Proceedings of the 5th International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/2026.39165

Is ESG Performance Priced in the Chinese A-Share Market

Mengting Zhang[1,*], Yitong Shen[2,#], Xiangyu Zhu[3,#], Zihao Wang[4,#], Yichen Liu[5,#]

[1]*School of Business and Management, The Hong Kong University of Science and Technology, Hong Kong, China*
[2]*Department of Economics, University College of London, London, UK*
[3]*School of Business, University of Nottingham Ningbo China, Ningbo, China*
[4]*School of Data Science, The Chinese University of Hong Kong (Shenzhen), Shenzhen, China*
[5]*School of Economics, Zhejiang University, Hangzhou, China*
**Corresponding Author. Email: mzhangct@connect.ust.hk*
#These authors contributed equally to this work and should be considered as co-second authors

Abstract. This paper examines whether ESG performance is priced in China's A-share market using a comprehensive framework by changing ESG measure, empirical approach, and pricing horizon. Unadjusted raw ESG scores do not have predictive power, as the gross return of the unadjusted long/short portfolio is insignificant and negative. When the ESG score is controlled for industry and size effect, the composite portfolio obtains a positive and significant annual return of 2.89% with a FF5+UMD alpha of 3.12% (t=4.02). The predictive power of ESG score comes from governance (G) dimension (alpha=2.60%, t=3.57) while environmental and social dimension are insignificant. The Fama-MacBeth regression analysis shows that each year 1 standard deviation increase in ESG leads to a 0.76 percentage point higher return next year (p=0.0093), while next-month return is positive but insignificant. Event study around the 2024 April announcement of the guidelines for the compulsory sustainability information disclosure exhibits an asymmetric response among big firms, which is indicative of relative revaluation. These findings are robust even if the Wind ESG rating is substituted by Sino-Securities ESG rating.

Keywords: ESG, asset pricing, Chinese A-share market, portfolio sorting, Fama-MacBeth regression

1. Introduction

Whether corporate environmental, social, and governance (ESG) performance is priced in stock markets is a central question in sustainable finance. The answer is theoretically ambiguous: preference-based models predict that green assets should earn lower expected returns because investors are willing to accept lower financial returns in exchange for non-pecuniary benefits [1], while information-based explanations predict that ESG can generate positive subsequent returns when it contains underappreciated cash-flow information [2]. In the United States, ESG and stock return relation is a lasting concern, as the transition toward a lower-carbon economy exposes firms to changes in climate policy, technology and consumer preferences that can affect both expected cash flows and required returns. Consistent with the pricing transition risk, Bolton and Kacperczyk [3] document a carbon premium among U.S. firms, while Ramelli et al. [4] show that changes in U.S. climate-policy

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.39165

expectations generate different stock-price responses across firms. However, this evidence is sensitive to the measurement of environmental performance [5]. Its interpretation is further complicated by greenwashing, through which corporate disclosures may present environmental performance more favorably than firms' underlying actions justify [6], as well as by substantial disagreement among ESG-rating providers [7]. Taken together, the U.S. evidence highlights a broader challenge for sustainable finance: whether stock markets price substantive ESG performance or instead respond to ESG signals that are noisy or strategic disclosed.

China, given the scale of its manufacturing sector and population, provides a particularly consequential setting for examining these questions. However, empirical evidence on the pricing of ESG performance in China's A share market remains mixed. Using Refinitiv ESG data on a sample of large-cap firms, Ni and Sun [8] find that low-ESG firms earn higher subsequent returns after the Paris Agreement, consistent with a risk-premium explanation. By contrast, two studies covering a broader set of A-share firms document a positive ESG–return relation. Liu and Wan [9], using Sino-Securities ratings, link this relation to the limited incorporation of ESG information by retail investors, while Chen et al. [10], using a China-specific composite ESG measure, attribute it to investor underreaction to ESG-implied cash-flow information. These contrasting findings suggest that conclusions about ESG pricing in China may depend on sample coverage, data provider, and empirical design.

This paper argues that the disagreement between these two conclusions is not merely a sampling or measurement artifact—it reflects a deeper empirical ambiguity that existing studies have not adequately addressed. The ambiguity arises from three sources. First, raw ESG scores are systematically correlated with firm size and industry affiliation. Large firms tend to have higher ESG scores simply because they have more resources for disclosure [11], and the materiality of ESG issues varies by industry [12]. A positive spread between high- and low-ESG portfolios may therefore reflect size or industry effects rather than ESG information per se. Conversely, a negative spread may obscure a genuine ESG signal that is masked by these same confounding factors. Second, ESG ratings differ substantially across providers in scope, measurement, and weighting [7]. A significant ESG-return relation based on one rating system may reflect provider-specific measurement choices rather than a general pricing pattern. Third, ESG pricing operates at different horizons—monthly portfolio returns, annual firm-level predictability, and short-run market reactions to policy announcements—each of which answers a different economic question and need not produce identical results. Existing studies typically examine one approach or one horizon in isolation, making it difficult to distinguish between a genuine sustainability premium and a series of distinct empirical phenomena.

The paper addresses these gaps through a research design that systematically varies the ESG measure, the empirical method, and the pricing horizon, while keeping the underlying sample of Chinese A-share firms fixed. Specifically, four related questions are asked. First, does the return difference between high- and low-ESG portfolios survive after adjusting for industry and size effects? Second, which ESG pillar —environmental (E), social (S), or governance (G)—drives the relation? Third, is the result robust to replacing Wind ESG ratings with Sino-Securities ESG ratings? Fourth, does the pricing pattern also appear in short-run market reactions to a major regulatory shock—the mandatory sustainability disclosure guidelines jointly issued by China's three stock exchanges on April 12, 2024? By answering these questions within a unified framework, a more complete picture of ESG pricing in China is provided than any single-method or single-horizon study can offer.

The empirical strategy combines three complementary approaches. First are industry- and size-neutralized portfolio sorts, which isolate the pure ESG premium from the confounding effects of industry and size. This is the primary specification because it provides the cleanest estimate of whether ESG information is independently priced. Then Fama-MacBeth regressions are used to test whether ESG retains incremental explanatory power in a multivariate setting that controls for conventional return predictors such as size, book-to-market, momentum, and profitability. Finally, there is a short-window

DOI: 10.54254/2754-1169/2026.39165

event study around the April 2024 sustainability disclosure mandate to examine whether the market instantaneously reprices firms based on their pre-existing ESG characteristics when ESG information becomes suddenly salient. The aggregate ESG score is further decomposed into its three pillars to assess whether the ESG premium is a broad sustainability effect or is driven by a specific dimension—and the entire analysis is repeated using Sino-Securities ratings to test whether the conclusions are robust to provider choice.

Main findings are as follows. First, raw ESG scores have no predictive power; the unadjusted long-short portfolio earns a negative and insignificant return. However, once ESG scores are neutralized for industry and size effects, the composite ESG portfolio generates a significant annualized return of 2.89% with a FF5+UMD alpha of 3.12% ($t = 4.02$). This contrast shows that the ESG signal is real but obscured by size and industry noise. Second, this predictive power is concentrated in the governance (G) pillar (alpha = 2.60%, $t = 3.57$), while the environmental and social pillars are weaker and insignificant. Third, the neutralized ESG premium is robust to replacing Wind with Sino-Securities ratings and holds across both the overlapping 2018–2025 sample and the extended 2013–2025 history. Fourth, in the annual Fama-MacBeth regressions, a one-standard-deviation increase in ESG is associated with a 0.76 percentage-point higher next-year return ($p = 0.0093$). Fifth, the event study reveals an asymmetric market response: the positive ESG-ACAR relation is driven entirely by large-cap firms, while small and mid-cap firms show no significant effect.

The results in this study have three contributions to the ESG pricing literature. First, an explanation is provided for inconsistent findings regarding the ESG premium in China by providing evidence that the direction and magnitude of ESG return predictability vary depending on whether the exposures included in the ESG scores to industries and sizes are neutralized. Second, using the same neutralization and portfolio strategies for Wind and Sino-Securities ESG ratings, it is found that the positive ESG premium after neutralization is generalizable to all providers. Third, the horizon dependence and concentration effects of ESG pricing in one setting are documented. Beyond its academic contribution, our findings also have practical implications for investors, listed firms, and policymakers. For investors, the results caution against using raw ESG ratings as mechanical trading signals and instead point to the value of relative comparisons across similar firms, particularly with respect to governance and longer investment horizons. For listed firms, the evidence highlights the relevance of governance quality and disclosure readiness. For policymakers, the heterogeneous response to the 2024 disclosure guidelines suggests that improving the comparability and reliability of ESG information, while recognizing differences in firms' reporting capacity, may help ESG information be incorporated more effectively into market prices.

