Available Online: 26 August 2026 DOI: 10.54254/2977-5701/2026.36367

Model selection, market capitalization heterogeneity and ESG asset pricing: an empirical study of Chinese A-shares under the LSY four-factor framework

Zitao Wei

Inner Mongolia University of Finance and Economics, Hohhot, China

weizitao0317@163.com

Abstract. This study investigates whether factor-model selection drives the mixed evidence on ESG pricing in China's A-share market, where "green discount" and "green premium" coexist. Taking the Liu–Stambaugh–Yuan (LSY) four-factor model as the benchmark pricing framework and combining Shangdao Ronglv ESG ratings with CSMAR data, it examines model dependence, market capitalization heterogeneity, policy asymmetry, and risk transmission. The results show that ESG pricing conclusions are strongly model-dependent: the estimated ESG premium reverses sign once the local LSY factors and firm fundamentals are controlled, and statistically significant evidence of a premium is concentrated among large-cap stocks. The 2016 green finance policy is associated with dimensionally asymmetric changes in corporate ESG performance, although pre-existing trends limit strict causal identification. The environmental and governance dimensions transmit through opposite stock-price-risk paths, a pattern consistent with a dimensional hedging interpretation. These results imply that credible ESG pricing evidence in the A-share market requires an empirical specification—and above all a pricing benchmark—suited to the Chinese market.

Keywords: ESG asset pricing, LSY four-factor model, market capitalization heterogeneity

1. Introduction

Sustainable investment has moved into the mainstream of global capital markets: institutional ESG assets climbed from $13.3 trillion in 2012 to over $30 trillion in 2022. Whether better ESG performance earns higher risk-adjusted returns is among the most debated questions in empirical asset pricing. Evidence from China's A-share market is far less settled—some studies find a "green premium", others a "green discount"—a divergence attributed to sample periods, data vendors, and model settings, yet insufficiently understood.

This paper argues that factor-model selection is a key but neglected source of the divergence. Prior discussions of the mixed A-share evidence emphasize ESG measurement and sample differences—rating vendors, sample periods, and index construction—but pay less attention to whether the asset-pricing benchmark itself fits the Chinese market. The A-share market differs from European and U.S. markets in investor structure, trading rules, and property rights, and classic value and momentum factors are significantly weakened there. Liu et al. [1] demonstrate the deficiencies of the Fama–French three-factor model in China and propose the Liu–Stambaugh–Yuan (LSY) four-factor model, replacing book-to-market with VMG and momentum with PMO. Most A-share ESG studies nonetheless apply overseas factor models directly, so their ESG coefficients may absorb specification bias. Once the pricing model is chosen to fit the A-share market, does the mixed ESG evidence survive?

Accordingly, this study addresses four research questions: (1) Is the estimated ESG premium in China's A-share market sensitive to the specification of the asset-pricing model? (2) Does ESG pricing differ between large-cap and small-cap stocks? (3) Is the 2016 green finance policy associated with asymmetric changes across the environmental, social, and governance dimensions? (4) Through which risk-transmission channels do the individual ESG dimensions affect stock returns?

Combining Shangdao Ronglv ESG ratings with CSMAR data—a quarterly panel of 22,914 firm-quarters (2022Q2–2024Q2) and an annual panel of 21,894 firm-years (2015–2022)—this paper takes the LSY model as the benchmark framework throughout. Progressive regressions identify the model dependence of ESG pricing; a difference-in-differences design around the 2016 *Guiding Opinions on Building a Green Financial System* examines dimension-level changes associated with the policy; mediation and Sobel tests trace the risk channel; and portfolio sorts test heterogeneity across market capitalization.

The results are fourfold. First, ESG pricing in A-shares is strongly model-dependent: the ESG coefficient is significantly negative in a univariate specification but turns significantly positive once the LSY factors and firm fundamentals are included. Second, the 2016 green finance policy is associated with improved environmental performance but weaker governance, although pre-existing trends limit strict causal identification. Third, the environmental and governance dimensions transmit through opposite stock-price-risk paths, so the composite index exhibits no significant overall mediation. Fourth, statistically significant evidence of an ESG premium is concentrated among large-cap stocks.

This paper makes three contributions. First, it shifts the explanation for the mixed A-share ESG evidence from ESG measurement and sample issues to the pricing model itself, providing—to the best of the authors' knowledge—the first systematic evaluation of ESG pricing under the China-specific LSY four-factor benchmark. Second, it establishes market capitalization as a boundary condition for ESG pricing in China, showing that aggregate estimates can mask economically important cross-sectional heterogeneity. Third, it traces risk transmission at the dimension level, demonstrating that aggregating the environmental, social, and governance dimensions into a composite score can conceal offsetting channels—a lesson for both empirical design and ESG rating construction.

Section 2 reviews the literature; Section 3 develops the hypotheses; Section 4 describes the data; Section 5 reports the results; Section 6 discusses robustness and interpretation; Section 7 concludes.

2. Literature review

The theoretical literature offers two competing benchmarks for how ESG should be priced. Under the classical shareholder-value view [2], corporate ESG spending is at best a cost that the capital market should not reward. The theory of sustainable investment instead holds that ESG enters asset pricing through heterogeneous investor preferences: Pástor et al. [3] show that investors with sustainability tastes accept lower expected returns for holding green assets, so ESG generates differentiated equilibrium pricing even without incremental cash-flow information. These views yield opposite predictions, and the empirical debate largely attempts to discriminate between them.

A related literature prices ESG as risk rather than taste. Studies in this tradition link ESG to climate transition risk [4], document a carbon premium in the cross-section of stock returns [5], and show that high-

ESG firms suffer less severe stock price crashes [6], consistent with ESG as tail-risk insurance. The sign of the ESG–return relation thus depends on whether risk compensation or risk protection dominates in a given market and period.

