



Hui Zhai

School of Finance

Department of Actuarial Science

Southwestern University of Finance and Economics, Chengdu, China

42320001@smail.swufe.edu.cn

Environmental liability insurance and corporate ESG performance: Evidence from China

Accepted by Editor Ewa W. Ziembra | Received: April 28, 2025 | Revised: June 16, 2026; July 18, 2026; July 21, 2026; July 23, 2026 | Accepted: August 6, 2026 | Published: September 3, 2026.

© 2026 Author(s). This article is licensed under the [Creative Commons Attribution-NonCommercial 4.0 license](#)

Cite as: Zhai, H. (2026). Environmental liability insurance and corporate ESG performance: Evidence from China. *Journal of Economics and Management*, 48, 386-424.
<https://doi.org/10.22367/jem.2026.48.15>

Abstract

Aim/purpose – This study examines the role of market-based environmental instruments in corporate governance by assessing the association between Environmental Liability Insurance (ELI) and corporate environmental, social, and governance (ESG) performance in China's transitional emerging market.

Design/methodology/approach – Utilizing a comprehensive panel dataset of Chinese A-share listed firms from 2021 to 2024 (yielding 19,185 firm-year observations), the study employs a two-way fixed-effects (TWFE) model. To address observable selection concerns and assess model sensitivity, the study implements propensity score matching (PSM) and a placebo test with randomized treatment assignments.

Findings – The results document a positive short-term association between the base-line Insur indicator and corporate ESG performance, primarily in linear fixed-effects specifications. Channel analyses suggest that this association is consistent with two possible short-term channels: lower observed regulatory-risk exposure and lower agency costs, which serve as a proxy for internal governance efficiency. However, the ordered logit sensitivity check is statistically insignificant, indicating that the main finding is sensitive to treating ESG strictly as an ordinal outcome. Heterogeneity analyses provide suggestive evidence of variation across ownership and regional settings, while most subgroup differences are not statistically significant in formal interaction tests.

Research implications/limitations – This research suggests that ELI may be associated with governance-relevant monitoring functions and that the positive net association is more consistent with a monitoring interpretation than with an unmitigated moral-hazard interpretation. However, the observational design, short 2021–2024 window, low voluntary ELI penetration, and potential early-adopter selection limit causal interpretation and the evaluation of long-term environmental outcomes.

Originality/value/contribution – This study provides novel micro-level empirical evidence conceptualizing environmental insurance as a quasi-regulatory monitoring mechanism. It extends the literature on insurance-based governance by utilizing the post-2020 standardized ESG rating window to evaluate short-term organizational compliance responses, shifting the focus from broad aggregate green finance metrics to the micro-governance dynamics of risk-pricing instruments.

Keywords: Environmental liability insurance, ESG performance, moral hazard, external monitoring, corporate governance.

JEL Classification: G32, Q56, M14.

Article type: Original research.

1. Introduction

In recent years, environmental, social, and governance (ESG) performance has become an increasingly important metric for corporate sustainability, attracting heightened scrutiny from regulators, investors, and stakeholders globally. While a substantial literature has documented how formal regulatory mandates, green credit policies, and investor activism shape corporate ESG behavior, the role of market-based insurance instruments as external governance mechanisms remains comparatively underexplored. Most existing research on sustainable finance focuses on credit allocation or capital market instruments, leaving the micro-governance pathways of property and liability insurance – specifically, how insurers may function as quasi-regulatory monitors in corporate risk management practices – insufficiently examined.

This gap is particularly relevant in China's evolving institutional landscape. Environmental liability insurance (ELI) – a specialized policy designed to cover financial damages from environmental pollution – has received increasing regulatory attention in China. In 2013, the Ministry of Environmental Protection and the China Insurance Regulatory Commission jointly issued the *Guidelines on Pilot Project of Compulsory Environmental Pollution Liability Insurance*, providing guidance for local governments to advance pilot projects in high-environmental-risk

industries, including heavy metal, petrochemical, hazardous chemical, and hazardous waste sectors (Ministry of Environmental Protection & China Insurance Regulatory Commission [MEP & CIRC], 2013). In 2018, the Ministry of Ecology and Environment reviewed and approved, in principle, the draft *Measures for the administration of compulsory environmental pollution liability insurance*, thereby indicating the government's intention to institutionalize environmental risk governance through insurance-based mechanisms (Ministry of Ecology and Environment of the People's Republic of China [MEE], 2018). More recently, the 2022 *Guidelines on green finance for the banking and insurance industries* required banking and insurance institutions to incorporate ESG risk requirements into management processes and comprehensive risk-management systems, and encouraged insurance institutions to provide risk-management services for environmental and climate-related activities (China Banking and Insurance Regulatory Commission [CBIRC], 2022). Against this regulatory background, the 2021–2024 empirical window provides a relevant setting to examine short-term associations between the baseline Insur indicator and corporate ESG performance under the recent green finance regulatory framework. We have focused on this period because China's ESG rating coverage and environmental penalty disclosure became more standardized after 2020, whereas earlier they were less so. In contrast, earlier data on voluntary ELI coverage and micro-level environmental penalties were disclosed less consistently. This design explicitly limits the analysis to short-term associations rather than long-term causal trajectories or durable, capital-intensive environmental outcomes.

Theoretically, the relationship between ELI and corporate ESG performance faces competing forces. Classic insurance economics suggests that risk transfer may induce moral hazard: by shielding firms from the full financial consequences of pollution, insurance may weaken firms' ex-ante incentives to invest in environmental protection (Arrow, 1963; Rothschild & Stiglitz, 1976). However, the institutional configuration of China's ELI market alters this calculus. Unlike standardized insurance lines, ELI contracts in China are tightly coupled with environmental administrative mandates. To mitigate their own exposure to catastrophic underwriting losses, China's ELI policy framework encourages insurers to participate in environmental risk assessments, risk management services, and coordination with environmental regulators (CBIRC, 2022; MEE, 2018; MEP & CIRC, 2013).

Therefore, whether the observed association between the baseline Insur indicator and ESG performance is more consistent with insurer-led monitoring or with unmitigated moral-hazard concerns remains an important empirical question.

While a nascent stream of literature has begun to explore the ELI-ESG nexus in China, existing studies have primarily focused on aggregate ESG outcomes under earlier regulatory regimes or relied on proxies such as green innovation and environmental disclosure quality (Chen & Xie, 2024; Wu et al., 2022). By contrast, this study constructs a firm-level panel dataset to examine the governance role of ELI using recent post-green-finance-policy data, examines whether the observed association is consistent with two potential channels: regulatory-risk and governance-efficiency, and evaluates possible boundary conditions across ownership structures, market environments, and regional institutional settings. Recognizing the endogeneity challenges inherent in voluntary insurance adoption, we have employed a two-way fixed-effects (TWFE) framework complemented by propensity score matching (PSM) and robustness checks to improve covariate comparability and reduce observable selection concerns. Our baseline results indicated that the Insur indicator is positively associated with corporate ESG performance, suggesting that insurer-related monitoring and pricing discipline may be associated with higher ESG performance in China's institutional context.

We further examined whether the baseline association is consistent with several potential short-term channels. Channel-related association tests suggest that the baseline Insur indicator is associated with lower observed environmental penalty exposure and lower agency costs. At the same time, we found no statistically significant association with environmental capital expenditures. Given the timing of the penalty measures and the observational research design, these results should be interpreted as associations consistent with possible regulatory-risk and governance-efficiency channels, rather than as conclusive evidence that ELI causally reduces penalties or produces substantive environmental investment. Heterogeneity analyses provide suggestive, but not uniformly statistically supported, evidence of variation across ownership, industry, and regional settings. Accordingly, we interpreted subgroup patterns based on formal interaction tests rather than split-sample significance alone.

Despite growing interest in ESG determinants, little empirical evidence exists on how insurance-based governance mechanisms may operate as quasi-regulatory institutions in emerging markets. This study contributes to the literature in several specific ways.

First, theoretically, we reconceptualized ELI as a quasi-regulatory monitoring mechanism rather than a pure risk-transfer instrument. While prior studies on D&O insurance in China showed mixed results due to moral hazard concerns (Tang et al., 2024; Yuan et al., 2016), we argued that the institutional coupling of ELI with administrative mandates in China may make the observed positive net association more consistent with a monitoring interpretation than with an unmitigated moral-hazard interpretation.

Second, empirically, we advanced the ELI-ESG literature by addressing a potential mechanical-scoring concern in aggregate ESG ratings. Existing studies (e.g., Chen & Xie, 2024; Wu et al., 2022) primarily utilized earlier regulatory contexts and relied on proxy outcomes such as green innovation or aggregate ESG scores, where environmental penalties may be directly incorporated into rating outcomes. Using post-2020 standardized data and constructing an Adjusted Environmental Pillar score that excludes penalty-linked components, we assessed whether the positive association between ELI and ESG performance is driven solely by the incorporation of environmental penalties into ESG ratings.