2. Background and literature review

2.1. ESG disclosure and pricing in China

China's ESG information environment has developed through a sequence of policy commitments, disclosure initiatives, and market-based rating systems. In September 2020, China announced that it would aim to peak carbon dioxide emissions before 2030 and achieve carbon neutrality before 2060. The commitment was subsequently embedded in a broader "1+N" policy framework covering energy, industry, transportation, construction, and green finance [13]. For capital-market participants, these policies increased the expected relevance of firms' environmental exposure, transition capacity, and sustainability governance. They also created the possibility that the pricing of ESG information would change when investors revised expectations about regulatory costs, financing access, and demand for sustainable assets.

Despite this growing demand, ESG information in China has historically been difficult to compare across listed firms until the introduction of the 2024 sustainability-reporting guidelines. On 12 April 2024, the Shanghai Stock Exchange released Guidelines No. 14 for Self-Regulation of Listed

DOI: 10.54254/2754-1169/2026.39165

Companies - Sustainability Report (Trial), the Shenzhen Stock Exchange released Guidelines No. 17, and the Beijing Stock Exchange released Continuous Supervisory Guidelines No. 11. The rules took effect on 1 May 2024. They established a more standardized domestic framework for sustainability reporting and required specified index-constituent and cross-listed firms on the Shanghai and Shenzhen exchanges to begin publishing sustainability reports for the 2025 reporting year, while encouraging broader voluntary reporting; the Beijing framework was primarily voluntary [14-16]. The guidelines therefore represent an important step toward improving the consistency and comparability of ESG information in China's capital market.

These institutional features have two implications for ESG asset pricing. First, an observed return difference between high- and low-ESG stocks may reflect economically relevant ESG information, but it may also arise because ESG ratings capture firm size, industry affiliation, or provider-specific measurement choices. The central empirical question is therefore whether ESG-related return differences remain after firms are compared with similar firms. Second, the 2024 sustainability-reporting guidelines provide an additional setting in which to examine whether the market reacts differently to firms with different pre-existing ESG characteristics when sustainability information becomes more salient. Together, the rating environment and the regulatory development provide the institutional foundation for examining ESG pricing in China's A-share market.

2.2. Literature review on ESG asset pricing, measurement, and pillar heterogeneity

There exist conflicting predictions in the literature of asset pricing regarding the connection between ESG and returns. In preference models, investors demand lower expected returns in exchange for investing in sustainable companies; hence, even if green stocks exhibit lower expected returns, they may outperform unexpectedly because of a change in tastes of investors [1]. According to Pedersen et al. [2], the value of ESG varies depending on the kind of investors who use the information since the former helps investors make good choices in portfolios only if there is risk or cash flow information present in it. Otherwise, investor preferences will increase prices and decrease expected returns. In terms of risk and information, both approaches predict opposite signs. Companies which do not engage in sustainability activities will have higher risks related to regulatory actions, litigations, and transition, and therefore, investors will want higher returns; this was proven by Bolton and Kacperczyk [3], where carbon premium exists in the U.S. In turn, ESG may be positively associated with future returns if it conveys unexplored cash flow information or skills of organizations [12].

Results in the Chinese A-share market reveal this ambivalence. With respect to positive signals, for instance, Chen et al. [17] show asymmetric market responses when China-specific ESG scores were introduced on Bloomberg: good scores led to positive abnormal stock returns, whereas bad scores yielded negative returns. Based on a larger sample of A-shares, Chen et al. [10] observe superior performance of high-ESG companies compared to low-ESG companies, with more predictable results for low-attention and high-arbitrage firms, which is regarded by the authors as evidence of underreaction to ESG-based cash flow news. Also, Liu and Wan [9] discover positive risk-adjusted excess returns to an ESG high-minus-low investment strategy that fades away with the increase of retail trading activity. Meanwhile, based on ESG data from Refinitiv for large-cap companies, Ni and Sun [8] find that low-ESG firms have higher future returns post-Paris Agreement deal, which can be explained by risk premiums.

There are two problems of measurement in this context. First, the ESG ratings include firm attributes which are independent of returns. As pointed out by Drempetic et al. [11], firm size affects the ESG score due to resource and data availability. The valuation relevance of ESG performance differs from industry to industry because materiality is industry-specific [12]. In that regard, an ESG spread could be affected by firm size or industry and not ESG information. Second, ratings differ significantly from one rating agency to another. According to Berg et al. [7], agencies vary in terms of scope, measurement, and

DOI: 10.54254/2754-1169/2026.39165

weights, with measurement being the most significant source of disagreement. Additionally, Avramov et al. [18] have shown that rating uncertainty has implications for asset pricing and portfolios. Nevertheless, almost all studies from China rely on a single measure, either Refinitiv [8], Sino-Securities [9] or a Chinese composite [10].

Another potential issue is the pillar heterogeneity. The combination of the performance on the pillars of environment, society, and governance in one score may mask the differences between them. According to Giese et al. [19], governance variables are correlated with more immediate event risks, whereas the environment and social factors appear gradually and have a more prolonged impact on the financials of the company. In China, there is a contradicting set of evidence. While Ni and Sun [8] find that E, S, and G contribute differently depending on regulatory periods, the ESG discount in the A-share market reported by Wu et al. [20] is mainly driven by the governance pillar. Most studies investigate only one metric, one methodology, or one horizon at once and hence do not allow identifying whether a sustainability premium exists as opposed to several different empirical effects. Our paper fills this research gap by varying the ESG measure, methodology, and pricing horizon while keeping the sample of A-shares fixed.

3. Data

Data are from four main sources:

Wind ESG data provides quarterly ESG ratings and scores for Chinese publicly listed companies, including the aggregate ESG score and the environmental, social, and governance pillar scores. The subset used in our analysis consists of Chinese A-share firms covered by Wind from 2018 to 2025. These data are used in the primary portfolio-sorting tests beginning in October 2018, the Fama-MacBeth regression, the pillar-decomposition analysis and the April 2024 event study.

Wind A-share market and fundamental data contain daily stock returns, market capitalization, book-to-market ratios, momentum, one-month reversal, turnover, return on equity, leverage, CITICS industry classifications and ST and trading-status indicators. The subset used in our analysis consists of non-ST A-share firms with valid trading records. These data are used to construct portfolio returns, firm-level control variables, annual compounded returns, and the Fama-French-Carhart style risk factors employed in the long-short portfolio analysis.

Sino-Securities ESG rating data provide an independent assessment of firms' aggregate and pillar-level ESG performance. The annual Sino-Securities ESG data are used for the robustness test They support (i) a provider-swap test including the overlapping Wind–Sino-Securities sample beginning in 2018 onward and the longer Sino-Securities sample covering 2013–2025, and (ii) a dual-carbon sub-period split comparing pre- and post-October 2020 returns (pre: formation months ≤ September 2020; post: ≥ October 2020), corresponding to China's carbon-neutrality policy shift. In all Sino-Securities robustness tests, the annual (t) return rate is paired with the annual (t-1) ESG score to avoid look-ahead bias.

Public announcements by the Shanghai, Shenzhen, and Beijing Stock Exchanges provide the official release dates and contents of the sustainability reporting guidelines for Chinese listed companies. The policy document used in our event study is the final set of sustainability reporting guidelines jointly released by the three exchanges on 12 April 2024 and effective from 1 May 2024. 12 April 2024 is defined as the event date and use daily stock returns around this date to examine the capital-market response to the new sustainability disclosure requirements.

The latest Wind ESG score is correlated with the observation date at or before the end of the month (t) and form an investment portfolio that can generate subsequent monthly returns without foresight. Industry- and size-neutralized sorts residualize the lagged ESG scores on the CITICS industry dummies and ln (MV) within each month, and then sort the residuals. The ESG annual score of Sino-Securities matches the stock code, but there is a one-year lag: the pricing month of calendar year (t) uses the Sino-

DOI: 10.54254/2754-1169/2026.39165

Securities annual score (t-1). For event research, as of March 31, 2024, Wind ESG scores are combined by stock code and daily stock returns, while CSI 300 returns are matched by trading date.