A third group of studies traces how environmental behavior transmits to investment decisions and portfolio outcomes: environmental engagement shapes mutual fund allocation [7] and fund-level performance [8], ESG incidents propagate across borders and trigger price reactions in connected markets [9], and the portfolio gains from ESG integration can be nonlinear [10]. This literature also warns that measurement is a first-order problem: Berg et al. [11] show that measurement differences drive the low correlation among rating agencies, so the choice of data vendor can by itself distort estimated pricing relations.

Research on China documents an even less settled picture. A-share studies report a "green rather than red" anomaly [12] and systematic deviations from overseas pricing patterns [13], attributed to ESG uncertainty and rating noise [14] and to inconsistencies in sample, model, and index selection [15]. A related policy literature shows that market-oriented environmental regulation and green finance reshape corporate ESG behavior [16, 17], implying that Chinese firms' ESG profiles are partly policy-driven. Because investor structure, trading rules, and policy environment differ sharply from mature markets, evidence based on overseas factor models may not transport cleanly to the A-share cross-section, where retail dominance and policy shocks loom large.

Taken together, the literature shows that ESG pricing is sensitive to preferences, risk exposures, measurement, and policy shocks. Yet this work emphasizes ESG measurement and sample differences while paying less attention to whether the asset-pricing benchmark itself fits the Chinese market. If the benchmark factor model is misspecified for China, ESG coefficients absorb the resulting bias regardless of which channel truly operates. This paper addresses that gap by re-estimating the ESG–return relation under the A-share-specific LSY four-factor framework and tracing how inference changes relative to conventional specifications.

3. Theoretical foundation and hypotheses

The empirical design is reduced-form: each specification follows directly from one argument about how ESG relates to A-share stock returns. This section states the four arguments and the hypotheses they imply; the specifications are presented together with the results in Section 5.

First, ESG performance is systematically linked to traditional risk-factor exposures: high-ESG A-share firms are typically large-cap, growth-style companies. A regression that omits risk factors therefore blends ESG pricing with style effects, and the estimated coefficient need not reflect the underlying ESG–return relation. This reasoning yields Hypothesis H1: the estimated ESG premium in China's A-share market is sensitive to asset-pricing model specification, so estimates obtained without adequate factor controls can differ materially from those obtained under a locally appropriate factor benchmark.

Second, the cross-sectional strength of ESG pricing should vary with the information environment. Large-cap stocks are held mainly by institutional investors, attract intensive analyst coverage, and exhibit high information transparency, so ESG signals are quickly incorporated into prices; small-cap stocks are dominated by retail investors with limited attention to ESG and sparse analyst coverage, so ESG signals enter prices slowly or not at all. This yields Hypothesis H2: the ESG premium is significant among large-cap stocks but insignificant among small-cap stocks.

Third, the *2016 Guiding Opinions* concentrate on environmental instruments such as green credit and green bonds. Under the innovation-compensation view, environmental regulation can induce efficiency-improving innovation; under the resource-crowding-out view, the environmental investment demanded by the policy diverts limited resources away from other domains, including corporate governance. Because governance improvement requires long-term, sustained resources, the policy is expected to lift the environmental dimension while squeezing the governance dimension. This yields Hypothesis H3: following the 2016 green finance policy, treated firms exhibit differential gains in environmental performance and the composite ESG level relative to control firms, alongside a relative decline in governance performance.

Fourth, the environmental and governance dimensions affect stock price risk in opposite directions. Environmental compliance reduces tail risks such as regulatory penalties and production suspensions, lowering stock price risk; governance deterioration—for example, weakened constraints on controlling shareholders under resource pressure—raises managerial uncertainty and stock price risk. When the two paths are aggregated into a composite ESG index, they offset each other. This yields Hypothesis H4: the overall mediating effect of the composite ESG index through stock price risk is insignificant, while the environmental and governance dimensions exhibit significant mediating paths in opposite directions.

4. Data

This study constructs two panel datasets matched to different empirical objectives. The first is a quarterly panel from 2022Q2 to 2024Q2—the period over which Shangdao Ronglv publishes quarterly ESG scores—containing 22,914 A-share non-financial firm-quarter observations; it supports the short-term pricing tests, including the baseline, portfolio, robustness, and identification analyses. The second is an annual panel from 2015 to 2022 containing 21,894 firm-year observations, which fully covers the periods before and after the 2016 green finance policy and is used for the DID, parallel-trend, and mediation analyses.

Four screening rules ensure data quality: (i) financial firms (monetary finance, capital market services, and insurance, per the 2012 CSRC industry classification) are excluded because their accounting rules and risk characteristics differ from real-sector firms; (ii) ST and *ST special-treatment firms are excluded; (iii) observations with missing core variables (ESG score, return, market value, price-earnings ratio, or ROE) are dropped; and (iv) all continuous variables are winsorized at the 1st and 99th percentiles. ESG ratings come from the Shangdao Ronglv database, stock returns and financial data from CSMAR, and the LSY risk factors from Liu et al. [1]; the data sources are summarized in Appendix Table A1. The rating system covers about 130 underlying indicators across the environmental, social, and governance dimensions; the quarterly numerical score (ESG_Num) takes integer values from 1 to 7, and the annual dimension scores (E_Score, S_Score, G_Score) range from 0 to 100.

The dependent variable is the quarterly excess return, defined as the difference between the firm's quarterly return and the risk-free rate (the three-month Treasury yield to maturity), as shown in Equation (1):

$$ExRet_{i,t} = R_{i,t} - RF_t \tag{1}$$

The core explanatory variable is ESG_Num. Two alternative measures are used in robustness tests: ESG_High, a dummy equal to one when ESG_Num is at least 5, and ESG_Pctl, the firm's ESG_Num percentile within year-industry. The annual dimension scores E_Score, S_Score, and G_Score (0–100) enter the policy and mediation analyses. For the DID design, Treat equals one for firms in the top 30% of pre-policy (2015–2016) E_Score and zero for the bottom 30%, with the middle 40% removed to sharpen the contrast; Post equals one from 2017 onward; and their interaction DID = Treat × Post captures the post-2016 differential change in the treatment group relative to the control group. The mediating variable ST_RISK measures firm-level stock price risk, constructed as the standard deviation of daily stock returns within the year, with higher values indicating higher stock price risk. Control variables are ln(MarketCap), PE_TTM, and ROE. Risk adjustment relies on the LSY four factors: MKT (market), VMG (value-growth, replacing the traditional

HML), SMB (size), and PMO (momentum, replacing the traditional MOM). All panel regressions include firm fixed effects and time (quarter or year) fixed effects, with standard errors clustered at the firm level.

