zhang, J. FinTech, lending and payment innovation: A review. *Asia-Pac. J. Financ. Stud.* **2020**, *49*, 353–367. [[CrossRef](#)]

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