The following variables are constructed:

Dependent variables of study:

$\boldsymbol{AR_{i,y+1}}$ (Annual stock return) is calculated by compounding the twelve-monthly returns from January to December of one year. Annual stock return is obtained from Wind A-share market and fundamental data and is used in the annual Fama-MacBeth regression.

$\boldsymbol{ACAR}$ (Annualized cumulative abnormal return) is the cumulative abnormal returns that are converted to an annualized measure. The formula to calculate $ACAR$ is

$$ACAR = CAR \times \frac{255}{L} \tag{1}$$

where CAR is the sum of daily abnormal returns which measures as the difference between actual stock returns and expected returns estimated using the CSI 300 Index as the market benchmark, and L denotes the number of trading days in the corresponding event window. The data is obtained from Wind A-share market and fundamental data.

Independent variable of the study:

$\boldsymbol{ESG_{i,t}}$ (ESG performance) is the aggregate ESG score assigned to firm (i) at time (t).

$\boldsymbol{zESG}$ (standardized ESG performance) refers to $ESG_{i,t}$ that is standardized by subtracting the cross-sectional mean ESG score and dividing by the corresponding standard deviation.

$\boldsymbol{zSize}$ (standardized firm size) is the natural logarithm of market capitalization standardized by subtracting its cross-sectional mean and dividing by its cross-sectional standard deviation.

Interaction variable of the study:

$\boldsymbol{zESG \times zSize}$ is the interaction term between standardized ESG performance and standardized firm size.

Control variables of the study:

Firm size is the natural logarithm of the firm's market capitalization at the end of the predictor year.

Book-to-market is the book value of equity divided by the market value of equity.

Momentum is the firm's cumulative stock return from month (t-12) to month (t-2).

One-month reversal is the firm's stock return in month (t-1).

Turnover is the ratio of shares traded to shares outstanding.

Return on equity is the ratio of net income to shareholders' equity.

Leverage is the ratio of total liabilities to total assets.

The Fama-French five factors (MKT, SMB, HML, RMW, CMA) and the momentum factor (UMD) are constructed following Fama and French [21] and Carhart [22], respectively, using Wind A-share market data. All factor returns are calculated on a monthly basis over the sample period.

Fixed effect:

$\boldsymbol{Industry_{i,k,t}}$ (CITICS Industry Classification) identifies the industry to which a firm belongs according to the CITICS Industry Classification. For each industry, an indicator variable is constructed that equals one if firm belongs to that industry and zero otherwise.

All explanatory variables are defined using data for predictor year t. The raw ESG is the Wind composite score on 31 December, winsorized at the 1st and 99th percentiles in each predictor-year cross section but not standardized, and then the Z_ESG is constructed based on the winsorized raw ESG using the mean and standard deviation of the within-year cross section. Annual Return, Raw ESG, and all continuous control variables are winsorized at the same percentiles in the final annual predictor-year

DOI: 10.54254/2754-1169/2026.39165

cross section. Wind ESG provides ratings and scores, whereas Wind A-share market and fundamentals provide returns and market capitalization.

Table 1. Descriptive statistics for the Fama–MacBeth estimation sample

Variable	N	Mean	Std. Dev.	Median	Min	Max
Annual Return	18,774	0.126	0.407	0.046	-0.553	2.313
Raw ESG	18,774	6.056	0.824	5.980	4.300	8.671
Z_ESG	18,774	0.000	1.000	-0.126	-2.088	2.976
Size	18,774	13.541	1.124	13.281	11.738	17.464
Book-to-market	18,774	0.509	0.343	0.427	0.044	2.148
Momentum	18,774	0.056	0.388	-0.013	-0.608	2.528
One-month reversal	18,774	0.016	0.101	0.030	-0.443	0.298
Turnover	18,774	2.117	2.610	1.249	0.070	19.923
ROE	18,774	5.955	6.529	5.527	-21.264	33.954
Leverage	18,774	40.096	20.229	4.547	39.042	91.955

Notes: *This table reports descriptive statistics for the firm-year observations used in the baseline annual Fama–MacBeth regressions. The estimation sample contains 18,774 firm-year observations from 4,234 unique Chinese A-share firms. Predictor years span 2018–2024, and realised Annual Return is measured over 2019–2025. All continuous variables are winsorised at the 1st and 99th percentiles within predictor-year cross-sections. Z_ESG is the within-year standardised Wind ESG composite score, constructed after winsorisation. Annual Return is the subsequent calendar-year compounded raw return.*

Table 2. Descriptive statistics for the event-study sample

Variable	N	Mean	Std. Dev.	Median	Min	Max
Raw Wind ESG	4,023	6.070	0.791	5.950	3.580	9.170
z(ESG)	4,023	0.000	1.000	-0.152	-3.146	3.917
Market capitalisation (¥bn)	4,023	19.55	92.09	5.169	0.427	2,262.9
ln(MV)	4,023	13.40	1.124	13.16	10.66	19.24
z(Size)	4,023	-0.000	1.000	-0.219	-2.437	5.190
CAR [-5, +5]	4,023	-0.057	0.092	-0.059	-0.482	0.714
ACAR [-5, +5]	4,023	-1.332	2.144	-1.374	-11.17	16.55

Notes:This table reports sample construction and event-window characteristics for the event study of the 12 April 2024 release of China's sustainability-reporting guidelines. ESG information is measured as of 31 March 2024. Abnormal returns are estimated using a market model relative to the CSI 300 Index over the trading-day estimation window [-120, -11], requiring at least 60 valid observations. CAR [-5, +5] is the cumulative abnormal return over the 11-trading-day event window. ACAR [-5, +5] = CAR [-5, +5] × 255 / 11.

4. Research design

4.1. Portfolio sorting construction

To evaluate the inclusion of ESG performance in long-run A-share returns, three different sorting frameworks are introduced that refine resolution to different biases and complications to estimate the ESG pricing effect. The first sorting framework measures the ESG-return relation in the raw ESG score sorting. This serves as a baseline. The second framework addresses the strong, significant bias from industry allocation and market capitalization by the raw ESG scores by cross-sectional regression. This results in a size- and industry-neutral framework in which the aim is to measure the ESG risk premium.

DOI: 10.54254/2754-1169/2026.39165

This is referred to as our primary specification. In the third framework, a conditional double sort of size and ESG scores is used to test if the measured premium is valid for all sizes of firms.

4.1.1. Raw ESG score sort

At the end of a given month t, the sample firms are sorted based on their Wind ESG scores (composite / E / S / G scores) into 10 equal size groups (G1 to G10). The returns in month t+1 are computed based on equal weighting across all firms in the respective groups. The long-short portfolio return is computed as follows:

$$\mathrm{LS}_{t+1} = \frac{R_{t+1}^{\mathrm{High}} - R_{t+1}^{\mathrm{Low}}}{2} \tag{2}$$

where R_{t+1}^{High} and R_{t+1}^{Low} are the equal-weighted month-(t+1) returns of Decile 10 and Decile 1, respectively. The 1/2 factor is consistent with prior methodologies for A-share factor research. It ensures that long–short returns indicate the average return difference with respect to the decile spread. Portfolios undergo monthly rebalancing with the subsequent one month return recorded.

4.1.2. Industry- and size-neutralized ESG sort

To extract the pure ESG risk premium from industry and size factors, the ESG scores are removed by conducting a cross-sectional regression at each month-end t:

$$ESG_{i,t} = \alpha_t + \sum_{k=1}^{K-1} \beta_{k,t}\mathrm{Industry}_{i,k,t} + \gamma_t \ln(MV)_{i,t} + \varepsilon_{i,t} \tag{3}$$

where $\mathrm{Industry}_{i,k,t}$ designates CITICS Level-1 industry dummy variables (k = 1, …, K−1, with the first category omitted as the baseline) and $\ln(\mathrm{MV})_{i,t}$ is the natural logarithm of firm i's market capitalization. The estimated residual $\varepsilon_{i,t}$ is the neutralized ESG score for firm i at month t, which shows relative ESG performance controlling for sector and market-cap. Decile portfolios are created based on the ranked values of $\varepsilon_{i,t}$ for the long-short portfolio. The (G10 − G1) / 2 rule is applied as in the raw sort for this specification. This is our main long-term specification.

4.1.3. Size × ESG double sort

There is also a conditional (within-size) double sort based on Fama and French [23] to check whether the ESG premium differs by firm size. At the end of each month, stocks are split into three Size groups (Small, Medium, Large) using 30%/40%/30% market cap splits. Within each Size group, stocks are sorted into three ESG groups (Low, Medium, High) based on 30%/40%/30% ESG score splits to form a long-short portfolio:

$$\mathrm{LS}_{t+1}^{\mathrm{Size}} = R_{t+1}^{\mathrm{High,Size}} - R_{t+1}^{\mathrm{Low,Size}} \tag{4}$$

The overall size-neutralized ESG factor refers to the equal-weighted average of the long-short returns in three size buckets:

DOI: 10.54254/2754-1169/2026.39165

$$ESG\text{Factor}_{t+1} = \frac{1}{3}\left(\text{LS}_{t+1}^{\text{Small}} + \text{LS}_{t+1}^{\text{Medium}} + \text{LS}_{t+1}^{\text{Large}}\right) \tag{5}$$

This design has two goals. First, the neutralization approach is tested for robustness. If the neutralized ESG premium is concentrated in a specific size segment, the double sorting will bring this heterogeneity to light. Second, it aims to answer concerns about the different impacts of ESG concentrations of large-cap stocks (which are prevalent in the A-share equity indices) vs. small-cap stocks (which have a greater number but are more illiquid).

4.1.4. Performance metrics and risk adjustment

The following performance metrics are computed for each long-short portfolio:

RankIC: This evaluates the monthly Spearman rank correlation between ESG scores and one-month ahead returns, which examines firm level stock returns in relation to ESG scores and the respective returns. The RankIC tables present the time-series means of these monthly rank correlations.

Annualized Return and Volatility: Annualized metrics are derived from the mean monthly return and standard deviation. These metrics refer to the *LS Ann Ret* and *LS Ann Vol* columns in the empirical tables.

Sharpe Ratio: Annualized return divided by annualized volatility, reported as *LS Sharpe*.

Maximum Drawdown: The maximum value of cumulative net loss from a peak to a trough in the sample, reported as *LS MaxDD*.

Newey-West t-statistic: This estimates the mean long-short return using Newey-West [24] t-statistics with 6 lags and tests if this return is significantly different from zero, reported as *NW t*.

Risk-Adjusted Alpha: Computes long-short monthly portfolio returns and regress these returns on standard asset pricing factor models augmented with a momentum factor [22] denoted as *FF5+UMD*. Four models are evaluated:

Model 1 (FF5):

$$R_t^{\text{LS}} = \alpha + \beta_1 \text{MKT}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 \text{CMA}_t + \varepsilon_t \tag{6}$$

Model 2 (FF5 + ESG):

$$R_t^{\text{LS}} = \alpha + \beta_1 \text{MKT}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 \text{CMA}_t + \gamma ESG_t + \varepsilon_t \tag{7}$$

Model 3 (FF5 + UMD):

$$R_t^{\text{LS}} = \alpha + \beta_1 \text{MKT}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 \text{CMA}_t + \beta_6 \text{UMD}_t + \varepsilon_t \tag{8}$$

Model 4 (FF5 + UMD + ESG):

$$R_t^{\text{LS}} = \alpha + \beta_1 \text{MKT}_t + \beta_2 \text{SMB}_t + \beta_3 \text{HML}_t + \beta_4 \text{RMW}_t + \beta_5 \text{CMA}_t + \beta_6 \text{UMD}_t + \gamma ESG_t + \varepsilon_t \tag{9}$$

DOI: 10.54254/2754-1169/2026.39165

The intercept α captures the risk-adjusted abnormal return of the ESG long-short strategy utilizing Newey-West standard errors with six lags and measuring the abnormal returns.

4.2. Fama-MacBeth regression construction

While portfolio sorting analysis examines whether ESG-based strategies generate abnormal returns, it does not directly test whether ESG information provides incremental explanatory power after controlling for established firm characteristics. Therefore, the Fama-MacBeth [25] regression framework is employed to examine whether ESG performance remains associated with future stock returns in a multivariate cross-sectional setting.

4.2.1. Sample construction and data alignment

The Fama-MacBeth analysis uses the A-share sample and ESG measures described in Section 3. The unit of observation is a firm-year. For predictor year t, December predictor record, the 31 December Wind ESG composite score, and contemporaneous firm characteristics are used and linked to the firm's realised return in calendar year t+1.

The Fama-MacBeth regressions use predictor years 2018–2024, with corresponding realised returns measured over 2019–2025. Observations are retained with valid year-end ESG scores, complete control variables, industry classifications, and twelve monthly return observations in the subsequent year. These restrictions ensure that the outcome is a fully observed next-calendar-year compounded return. The final sample contains 18,774 firm-year observations from 4,234 unique firms.

4.2.2. ESG measurement and variable construction

The main ESG measure is Wind's composite score observed on 31 December of predictor year t. The dependent variable is the next-calendar-year compounded raw stock return, calculated as the product of one plus each monthly raw return in t+1, minus one. Requiring all twelve monthly observations avoids approximating a buy-and-hold annual return with a partial-year return or a sum of monthly returns.

Variable construction proceeds in three steps. First construct raw annual return, raw ESG, and the control variables. Then winsorise all continuous variables, including raw ESG and annual return, at the 1st and 99th percentiles within each final annual predictor-year cross-section. Finally, Z_ESG is calculated from the winsorised raw ESG score using the within-year cross-sectional mean and sample standard deviation. The baseline coefficient therefore compares firms with relatively higher Wind ratings within a given annual cross-section rather than relying on pooled standardisation across years.

The control vector includes Size and Book-to-market, conventional cross-sectional asset-pricing characteristics [26], together with Momentum [27], One-month reversal [28], Turnover, ROE, and Leverage. These variables proxy for scale, valuation, prior performance, trading activity, profitability, and financial risk—firm attributes that may be related to both ESG ratings and future returns.

4.2.3. Fama-MacBeth regression design

To test whether ESG performance contains incremental information about future stock returns, the Fama-MacBeth [25] two-stage regression framework is employed. The first-stage regressions estimate the conditional ESG-return relation within each predictor year, and the second stage averages these annual coefficients to obtain the overall estimate. It is therefore better suited to the question at hand than a pooled specification that imposes a common slope on all firm-year observations.

In the first stage, annual cross-sectional regressions are estimated for predictor years 2018–2024. Then regress next-calendar-year raw return on Z_ESG, firm characteristics measured in year t, and CITIC level-1 industry fixed effects. The ESG coefficient from each annual regression captures the

DOI: 10.54254/2754-1169/2026.39165

conditional ESG-return relation for that year. These annual coefficients are averaged and equal weight is assigned to each predictor-year cross-section. The main specification is:

$$R_{i,t+1}^{\mathrm{Ann}} = \alpha_t + \beta_{\mathrm{ESG},t} Z_{\mathrm{ESG},i,t} + \gamma_t^{'} \mathrm{Controls}_{i,t} + \mathrm{IndustryFE}_{i,t} + \varepsilon_{i,t+1} \tag{10}$$

Inference is based on Newey-West heteroskedasticity- and autocorrelation-consistent standard errors [24], calculated with a Bartlett kernel and three lags. This correction affects the precision assigned to the average annual coefficient rather than the coefficient itself. With only seven annual observations in the second stage, the estimates are interpreted as predictive associations conditional on observable firm characteristics and industry effects, rather than causal effects.

4.3. Event study design

Long-run portfolio evidence is only half the story. A short-run, policy-anchored test is added: an event study built around the mandatory ESG information disclosure regime that the Shanghai Stock Exchange, the Shenzhen Stock Exchange and the Beijing Stock Exchange released on April 12, 2024 under the unified guidance of the China Securities Regulatory Commission (CSRC). The mandate pushed disclosure up sharply, and by May 1 of the same year the share of A-share firms publishing ESG information for 2023-2024 had risen by 21.5%. A nationwide policy at a single point in time leaves no untreated control group, so a DID design is off the table; what remains feasible is to trace how each firm's ESG standing relates cross-sectionally to its abnormal returns around the announcement, asking whether publishing an ESG report feeds into stock prices at all.

4.3.1. Event, sample and data

The sample takes in every A-share firm that discloses ESG information, and for each the cumulative abnormal return (CAR) over the trading days bracketing the announcement is tracked. Financial data for 2023-2024 come from Wind: abnormal returns (AR) of A-share ESG disclosers, their ESG scores and the count of disclosing firms. The unstructured event material (social-media text) traces back to the joint release issued by the three exchanges under CSRC coordination.

4.3.2. Variable construction

Independent variable: $z(ESG)_i$, the standardized value (Z-score) of the Wind ESG composite score measured on March 31, 2024, which fixes each firm's ESG standing ahead of the event. Standardization puts 0 at the sample average; +1 reads as one standard deviation above the mean, −1 as one below.

Dependent variable: $\mathrm{ACAR}_{i,[-5,+5]}$, the cumulative abnormal return over the five trading days before and after the announcement, annualized. CAR is the realized return net of the expected return. The window [-5, +5] spans eleven trading days around event day and catches the market's short-run reaction on both sides; the annualization, ACAR = CAR × 255/11, rescales the eleven-day figure to a 255-trading-day year so magnitudes stay comparable.

4.3.3. Model specification

A baseline cross-sectional regression of ACAR on z(ESG) is as followed:

DOI: 10.54254/2754-1169/2026.39165

$$\mathrm{ACAR}_{i,[-5,+5]} = \alpha + \beta \cdot z(ESG)_i + \varepsilon_i \tag{11}$$

and read off the slope β with its p-value. To probe size heterogeneity, the standardized size control z_Size and industry fixed effects (Ind FE) are added in a multivariate specification.

5. Empirical results and mechanisms

5.1. Portfolio sorting results

Table 3 and Table 4 show the performance of long-short portfolios across three sorting methods.

Panel A reveals that unsorted ESG sorts give rise to negative and insignificant long-short returns for all four factors, Sharpe ratios are negative, and the overall MaxDD is −19.12%. Similarly, FF5+UMD alphas are negative and insignificant, while the overall RankIC of 0.05% indicates that unsorted ESG scores have no predictive ability.

Panel B demonstrates that neutralization leads to a significant improvement in results; specifically, the portfolio earns an annual return of 2.89% with a Sharpe ratio of 1.02 and an FF5+UMD alpha of 3.12% (t=4.02) with a RankIC of 1.71%. Note that Models 2 and 4 have been excluded from consideration since there is perfect collinearity between the neutralized portfolio and the ESG factor. As for pillars, governance (G) plays a leading role with a 2.60% alpha (t=3.57) while the other two pillars are insignificant.

Panel C provides the size × ESG double sort result. The size-adjusted composite produces an annualized return of 2.51% with RankIC equaling 1.20%. If ignoring the role of the ESG factor, the size-sorted portfolio yields a statistically significant alpha of 2.45% (t = 1.83), but after considering the ESG factor, the alpha falls between 0.79% and 0.85%, making it statistically insignificant. At the same time, R^2 increases from 0.23 to 0.29. This suggests that ESG factor captures returns on the size-sorted portfolio.

Table 3. Long-short portfolio performance (three sorting designs)

Factor	RankIC	AnnRet	Sharpe	MaxDD	NW t	a FF5	t(a)	a+UMD	t(+UMD)
Panel A. Raw ESG Score Sort									
Composite	0.05	-1.09	-0.23	-19.12	-0.64	-0.35	-0.32	-0.35	-0.31
E	-0.51	-1.40	-0.33	-16.91	-0.94	-0.45	-0.43	-0.45	-0.43
S	0.20	-0.05	-0.01	-14.05	-0.03	-0.50	-0.44	-0.51	-0.43
G	-0.58	-2.24	-0.54	-18.72	-1.68*	-1.03	-0.79	-1.03	-0.78
Panel B. Industry- and Size-Neutralized ESG Sort									
Composite	1.71	2.89	1.02	-5.00	2.99***	3.16	4.05***	3.16	4.02***
E	0.39	0.73	0.27	-5.68	0.75	1.29	1.27	1.29	1.32
S	0.63	0.96	0.43	-2.56	1.34	1.30	1.86*	1.30	1.87*
G	1.46	2.22	0.75	-4.54	2.36**	2.60	3.57***	2.60	3.57***
Panel C. Size × ESG Double Sort									
Composite	1.20	2.51	0.52	-6.07	1.59	2.48	1.83*	2.48	1.81*
E	0.82	0.44	0.12	-6.33	0.42	1.47	1.36	1.47	1.34
S	1.03	2.56	0.48	-10.04	1.51	1.87	1.16	1.87	1.14
G	0.34	-0.26	-0.07	-8.92	-0.22	0.81	0.49	0.81	0.49

DOI: 10.54254/2754-1169/2026.39165

Note: **, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.*

Table 4. Factor regression

Model	a ann (%)	t(a)	R^2	β_MKT	β _SMB	β_HML	β_RMW	β_CMA	β_UMD	γ _ESG
Panel A. Raw ESG Score Sort										
M1 FF5	-0.35	-0.32	0.36	0.06 (1.10)	-0.12 (-3.56***)	-0.07 (-1.53)	-0.03 (-0.35)	-0.09 (-0.84)	-	-
M2+ESG	-1.45	-0.92	0.38	0.06 (1.17)	-0.13 (-3.79***)	-0.11 (-2.14**)	0.00 (0.04)	-0.06 (-0.57)	-	0.35 (1.15)
M3+UMD	-0.35	-0.31	0.36	0.06 (1.10)	-0.11 (-3.13***)	-0.08 (-1.22)	-0.01 (-0.10)	-0.09 (-0.79)	-0.03 (-0.57)	-
M4+UMD +ESG	-1.43	-0.90	0.39	0.06 (1.17)	-0.13 (-3.27***)	-0.12 (-1.64)	0.02 (0.18)	-0.06 (-0.54)	-0.02 (-0.42)	0.35 (1.12)
Panel B. Industry- and Size-Neutralized ESG Sort										
M1 FF5	3.12	4.05***	0.32	-0.01 (-0.35)	0.05 (2.72***)	0.10 (3.77***)	-0.11 (-1.88*)	-0.09 (-1.59)	-	-
M2+ESG	-	-	-	-	-	-	-	-	-	-
M3+UMD	3.12	4.02***	0.32	-0.01 (-0.38)	0.05 (2.66**)	0.10 (3.71***)	-0.09 (-1.48)	-0.09 (-1.61)	-0.01 (-0.61)	-
M4+UMD+ ESG	-	-	-	-	-	-	-	-	-	-
Panel C. Size × ESG Double Sort										
M1 FF5	2.45	1.83*	0.20	0.01 (0.26)	-0.02 (-0.48)	-0.07 (-1.69*)	0.01 (0.10)	-0.14 (-1.56)	-	-
M2+ESG	0.79	0.45	0.26	0.02 (0.33)	-0.05 (-0.90)	-0.12 (-2.52**)	0.07 (0.91)	-0.09 (-1.19)	-	0.53 (1.18)
M3+UMD	2.45	1.81*	0.23	0.01 (0.22)	-0.01 (-0.13)	-0.11 (-1.91*)	0.08 (0.64)	-0.14 (-1.42)	-0.08 (-1.71*)	-
M4+UMD +ESG	0.85	0.47	0.29	0.01 (0.30)	-0.03 (-0.55)	-0.16 (-2.89***)	0.13 (1.21)	-0.09 (-1.07)	-0.08 (-1.35)	0.51 (1.13)

Note: *The long-short Industry- and Size-Neutralized ESG portfolio is the foundational level of ESG factor. Therefore, Models 2 and 4 on this portfolio are assessed, which results in perfect collinearity (RtLS = ESGt), thus these parameters (indicated as - in Table 3) cannot be estimated. To examine if ESG factors capture a cross effect on portfolio returns, Models 2 and 4 are run on other test portfolios (the Raw Composite and Size × ESG Double Sort Composite).*

An examination of Panel A and Panel B demonstrates that without accounting for known risk factors, specifically Size and Industry, single-variable sorts are generally susceptible to the masking of a signal or spurious regression [26, 27]. This is also consistent with the results of Avramov et al. [18] and Berg et al. [7]. These authors show that the raw ESG (Environmental, Social and Governance) metrics are highly contaminated by biases related to a firm's size and its industry. After the biases are accounted for using cross-sectional regression, the neutral ESG factors exhibit sizable and significant independent pricing.

The ESG premium is mainly attributed to the G pillar, in view of the significant importance of governance structures in mitigating agency problems and information asymmetries within the institutional context of the China A-share market. The pricing power of the E and S pillars is rather low, which might result from their longer-term and harder-to-quantify effects on companies' financial performance in the long run, which have not been priced into the retail-dominated market.

DOI: 10.54254/2754-1169/2026.39165

Most importantly, the ESG premium is not an isolated phenomenon restricted to small cap and large cap stocks. It is a systematic factor of pricing that allows explaining the returns of portfolios based on the classification of firms by other attributes. An ESG exposure can be considered as an indicator of systemic risks arising from global megatrends like low-carbon transition, regulation shifts, and supply chain resilience, which cannot be completely eliminated through portfolio diversification.

5.2. Fama-MacBeth regression results

5.2.1. Baseline Fama-MacBeth results

Complementing the portfolio sorting evidence, Table 5 reports baseline Fama-MacBeth regressions of subsequent stock returns on ESG performance. The specification includes conventional firm characteristics and industry fixed effects, so the Z_ESG coefficient is interpreted as a conditional cross-sectional relation rather than a simple ESG-return correlation.

The Z_ESG estimate is 0.007578, with a Newey-West HAC standard error of 0.002011. Under the Bartlett lag-three specification, this produces a t-statistic of 3.769 and a two-sided p-value of 0.0093. The relation remains positive and statistically significant once the regression accounts for size, book-to-market, momentum, one-month reversal, turnover, ROE, leverage, and industry fixed effects.

Because Z_ESG is standardised within each predictor-year cross-section, a one-standard-deviation increase in ESG is associated with a 0.758 percentage-point increase in next-calendar-year compounded raw stock return. This is an economically meaningful conditional predictive association. It does not imply that a change in ESG performance causes a change in returns, because unobserved firm characteristics, investor demand, measurement error, and omitted risk exposures may still matter.

Table 5. Baseline Fama-MacBeth regression results

Variable	Full Specification	
Constant	0.421888***	(3.972)
Standardised Wind ESG (Z_ESG)	0.007578***	(3.769)
Size	-0.022537	(-1.894)
Book-to-market	0.015407	(0.356)
Momentum	0.005160	(0.445)
One-month reversal	-0.097191**	(-3.214)
Turnover	-0.014926***	(-9.959)
ROE	0.000307	(1.290)
Leverage	-0.000008	(-0.068)
Industry fixed effects	Yes	
Observations	18,774	
Average adjusted R^2	0.1698	
Annual periods	7	

Notes: *The dependent variable is next-calendar-year compounded raw stock return. Coefficients are averages of seven annual cross-sectional Fama-MacBeth regressions. Newey-West HAC standard errors use a Bartlett kernel with lag 3; t-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively. The specification includes size, book-to-market, momentum, one-month reversal, turnover, ROE, leverage, and CITIC level-1 industry fixed effects. All continuous variables are winsorised at the 1st and 99th percentiles within final predictor-year cross-sections; Z_ESG is calculated after raw-ESG winsorisation.*

DOI: 10.54254/2754-1169/2026.39165

5.2.2. Interpretation of Fama-MacBeth evidence

The baseline regression indicates that Wind ESG ratings contain incremental predictive information for subsequent annual returns after conditioning on conventional firm characteristics and industry fixed effects. The positive ESG coefficient is consistent with the view that ESG performance may capture favourable firm-quality or cash-flow-related information that is not yet fully incorporated into stock prices. This interpretation is consistent with Chen et al. [10], who find that high-ESG stocks outperform low-ESG stocks in China and attribute this return predictability to investor underreaction to cash-flow information embedded in ESG performance. Together, the evidence suggests that ESG-related information may be incorporated into stock prices only gradually.

The estimates should nevertheless be interpreted as predictive associations rather than causal effects. Unobserved firm characteristics, investor demand, measurement error, and omitted risk exposures may affect both ESG ratings and subsequent returns. Accordingly, the Fama–MacBeth evidence supports the view that ESG contains incremental information about future returns, but it does not by itself identify the specific mechanism through which this information is incorporated into prices.

5.3. Event study results and mechanisms

5.3.1. Baseline cross-sectional results

Table 6 reports the cross-sectional regression of ACAR[-5,+5] on z(Wind ESG). The slope, b = 0.345886, says a one-standard-deviation higher ESG score comes with an average 34.59% higher annualized CAR. Around this event, in other words, the market rewarded high-ESG firms richly and punished low-ESG ones. A p-value of 3.32e-28, far below 0.01, and a t-statistic of 11.09 put the relation well past the 1% level; it is not a chance artifact. $R^2 = 0.0258$ means the ESG score on its own accounts for roughly 2.58% of the variation in ACAR. A single regressor reaching 2-3% is respectable in a cross-sectional return regression, yet it also says roughly 97% of the variation belongs to forces outside the model, firm size, momentum, industry traits and the like, which is exactly why it turns to size-heterogeneity tests next.

Table 6. Cross-sectional regression of annualized CAR [-5, +5] on z(Wind ESG) around the April 12, 2024 mandatory disclosure release

Metric	ACAR [-5, +5]
Z(ESG) coefficient (b)	0.345886
HC1 robust SE	0.031176
t-statistic	11.09
p-value	3.316E-28
Significance	***
Intercept (a)	-1.366291
N (firms)	4,148
R^2	0.0258
+1 SD ESG $\rightarrow$ Δ ACAR	34.59%
Mean ACAR	-136.63%
Mean z(ESG) / SD	-0.000/1.000
Ann. scale	255/1=23.1818
ESG scaling	z-score within regression sample

DOI: 10.54254/2754-1169/2026.39165

5.3.2. Size heterogeneity: double sort and sub-sample evidence

A non-parametric double sort tests whether the ESG premium varies by firm size. Stock returns answer to many forces at once—market-wide moves, industry swings, firm Beta—so within each size bucket there are high-ESG portfolio and short the low-ESG one, letting the remaining High-Low gap isolate the excess return tied to ESG performance itself. Firms are split by market value into low, mid and high terciles (bottom 33%, middle 33%, top 33%), then into four ESG tiers inside each bucket (no disclosure, low, medium, high); Figure 1 plots the mean event-window CAR of every cell. The gaps tell a clean story. The High-Low gap runs −0.1273 (t = −0.91) for low caps and +0.0644 (t = 0.44) for mid caps, both miles from significance, then jumps to +0.4092 for high caps, where t = 3.04 passes the 1% test.

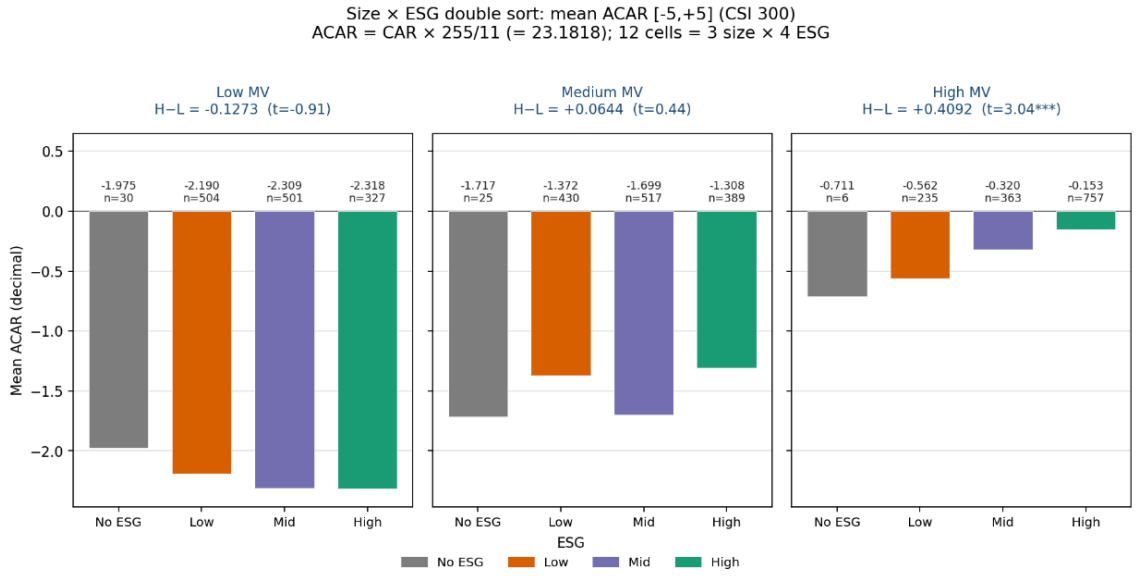


Figure 1. Size × ESG double sort

Note: *This graph mean event-window ACAR[-5,+5] by market-cap tercile and ESG tier. Firms are split into low, medium and high market-value terciles (left to right); within each tercile they are further sorted into four ESG tiers (no disclosure, low, medium, high). Bars report mean cumulative abnormal returns (annualized, ACAR = CAR × 255/11). The High–Low gap and its t-statistic are reported above each panel. Only the high-market-cap group posts a significant positive ESG premium.*

Running the same ACAR-on-z(ESG) regression separately inside each size tercile traces a clear gradient: a slope of −0.0515 (t = −0.93) for small caps, +0.0931 (t = 1.67, 10% level) for mid caps and +0.1273 (t = 2.60, significant at 1%) for large caps, the only group clearing the 1% bar. The ESG premium peaks at large firms, decays sharply as size shrinks, and disappears among small caps.

Taken together, the double-sort and sub-sample evidence reveal a stark size dependence: the ESG premium is concentrated almost entirely in large-cap stocks. In the low- and mid-cap terciles, ESG scores generate no significant abnormal returns—the slope is negative and insignificant for small caps and only marginally positive for mid caps. It is only among large caps that a high ESG score translates into a robust, statistically significant market return. This pattern implies that investors currently award an ESG premium mainly to large enterprises with established governance frameworks and greater capacity to absorb compliance costs, while treating the ESG signal as noise for smaller, more opaque firms.

DOI: 10.54254/2754-1169/2026.39165

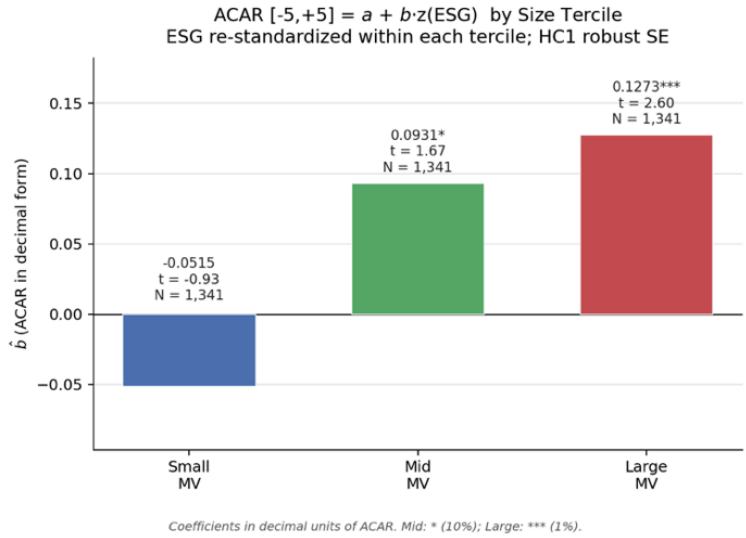


Figure 2. Slope of ACAR[-5,+5] on z(ESG), estimated separately within each market-cap tercile

5.3.3. Mechanisms: compliance costs, signal recognition and relative repricing

Why is the premium limited to large caps? Following the mandate, each firm's compliance cost and ESG risk were repriced by investors. Since high-ESG firms already have developed ESG programs, their incremental compliance cost would be minimal, along with low regulatory risk. On the other hand, low-ESG firms require more effort for rectification and implementation of ESG reporting systems. Small and medium caps have insignificant High-Low gaps for two reasons: First, some of them are not subject to the mandate and, secondly, lack of information environments renders the processing of ESG signals difficult [7]. This premium therefore is the result of relative repricing of high-ESG large caps since they were undervalued less than others, not rewarded directly as seen in prior studies on mandatory non-financial disclosure.

6. Discussion on robustness

6.1. Robustness of the neutral ESG premium

The robustness test shows that the neutralized ESG premium is not due to any particular data source, holding horizon, or policy environment. First, using Sino-Securities' ESG rating instead of Wind rating following the same process of neutralization and sorting generates a stable premium in the overlapping 2018–2025 sample and the extended 2013–2025 sample. Second, the holding period is further extended to three and six months, and annualized long-short returns are found to decline slightly from 2.89% to 2.11% and 1.82%, respectively, whereas RankIC is observed to rise from 1.71% to 3.58%. Third, separating the sample at the time of China's Dual Carbon announcement reveals that the neutralized ESG premium is prior to the policy change and strengthened afterwards, especially for the Wind sample.

Table 7. Robustness checks for portfolio sorting approach

Panel A, Alternative ESG Provider								
Sample	Provider	N	Period	AnnRet	NW t	RankIC	FF5+UMD a	t(a)
Wind baseline	Wind	86	2018-10-2025-11	2.89	2.97***	1.71	3.12	3.75
Sino overlap	Sino-Securities	86	2018-10-2025-11	1.89	2.07***	1.78	2.24	2.19

DOI: 10.54254/2754-1169/2026.39165

Table 7. (continued)

Sino long	Sino-Securities	155	2013-01-2025-11	2.26	3.77***	1.67	1.59	2.61

Panel A of shows how the Sino-Securities data yields an annualized return of 1.89% (NW t = 2.07) for the industry- and size-neutralized ESG long-short portfolio, with the overlapping period having an FF5+UMD alpha of 2.24% (t = 2.19). When considering the full period (2013-01 to 2025-11) the annualized return is 2.26% (NW t = 3.77) and the FF5+UMD alpha is 1.59% (t = 2.61), with RankIC showing a similar profile of 1.67% - 1.78%.

Table 8. Robustness checks for portfolio sorting approach

Panel B, Holding Horizons (Wind Neutral)							
Horizon	N	Period	AnnRet	NW t	RankIC	a FF5+UMD	t(a)
t+1	86	2018-10-2025-11	2.89	2.97***	1.71	3.12	3.75
t+3	86	2018-10-2025-09	2.11	2.73***	2.63	2.19	3.32
t+6	81	2013-01-2025-06	1.82	2.39***	3.58	1.53	1.85

Wind Neutral Composite; holding horizon t+h. AnnRet/a in %. NW lag = 6

Panel B examines results based on varying holding horizons. Annualized long-short return grades exhibit mild decay (2.89% at t+1, 2.11% at t+3, and 1.82% at t+6). However, the decay is mostly an artifact of risk. As decay occurs, RankIC, and therefore risk-adjusted return, trends up given longer holding horizons. Although risk-adjusted returns weaken, RankIC (increasing from 1.71% at t +1 to 3.58% at t +6) and risk-adjusted returns remain positive and statistically significant.

Table 9. Robustness checks for portfolio sorting approach

Panel C, Dual-Carbon Sub-periods									
Provider	Sample	Subperiod	N	Window	AnnRet	NW t	RankIC	FF5+UMD a	t(a)
Wind	Overlap	Full	86	2018-10- 2025-11	2.89	2.97***	1.71	3.12	3.75
Wind	Overlap	Pre	24	2018-10-2020-09	1.20	1.00	1.26	4.86	3.63
Wind	Overlap	Post	61	2020-11-2025-11	3.61	3.04***	1.93	3.02	2.57
Sino	Overlap	Full	86	2018-10-2025-11	1.89	2.07**	1.78	2.24	2.19
Sino	Overlap	Pre	24	2018-10-2020-09	1.75	1.34	1.30	4.62	2.77
Sino	Overlap	Post	61	2020-11-2025-11	2.02	1.70*	1.99	1.25	0.95
Sino	Long	Full	155	2013-01-2025-11	2.26	3.77***	1.67	1.59	2.61
Sino	long	Pre	93	2013-01-2020-09	2.46	3.89***	1.47	1.50	3.02
Sino	long	Post	61	2020-01-2025-11	2.02	1.70*	1.99	1.25	0.95

6.2. Robustness tests of Fama-MacBeth regression

Several robustness checks are conducted to examine whether the baseline ESG-return relation is sensitive to alternative ESG measurements, ESG providers, inference procedures, and return horizons.

DOI: 10.54254/2754-1169/2026.39165

Table 10. Additional robustness tests for Fama-MacBeth approach

Panel A. Alternative ESG Measurement				
Measure	Coefficient	Standard Error	t-statistic	p-value
Raw Wind ESG score	0.009319**	0.002899	3.214	0.0183
Panel B. Alternative ESG Provider				
Measure	Coefficient	Standard Error	t-statistic	p-value
Sino-Securities ESG Z-score	0.007224	0.003850	1.876	0.1097
Panel C. Newey-West Sensitivity				
Inference	Coefficient	Standard Error	t-statistic	p-value
Conventional	0.007578*	0.003786	2.002	0.0922
NW(1)	0.007578**	0.002621	2.891	0.0277
NW(2)	0.007578**	0.002305	3.287	0.0167
NW(3)	0.007578***	0.002011	3.769	0.0093
Panel D. Alternative Return Horizon				
Specification	Coefficient	Standard Error	t-statistic	p-value
Next-month return (NW(3))	0.000523	0.000345	1.515	0.1342

Notes:*Panel A replaces Z_ESG with the winsorised raw Wind ESG score and retains the baseline annual sample, controls, industry fixed effects, and NW(3) inference. Panel B uses Sino-Securities ESG ratings with the same sample period, preprocessing, controls, industry fixed effects, Fama-MacBeth estimation, and NW(3) inference. Panel C applies alternative inference procedures to the same annual Fama-MacBeth coefficient sequence. Panel D uses next-month returns and 72 monthly periods. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.*

Robustness checks indicate that the Fama-MacBeth relationship is not driven by any specific definition or estimation of ESG. Using winsorised raw Wind ESG score in place of Z_ESG results in a coefficient of 0.009319 (SE = 0.002899; t = 3.214; p = 0.0183). This indicates that the finding is not simply a function of standardisation. Using the Sino-Securities ratings in place of Wind, with the same sample, controls, and inference procedures, generates a coefficient of 0.007224 (SE = 0.003850; t = 1.876; p = 0.1097), which is positive and of similar size to the coefficient of the Wind ESG rating but less precisely estimated, reflecting continued measurement heterogeneity across providers. The point estimate stays constant at 0.007578 across ordinary least squares (OLS), Newey-West (1-3 lags) at 0.0922, 0.0277, 0.0167, and 0.0093, respectively. So, the economic magnitude is robust to the choice of inference procedure whereas the statistical significance is sensitive to the treatment of the short annual coefficients time series. Using next month returns for 72 months produces a positive but statistically insignificant coefficient of 0.000523 (SE = 0.000345; t = 1.515; p = 0.1342), in line with the greater predictability of ESG at the annual level.

The main findings remain broadly stable across these specifications. The alternative ESG measurement produces a positive and significant coefficient, suggesting that the baseline result is not

DOI: 10.54254/2754-1169/2026.39165

driven by the standardisation procedure. Using Sino-Securities ESG ratings yields a coefficient of similar economic magnitude but with weaker statistical precision, which is consistent with potential measurement differences across ESG providers. The ESG coefficient remains stable under alternative Newey-West specifications, while the monthly return specification shows a positive but statistically insignificant relation, suggesting stronger evidence at the annual horizon. Overall, these robustness tests support the interpretation that ESG performance contains return-relevant information in the Chinese A-share market, while measurement choices and investment horizons affect the precision of the estimated relation.

6.3. Robustness of the event-window ESG premium

There are two more tests which further substantiate the event window approach. The first one involves the re-estimation of the ESG–ACAR regression in the pre-event estimation window [−120, −1], resulting in a placebo slope coefficient equal to 0.019, or 1.90% ACAR per SD of ESG, compared to 34.59% in the event window; despite being statistically significant, the placebo t-statistic of 2.35 is clearly insufficient to account for the event-window premium, linking the premium not to some ongoing pre-event trend on the market but to the April 12 event. The second test, involving non-parametric High-Low gaps from the size × ESG double-sort, is supportive of the parametric findings – the gaps are insignificant for low-cap (t = −0.91) and mid-cap (t = 0.44) companies, while significant for large caps (t = 3.04), which is shown in Figure 1.

Table 11. Pre-event placebo regression of annualized CAR [-120. -1] on z (Wind ESG)

Metric	ACAR [-120, -1]
Z(ESG) coefficient (b)	0.019021
HC1 robust SE	0.008089
t-statistic	2.35
p-value	1.874E-02
Significance	**
Intercept (a)	-0.182163
N (firms)	4,071
R^2	0.0013
+1 SD ESG → Δ ACAR	1.90%
Mean ACAR	-18.22%
Mean z(ESG) / SD	-0.000/1.000
Ann. scale	255/120=2.1250
ESG scaling	z-score within regression sample

DOI: 10.54254/2754-1169/2026.39165

7. Conclusion

7.1. Summary of empirical findings

This paper studies if and how the effect of ESG performance is priced in China's A-share market. ESG is priced, but only when industry and size factors are controled: The returns from sorting on ESG are insignificant, but the neutralized ESG composite generates a 2.89% per year return with an FF5+UMD alpha of 3.12% (t = 4.02). The return premium is systematic rather than an anomaly localized in China, as it captures the extra return from sorting on size. It is mainly attributable to the governance component (alpha = 2.60%, t = 3.57), whereas the environment and social components are weak and insignificant. ESG information is incorporated progressively, as the coefficient from Fama-MacBeth regression per year is positive and significant (0.76 percentage point per standard deviation, p = 0.0093), but that for the following month is insignificant. The ESG disclosure requirement in April 2024 gives rise to a one-sided reaction, which is especially pronounced among large-cap stocks.

7.2. Practical and policy implications

The implications for investors, listed companies, and policymakers follow from the findings of the research. First of all, investors should not take ESG scores as automatic trading signals, but rather as signals about ESG performance compared to similar firms by size and industry, with more focus on the governance pillar and a wider time horizon. Listed companies should increase the efficiency of data collection, monitoring of the governance pillar at the level of the board of directors, and consistency of reporting because the highest risk-adjusted spread comes from good governance, while large firms gain most of the reaction in the announcement window.

7.3. Limitations and future research

This study is subject to several limitations. First, the main analysis uses Wind ratings; while the portfolio returns are still favorable in Sino-Securities ratings, the precision of the Fama-MacBeth regression decreases, and the scope, indicators, and weights used by different rating providers can lead to discrepancies in the ratings [7, 29]. It could be worth considering future research that would involve several rating providers using pillar or issue level data in order to distinguish the true ESG performance from measurement and reporting effects. Second, the paper concentrates on firm size, but the ESG premium may depend on the ownership structure, materiality of the industry, policy exposure, and investor types as well. The research could consider the state-owned enterprises and the private companies separately and check whether the difference in the information environment causes the premium to occur. Third, the economic mechanisms are not captured directly: the investor attention, the capital flows, the information asymmetry, the expenditures on compliance, and the quality of disclosures are not present in the data set; the April 2024 reform does not have an untreated comparison group. The further research could leverage the eligibility thresholds and implementation dates for testing these channels.

Author contribution

Conceptualization, M.Z. and Y.S.; methodology, M.Z., X.Z. and Z.W.; software, M.Z. and X.Z.; validation, M.Z., X.Z. and Z.W.; formal analysis, M.Z., X.Z., Z.W. and Y.L; investigation, M.Z., Z.W. and Y.L.; resources, M.Z. and Y.S.; data curation, M.Z. and Z.W.; writing—original draft preparation, M.Z., Y.S., X.Z., Z.W. and Y.L.; writing—review and editing, M.Z., Y.S., X.Z., Z.W. and Y.L.; visualization, M.Z.; supervision, M.Z.; project administration, M.Z. All authors have read and agreed to the published version of the manuscript.

DOI: 10.54254/2754-1169/2026.39165

References

[1] Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics, 142*(2), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011

[2] Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics, 142*(2), 572–597. https://doi.org/10.1016/j.jfineco.2020.11.001

[3] Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics, 142*(2), 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

[4] Ramelli, S., Wagner, A. F., Zeckhauser, R. J., & Ziegler, A. (2021). Investor rewards to climate responsibility: Stock-price responses to the opposite shocks of the 2016 and 2020 U.S. elections. *The Review of Corporate Finance Studies, 10*(4), 748–787. https://doi.org/10.1093/rcfs/cfab010

[5] Aswani, J., Raghunandan, A., & Rajgopal, S. (2024). Are carbon emissions associated with stock returns? *Review of Finance, 28*(1), 75–106. https://doi.org/10.1093/rof/rafd013

[6] Lyon, T. P., & Montgomery, A. W. (2015). The means and end of greenwash. *Organization & Environment, 28*(2), 223–249. https://doi.org/10.1177/1086026615575332

[7] Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance, 26*(6), 1315–1344. https://doi.org/10.1093/rof/rfac033

[8] Ni, Y., & Sun, Y. (2023). Environmental, social, and governance premium in Chinese stock markets. *Global Finance Journal, 55*, 100811. https://doi.org/10.1016/j.gfj.2023.100811

[9] Liu, X., & Wan, D. (2023). Retail investor trading and ESG pricing in China. *Research in International Business and Finance, 65*, 101911. https://doi.org/10.1016/j.ribaf.2023.101911

[10] Chen, S., Han, X., Zhang, