1 ADDITIONAL EMPIRICAL RESULTS

1.1 Comprehensive Robustness Checks

Table 1 presents an extensive battery of robustness checks designed to address potential concerns about our identification strategy, sample selection, and variable construction.

Table 1. Comprehensive Robustness Analysis

Robustness Check:	(1) Balanced Panel	(2) Winsorized Variables	(3) Bootstrap SE	(4) SIMEX Correction	(5) Alternative Green ID	(6) Spatial Lag	(7) Temporal Subsample
Green Credit	0.016***	0.012**	0.013**	0.015***	0.011**	0.014***	0.013**
	(3.21)	(2.47)	(2.54)	(2.81)	(2.31)	(2.94)	(2.67)
Bootstrap p-value			0.008				
SIMEX-corrected coef				0.017			
Alternative ID match rate					87.3%		
Observations	16,848	23,674	23,674	23,674	21,947	22,106	15,729
R-squared	0.194	0.184	0.187	0.187	0.179	0.191	0.201
Sample Period	2013-2023	2013-2023	2013-2023	2013-2023	2013-2023	2014-2023	2015-2020
Special Features	Full years	1%/99%	wild-cluster	Meas error	Bank reports	Spatial FE	Pre-reform

Notes: Robustness checks address various potential concerns. Column (1) requires firms present in all years. Column (2) winsorizes continuous variables at 1%/99% levels. Column (3) uses wild-cluster bootstrap with 999 replications. Column (4) implements SIMEX correction for green credit measurement error. Column (5) uses alternative identification based on bank disclosure data. Column (6) includes spatial fixed effects. Column (7) restricts to pre-reform period. All specifications include full controls and standard fixed effects except where noted. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The robustness analysis confirms that our core findings are not driven by sample selection, outliers, or measurement concerns. The balanced panel specification in Column (1) requires firms to appear in all sample years, addressing concerns that entry and exit dynamics might bias our results. The coefficient remains highly significant and economically meaningful despite the reduced sample size.

Winsorization of continuous variables at the 1% and 99% levels in Column (2) addresses outlier concerns while maintaining the full sample. wild-cluster bootstrap standard errors in Column (3) provide conservative inference when the number of clusters may be too small for asymptotic approximations. The SIMEX correction in Column (4) directly addresses measurement error in our green credit identification, suggesting that our main estimates may actually understate the true effects.

Column (5) employs an alternative green credit identification strategy based solely on bank disclosure data, achieving an 87.3% match rate with our primary identification while maintaining statistical significance. Spatial fixed effects in Column (6) control for unobserved geographic clustering, while the pre-reform subsample in Column (7) confirms that our effects are not driven by post-2018 changes in environmental regulation.

1.2 Extended Heterogeneity Analysis

Table 2 provides additional heterogeneity analysis across dimensions relevant for policy design and external validity.

Table 2. Extended Heterogeneity Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample Split:	Bank Ownership		Media Attention		Analyst Coverage		Bond Rating	
	State	Non-State	High	Low	High	Low	Rated	Unrated
Green Credit	0.008	0.021***	0.006	0.024***	0.009*	0.019***	0.007	0.018***
	(1.34)	(3.47)	(0.94)	(3.84)	(1.67)	(3.21)	(1.28)	(2.94)
Sample Split:	Green Bond Issue		Carbon Intensive		Export Intensity		Innovation	
	Yes	No	High	Low	High	Low	High	Low
Green Credit	0.003	0.016***	0.022***	0.008*	0.011**	0.017***	0.010*	0.019***
	(0.47)	(3.14)	(3.27)	(1.74)	(2.14)	(2.87)	(1.81)	(3.08)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Extended heterogeneity analysis across additional firm characteristics. Bank ownership distinguishes loans from state-owned versus private banks. Media attention based on environmental news coverage. Analyst coverage from financial database. Bond ratings from credit rating agencies. Green bond issuers identified from bond prospectuses. Carbon intensity based on scope 1+2 emissions per unit revenue. Export intensity from customs data. Innovation measured by R&D intensity and patent stocks. Sample splits at median except for bond ratings and green bonds. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The extended heterogeneity analysis reveals several additional patterns consistent with regulatory arbitrage. Firms borrowing from non-state banks show larger effects, suggesting that private lenders may exercise less oversight over green credit usage. Companies with low media attention and analyst coverage exhibit stronger effects, consistent with reduced external monitoring enabling opportunistic behavior.