Table A2 reports descriptive statistics for the quarterly panel. The mean quarterly excess return is −0.0303, reflecting the weak market conditions over the sample window; the mean ESG_Num is 3.69 on the 1–7 scale; and the wide dispersion of ln(MarketCap) confirms that the sample covers large-, mid-, and small-cap firms. The Pearson correlation matrix, reported in Table A3, shows that ESG_Num is negatively correlated with excess returns (−0.05, significant at the 10% level) before any risk adjustment, and that the strong negative correlation between VMG and SMB (−0.93) is an inherent feature of the LSY factor construction rather than a data-quality problem.

5. Empirical results

5.1. Model dependence of ESG pricing

To test H1, this study estimates the baseline panel model in Equation (2):

$$ExRet_{i,t} = \alpha + \beta_1 ESG_{i,t} + \gamma' Controls_{i,t} + \lambda_i + \mu_t + \varepsilon_{i,t} \tag{2}$$

where Controls includes ln(MarketCap), PE_TTM, ROE, and the LSY four factors, λ_i denotes firm fixed effects, μ_t denotes quarter fixed effects, and standard errors are clustered at the firm level. Equation (2) is estimated in three steps: M1 includes only ESG_Num and the two-way fixed effects; M2 adds the LSY factors; and M3 adds the fundamental controls.

Specification M2 makes the LSY factor structure explicit (Equation (3)):

$$ExRet_{it} = \alpha + \beta_1 ESG_{it} + \gamma' LSY4_t + \delta' Controls_{it} + \mu_i + \varepsilon_{it} \tag{3}$$

Table A4 reports the progressive estimates. In the univariate specification M1, the ESG_Num coefficient is −0.0298 and significant at the 1% level: without any risk adjustment, better ESG performance is associated with lower quarterly excess returns—an apparent green discount. The fit of M1 is extremely poor (R^2 = 0.0068). Once the LSY four factors are introduced in M2, the coefficient flips to 0.0058 and becomes significantly positive at the 5% level; after the fundamental controls are added in M3, it rises further to 0.0098, significant at the 1% level, with R^2 reaching 0.3696. The LSY factors themselves are highly significant throughout. This sign reversal confirms H1.

5.2. Asymmetric effects of the 2016 green finance policy

To test H3, this study constructs the policy interaction term (Equation (4)):

$$DID_{i,t} = Treat_i \times Post_t \tag{4}$$

and estimates the DID model on the annual panel (Equation (5)):

$$Y_{i,t} = \alpha + \delta DID_{i,t} + \theta Treat_i + \rho Post_t + \lambda_i + \mu_t + \varepsilon_{i,t} \tag{5}$$

where the dependent variable Y is in turn E_Score, ESG_Score, S_Score, and G_Score, and δ measures the net policy effect on the treatment group relative to the control group.

Table A5 reports the estimates. The policy coefficient is 5.0729 for E_Score and 1.9697 for the composite ESG score, both significant at the 1% level; it is 0.4256 and insignificant for S_Score; and it is −1.0465, significant at the 1% level, for G_Score. The estimates thus indicate dimensionally asymmetric post-2016 changes in the treatment group relative to the control group—stronger environmental improvement, no discernible social change, and a relative governance decline—associated with the green finance policy environment, consistent with H3. Because the event-study and placebo tests (Section 6.1) indicate pre-existing trends, these differential changes should not be interpreted as the isolated causal effect of the 2016 policy.

5.3. Risk transmission: mediation tests

To test H4, this study follows the Baron–Kenny three-step procedure (Equations (6a)–(6c)):

$$ExRet = c \cdot ESG + Controls + \varepsilon \tag{6a}$$

$$ST\_RISK = a \cdot ESG + Controls + \varepsilon \tag{6b}$$

$$ExRet = c' \cdot ESG + b \cdot ST\_RISK + Controls + \varepsilon \tag{6c}$$

The total effect decomposes into direct and indirect parts (Equation (7)):

$$c = c' + a \times b \tag{7}$$

and the indirect effect $a \times b$ is evaluated with the Sobel statistic in Equation (8):

$$Z = \frac{a \cdot b}{\sqrt{a^2 \cdot SE_b^2 + b^2 \cdot SE_a^2}} \tag{8}$$

For the composite ESG index (Table A6), the total effect of ESG_Num on excess returns is −0.0057 (significant at the 5% level); ESG_Num significantly reduces stock price risk (path a = −0.002484); and the risk variable enters the return equation positively (path b = 0.1167, significant at the 10% level). The Sobel statistic, however, is insignificant (p = 0.1594), so no overall mediating effect exists for the composite index.

Table A7 decomposes the mediation by dimension; the corresponding path-a regressions are reported in Table A8. In the annual panel (Table A8, Panel A), the environmental and governance dimension scores are significantly associated with firm-level stock price risk (0.000096 and 0.000891 respectively, both significant at the 1% level), while the social dimension and the composite score show no significant risk association. Table A8, Panel B reports the DID policy effects by dimension, which reproduce the dimension-level DID estimates of Table A5. For the environmental dimension, the indirect effect is significant at the 5% level (Sobel p = 0.0206). For the governance dimension, the indirect effect is again significant but runs in the opposite direction (Sobel p = 0.0002). The social dimension exhibits no significant risk transmission (Sobel p = 0.4721). The two significant paths offset each other, so the composite index shows no overall mediation—a pattern consistent with the dimensional hedging interpretation of H4.