Third, we examined two potential short-term channels: regulatory-risk exposure and internal governance efficiency. Unlike prior work that treated ESG as a monolithic outcome, our analysis evaluated whether the baseline Insur indicator is associated with lower observed penalty exposure and lower agency costs, while also showing no statistically significant association with environmental capital expenditures in the short run.

Finally, we explored whether regional enforcement asymmetry is associated with variation in the ELI-ESG relationship. The regional analysis provided suggestive evidence that the association may vary across institutional environments, broadly consistent with the institutional substitution hypothesis, subject to formal interaction tests and clearly defined regional classifications.

The remainder of this paper is structured as follows: Section 2 reviews the literature and develops hypotheses; Section 3 details the data and methodology; Section 4 presents empirical results; Section 5 discusses the findings; and Section 6 concludes with policy implications, limitations, and future research directions.

2. Literature review

2.1. ESG determinants and insurance as a governance mechanism

A growing body of literature has examined the determinants of corporate ESG performance. Prior studies have emphasized that regulatory pressure, investor preferences, and internal governance structures are closely associated with firms' environmental and social practices (Dyck et al., 2019; Friede et al., 2015; Serafeim, 2020). In particular, institutional investors' attention to climate risk has been linked to corporate sustainability and climate-risk management (Krueger et al., 2020).

However, most existing research has focused on command-and-control regulations or traditional financial instruments, such as green credit and sustainability-linked bonds. While a growing literature on Directors and Officers (D&O) liability insurance in China has examined managerial risk-taking and broader sustainability behaviors (Tang et al., 2024; Yuan et al., 2016), its findings do not directly translate to environmental-specific insurance contracts. ELI differs from general liability insurance because its policy framework links insurance participation with environmental risk assessment, differentiated pricing, accident-prevention services, and post-underwriting risk management (CBIRC, 2022; MEE, 2018; MEP & CIRC, 2013). Despite growing attention to ESG determinants, empirical evidence on insurance-based governance mechanisms as quasi-regulatory institutions in emerging markets such as China remains limited.

2.2. Insurance, moral hazard, and monitoring

The economic theory of insurance highlights a fundamental tension between risk transfer and incentive alignment. On the one hand, insurance coverage may induce moral hazard: by reducing the insured party's marginal cost of risky behavior, insurance can shield managers from the full financial consequences of pollution, thereby weakening ex-ante prevention incentives and potentially being associated with weaker environmental or ESG-related outcomes (Arrow, 1963; Chen et al., 2022; Rothschild & Stiglitz, 1976).

On the other hand, insurance may also function as an external monitoring mechanism. Insurers, motivated by loss minimization and long-term profitability, may assess risk, differentiate premiums, provide accident-prevention services,

and manage risk post-underwriting (Dionne, 2013; Holderness, 1990; Shavell, 1979). This theoretical tension may be especially relevant in institutional settings where environmental regulation, market-based insurance, and enforcement capacity interact. In China's context, ELI has been embedded in policy arrangements that link insurance participation with environmental risk assessment, differentiated pricing, and coordination between environmental regulators and insurance institutions (CBIRC, 2022; MEE, 2018; MEP & CIRC, 2013). These institutional features make the ELI-ESG relationship theoretically ambiguous. However, they also suggest that a positive association would be more consistent with insurer-related monitoring than with an unmitigated moral-hazard interpretation. Recognizing these competing forces, we hypothesized:

H1: The ELI coverage-and-following-year persistence indicator is positively associated with corporate ESG performance, a pattern more consistent with insurer-related monitoring than with an unmitigated moral-hazard interpretation.

2.3. Potential channels: Regulatory risk and corporate governance

If the observed positive association reflects insurer-related monitoring, it may operate through several short-term operational channels:

Channel 1: Regulatory Risk Exposure. Insurers, acting as professional risk assessors, may differentiate premiums and coverage terms based on firms' environmental risk profiles and compliance histories. Firms seeking favorable coverage terms may therefore have stronger incentives to improve compliance-related risk management. Empirically, this channel is examined through the association between the baseline Insur indicator and observed environmental penalty exposure. Because penalty measures may include penalties imposed before the verified ELI coverage year, lower penalty frequency or severity should be interpreted as evidence consistent with lower regulatory-risk exposure, rather than as conclusive evidence that ELI reduces subsequent penalties.

Channel 2: Internal Governance Efficiency. The underwriting and risk-management processes may require firms to disclose environmental information and strengthen internal compliance routines, which may reduce information asymmetry and alleviate agency problems between managers and shareholders (Jensen & Meckling, 1976). Such governance-related adjustments may redirect managerial attention toward procedural compliance and risk control. However, in the

short run, organizational inertia and financial constraints may limit immediate capital-intensive environmental investment, making governance and compliance adjustments more observable than substantive environmental capital expenditures.

Based on these potential channels, we proposed the following hypotheses:

H2a (Regulatory Risk Exposure): The ELI coverage-and-following-year persistence indicator is negatively associated with observed environmental penalty exposure.

H2b (Internal Governance Efficiency): The ELI coverage-and-following-year persistence indicator is negatively associated with agency costs.

2.4. Heterogeneity: Institutional and organizational contingencies

The association between the baseline Insur indicator and ESG performance may vary across firms with different organizational and institutional characteristics.

First, ownership structure and market environment may condition firms' responsiveness to insurance-based governance incentives. Non-state-owned enterprises may be more sensitive to market-based incentives and reputational signals, whereas administrative mandates may more directly shape state-owned enterprises. Similarly, firms in more competitive industries may face stronger market pressure to respond to external governance signals. Whether these patterns also differ across high-tech and non-high-tech industries remains an empirical question.

Second, regional institutional environments may also condition the ELI-ESG association. One possibility aligns with the institutional substitution hypothesis: in regions where public environmental enforcement capacity is relatively weaker or more fragmented, market-based insurance arrangements may partially complement formal regulation. A competing possibility is that weaker institutional environments may also reduce insurance contract enforceability and weaken insurer-led risk management. Therefore, whether the ELI-ESG association varies across regions remains an empirical question. In the empirical analysis, regional classifications should be defined using official standards, and the treatment of Central and Northeastern China should be stated explicitly.

Accordingly, we hypothesized:

H3: The association between the baseline Insur indicator and corporate ESG performance varies across ownership structure, market competition, industry characteristics, and regional institutional environments.

3. Research methodology

3.1. Data and sample selection

This study used an unbalanced panel dataset of Chinese firms listed on the Shanghai and Shenzhen A-share markets from 2021 to 2024. We restricted the sample to this post-2020 period to improve measurement consistency and data comparability. Although ELI pilot programs existed before this period, earlier data have been less comparable because ESG rating coverage, environmental penalty disclosure, and voluntary ELI coverage information were less consistently available. By focusing on 2021–2024, we used a more consistent post-2020 data environment to examine short-term associations between the baseline Insur indicator and ESG performance under China's recent green finance regulatory framework. We applied the following sample selection criteria: (1) firms in the financial sector were excluded due to their distinct regulatory regimes and accounting standards; (2) Special Treatment (ST and *ST) firms were removed because of their abnormal financial status; and (3) firm-year observations with missing values for key dependent, independent, or control variables were dropped. Excluding observations with missing values removed 224 firm-year observations, representing 1.1% of the initial sample. Finally, to reduce the influence of extreme outliers, we winsorized all continuous variables at the 1st and 99th percentiles. The sample selection process began with 20,426 initial firm-year observations and yielded a final baseline sample of 19,185 firm-year observations, as reported in Table A1.

We obtained ESG ratings from the Huazheng ESG database, whose rating system classifies listed firms into nine ESG grades (Shanghai Huazheng Index Information Service Co., 2024). Alternative ESG ratings were obtained from the Shangdao Ronglü database (SynTao Green Finance, n.d.). Industry classifications followed the Guidelines for the Industry Classification of Listed Companies issued by the China Securities Regulatory Commission (CSRC, 2012). Heavily polluting industries were identified using the Ministry of Ecology and Environment's classified management catalog for environmental verification of listed companies

(MEE, 2008). Regional classifications followed the National Bureau of Statistics [NBS] (2024) division of eastern, central, western, and northeastern China. We coded and interpreted Huazheng and Shangdao Ronglü ESG ratings in accordance with their published methodology documents (Shanghai Huazheng Index Information Service Co., 2024; SynTao Green Finance, n.d.).

3.2. Defining variables

Dependent variable: Corporate ESG Performance (ESG). We measured corporate ESG performance using the Huazheng ESG Rating Index, a widely used proxy for corporate ESG performance in the Chinese market in recent empirical research (e.g., Fu & Li, 2023; Li et al., 2023). The Huazheng rating system assigned firms to nine ordinal grades, which were converted into numerical scores in ascending order: C=1, CC=2, CCC=3, B=4, BB=5, BBB=6, A=7, AA=8, AAA=9. To construct the annual ESG variable, we first calculated the arithmetic mean of the available quarterly numerical scores within each fiscal year. Then we rounded the annual average to the nearest integer. Therefore, the baseline ESG variable remained an integer-valued annual rating score on the 1–9 scale, rather than a non-integer annual average. In the final sample, the annual ESG variable contained nine unique values. We used this rounded annual ESG score in the baseline TWFE regressions (Ferrer-i-Carbonell & Frijters, 2004; Robitzsch, 2020). Because the resulting measure preserved the ordered structure of the original Huazheng rating categories, we also used the same rounded annual ESG score in the ordered logit robustness check. We interpreted the ordered logit results as a sensitivity analysis for treating ESG as ordinal rather than quasi-continuous.

Independent variable: ELI Coverage and Following-year Persistence (Insur). Insur is a binary indicator capturing verified ELI contractual coverage and assumed following-year post-coverage persistence. For each disclosed ELI policy, we identified the fiscal year the policy was contractually effective using the Qinglü Data Platform and corporate annual reports. Because most ELI contracts provided one-year coverage, we generally treated the policy purchase year as the contractual coverage year. In the baseline specification, Insur equaled 1 in the verified contractual coverage year and in the immediately following year. We included the following year only to capture potential short-term persistence in

compliance routines, risk-management procedures, and organizational adjustments after insurance coverage; we did not interpret it as a year in which insurance remained in force unless policy renewal was explicitly disclosed.

To further document the timing of the baseline Insur indicator, Table A6 reports the annual dynamics of first-time adoption, following-year observations under the baseline treatment definition, discontinuation, and total treated firm-year observations. Table A6 reports 420 total treated firm-year observations, consistent with the treatment composition used in the baseline regression. Because annual reports do not always distinguish newly disclosed or renewed contractual coverage from following-year persistence under the baseline coding rule, this study did not separately identify policy renewal unless renewed coverage was explicitly disclosed.

In this study, the policy purchase year referred to the fiscal year in which a firm disclosed an ELI policy. The coverage year referred to the year in which the disclosed policy was contractually effective, which was generally the same as the purchase year because most ELI policies provided one-year coverage. The following year was not treated as a year of current insurance coverage; instead, it was included in the baseline Insur variable only to capture potential short-term post-coverage persistence in compliance routines and risk-management procedures. First-time adoption referred to the first fiscal year in which a firm disclosed verified ELI contractual coverage during the sample period. Policy renewal referred only to a subsequent disclosed ELI policy after prior coverage and was not inferred from the following-year persistence rule. Following-year persistence referred to the immediately following year coded as treated under the baseline Insur definition; it captured assumed short-term persistence in compliance routines and should not be interpreted as verified current coverage. To distinguish these timing assumptions, Table A2 reports alternative treatment-timing specifications. The positive coefficient on ELI remained statistically significant with a contemporaneous coverage-year indicator, a one-year lagged indicator, and a post-adoption indicator, suggesting the baseline association was not driven solely by coding the coverage year and the immediately following year.

Control, Channel, and Heterogeneity variables. We included a set of firm-level controls commonly used in the ESG literature: listing age (ListAge), firm size (Size), leverage (Lev), profitability (ROA), growth opportunity (TobinQ), largest

shareholder ownership (Top1), board size (Board), independent director ratio (Indep), and institutional ownership (Inst). We measured ListAge as the natural logarithm of the number of years since listing plus one, as reported in Table 1. For Channel analyses, we constructed variables for environmental penalty exposure (Penalty and Fine), agency costs (Mfee), and environmental investment (EPIInvest). Penalty and Fine were based on environmental penalties observed over the previous three years; therefore, we interpreted these variables as measures of observed regulatory-risk exposure rather than post-treatment outcomes unless the temporal order was explicitly established. For heterogeneity analyses, we classified firms using official, reproducible criteria. We identified high-tech industries using the China Securities Regulatory Commission (CSRC) Industry Classification Guide (2012); we measured market competition with the industry-year Herfindahl–Hirschman Index (HHI), classifying the firms below the annual industry median as operating in more competitive industries. We defined pollution intensity using the Ministry of Ecology and Environment’s target list for heavily polluting sectors. We identified heavily polluting industries using the MEE-classified management catalog for environmental verification of listed companies. We defined regional groups using the National Bureau of Statistics (NBS) regional classification. To avoid overlapping regional categories, the regional heterogeneity analysis used a binary East-versus-non-East classification. We coded firms in the eastern region as East, while firms located in central, western, and northeastern China were grouped as non-East. This classification is mutually exclusive and compares firms in China’s more developed eastern region with those outside it.

Table 1. Definitions of variables

Category	Name	Sign	Definition
Dependent variables	Corporate ESG performance	ESG	Quarterly Huazheng ESG ratings are first mapped onto numerical scores from 1 to 9 in ascending order (C = 1, CC = 2, CCC = 3, B = 4, BB = 5, BBB = 6, A = 7, AA = 8, AAA = 9). The annual ESG variable is calculated as the arithmetic mean of available quarterly scores within each fiscal year and then rounded to the nearest integer; the final annual measure contains nine integer values.

Table 1 cont.

Category	Name	Sign	Definition
Independent variables	Environmental liability insurance	Insur	Set to 1 in the verified ELI contractual coverage year and the immediately following year to capture assumed short-term post-coverage persistence; otherwise 0. The following year is not interpreted as verified current insurance coverage unless policy renewal is explicitly disclosed.
Control variables	Age at listing	ListAge	The natural logarithm of the number of years since listing plus one.
	Company size	Size	The natural logarithm of the company's total assets at year-end.
	Leverage	Lev	The ratio of total liabilities to total assets.
	Profitability	ROA	The company's return on total assets for the current year.
	Major shareholder ownership	Top1	The ownership percentage of the company's largest shareholder.
	Growth potential	TobinQ	The ratio of the company's market capitalization to total assets.
	Board size	Board	The natural logarithm of the number of board members.
	Proportion of independent directors	Indep	The proportion of independent directors among the total number of board members.
	Institutional ownership	Inst	The proportion of shares held by institutional investors relative to the total shares of the listed company.
Other variables	Whether subject to environmental penalties	Penalty	Whether the company has received environmental penalties in the past three years; 1 if it has, 0 otherwise.
	Amount of environmental penalties	Fine	The natural logarithm of the company's annual average environmental penalties over the past three years, plus 1.
	Environmental investments	EPInvest	The natural logarithm of the company's environmental investment for the current year, plus 1.
	Agency costs	Mfee	The ratio of administrative expenses to operating revenue.

Source: Author's own compilation based on Huazheng ESG, Qinglü Data Platform, CCER, and CSMAR databases.

3.3. Model specification

The primary empirical framework was a TWFE model that estimates the association between the baseline Insur indicator and corporate ESG performance using within-firm variation over time. Specifically, we estimated the following baseline model:

$$ESG_{i,t} = \beta_0 + \beta_1 Insur_{i,t} + \gamma X_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

where $ESG_{i,t}$ denotes the ESG rating of firm i in year t , and $Insur_{i,t}$ is the baseline binary indicator of verified ELI contractual coverage and following-year post-coverage persistence, as defined above. $X_{i,t}$ represents the vector of firm-level characteristics described above. μ_i captures firm fixed effects, which absorb time-invariant unobserved firm characteristics, and λ_t captures year fixed effects, which account for common macroeconomic shocks and regulatory changes affecting all firms in a given year. $\varepsilon_{i,t}$ is the idiosyncratic error term. Standard errors were clustered at the firm level to account for serial correlation within firms over time. The coefficient of interest was β_1 , which should be interpreted as the conditional association between the baseline Insur indicator and ESG performance, not as a causal treatment effect. Although firm and year fixed effects reduce concerns about time-invariant firm heterogeneity and common time shocks, they do not eliminate time-varying omitted variables, reverse causality, or early-adopter selection.

3.4. Identification strategy and endogeneity concerns

Empirically estimating the association between the baseline Insur indicator and ESG performance is challenging because ELI adoption is not random. Firms with stronger ESG preferences, higher environmental risk exposure, or greater regulatory attention may be more likely to purchase ELI. In addition, both the baseline Insur indicator and ESG performance may be jointly affected by unobserved time-varying firm characteristics.

Ideally, a quasi-natural experiment, such as a staggered difference-in-differences (DiD) design based on policy shocks, would provide stronger identification. However, the 2022 green finance regulatory framework was implemented broadly and did not provide sufficient staggered regional or temporal variation for a credible policy-shock design. The short 2021–2024 sample window also lim-

ited the feasibility of pre-trend and event-study analyses. Given these constraints, this study relied on firm fixed effects, year fixed effects, and matching procedures to improve comparability and reduce observable selection concerns. We therefore interpreted the results as conditional short-term associations rather than causal treatment effects.

Matching was conducted at the firm-year level using 1:1 nearest-neighbor matching with a caliper of 0.05 and with replacement. Before matching, the baseline sample contained 420 treated firm-year observations and 18,765 control firm-year observations. After matching, all 420 treated firm-year observations obtained valid matches within the caliper. Because matching was conducted with replacement, some control observations matched more than one treated observation. As a result, the matched sample used in the PSM regression included 839 unique firm-year observations: 420 treated observations and 419 unique matched control observations. Table A3 reports balance diagnostics, including standardized mean differences for all covariates used in the propensity score model. We estimated the propensity score using a logit model in which *Insur* is regressed on *ListAge*, *Size*, *Lev*, *ROA*, *Top1*, *TobinQ*, *Board*, *Indep*, and *Inst*, corresponding to the covariates reported in the balance diagnostics.

To ensure comparability with the baseline specification, we re-estimated the matched-sample regression with the same TWFE structure as the baseline model, including firm and year fixed-effects and firm-level clustered standard errors. The PSM analysis evaluated whether the baseline association was sensitive to observable covariate imbalance; it should not be interpreted as eliminating unobserved selection, reverse causality, or time-varying omitted variables.

The main matched-sample regression was estimated without matching-frequency weights to keep the unit of analysis at the unique firm-year level and to avoid giving repeated control observations disproportionate influence in the TWFE specification. As a sensitivity check, we also estimated a weighted PSM specification using matching-frequency weights. The matched sample included 839 unique firm-year observations and 613 firm-level clusters. Within the matched regression sample, 20 firms exhibited within-sample changes in the baseline *Insur* indicator. Therefore, the fixed-effect estimate should be interpreted as being identified by within-firm variation in treatment status, rather than by the total number of treated firm-year observations mechanically. Not all 420 treated firm-year observations necessarily provided independent within-firm identifying variation.

We also conducted additional robustness checks, including alternative ELI timing definitions, alternative ESG rating measures, and placebo tests based on randomized treatment assignments. These tests showed the sensitivity of the baseline association, but they did not rule out all forms of unobserved time-varying confounding.

3.5. Descriptive statistics

Table 2 presents summary statistics for the main variables. The mean ESG score was 4.095 (SD=1.200) on the rounded annual 1–9 Huazheng scale, indicating moderate ESG performance across the sample. The Insur mean was 0.022, meaning approximately 2.2% of firm-year observations were coded as verified ELI coverage or as following-year post-coverage persistence observations under the baseline treatment definition. This low treatment prevalence reflects the early-stage development of China's voluntary ELI market, but it also limits statistical power and the generalizability of subgroup analyses. Variance inflation factors (VIFs) for all continuous control variables were below conventional thresholds, and the Pearson correlation matrix does not indicate severe multicollinearity among the main variables.

Table 2. Descriptive statistics

Variable	Observations	Mean	Std.Dev.	Min	Max
ESG	19,185	4.095	1.200	1	9
Insur	19,185	0.022	0.146	0	1
ListAge	19,185	2.029	0.983	0	3.555
Size	19,185	22.274	1.325	19.796	26.452
Lev	19,185	0.402	0.205	0.049	0.910
ROA	19,185	0.027	0.067	−0.264	0.222
Top1	19,185	0.322	0.146	0.075	0.740
TobinQ	19,185	1.937	1.191	0.795	9.598
Board	19,185	2.079	0.195	1.609	2.565
Indep	19,185	0.381	0.0538	0.286	0.571
Inst	19,185	0.406	0.250	0.003	0.920
Penalty	19,185	0.171	0.377	0	1
Fine	19,185	0.458	1.188	0	8.485

Table 2 cont.

Variable	Observations	Mean	Std.Dev.	Min	Max
EPInvest	19,185	1.233	2.803	0	21.535
Mfee	19,185	0.076	0.061	0.007	0.400

Source: Author's own calculations using data from Huazheng ESG, Qinglü Data Platform, CCER, and CSMAR.

4. Research findings

4.1. Baseline results

Table 3 shows baseline TWFE estimation results for the association between the baseline Insur indicator and corporate ESG performance. To clarify the treatment composition, the sample includes exactly 420 firm-year observations coded as Insur = 1 under the baseline treatment definition across the 2021–2024 sampling window. During the sample period, 236 unique firms disclosed verified ELI contractual coverage, generating 420 treated firm-year observations under the baseline Insur definition after including following-year persistence observations. Among these firms, 224 experienced within-sample treatment-status changes, thereby providing the main within-firm variation used to estimate the coefficient on Insur in the fixed-effects specifications.

Table 3. Baseline TWFE results: ELI and corporate ESG performance

Variables	ESG (1)	ESG (2)
Insur	0.319*** (0.069)	0.266*** (0.075)
Size	—	0.689*** (0.061)
Lev	—	−1.909*** (0.148)
ROA	—	−1.057*** (0.187)
Board	—	0.237* (0.135)
Indep	—	0.734* (0.385)
Top1	—	1.553*** (0.316)

Table 3 cont.

Variables	ESG (1)	ESG (2)
TobinQ	—	−0.045*** (0.013)
ListAge	—	0.812*** (0.055)
Inst	—	0.559*** (0.169)
Constant	4.100*** (0.007)	−13.304*** (1.321)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Observations	19,185	19,185
R-squared	0.587	0.609

Note: Standard errors clustered at the firm level are reported in parentheses. Columns 2 and 3 include firm fixed effects and year fixed effects. Column 3 additionally includes the full set of firm-level controls defined in Table 1. *, **, and *** indicate significance at the 10%, 5%, and 1% significance levels, respectively.

Source: Author's estimations based on TWFE models using firm-level panel data (2021–2024).

Column 2 of Table 3 presents the model without control variables, yielding a positive and statistically significant coefficient on *Insur* ($\beta = 0.319$, $p < 0.01$). In Column 3, after incorporating the full vector of firm-level control variables, the coefficient on *Insur* remains positive and statistically significant ($\beta = 0.266$, $p < 0.01$). Economically, conditional on firm fixed effects and common year shocks, the baseline *Insur* indicator was associated with an average difference of approximately 0.27 points on the 1–9 Huazheng ESG scale. Given that the standard deviation of the ESG measure was 1.200, this estimated association corresponds to approximately 22.2% of one standard deviation. This finding supports H1, as a positive conditional association better fits an insurer-related monitoring interpretation than an unmitigated moral-hazard interpretation.

4.2. Endogeneity and robustness checks

To assess whether the baseline association is sensitive to observable selection, we employed propensity score matching (PSM) and re-estimated the baseline TWFE specification on a matched sample constructed using 1:1 nearest-neighbor

matching at the firm-year level with a caliper of 0.05 and with replacement. Before matching, the baseline sample contains 420 treated firm-year observations and 18,765 control firm-year observations. All treated observations obtained valid matches within the caliper. Because matching was conducted with replacement, some control observations matched more than one treated observation; therefore, the matched regression sample included 839 unique firm-year observations: 420 treated observations and 419 unique matched control observations. Balance diagnostics, reported in Table A3, show that post-matching standardized mean differences (SMD) for all covariates used in the propensity score model fell below the conventional 5% threshold. The matched sample contained 613 firm-level clusters, and 20 firms changed the baseline Insur status within the matched regression sample.

Table 4. Robustness check: PSM-matched sample

Variables	PSM-matched sample	Weighted PSM sensitivity
Insur	0.870** (0.372)	0.854** (0.378)
Controls	YES	YES
Year FE	YES	YES
Firm FE	YES	YES
Observations	839	839
Firm-level clusters	613	613
R-squared	0.128	0.129

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% significance levels, respectively. We conducted matching at the firm-year level using 1:1 nearest-neighbor matching with replacement and a caliper of 0.05. We estimated the propensity score using ListAge, Size, Lev, ROA, Top1, TobinQ, Board, Indep, and Inst. The matching procedure produced 420 matched treatment-control pairs, corresponding to 840 observation uses. Because matching was conducted with replacement, repeated use of some control observations resulted in 839 unique firm-year observations in the matched regression sample: 420 treated firm-year observations and 419 unique matched control firm-year observations. The matched sample included 613 firm-level clusters, and 20 firms changed the baseline Insur status within the matched regression sample. Column 2 reports the unweighted matched-sample regression. Column 3 reports a weighted PSM sensitivity specification using matching-frequency weights. The coefficient on Insur remained positive and statistically significant, suggesting that the PSM result was not driven solely by the unweighted treatment of repeatedly matched control observations.

Source: Author's estimations on propensity score matched sample (1:1 nearest-neighbor, caliper = 0.05).

As reported in Table 4, re-estimating the baseline specification on the matched sample yielded a positive coefficient on Insur ($\beta = 0.870$), which was statistically significant at the 5% level. The larger coefficient in the matched sample should be interpreted cautiously because it may reflect differences in matched-sample composition and the small number of treated observations. The PSM analysis improved balance on observable covariates, but it did not address unobserved selection, reverse causality, or time-varying omitted variables. Therefore, we interpreted the PSM result as supportive sensitivity evidence rather than as a stronger or more unbiased estimate of the ELI-ESG association.

Because the matched-sample regression included firm fixed effects, the coefficient was identified by within-firm changes in Insur rather than by cross-sectional differences between treated and control firms. Accordingly, the PSM result should be interpreted as a sensitivity check based on improved observable covariate balance, not as evidence that matching eliminates unobserved selection.

Table 5. Sensitivity check: Alternative model specifications

Variables	Basic FE ESG	High-dimensional FE ESG	Ordinal Logit ESG	Random ESG
Insur	0.266*** (0.075)	0.285*** (0.063)	0.022 (0.108)	0.203*** (0.067)
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	No	No
Firm RE	No	No	No	Yes
Industry × Year FE	No	Yes	No	No
<i>N</i>	19,185	19,185	19,185	19,185
<i>R-squared</i>	0.609	0.689	–	0.262
<i>Pseudo R²</i>	–	–	0.072	–

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column 2 reports the baseline fixed-effects model; Column 3 adds industry-by-year fixed effects, and the industry-by-year fixed effects absorb the common year effects; Column 4 reports an ordered logit model using the rounded annual ESG score as an ordinal dependent variable; Column 5 reports a random effects model. R-squared values are reported for linear specifications, while pseudo R-squared is reported for the ordered logit model.

Source: Author's own estimations using high-dimensional FE, random effects, and ordered logit specifications.

To assess robustness and alternative functional forms and sample imbalance, Table 5 reports additional specification checks. Column 2 and 3 present the baseline fixed-effects model and a high-dimensional fixed-effects model incorporating industry-by-year effects, respectively. In Column 3, the coefficient on *Insur* remains positive and statistically significant ($\beta = 0.285$, $p < 0.01$), suggesting that common industry-year shocks do not drive the baseline association.

Column 4 reports an ordered logit model using the rounded annual Huazheng ESG score as an ordinal dependent variable. The coefficient on *Insur* is positive but statistically insignificant ($\beta = 0.022$, $SE = 0.108$, $p = 0.84$, pseudo $R^2 = 0.072$). This result indicates that the main finding is sensitive to treating ESG strictly as an ordinal outcome. The insignificant ordered logit result may reflect the sparse treatment distribution, the concentration of observations in a limited range of ESG categories, and the ordered logit model's restrictive proportional-odds structure. Therefore, we did not interpret the ordered logit result as confirming robustness. Instead, we reported an important sensitivity check showing that the positive association is clearer in the quasi-continuous fixed-effects framework than in the ordinal specification.

Column 5 reports a random effects specification. The coefficient on *Insur* remains positive and statistically significant ($\beta = 0.203$, $p < 0.01$), consistent with the baseline fixed-effects result. However, because unobserved firm-level heterogeneity is likely to be correlated with the baseline *Insur* indicator, the fixed-effects specification remains the preferred baseline model.

To assess whether the baseline association depends on the Huazheng rating methodology, we reconstructed the dependent variable using the alternative Shangdao Ronglü ESG database. As reported in Table A4, the coefficient on *Insur* remains positive and statistically significant at the 1% level ($\beta = 0.143$, $p < 0.01$). This result suggests that the positive association is not specific to a single ESG rating provider.

We further examined whether the baseline association is sensitive to the annual aggregation method used for Huazheng ESG ratings. As reported in Table A7, the coefficient on *Insur* remains positive and statistically significant when ESG is measured using the year-end Huazheng rating ($\beta = 0.251$, $p < 0.01$) and the annual median of quarterly Huazheng ratings ($\beta = 0.238$, $p < 0.01$). These results suggest that the specific use of the rounded annual average ESG measure does not drive the baseline association.

We also conducted a placebo test by randomly assigning ELI treatment status 500 times across the panel and re-estimating the baseline model. The distribution of pseudo-treatment coefficients is centered close to zero, whereas the actual baseline coefficient lies in the right tail of the placebo distribution. This pattern suggests that the estimated association is unlikely to arise from random treatment assignment or mechanical spurious correlation. However, the placebo test did not eliminate concerns about reverse causality, early-adopter selection, or time-varying omitted variables.

4.3. Channel evidence: Regulatory risk exposure and governance efficiency

To examine regulatory risk exposure as the first channel, Table 6 assesses whether the baseline Insur indicator is associated with firms' observed environmental regulatory-risk exposure. Columns 1 and 2 show that the baseline Insur indicator is negatively associated with both the probability of environmental penalties (Penalty: $\beta = -0.057$, $p < 0.01$) and the log penalty amount (Fine: $\beta = -0.185$, $p < 0.01$). Because Penalty and Fine are constructed using environmental penalties observed over the previous three years, these results should not be interpreted as evidence that ELI reduces subsequent penalties. Instead, they indicate that firms coded as treated under the baseline Insur definition tend to have lower observed penalty exposure, consistent with the regulatory-risk channel proposed in H2a.

Table 6. Channel evidence: Environmental regulatory-risk exposure

Variables	Penalty	Fine
Insur	-0.057*** (0.021)	-0.185*** (0.065)
Controls	Yes	Yes
Constant	-0.205 (0.347)	-0.690 (0.962)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Observations	19,185	19,185
R-squared	0.698	0.743

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: Author's own estimations using environmental penalty data from the CCER database.

In Table 7, Columns 2 and 3 include Penalty and Fine in the baseline ESG model to examine whether penalty exposure is related to ESG performance. The coefficients on Penalty and Fine are negative and statistically significant, indicating that environmental penalty exposure is strongly associated with lower ESG ratings. After including these penalty-related variables, the coefficient on Insur remains positive and statistically significant. This pattern suggests that observed penalty exposure does not fully explain the positive association between ELI and ESG performance. However, given the timing of the penalty variables, we interpreted these results as associative channel evidence rather than as formal mediation evidence.

Table 7. Channel evidence: Penalty exposure, governance efficiency, and environmental investment

Variables	ESG (1)	ESG (2)	Mfee	EPInvest
Insur	0.201*** (0.074)	0.218*** (0.074)	−0.004*** (0.001)	0.213 (0.216)
Penalty	−0.892*** (0.089)	—	—	—
Fine	—	−0.156*** (0.031)	—	—
Controls	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	19,185	19,185	19,185	19,185
R-squared	0.612	0.611	0.893	0.515

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Columns 2 and 3 illustrate the association between penalty exposure and ESG performance; Column 4 displays agency costs; Column 5 shows environmental investment. These estimates reflect evidence of associative channels rather than formal mediation effects.

Source: Author's own estimations.

To address the possibility that the positive ELI–ESG association partly reflects the mechanical incorporation of environmental penalties into aggregate ESG ratings, we conducted two additional checks. First, we examined whether verified ELI coverage was associated with the Social and Governance pillar scores, which were conceptually less directly tied to environmental penalty records. The results showed positive associations with both the Social pillar ($\beta = 0.142$, $p < 0.05$) and

the Governance pillar ($\beta = 0.188$, $p < 0.01$). These findings alleviate, but do not fully eliminate, concerns that the baseline ESG association merely reflects the mechanical incorporation of environmental penalties into the aggregate ESG rating methodology.

Second, to reduce concerns about potential mechanical scoring overlap, we constructed an Adjusted Environmental Pillar (Adjusted-E) score that excludes environmental-rating information directly related to regulatory sanctions, environmental violations, pollution accidents, or negative environmental screening items. Adjusted-E is used only as a sensitivity measure and should not be interpreted as a reconstruction of the official proprietary Huazheng Environmental Pillar rating. Re-estimating the baseline model using Adjusted-E yielded a positive and statistically significant coefficient on Insur ($\beta = 0.165$, $p < 0.01$), as reported in Table A5. This test should be interpreted as supportive sensitivity evidence rather than definitive proof of substantive environmental improvement.

Table 7 examines internal governance efficiency and environmental investment as the second potential channel. Table 7, Column 4, shows that the baseline Insur indicator is negatively associated with agency costs, measured by the ratio of administrative expenses to operating revenue (Mfee: $\beta = -0.004$, $p < 0.01$). This finding is consistent with the possibility that the baseline Insur indicator has been associated with improved internal governance efficiency. However, it should be interpreted as an association rather than as evidence that insurer monitoring directly disciplines managerial behavior.

Column 5 shows that the association between the baseline Insur indicator and environmental capital expenditure is positive but statistically insignificant (EPIInvest: $\beta = 0.213$, $p = 0.32$). This result indicated that the observed ELI-ESG association was not accompanied by a statistically significant increase in substantive environmental investment during the short sample period. One possible interpretation is that compliance-related and governance-related adjustments may appear earlier than capital-intensive environmental investments. Overall, the channel results support H2a and H2b, while also highlighting the limits of interpreting these associations as causal mechanisms.

4.4. Heterogeneity analysis

To avoid relying solely on split-sample significance, we tested cross-group differences by estimating pooled full-sample models with interaction terms, such as

Insur $\times$ SOE and Insur $\times$ HighTech. We used the p-values of the interaction coefficients as formal tests of whether the ELI-ESG association differs across groups. We reported these p-values as difference test p-values at the bottom of the corresponding tables.

Industry-Level Heterogeneity. Table 8 presents industry-level heterogeneity results based on technological intensity and market competition. For technological intensity, the coefficient on Insur was positive and statistically significant in both high-tech industries ($\beta = 0.230$, $p < 0.01$) and non-high-tech industries ($\beta = 0.268$, $p < 0.05$). The coefficient was numerically larger for non-high-tech firms, and the formal difference test was statistically insignificant ($p = 0.345$). Therefore, the results do not support the claim that the ELI-ESG association is concentrated among high-tech firms.

Table 8. Heterogeneity analysis: Industry-level

Grouping criterion	Industry technical level		Industry competitiveness	
Variables	High-tech	Non-high-tech	Competitive	Non-competitive
Insur	0.230*** (0.078)	0.268** (0.133)	0.304*** (0.074)	0.082 (0.171)
Controls	Yes	Yes	Yes	Yes
Constant	-11.122*** (1.778)	-15.583*** (2.128)	-13.020*** (1.579)	-14.294*** (2.619)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	12,463	6,683	15,249	3,901
R-squared	0.605	0.634	0.606	0.645
Difference Test (p-value)	0.345	—	0.298	—

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Difference test p-values were obtained from pooled full-sample interaction models. Observation counts may not sum exactly to the full baseline sample because we excluded observations without valid classification information from the corresponding subgroup analysis.

Source: Author's own subgroup estimations based on industry technological intensity and market competition.

For market competition, the coefficient on Insur was larger and statistically significant among firms in competitive industries ($\beta = 0.304$, $p < 0.01$), but smaller and statistically insignificant among firms in non-competitive industries ($\beta = 0.082$).

However, the formal interaction test indicated that the difference between competitive and non-competitive industries was not statistically significant ($p = 0.298$). Accordingly, the competition-based heterogeneity should be interpreted as suggestive rather than conclusive evidence.

Table 9 presents firm-level subgroup results by ownership structure and pollution intensity. For ownership, the coefficient on Insur was statistically insignificant for SOEs ($\beta = 0.180$) but positive and statistically significant for non-SOEs ($\beta = 0.268$, $p < 0.01$). The formal interaction test indicated a marginally significant difference between SOEs and non-SOEs at the 10% level ($p = 0.093$). This result provides limited evidence that the ELI-ESG association may be stronger among non-SOEs, which is consistent with the view that non-SOEs may be more responsive to market-based governance incentives.

Table 9. Heterogeneity analysis: Firm level

Grouping criterion	Type of ownership		Level of pollution	
Variables	State-owned	Non-state-owned	Heavily polluted	Not heavily polluted
Insur	0.180 (0.140)	0.268*** (0.076)	0.141 (0.097)	0.324*** (0.095)
Controls	Yes	Yes	Yes	Yes
Constant	-14.618*** (2.365)	-12.629*** (1.584)	-11.779*** (3.110)	-13.206*** (1.476)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	5,130	13,975	3,807	15,356
R-squared	0.619	0.618	0.625	0.607
Difference test (p-value)	0.093	—	0.533	—

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Difference test p-values were obtained from pooled full-sample interaction models. Observation counts may not sum exactly to the full baseline sample because we excluded observations without valid classification information from the corresponding subgroup analysis.

Source: Author's own subgroup estimations based on ownership structure and pollution intensity.

Pollution-intensity heterogeneity was based on the heavily polluting industry classification used by MEE. For pollution intensity, the coefficient on Insur was statistically insignificant among heavily polluting firms ($\beta = 0.141$) and positive

and statistically significant among non-heavily polluting firms ($\beta = 0.324$, $p < 0.01$). However, the formal difference test was statistically insignificant ($p = 0.533$). Therefore, the pollution-intensity results should not be interpreted as evidence of a statistically supported subgroup difference.

Table 10 presents regional heterogeneity based on a mutually exclusive East versus non-East classification. The coefficient on Insur was positive and statistically significant for both eastern firms ($\beta = 0.183$, $p < 0.05$) and non-eastern firms ($\beta = 0.509$, $p < 0.01$). The formal interaction test indicated that the coefficient for non-eastern firms differed significantly from the coefficient for eastern firms ($p = 0.035$). This finding provided the strongest subgroup evidence in the heterogeneity analysis. It is consistent with the possibility that insurance-based governance mechanisms may be more closely associated with ESG performance outside China's more developed eastern region. However, this result should be interpreted cautiously because the non-East group combines central, western, and northeastern regions, which may differ in institutional conditions and local environmental enforcement capacity.

Table 10. Heterogeneity analysis: Regional level

Variables	East	Non-East
Insur	0.183** (0.075)	0.509*** (0.141)
Controls	Yes	Yes
Constant	-13.314*** (1.537)	-13.826*** (2.675)
Year FE	Yes	Yes
Firm FE	Yes	Yes
Observations	14,148	5,027
R-squared	0.604	0.619
Difference Test (p-value)	0.035	—

Note: Difference test p-values were obtained from pooled full-sample interaction models. East was defined according to the National Bureau of Statistics (NBS) regional classification. Non-East included central, western, and northeastern China. Observations with missing or unclassified regional information were excluded. Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: Author's own subgroup estimations based on geographic regions (East/non-East).

Overall, the heterogeneity analysis provided mixed evidence. The regional difference between eastern and non-eastern firms was statistically supported, and the ownership difference between SOEs and non-SOEs was only marginally significant. By contrast, the differences across technological intensity, market competition, and pollution intensity were not statistically significant in the formal interaction tests. Therefore, we interpreted the heterogeneity results as suggestive boundary-condition evidence rather than definitive proof of structural differences across all subgroups.

5. Discussion

5.1. Interpretation of channel evidence

This study examines the role of market-based environmental instruments in corporate governance by analyzing the association between ELI and corporate ESG performance. Although the empirical design uses firm fixed effects, PSM, alternative specifications, and placebo tests, it remains fundamentally observational. Therefore, the results should be interpreted as conditional associations rather than causal effects.

Within this framework, the evidence is consistent with two possible channels. First, the baseline *Insur* indicator is associated with lower observed regulatory-risk exposure, as reflected in lower three-year environmental penalty incidence and penalty amounts. However, because the penalty variables are constructed using penalties observed over the previous three years, they may include events that occurred before the verified ELI coverage year. Therefore, these results should not be interpreted as evidence that ELI reduces subsequent penalties. Instead, they suggest that firms coded as treated under the baseline *Insur* definition tend to have lower observed penalty exposure, consistent with a regulatory-risk channel.

Second, the baseline *Insur* indicator is associated with lower agency costs, suggesting a possible link between the baseline *Insur* indicator and internal governance efficiency. However, this association does not establish that insurer monitoring directly improves managerial behavior. Alternative explanations, including early-adopter characteristics, regulatory pressure, reputational incentives, and self-selection into ELI, remain plausible. The positive net association between ELI

and ESG performance is therefore more consistent with a monitoring interpretation than with an unmitigated moral-hazard interpretation. However, it does not separately identify monitoring and moral-hazard effects.

Importantly, these governance-related associations are not accompanied by a statistically significant increase in environmental capital expenditures during the short sample period. This result highlights a distinction between compliance- or governance-related ESG improvements and substantive environmental investment. The findings align more with observable compliance and governance adjustments than with immediate, capital-intensive environmental upgrading. This limitation is important because improvements in ESG ratings may partly reflect procedural compliance, penalty avoidance, or rating-related dimensions rather than fundamental changes in environmental impact.

Compared with prior studies that report a positive relationship between ELI and ESG-related outcomes using broader ESG, environmental-performance, or green-innovation measures, our findings provide more cautious evidence based on a recent post-2020 listed-firm panel. The positive association is broadly consistent with Wu et al. (2022) and Chen and Xie (2024), but this study differs in five respects. First, it focuses on the post-2020 period, when ESG rating coverage and environmental penalty disclosure became more standardized in China. Second, it examines ELI as a firm-level insurance-based governance arrangement, rather than treating it only as a risk-transfer instrument or a broad green-finance policy proxy. Third, it explicitly tests the sensitivity of the result to treating Huazheng ESG ratings as ordinal outcomes; the statistically insignificant ordered-logit result indicates that the baseline finding is clearer in the quasi-continuous fixed-effects framework than in the ordinal specification. Fourth, the channel evidence distinguishes governance- and compliance-related associations from substantive environmental investment: ELI is associated with lower observed penalty exposure and lower agency costs, but not with a statistically significant increase in environmental capital expenditure. Fifth, the interpretation frames ELI as a potential quasi-regulatory monitoring arrangement, while avoiding the claim that the study conclusively establishes a new causal mechanism. Therefore, this study should be viewed as providing cautious firm-level evidence on insurance-based environmental governance rather than definitive proof of substantive environmental improvement.

5.2. Boundary conditions and heterogeneity

The heterogeneity results provide mixed evidence on the boundary conditions of the ELI-ESG association. Split-sample estimates suggest that the association is positive in several subgroups, but formal interaction tests show that not all subgroup differences are statistically significant. Therefore, the heterogeneity results should be interpreted cautiously.

The strongest subgroup evidence appears in the regional analysis. The coefficient for non-eastern firms is larger than that for eastern firms, and the formal difference test is statistically significant. This pattern is consistent with the possibility that insurance-based governance mechanisms may be more salient outside China's more developed eastern region. However, this interpretation remains suggestive because the non-East group combines central, western, and north-eastern regions and may reflect unobserved local policy conditions, firm composition, or differences in ELI adoption patterns.

The ownership result provides weaker evidence. The coefficient is statistically significant for non-SOEs but not for SOEs, and the formal interaction test is only marginally significant. This pattern is consistent with the possibility that non-SOEs may be more responsive to market-based governance incentives. However, it should not be interpreted as definitive evidence of a structural ownership difference.

By contrast, the differences across technological intensity, market competition, and pollution intensity are not statistically significant in the formal interaction tests. In particular, the results do not support the claim that the ELI-ESG association is concentrated among high-tech firms. Overall, the heterogeneity analysis suggests possible boundary conditions, especially across regions and, to a lesser extent, ownership types, but it does not provide uniform evidence of statistically significant subgroup differences.

6. Conclusions

6.1. Main findings

This study contributes to the literature on corporate sustainability and insurance-based governance by examining ELI as a micro-level quasi-regulatory instrument in an emerging-market context. Using a panel dataset of Chinese A-share listed

firms from 2021 to 2024, we document a positive short-term association between the baseline Insur indicator and corporate ESG performance, mainly in linear fixed-effects specifications.

The main analysis maps Huazheng's ordered ESG rating categories to a rounded annual 1–9 score and treats this measure as quasi-continuous in the baseline fixed-effects model. However, the ordered logit sensitivity check is statistically insignificant, indicating that the main finding is sensitive to treating ESG strictly as an ordinal outcome. Therefore, the evidence should be interpreted as supportive but not conclusive.

The findings are consistent with the view that the positive net association between ELI and ESG performance is more consistent with an insurer-related monitoring interpretation than with an unmitigated moral-hazard interpretation. However, the study does not separately identify monitoring and moral-hazard effects. Alternative explanations, including self-selection into ELI, early-adopter characteristics, regulatory pressure, and reputational incentives, remain plausible.

The channel analyses suggest that the baseline Insur indicator is associated with lower observed environmental penalty exposure and lower agency costs. These patterns are consistent with possible regulatory-risk and governance-efficiency channels. However, because the penalty measures use environmental penalties observed over the previous three years, they should not be interpreted as evidence that ELI reduces subsequent penalties. In addition, the association between the baseline Insur indicator and environmental capital expenditure is statistically insignificant, limiting the extent to which the results can be interpreted as evidence of substantive environmental improvement.

Regarding heterogeneity, the regional difference between eastern and non-eastern firms is statistically supported, while the ownership difference between SOEs and non-SOEs is only marginally significant. Differences across technological intensity, market competition, and pollution intensity are not statistically significant in the formal interaction tests. Therefore, the heterogeneity results provide suggestive boundary-condition evidence rather than uniform evidence of structural subgroup differences.

6.2. Contributions and limitations

This study makes three contributions. Theoretically, it extends the insurance-governance literature by conceptualizing ELI as a potential quasi-regulatory monitoring mechanism rather than only as a risk-transfer instrument. Empirically, it provides firm-level evidence on the association between the baseline Insur indicator and ESG performance using recent post-2020 Chinese listed-firm data. Contextually, it highlights how market-based environmental insurance may be related to corporate ESG outcomes in China's evolving green finance and environmental governance framework.

Several limitations should be acknowledged. First, the empirical design is observational. Firm fixed effects, PSM, lagged specifications, and placebo tests reduce some concerns but do not eliminate reverse causality, early-adopter selection, or time-varying omitted variables. Second, ELI penetration remains low in the sample, limiting statistical power and making subgroup results especially sensitive to small treated-cell sizes. Third, the 2021–2024 sample period captures only short-term associations and cannot evaluate long-term sustainability trajectories. Fourth, the insignificant environmental investment result suggests that the observed ESG association may reflect compliance- or governance-related adjustments rather than immediate substantive environmental upgrading.

Future research could extend the observation window, more precisely distinguish first-time adoption, renewal, continuation, and discontinuation of ELI, and exploit stronger exogenous variation to address early-adopter selection and identify long-term causal effects more credibly. Future work could also examine whether verified ELI coverage is associated with real environmental outcomes, such as emissions reductions, pollution intensity, or verified environmental investment.

6.3. Policy implications

The findings suggest that ELI may complement traditional command-and-control regulation, but only under appropriate institutional conditions. Under appropriate institutional conditions, policymakers could strengthen ELI's potential governance role by improving the transparency of underwriting standards, encouraging insurers to conduct credible environmental risk assessments, and enhancing coordination between insurance institutions and environmental regulators.

For regulators, the results imply that ELI should not be treated merely as a compensation instrument after pollution accidents. It may also provide information and monitoring functions that are relevant to corporate environmental governance. However, because this study does not establish causal effects, policy implications should be interpreted cautiously and should be supported by further evidence from longer panels or stronger quasi-experimental designs.

For corporate managers, ELI should be understood not only as a risk-transfer instrument but also as a governance-relevant arrangement associated with external risk assessment, compliance routines, and ESG evaluation. Nevertheless, the lack of a significant environmental investment effect suggests that firms should not treat ELI coverage as a substitute for substantive environmental improvement.

Finally, the policy implications of this study should be interpreted within appropriate institutional boundaries. The findings may be most relevant to institutional settings where environmental insurance is linked to regulatory oversight, insurers have both the incentives and capacity to conduct environmental risk assessment, ESG disclosure is sufficiently standardized, and firms face reputational or financing pressure related to sustainability performance. In countries or regions where environmental insurance functions mainly as passive compensation, where insurers lack monitoring capacity, or where ESG and penalty disclosure are weak, the association documented in this study may not generalize directly.

Funding

This research received no external funding.

Conflict of interest

The author declares that there is no conflict of interest.

Ethics statement

This study did not involve human participants and did not use personal, sensitive, or identifiable data; therefore, ethical approval was not required.

Data availability statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

References

- Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5), 941–973. <https://www.jstor.org/stable/1812044>
- Chen, S., Ding, X., Lou, P., & Song, H. (2022). New evidence of moral hazard: Environmental liability insurance and firms' environmental performance. *Journal of Risk and Insurance*, 89(3), 581–613. <https://doi.org/10.1111/jori.12380>
- Chen, Y., & Xie, Y. (2024). Environmental liability insurance, green innovation, and mediation effect study. *Frontiers in Environmental Science*, 12, 1363199. <https://doi.org/10.3389/fenvs.2024.1363199>
- China Banking and Insurance Regulatory Commission [CBIRC]. (2022). Zhōngguó yín bǎojiānhuì guānyú yìnfā yínhángyè bǎoxiǎnyè lǜsè jīnróng zhǐyǐn de tōngzhī yín bǎo jiān fā [Notice by the China Banking and Insurance Regulatory Commission of issuing the Guidelines on Green Finance for the Banking and Insurance Sectors]. *CBIRC Notice*, 15. https://www.gov.cn/zhengce/zhengceku/2022-06/03/content_5693849.htm
- China Securities Regulatory Commission [CSRC]. (2012). Annex: *Guidelines for the industry classification of listed companies* (2012 revision). http://www.csrc.gov.cn/csrc_en/c102034/c1371375/content.shtml
- Dionne, G. (2013). Risk management: History, definition, and critique. *Risk Management and Insurance Review*, 16(2), 147–166. <https://doi.org/10.1111/rmir.12016>
- Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3), 693–714. <https://doi.org/10.1016/j.jfineco.2018.08.013>
- Ferrer-i-Carbonell, A., & Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? *The Economic Journal*, 114(497), 641–659. <https://doi.org/10.1111/j.1468-0297.2004.00235.x>
- Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233. <https://doi.org/10.1080/20430795.2015.1118917>
- Fu, T., & Li, J. (2023). An empirical analysis of the impact of ESG on financial performance: The moderating role of digital transformation. *Frontiers in Environmental Science*, 11, 1256052. <https://doi.org/10.3389/fenvs.2023.1256052>

- Holderness, C. G. (1990). Liability insurers as corporate monitors. *International Review of Law and Economics*, 10(2), 115–129. [https://doi.org/10.1016/0144-8188\(90\)90018-O](https://doi.org/10.1016/0144-8188(90)90018-O)
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies*, 33(3), 1067–1111. <https://doi.org/10.1093/rfs/hhz137>
- Li, H., Guo, H., Hao, X., & Zhang, X. (2023). The ESG rating, spillover of ESG ratings, and stock return: Evidence from Chinese listed firms. *Pacific-Basin Finance Journal*, 80, 102091. <https://doi.org/10.1016/j.pacfin.2023.102091>
- Ministry of Ecology and Environment of the People's Republic of China [MEE]. (2008). Guānyú yīnfā “shàngshì gōngsī huánbǎo héchá hángyè fēnlèi guǎnlǐ mínglù” de tōngzhī [Circular on issuing the classified management catalogue for environmental verification of listed companies]. https://www.mee.gov.cn/gkml/hbb/bgth/200910/t20091022_174891.htm
- Ministry of Ecology and Environment of the People's Republic of China [MEE]. (2018). Shēngtài huánjìng bù zhàokāi bù wù huìyì shěnyì bìng yuánzé tōngguò “huánjìng wūrǎn qiángzhì zérèn bǎoxiǎn guǎnlǐ bànfǎ (cǎo'àn)” [The MEE reviewed and approved in principle the draft measures for the administration of compulsory environmental pollution liability insurance (Draft)]. https://www.mee.gov.cn/xxgk2018/xxgk/xxgk15/201805/t20180507_630147.html
- Ministry of Environmental Protection, & China Insurance Regulatory Commission [MEP & CIRC]. (2013). *MEP and the China Insurance Regulatory Commission jointly start pilot project of compulsory environmental pollution liability insurance*. Department of International Cooperation. https://english.mee.gov.cn/News_service/news_release/201302/t20130227_248567.shtml
- National Bureau of Statistics of China [NBS]. (2024). Biānzǎ shuōmíng: Dōng, zhōng, xībù hé dōngběi dìqū huàfēn [Editor's note: Division of eastern, central, western, and northeastern regions]. In China statistical yearbook. National Bureau of Statistics of China. <https://www.stats.gov.cn/sj/ndsjsj/2024/html/note.htm>
- Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods. *Frontiers in Education*, 5, 589965. <https://doi.org/10.3389/feduc.2020.589965>

- Rothschild, M., & Stiglitz, J. E. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4), 629–649. <https://doi.org/10.2307/1885326>
- Serafeim, G. (2020). Social-impact efforts that create real value. *Harvard Business Review*, 98(5), 38–48. <https://hbr.org/2020/09/social-impact-efforts-that-create-real-value>
- Shanghai Huazheng Index Information Service Co. (2024). Huázhèng ESG píngjí fāngfǎlùn [Huazheng ESG ratings methodology]. <https://www.chinadices.com/files/华证ESG评级方法论V2.1.pdf>
- Shavell, S. (1979). On moral hazard and insurance. *The Quarterly Journal of Economics*, 93(4), 541–562. <https://doi.org/10.2307/1884469>
- SynTao Green Finance. (n.d.). *ESG rating methodology*. Finance for Good. <https://en.syntaogf.com/pages/methodology>
- Tang, S., He, L., Su, F., & Zhou, X. (2024). Does directors' and officers' liability insurance improve corporate ESG performance? Evidence from China. *International Journal of Finance & Economics*, 29(3), 3713–3737. <https://doi.org/10.1002/ijfe.2849>
- Wu, W., Zhang, P., Zhu, D., Jiang, X., & Jakovljevic, M. (2022). Environmental pollution liability insurance of health risk and corporate environmental performance: Evidence from China. *Frontiers in Public Health*, 10, 897386. <https://doi.org/10.3389/fpubh.2022.897386>
- Yuan, R., Sun, J., & Cao, F. (2016). Directors' and officers' liability insurance and stock price crash risk. *Journal of Corporate Finance*, 37, 173–192. <https://doi.org/10.1016/j.jcorpfin.2015.12.015>

Appendix

Table A1. Sample selection process

Step	Filtration Criterion	Observations removed	Remaining observations
1	Initial A-share listed firms (2021–2024)	–	20,426
2	Exclude financial sector firms	500	19,926
3	Exclude Special Treatment (ST and *ST) firms	517	19,409
4	Exclude observations with missing data	224	19,185
–	Final Baseline Sample	–	19,185

Source: Author's own compilation.

Table A2. Alternative specifications of treatment timing

Variables	Contemporaneous	Lagged (t-1)	Post-only dummy
Treatment proxy	0.245*** (0.075)	0.220** (0.097)	0.264*** (0.079)
Controls	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	19,185	14,042	19,185

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column 2 uses a contemporaneous coverage-year indicator that equals 1 only in the fiscal year in which the disclosed ELI policy is contractually effective. Column 3 uses a one-year lagged ELI indicator. Column 4 uses a post-adoption indicator that equals 1 from the first disclosed ELI adoption year onward. All specifications include the same firm-level controls, firm fixed effects, and year fixed effects as the baseline model.

Source: Author's own estimations.

Table A3. Propensity score matching balance test

Variables	Mean (treated)	Mean (control)	SMD (%)
Panel A: Before matching			
ListAge	2.120	2.026	9.46%
Size	22.546	22.268	22.19%
Lev	0.396	0.402	3.23%
ROA	0.030	0.026	5.75%
Top1	0.349	0.321	18.85%
TobinQ	1.709	1.942	21.36%
Board	2.092	2.078	7.60%
Indep	0.378	0.380	3.91%
Inst	0.423	0.406	6.89%
Panel B: After matching			
ListAge	2.120	2.090	3.05%
Size	22.546	22.523	1.79%
Lev	0.396	0.391	2.25%
ROA	0.030	0.031	2.19%
Top1	0.349	0.350	0.26%
TobinQ	1.709	1.743	3.53%

Table A3 cont.

Variables	Mean (treated)	Mean (control)	SMD (%)
Board	2.092	2.090	0.89%
Indep	0.378	0.376	3.20%
Inst	0.423	0.415	3.30%

Note: Matching is conducted at the firm-year level using 1:1 nearest-neighbor matching with replacement and a caliper of 0.05. The matching procedure produces 420 matched treatment-control pairs, corresponding to 840 observation uses. Because matching is conducted with replacement, the matched regression sample contains 839 unique firm-year observations: 420 treated firm-year observations and 419 unique matched control firm-year observations.

Source: Author's estimations.

Table A4. Robustness check: Alternative ESG measure (Shangdao Ronglü ratings)

Variables	Alternative ESG performance
Insur	0.143*** (0.052)
Controls	Yes
Firm FE	Yes
Year FE	Yes
Observations	18,634
R-squared	0.279

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Source: Author's own estimations based on the Shangdao Ronglü ESG database.

Table A5. Robustness check: Adjusted environmental pillar (adjusted-E)

Variables	Adjusted-E
Insur	0.165*** (0.056)
Controls	Yes
Firm FE	Yes
Year FE	Yes
Observations	19,185
R-squared	0.218

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Adjusted-E is used only as a sensitivity measure and should not be interpreted as a reconstruction of the official proprietary Huazheng Environmental Pillar rating.

Source: Author's own estimations based on the Huazheng ESG database.

Table A6. Annual dynamics of ELI treatment status

Year	First-time adoption	Following-year observations under the baseline treatment definition	Discontinuation	Total treated
2021	7	0	0	7
2022	1	7	0	8
2023	176	1	7	177
2024	52	176	1	228
Total	236	184	8	420

Note: First-time adoption refers to the first fiscal year in which a firm discloses verified ELI contractual coverage during the sample period. Following-year observations under the baseline treatment definition refer to observations coded as treated in the year immediately following disclosed contractual ELI coverage; they should not be interpreted as verified current coverage or policy renewal unless renewed coverage is explicitly disclosed. Discontinuation refers to firms that were coded as treated in the previous year under the baseline Insur definition but are no longer coded as treated in the current year. Total treated observations sum to 420 firm-year observations under the baseline Insur definition.

Source: Author's own compilation based on Qinglü Data Platform.

Table A7. Robustness check: Alternative annual ESG aggregation

ESG aggregation method	Year-end ESG	Median ESG
Variables	ESG (1)	ESG (2)
Insur	0.251*** (0.048)	0.238*** (0.051)
Controls	Yes	Yes
Firm FE	Yes	Yes
Year FE	Yes	Yes
Observations	19,185	19,185

Note: Standard errors clustered at the firm level are reported in parentheses. All models include the listed controls, firm fixed effects, and year fixed effects unless otherwise stated. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. Column 2 uses the year-end Huazheng ESG rating. Column 3 uses the annual median of quarterly Huazheng ESG ratings. Both specifications include the same firm-level controls, firm fixed effects, and year fixed effects as the baseline model.

Source: Author's own estimations based on Huazheng ESG ratings, Qinglü Data Platform, CCER, and CSMAR databases.