Firms with bond ratings show minimal green credit effects, while unrated firms drive most of our results. This pattern suggests that credit market discipline constrains regulatory arbitrage when firms depend on external capital markets. Similarly, firms that have issued green bonds show no significant effects, possibly because green bond frameworks typically impose stricter monitoring requirements than green loans.

Carbon-intensive industries show the largest effects, consistent with these firms having the most to gain from regulatory leniency. Export-intensive firms show smaller effects, possibly due to international supply chain pressures for environmental compliance. High-innovation firms also show reduced effects, suggesting that genuine technological capacity may correlate with legitimate green finance usage.

1.3 Temporal Dynamics and Policy Learning

Table 3 examines how green credit effects evolve over our sample period, providing insights into policy learning and adaptation.

Table 3. Temporal Evolution of Green Credit Effects

Period:	2013-2015 Early	2016-2017 Pre-Reform	2018-2019 Reform	2020-2021 post-Reform	2022-2023 Recent
Green Credit	0.032***	0.029***	0.018***	0.009*	0.006
	(4.21)	(3.84)	(2.94)	(1.74)	(1.28)
Green Credit × High Governance	-0.018**	-0.016**	-0.011*	-0.005	-0.002
	(-2.31)	(-2.14)	(-1.67)	(-0.84)	(-0.41)
Observations	6,847	4,729	4,918	4,736	2,444
R-squared	0.156	0.172	0.198	0.214	0.267
Mean Green Credit Rate	0.21	0.28	0.35	0.42	0.47
Mean Violation Rate	0.093	0.087	0.074	0.065	0.059
Trend Analysis:					
Effect Magnitude	Declining	Declining	Sharp Drop	Stabilizing	Minimal
Policy Learning	Limited	Limited	Active	Adapting	Effective

The temporal analysis reveals a clear learning curve in green finance policy implementation. Early period effects (2013-2015) are largest, suggesting that initial policy design created substantial arbitrage opportunities. Effects remain elevated through 2017 but begin declining with the 2018 environmental tax reform. By 2020-2023, effects become statistically and economically small, indicating successful policy adaptation.

The interaction with governance quality shows that learning occurred faster in well-governed jurisdictions, where institutional capacity enabled more rapid policy adjustment. This pattern suggests that successful green finance implementation requires not only good initial design but also adaptive capacity to respond to emerging challenges.

2 INTERNATIONAL CONTEXT AND POLICY FRAMEWORK

2.1 Comparative Analysis with International Green Finance Programs

While comprehensive firm-level data comparable to ours remains unavailable for most countries, aggregate evidence from international green finance programs provides valuable context for assessing the external validity of our findings. European green finance initiatives exhibit patterns broadly consistent with our results, with studies by the European Securities and Markets Authority documenting substantial variation in environmental outcomes across ostensibly similar programs, closely correlated with institutional quality and regulatory coordination. The United States presents an interesting contrast through its tax credit approach, where state-level implementation creates opportunities for regulatory arbitrage between federal tax authorities and state environmental agencies. Japan's voluntary approach through self-reporting has limited opportunities for regulatory arbitrage but also reduced program scale and environmental impact, highlighting the trade-off between program ambition and implementation challenges.

2.2 Policy Framework for Sustainable Finance Reform

Based on our empirical findings and international experience, we propose three core reforms for sustainable finance effectiveness. First, the transition from identity-based to performance-based allocation represents the most fundamental change needed globally. Performance-based systems tie financial benefits directly to measurable environmental outcomes through clawback provisions, third-party verification, and real-time monitoring, aligning private incentives with policy objectives. Pilot programs in Singapore and the Netherlands demonstrate this viability by tying green loan rates to achieved environmental improvements.

Second, regulatory coordination emerges as equally critical. Essential coordination mechanisms include unified compliance databases accessible to all relevant agencies, joint performance metrics preventing conflicting objectives, and ”no forbearance” rules preventing special treatment based on participation in other agencies' programs. The European Union's sustainable finance taxonomy represents an ambitious coordination attempt, though implementation challenges remain.

Third, enhanced monitoring and enforcement require capabilities matching the sophistication of potential gaming strategies. Modern systems should incorporate real-time data collection through IoT sensors and satellite monitoring, AI-powered analysis for greenwashing detection, and citizen reporting mechanisms. China's national environmental monitoring network and the European Space Agency's Copernicus program exemplify technological integration into environmental verification.

Successful implementation requires coordinated action prioritizing jurisdictions with strong institutional capacity as proof-of-concept models, followed by scaling through international standards and technical assistance for developing economies. The stakes have never been higher: with over $4 trillion in annual green finance flows anticipated by 2030, the choice between performance-based and identity-based systems will fundamentally shape finance's role in addressing climate change.

3 DATA APPENDIX

3.1 Environmental Violation Data Collection Architecture

Our environmental violation database construction represents one of the most comprehensive efforts to systematically collect corporate environmental compliance data in China. The technical architecture begins with a comprehensive website identification process that leverages the Ministry of Ecology and Environment's official directory, provincial environmental department link structures, and government portal hierarchies to ensure complete coverage of all prefecture-level environmental bureaus. This systematic approach yields 387 active bureau websites covering all 337 prefecture-level cities, with some jurisdictions maintaining multiple specialized bureaus for different environmental domains such as water pollution control, air quality management, and solid waste regulation.

The web scraping system employs a multi-layered technical approach designed to handle the heterogeneous nature of Chinese government websites. Our Python-based framework integrates BeautifulSoup for standard HTML parsing with Selenium WebDriver for JavaScript-rendered content, which proves essential given that many environmental bureaus have modernized their disclosure platforms using dynamic web technologies. For PDF documents containing penalty announcements, we implement optical character recognition using Tesseract OCR with Chinese language optimization, followed by regular expression-based structured data extraction that identifies key fields including company identifiers, violation dates, penalty amounts, and corrective measures. Quality control mechanisms include cross-validation procedures that compare overlapping records between city and provincial databases, completeness verification against aggregate statistics published in government environmental reports, and manual verification through research assistant review of randomly selected penalty records.

3.2 Green Credit Identification Through Natural Language Processing

The identification of green credit at the firm level presents unique methodological challenges given the absence of standardized reporting formats for environmental lending in Chinese corporate disclosures. Our approach addresses this challenge through a sophisticated natural language processing pipeline that analyzes multiple data sources simultaneously. We collect loan information from three primary sources: annual reports downloaded directly from exchange websites for all A-share listed firms covering 2010-2023, loan-related announcements collected from Shanghai and

Shenzhen stock exchange disclosure systems totaling 157,843 individual announcements, and green credit statistics gathered from social responsibility reports published by 42 major Chinese banks that participate in green finance programs.

The textual analysis algorithm employs a multi-stage identification process optimized for both Chinese and English financial terminology. Initial keyword searches identify loan-related sections using financial terms including loan, borrowing, financing, and credit in both languages, alongside their contextual variations. Within identified loan sections, our algorithm searches for green credit markers including purpose indicators such as green, environmentally friendly, energy-saving, emission reduction, clean energy, and circular economy, as well as explicit certification mentions referencing green credit certification programs and detailed environmental project descriptions that specify pollution control or sustainability objectives.

Context validation employs natural language processing techniques that require green keywords to appear within 50 characters of loan amount specifications, exclude negations indicating non-environmental purposes, and verify that project descriptions match established green finance categories published by the People's Bank of China. Our validation process achieves 91.8% accuracy through manual verification of 500 randomly selected firm-years, with false positive rates of 3.2% for regular loans misclassified as green and false negative rates of 5.0% for missed green loans, providing confidence in our identification methodology.

3.3 Variable Construction and Data Integration

Our dependent variables capture multiple dimensions of environmental compliance behavior using penalty data systematically extracted from government enforcement bulletins. The primary violation indicator represents a binary measure equaling one if firm i received any environmental penalty in year t, while secondary measures include the natural logarithm of total penalty amounts measured in thousand RMB, violation counts representing the number of separate infractions, and categorical indicators for serious violations exceeding 500,000 RMB thresholds and violation types spanning water, air, solid waste, and procedural infractions.

The key independent variable measuring green-credit exposure employs a binary indicator equaling one if firms maintain outstanding green credit in year t, determined through identification in either annual reports or loan announcements and assuming multi-year loans remain outstanding until contractual maturity. Control variables draw from multiple data sources to ensure comprehensive coverage of firm characteristics that might influence both green credit access and environmental compliance. Firm-level financial data sourced from CSMAR and Wind databases include logarithmic total assets, leverage ratios calculated as total debt divided by total assets, return on assets measured as net income divided by total assets, annual sales growth rates, firm age calculated as years since establishment, and state ownership indicators for firms with government ownership exceeding thirty percent.

Environmental investment proxies represent a particularly important methodological innovation, constructed by identifying environmental protection assets from fixed asset schedules that mention wastewater treatment facilities, exhaust gas treatment equipment, and environmental protection machinery. This variable, available for approximately 72% of firm-years, provides crucial controls for genuine environmental commitment that help distinguish regulatory arbitrage from resource misallocation mechanisms. City-level variables drawn from statistical yearbooks include GDP per capita, fiscal pressure measured as expenditure-to-revenue ratios, air quality indicators based on annual average PM2.5 concentrations from monitoring stations, and government environmental expenditure ratios calculated as environmental spending divided by total municipal expenditure, ensuring that our analysis accounts for local economic conditions and regulatory capacity that might influence both green credit allocation and environmental enforcement patterns.

4 ROBUSTNESS TEST APPENDIX

Table 4. Placebo and Falsification Tests for the IV Strategy

Test / Instrument:	First Stage Green Credit	Reduced Form Violation	IV–2SLS Violation	N
Panel A: Lead(2) Placebo Instrument				
Lead-2 Guidelines×Eligible	0.009	0.003	–	23,674
	(0.41)	(0.48)		
First-stage F	0.17			
Kleibergen–Paap F	0.19			
Panel B: Never-Eligible Sample				
Guidelines×Eligible	0.015	0.002	–	8,731
	(0.68)	(0.36)		
First-stage F	0.46			
Kleibergen–Paap F	0.44			
Panel C: SOE Subsample				
Guidelines×Eligible	0.052			9,800
	(1.21)			
Green Credit (fitted)			0.008	
			(0.29)	
K-P F / A-R p-value	1.19		0.62	
Panel D: Permutation Test (500 reassignments)				
Mean placebo $\hat{\beta}_{2SLS}$			0.001	
SD placebo $\hat{\beta}_{2SLS}$			0.013	
Empirical p-value			0.008	

Notes: Panel A uses a two-year lead of the policy$\times$eligibility interaction as a placebo instrument. Panel B restricts to firms predicted never eligible before the policy; Panel C limits to SOEs. Panel D reassigns eligibility within industry–year cells 500 times to form a placebo distribution of 2SLS coefficients. Dashes indicate not estimated or not applicable. The very weak first stages in Panels A–C and the extreme-tail placement of the actual 2SLS effect (about 0.098) in Panel D are consistent with instrument exogeneity and the exclusion restriction.

Table 5. Placebo Difference-in-Differences with Fake Reform Years

Dependent Variable: Violation (0/1)	(1) Post-2016	(2) Post-2017	N
Green Credit × Post(fake)	0.003	-0.002	23,674
	(0.54)	(-0.41)	
Green Credit	0.024***	0.022***	
	(3.68)	(3.47)	
Post(fake)	0.006	0.008	
	(1.11)	(1.34)	
Controls; Firm FE	Yes	Yes	
Industry×Year FE; City×Year FE	Yes	Yes	
R^2	0.298	0.301	

Notes: We intentionally mis-assign the reform to 2016 (col. 1) or 2017 (col. 2). The interaction Green Credit×Post(fake) is insignificant in both cases, consistent with no pre-trend treatment effect. Two-way clustered (firm and year) standard errors; t-statistics in parentheses. Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 6. Balance and Pre-trend Checks with Respect to the Instrument

Dependent Var. at $t-1$	$\ln$(Assets)	Leverage	ROA	Env Asset Ratio	Violation$_{t-1}$
Guidelines×Eligible	0.003	0.002	-0.001	-0.004	0.001
	(0.58)	(0.44)	(-0.31)	(-0.49)	(0.19)
Observations	17,084	17,084	17,084	17,084	23,674
Fixed effects; Controls	Firm FE; industry×year and city×year FE; full controls				
Joint test (all five coeffs = 0)	$\chi^2(5) = 3.27, p = 0.66$				
R^2	0.71	0.64	0.59	0.42	0.36

Notes: Each column regresses a lagged covariate (or lagged violation) on the instrument, testing whether the instrument loads on pre-treatment observables or prior violations. Coefficients are near zero and statistically insignificant, supporting balance and no pre-trend correlation. Standard errors are two-way clustered by firm and year; t-statistics in parentheses.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

SY conceived and designed the study, developed the theoretical framework, collected and analyzed all data, performed the empirical analysis, and wrote the manuscript. The author takes full responsibility for the content of the work and agrees to be accountable for all aspects of the research.

FUNDING

This research received no external funding.

ACKNOWLEDGMENTS

The author acknowledges the valuable feedback received from anonymous reviewers and conference participants that helped improve the manuscript.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this study contain sensitive information about individual firms' environmental violations and cannot be made publicly available due to confidentiality restrictions and data protection considerations. The data were collected from publicly accessible government websites, and the methodology for data collection and processing is fully described in the manuscript to ensure reproducibility.