5.4. Market capitalization heterogeneity

To test H2, firms are sorted into ESG quintiles and, independently, into two groups by the median of total market capitalization, forming 2 × 5 portfolios. For each portfolio p, the time-series regression in Equation (9) is estimated:

$$ExRet_{p,t} = \alpha_p + \beta_{MKT} MKT_t + \beta_{VMG} VMG_t + \beta_{SMB} SMB_t + \beta_{PMO} PMO_t + \varepsilon_{p,t} \tag{9}$$

and the risk-adjusted alphas of the high-minus-low ESG portfolios are compared across the two market-capitalization groups.

Table A9 reports the LSY four-factor alphas. Within the small-cap group, the alphas of the five ESG portfolios show no discernible pattern, and the high-minus-low ESG alpha spread is 0.0075 and insignificant: there is no statistically significant evidence of an ESG premium among small-cap stocks. Within the large-cap group, by contrast, the high-minus-low spread is 0.0140 and significant at the 5% level. This contrast supports H2.

6. Discussion

6.1. Robustness and identification

The core results are robust to alternative measurement and sampling choices (details in the Appendix). Replacing ESG_Num with the high-ESG dummy or the industry percentile in the baseline specification leaves coefficient signs unchanged, although significance declines as discretization discards information (Table A10). Specifically, the continuous ESG_Num specification (Table A10, Panel A) reproduces the benchmark estimates; the binary ESG_High coefficient (Panel B) is positive but statistically insignificant; and the ESG_Pctl coefficient under the full-control specification (Panel C) is 0.0161, significant at the 10% level. Re-estimating the model without firm fixed effects—pooled OLS with the same sample, controls, and clustered standard errors—continues to yield significantly positive ESG coefficients for all three measures (0.0065 to 0.0262), and the 2023 subsample gives 0.0103, also significant (Table A11). Relative to the fixed-effects benchmark, the pooled OLS coefficient on ESG_Num is smaller in magnitude (0.0065 versus 0.0098), indicating that part of the cross-sectional variation in ESG scores reflects time-invariant firm heterogeneity that the fixed effects absorb (Table A11, Panel C).

For the DID design, the parallel-trend assumption is examined with an event-study specification using 2016 as the reference year (Equation (10)):

$$Y_{i,t} = \alpha + \sum\nolimits_{\tau=-2,\tau\neq-1}^{5} \delta_\tau (Treat_i \times Year_t^\tau) + \lambda_i + \mu_t + \varepsilon_{i,t} \tag{10}$$

The event-study estimates (Table A12) show that the pre-policy coefficient for 2015 is 4.7265 and significant at the 1% level, so the treatment group's environmental scores were already rising before the policy; a placebo test with an artificial 2015 policy date (Table A13) likewise yields a significant FakeDID coefficient of 6.9536. The DID estimates should thus be read as the combined effect of the 2016 policy, pre-existing regulations, and firms' long-term green transformation—a characterization of structural change rather than the isolated causal effect of a single policy. This study therefore emphasizes the direction and economic meaning of the effects instead of strong causal claims.

Separately, the ESG coefficient is allowed to shift after 2016 (Equation (11)):

$$ExRet_{i,t} = \alpha + \beta_1 ESG_{i,t}^{dm} + \beta_2 (ESG_{i,t}^{dm} \times Post_t) + \theta Post_t + \gamma\prime Controls_{i,t} + \varepsilon_{i,t} \tag{11}$$

The estimates (Table A14) show that the post-2016 dummy is significantly positive (0.0398) while the ESG × Post2016 interaction is insignificant: the post-2016 period is associated with a higher absolute level of ESG performance, but the pricing relation linking ESG to returns is not detectably altered.

Finally, to address reverse causality and omitted variables, ESG_Num is instrumented with the leave-one-out mean ESG_Num of firms in the same CSRC first-level industry and year, and the two-stage model in Equations (12) and (13) is estimated:

$$ESG_{i,t} = \pi_0 + \pi_1 IV_{i,t} + \gamma' Controls + \upsilon_{i,t} \tag{12}$$

$$ExRet_{i,t} = \alpha + \delta \widehat{ESG}_{i,t} + \gamma' Controls + \varepsilon_{i,t} \tag{13}$$

Table A15 reports the 2SLS estimates. The first stage is strong: the instrument coefficient is 0.6493 with an F-statistic of 960.25, far above the weak-instrument threshold. In the second stage, however, the fitted ESG coefficient is −0.0072 and insignificant (p = 0.152). This paper reports this null result transparently, for three reasons. First, the exclusion restriction is structurally fragile: industry-wide common shocks—such as tightening environmental regulation or industry technology cycles—can affect returns directly while also moving industry-average ESG. Second, 2SLS estimates a local average treatment effect for firms whose ESG responds to peer pressure, which need not coincide with the average effect. Third, the first stage explains only about 17% of ESG variation (R^2 = 0.1683), so the fitted values dilute much of the useful variation. Future work could improve identification with regional or peer-group instruments. Taken together, the fixed-effects benchmark, the DID evidence, and the dimension-level mediation results leave the paper's core conclusions intact.

6.2. Interpretation

Why is A-share ESG pricing so sensitive to model choice? High-ESG A-share firms are systematically large-cap, growth-style companies, so any specification that omits risk factors loads style exposures onto the ESG coefficient and manufactures a spurious discount; once risk is stripped out with the local LSY factors, the estimated ESG premium turns positive. The comparison between pooled OLS and fixed-effects estimates (Table A11) reinforces this conclusion from a different angle. High-ESG firms are not randomly distributed: they tend to possess time-invariant quality attributes—superior management, standardized internal control, and stronger competitive positions—correlated with ESG scores. Pooled OLS loads part of this time-invariant heterogeneity onto the ESG coefficient, so controlling for it through firm fixed effects materially changes the estimated relationship: the ESG_Num coefficient rises from 0.0065 to 0.0098 and is identified from within-firm ESG improvements rather than cross-firm comparisons. Firm fixed effects thus serve as an essential "firm quality filter" in A-share ESG pricing research—strengthening, rather than reversing, the estimated premium. The larger specification effect, however, comes from the factor benchmark: it is the omission of the local LSY factors, not of firm fixed effects, that manufactures the spurious "green discount" reported in parts of the China literature [12, 13], reconciling the divergence with equilibrium-based premium predictions [3].

The dimension-level results are likewise consistent with a structural interpretation. Environmental improvement lowers tail risks such as regulatory penalties and production suspensions, forming an "environmental improvement → risk reduction → return improvement" path; policy-induced resource pressure, by contrast, weakens governance, raises managerial uncertainty and risk, and drags down returns—the mirror image of the crowding-out effect in the DID estimates. Because the two paths cancel when aggregated, composite-score mediation tests mask dimension-level transmission logic and can produce misleading inferences. Dimension-level analysis is therefore essential for understanding ESG risk transmission in emerging markets, and rating agencies should account for the internal offsetting among dimensions when constructing composite scores.

7. Conclusion

Combining a quarterly panel (2022Q2–2024Q2) with an annual panel (2015–2022) around the 2016 green finance policy under the LSY four-factor benchmark, this paper examines A-share ESG pricing along four dimensions: model dependence, market capitalization heterogeneity, policy asymmetry, and risk transmission. The central takeaway is that ESG pricing conclusions in China's A-share market are sensitive to empirical specification. The estimated ESG premium is significantly negative without risk controls but turns significantly positive once the local LSY factors and firm fundamentals are included, indicating that the apparent "green discount" largely reflects uncontrolled style exposures; a locally appropriate factor framework is therefore a prerequisite for credible ESG pricing evidence. The premium also exhibits key heterogeneity:

statistically significant evidence is concentrated among large-cap stocks, and dimension-level analysis reveals offsetting environmental and governance risk-transmission paths that the composite score conceals. For the 2016 green finance policy, the DID estimates indicate dimensionally asymmetric post-2016 changes associated with the policy environment; given the pre-existing trends documented by the parallel-trend and placebo tests, however, these estimates do not admit a strict causal interpretation and are best read as reflecting broader structural change.

These findings carry practical implications. Researchers should treat local factor models and firm fixed effects as essential elements of A-share ESG designs, and decompose ESG into dimensions before drawing mechanism inferences. Investors should condition ESG strategies on market capitalization rather than importing overseas premium conclusions mechanically. Regulators may consider pairing environmental incentives with governance safeguards, and rating agencies should account for dimensional offsetting when constructing composite scores.

The study has several limitations: the quarterly ESG scores span only from 2022Q2; the DID design cannot fully isolate the 2016 policy from pre-existing trends and overlapping regulations; and the industry-mean instrument cannot fully satisfy the exclusion restriction. Future research can exploit longer panels, examine ownership and pollution-intensity heterogeneity, and develop stronger instruments.

References

[1] Liu, J., Stambaugh, R. F., & Yuan, Y. (2019). Size and value in China. *Journal of Financial Economics, 134*(1), 48–69.

[2] Friedman, M. (1970, September 13). The social responsibility of business is to increase its profits. *The New York Times Magazine*, 32–33, 122, 124, 126.

[3] Pástor, Ľ., Nagel, S., & Taylor, A. L. (2020). Sustainable investing in equilibrium. *Journal of Financial Economics, 142*(2), 550–571. doi: 10.1016/j.jfineco.2020.12.011

[4] Fliegel, P. (2025). "Brown" risk or "green" opportunity? The dynamic pricing of climate transition risk on global financial markets. *Energy Economics, 145*, 108456. doi: 10.1016/j.eneco.2025.108456

[5] Hambel, C., & van der Sanden, F. (2025). Reevaluating the carbon premium: Evidence of green outperformance. *International Review of Financial Analysis, 102*, 104042.

[6] Chai, H., Cheng, Z., & Wu, W. (2025). Is ESG performance a protective umbrella for ESG violations? *International Review of Financial Analysis, 98*, 103858.

[7] Qian, S., You, H., & Wan, D. (2025). From pro-environmental behavior to ESG fund investing: Evidence from account-level data in China. *Energy Economics, 143*, 108271.

[8] Lin, J., Pan, D., & Sha, Y. (2025). The impact of ESG investment on fund performance: Evidence from mutual fund style drift. *Pacific-Basin Finance Journal, 91*, 102707.

[9] Orpiszewski, T., & Thompson, M. (2025). Beyond borders: Asset price reaction to ESG incidents at home and abroad. *International Review of Financial Analysis, 104*, 104246.

[10] Du, P., Gu, R., Luo, L., & Sun, Q. (2025). The enhanced gain effects of ESG's non-linearity on portfolios: An asset pricing tree model perspective. *International Review of Financial Analysis, 99*, 103971.

[11] Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance, 26*(6), 1315–1344.

[12] He, F., Du, H. Y., & Hao, J. (2025). "Green" rather than "Red": ESG investment anomalies in China's stock market. *Chinese Journal of Management Sciences, 28*(7), 75–95. doi: 10.19920/j.cnki.jmsc.2025.07.005

[13] Lee, L. (2025). Is there an ESG risk premium in China's A-share market? Empirical analysis based on factor model. *Journal of Financial Theory and Practice*, (4), 97–107.

[14] Yan, J. Z., Luo, Q., & Li, Z. F. (2025). Research on the impact of ESG uncertainty and noise on sustainable investment. *Systems Engineering — Theory and Practice*. Advance online publication.

[15] Zhuan, Z., & Wang, C. Y. (2025). Bibliometric analysis of hot spots and frontiers of ESG: A case study of UTD24 journal. *Friends of Accounting*, (7), 104–112.

[16] Sang, J. F., & Zhang, L. J. (2025). Market-driven environmental regulation and corporate green governance: A dual machine learning test based on ESG high rating. *Statistical Research, 42*(11), 86–101. doi: 10.19343/j.cnki.11-1302/c.2025.11.007

[17] Song, J. B., Dong, Z. N., & Cao, W. C. (2025). Research on the construction of valuation system with Chinese characteristics under the guidance of market value management assessment of central enterprises. *Finance and Accounting*, (2), 48–51.

Appendix

Table A1. Summary of data sources

Categories of Data	Specific data	Time frame	Sources
ESG Ratings	Shangdao Ronglv ESG Quarterly Digital Score (ESG_Num, level 1-7)	2022Q2-2024Q2	Shangdao Rong Green Database
ESG Year	E/S/G component score (E_Score/S_Score/G_Score)	2015-2022	Shangdao Rong Green database
Stock returns	Quarterly returns	2015-2024	CSMAR database
Financial data	Market value, PE, ROE, etc.	2022Q2-2024Q2	CSMAR database
Risk factors	LSY four factors (MKT/VMG/SMB/PMO)	2015-2024	Liu et al. [1]
Stock price risk (ST_RISK)	Daily stock returns (for computing ST_RISK)	2015-2022	CSMAR database
Policy Data	Green finance policy for 2016	2015-2022	Documents from seven ministries and commissions

Table A2. Descriptive statistics of variables

Variable	*N*	Mean	Std.Dev	Min	P25	Median	P75	Max
ExRet	22,914	-0.0303	0.1834	-0.8930	-0.1457	-0.0369	0.0622	1.6581
ESG_Num	22,914	3.6914	1.0614	1.0000	3.0000	3.0000	4.0000	7.0000
ESG_Pctl	22,914	0.5070	0.2732	0.0004	0.2866	0.3641	0.7284	0.9998
ln(MarketCap)	22,914	4.3056	1.0808	1.5495	3.5229	4.1286	4.9386	7.7083
PE_TTM	22,914	43.9689	88.2667	-414.85	17.24	29.00	51.17	649.97
ROE	22,914	5.4561	5.8663	-7.7983	1.5620	3.9073	7.9719	46.006

Note: The sample period is 2022Q2-2024Q2, with a total of 22,914 firm-quarter observations. ExRet is the quarterly excess return (adjusted for the risk-free rate); ESG_Num is the Shangdao Ronglv ESG numerical score (1-7); ESG_Pctl is the ESG rating percentile; ln(MarketCap) is the natural logarithm of total market capitalization; PE_TTM is the rolling P/E ratio; ROE is return on equity (%).

Table A3. Pearson correlation coefficient matrix of variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) ExRet	1.00									
(2) ESG_Num	-0.05*	1.00								
(3) ESG_Pctl	-0.02	0.96***	1.00							
(4) MKT	0.24***	0.03*	-0.01	1.00						
(5) VMG	0.33***	0.10***	-0.01	0.08***	1.00					

Table A3. Continued

(6) SMB	0.29***	0.08***	0.01	0.15***	-0.93***	1.00				
(7) PMO	0.21***	0.04*	0.01	0.26***	0.20***	0.32***	1.00			
(8) ln(MarketCap)	0.05***	0.43***	0.41***	0.01	0.06***	0.07***	0.04*	1.00		
(9) PE_TTM	0.06***	0.05*	0.05*	0.02	-0.02	0.01	-0.01	0.05*	1.00	
(10) ROE	0.06***	0.09***	0.10***	0.21***	0.36***	0.33***	0.09***	0.30***	0.14***	1.00

Note: Pearson correlation coefficients are reported in the table. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively. The sample period is 2022Q2-2024Q2.

Table A4. ESG performance and excess stock return (LSY four-factor model)

variable	M1 ESG Only	M2 + LSY4	M3 + Controls
ESG_Num	-0.0298***	0.0058**	0.0098***
	(0.0024)	(0.0024)	(0.0025)
MKT		0.9278***	0.1990***
		(0.0342)	(0.0368)
VMG		0.2508***	1.6707***
		(0.0741)	(0.0755)
SMB		0.7649***	0.8245***
		(0.0758)	(0.0771)
PMO		0.1426***	0.1841***
		(0.0285)	(0.0234)
ln(MarketCap)			0.4412***
			(0.0087)
PE_TTM			0.0001***
			(0.0000)
ROE			0.0049***
			(0.0003)
Constant term	0.0796***	0.0175*	1.8651***
	(0.0087)	(0.0094)	(0.0376)
Observed values of *N*	22,914	22,914	22,914
R^2	0.0068	0.1857	0.3696
Individual fixed effects	is	is	is
The standard error of clustering	Firm-level	The corporate level	The corporate level

Note: The dependent variable is the quarterly excess return ExRet. ESG_Num is the Shangdao Ronglv ESG numerical score. The four factors are the LSY factors MKT (market), VMG (value-growth), SMB (size), and PMO (momentum). Control variables include ln(MarketCap) (logarithm of market capitalization), PE_TTM (rolling P/E ratio), and ROE (return on equity). Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A5. Green financial reform of ESG performance influence

The dependent variable	E_Score	ESG_Score	S_Score	G_Score
DID (Treat × Post)	5.0729***	1.9697***	0.4256	-1.0465***
	(0.3296)	(0.2820)	(0.4340)	(0.2955)
Post	7.4018***	6.1011***	6.0633***	4.7211***
	(0.2416)	(0.1992)	(0.2998)	(0.2175)
N	21,894	21,894	21,894	21,894
R^2	0.1524	0.1955	0.0975	0.0618
Individual fixed effects	is	is	is	is

Note: The estimates are based on the quasi-natural experiment of the 2016 "*Guiding Opinions on Building a Green Financial System*". Treat=1 indicates firms in the top 30% of pre-policy (2015-2016) E_Score, and Post = 1 indicates the post-policy period (2017-2022). DID = Treat × Post. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A6. Mediating effect of ST_RISK (Baron-Kenny three-step method +Sobel test)

Path	Variables	Coefficients	Standard error	*p*-value	Significance
Step 1: ESG- > ExRet (total effect c)	ESG_Num	-0.0057	0.0021	0.0073	**
Step 2: ESG- > ST_RISK (path a)	ESG_Num	-0.002484	0.001085	0.0220	**
Step 3: ESG- > ExRet (direct effect c ')	ESG_Num	-0.0054	0.0021	0.0156	**
Step 3: ST_RISK- > ExRet (path b)	ST_RISK	0.1167	0.0654	0.0744	*
Sobel test	The indirect effect	-0.000290	0.000206	0.1594	No significant
Proportion of intermediaries		5.06%			

Note: Sample *N* = 5,538 (subsample with ST_RISK data). ST_RISK is firm-level stock price risk, measured as the standard deviation of daily stock returns within the year. Intermediary ratio = indirect effect/total effect. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A7. Sobel test results for each E/S/G dimension

Independent variable	Indirect effects	SobelZ	Sobelp value	Significance
ESG_Score	-0.000290	-1.4069	0.1594	Not significant
E_Score	-0.000190	-2.3145	0.0206	**
S_Score	0.000027	0.7191	0.4721	Not significant
G_Score	-0.001968	-3.6908	0.0002	***

Note: The mediation path is X- > ST_RISK- > ExRet. The mediating paths of E_Score and G_Score through ST_RISK are significant, whereas the mediating effect of the composite ESG score is not significant, possibly because the E and G dimensions mediate in opposite directions and offset each other. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A8. Effects across E/S/G dimensions

Panel A: E/S/G of ST_RISK of each dimension				
Independent variable	Effect of ST_RISK	Standard errors	N	R^2
E_Score	0.000096***	(0.000035)	28,180	0.0003
S_Score	0.000052	(0.000053)	28,180	0.0000
G_Score	0.000891***	(0.000059)	28,180	0.0080
ESG_Score	0.000083	(0.000058)	28,180	0.0001
Panel B: DID policy effects across E/S/G dimensions				
Dependent variable	DID coefficient	N	R^2	
E_Score	5.0729***	21,894	0.1524	
S_Score	0.4256	21,894	0.0975	
G_Score	-1.0465***	21,894	0.0618	

Note: Panel A uses annual panel data (N = 28,180), and the dependent variable is ST_RISK. Panel B is the same as Table 5. Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A9. ESG Portfolio ranking Alpha test (LSY four-factor model)

combination	Alpha	MKT	VMG	SMB	PMO	R^2	N
Small×ESG1	0.0153	0.5759	-0.6665	0.8318	-0.3172	0.2729	2,519
Small×ESG2	-0.0230***	1.1275	0.5396	1.8771	-0.1961	0.2983	2,984
Small×ESG3	0.0476***	0.2094	-1.4099	0.3070	-0.5423	0.2807	2,752
Small×ESG4	-0.0150	0.8890	-0.1133	1.2845	-0.1932	0.2787	2,181
Small×ESG5	-0.0119	0.9812	-0.1132	1.2941	-0.0989	0.2701	1,022
Big×ESG1	0.0340***	0.8961	-0.6771	-0.1012	-0.2318	0.1269	1,981
Big×ESG2	0.0105	0.9528	0.0459	0.7257	-0.0440	0.0860	1,516
Big×ESG3	0.0188	0.6555	-0.5618	0.1054	-0.2426	0.0972	1,786
Big×ESG4	0.0112	0.9293	-0.3196	0.1173	-0.1316	0.1095	2,479
Big×ESG5	0.0042	1.2261	-0.4058	-0.0176	0.0585	0.1584	3,694
Small×(ESG5-1)	0.0075	0.6986	-0.4987	0.9711	-0.2649	0.2702	3,541
Big×(ESG5-1)	0.0140**	1.1123	-0.4929	-0.0372	-0.0326	0.1420	5,675

Note: Portfolios are formed by sorting firms into ESG_Num quintiles (Q1 = lowest, Q5 = highest) and, independently, into two groups by the median of total market capitalization (Small/Big), yielding 2 × 5 = 10 portfolios. Within each portfolio, the time-series regression ExRet = Alpha + beta_MKT×MKT + beta_VMG×VMG + beta_SMB×SMB + beta_PMO×PMO + epsilon is estimated. The last two rows report the alpha spread between the high-ESG and low-ESG portfolios (ESG5-1). Robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A10. ESG metrics

Panel A: Continuous ESG metrics (ESG_Num)		
variable	M1 (+LSY4)	M2 (+Controls)
ESG_Num	0.0058**	0.0098***
	(0.0024)	(0.0025)
MKT	0.9278***	0.1990***
	(0.0342)	(0.0368)
VMG	0.2508***	1.6707***
	(0.0741)	(0.0755)
SMB	0.7649***	0.8245***
	(0.0758)	(0.0771)
PMO	0.1426***	0.1841***
	(0.0285)	(0.0234)
ln(MarketCap)		0.4412***
		(0.0087)
PE_TTM		0.0001***
		(0.0000)
ROE		0.0049***
		(0.0003)
N	22,914	22,914
R^2	0.1857	0.3696
Panel B: Virtual ESG metrics (ESG_High = 1 if ESG_Num ≥ 5)		
Variables	M1 (+LSY4)	M2 (+Controls)
ESG_High	0.0001	0.0080
	(0.0051)	(0.0055)
MKT	0.9131***	0.1776***
	(0.0339)	(0.0368)
VMG	0.2675***	1.6978***
	(0.0744)	(0.0759)
SMB	0.7309***	0.8772***
	(0.0754)	(0.0772)
PMO	0.1589***	0.2085***
	(0.0281)	(0.0232)
N	22,914	22,914
R^2	0.1854	0.3690
Panel C: Sorted ESG Metrics (ESG_Pctl)		
Variables	M1 (+LSY4)	M2 (+Controls)
ESG_Pctl	0.0077	0.0161*
	(0.0091)	(0.0095)

Table A10. Continued

MKT	0.9142***	0.1788***
	(0.0339)	(0.0368)
VMG	0.2648***	1.6883***
	(0.0743)	(0.0758)
SMB	0.7345***	0.8704***
	(0.0754)	(0.0772)
PMO	0.1585***	0.2107***
	(0.0280)	(0.0232)
N	22,914	22,914
R^2	0.1855	0.3690

Note: The dependent variable is the quarterly excess return ExRet. Three ESG measures are used: the continuous variable ESG_Num (1-7), the dummy variable ESG_High (equal to 1 if ESG_Num ≥ 5), and the sorted variable ESG_Pctl (percentile, 0-1). Control variables are the same as in Table 4. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A11. Replace ESG measure and 2023 subsample

Panel A: replace ESG measure				
ESG metrics	Coefficient	Standard error	*N*	R^2
ESG_Num (continuous)	0.0065***	(0.0012)	22,914	0.1693
ESG_High (virtual)	0.0174***	(0.0030)	22,914	0.1694
ESG_Pctl (sort)	0.0262***	(0.0045)	22,914	0.1695
Panel B: 2023 subsample				
Variables	Coefficients	Standard error	*N*	R^2
ESG_Num	0.0103***	(0.0016)	12,502	0.1144

Panel C: benchmark regression (Table 4 M3) vs robustness test (Table 11 Panel A) coefficient of contrast			
Dimensions of comparison	Table 4 M3 (benchmark regression)	Table 11 Panel A (Robustness test)	Explanation of differences
Model	xtreg, fe (plus Entity FE)	reg (OLS, without Entity FE)	FE switch
ESG_Num coefficient	+ 0.0098 ***	+ 0.0065 ***	Same sign, attenuated magnitude
Control variables	MKT+VMG+SMB+PMO +lnMcap+PE+ROE	Exactly the same	consistent
Clustered standard errors	cluster(Scode)	cluster(Scode)	Consistent
Sample size	22,914	22,914	Same

R^2	0.3696	0.1693	FE model of higher explanatory power

Notes: In Panel A and B, the explained variable is the quarterly excess return ExRet, and all specifications control the LSY four factors and the control variables. Both panels are estimated by pooled OLS without firm fixed effects: Panel A reports robustness tests with alternative ESG measures, and Panel B uses the 2023 subsample only. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively. Panel C reveals the key findings: the two regressions use the same sample (N = 22,914), the same combination of control variables (the LSY four factors + lnMcap + PE + ROE), and the same clustered standard errors (cluster(Scode)); the only difference is whether firm fixed effects are controlled. This single difference in model specification materially reduces the magnitude of the ESG coefficient, although it remains significantly positive.

Table A12. Parallel Trend Test (event study method, with 2016 as the reference group)

Variables	Coefficients	Standard error	*p*-value	Significance
Pre2 (2015)	4.7265	0.4015	0.0000	***
Post0 (2017)	6.7670	0.6045	0.0000	***
Post1 (2018)	7.8523	0.5146	0.0000	***
Post2 (2019)	7.9190	0.5397	0.0000	***
Post3 (2020)	10.0929	0.5546	0.0000	***
Post4 (2021)	11.5357	0.6080	0.0000	***
Post5 (2022)	34.7813	0.7815	0.0000	***
N	21,894			
R^2	0.2530			

Note: E_Score is the dependent variable. The reference group is 2016 (the year before the policy). Pre2 represents two periods before the policy (2015), and Post0-Post5 represent the periods after the policy (2017-2022). Firm fixed effects are controlled. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A13. Replacing the dependent variable and Placebo test

Panel A: Replacing the dependent variable			
Dependent variable	DID coefficient	N	$R2$
ESG_Score	1.9697***	21,894	0.1955
S_Score	0.4256	21,894	0.0975
G_Score	-1.0465***	21,894	0.0618

Panel B: Placebo test (fake policy time point = 2015)	
Variables	Coefficient of
FakeDID	6.9536***
	(0.5180)
N	21,894
R^2	0.1254

Note: Panel A uses ESG scores of different dimensions as dependent variables to examine the robustness of the policy effects. Panel B reports a placebo test that moves the policy date forward to 2015: if the treatment effect is genuine, FakeDID should not be significant. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A14. Continuous ESG x policy interaction effect

Variable	Coefficient of
ESG_Num (de-mean)	0.0139
	(0.0205)
ESG_Num x Post2016	0.0139
	(0.0205)
Post2016	0.0398***
	(0.0131)
MKT	0.8457***
	(0.0374)
VMG	0.4523***
	(0.0201)
SMB	0.0973***
	(0.0045)
PMO	0.2298***
	(0.0100)
N	1,967
R^2	0.2472

Note: The dependent variable is the quarterly excess return ExRet. The interaction term ESG_Num×Post2016 was constructed after ESG_Num was demeaned to mitigate multicollinearity. The sample size is reduced because the Treat/Post information from the DID sample must be matched. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A15. Instrumental variable estimation results (industry average ESG as IV)

Variables	Stage 1 (IV- > ESG)	Stage 2 (ESG_hat- > ExRet)
Industry_ESG_IV	0.6493***	
	(0.0193)	
ESG_Num_hat		-0.0072
		(0.0050)
MKT	0.0196***	0.1892***
	(0.0010)	(0.0372)
VMG	0.0335***	1.6674***
	(0.0021)	(0.0759)
SMB	0.0062***	0.8384***
	(0.0021)	(0.0773)

Table A15. Contiuned

PMO	0.0068***	0.1915***
	(0.0007)	(0.0237)
First stage F-statistic	960.25	
N	22,914	22,914
R^2	0.1683	0.1683

Note: The instrumental variable is the average ESG_Num of firms in the same industry and year (CSRC classification), excluding the firm itself. The first-stage F-statistic is 960.25 (> 10), indicating that there is no weak-instrument problem. The dependent variable is the quarterly excess return ExRet. Firm-level clustered robust standard errors are in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.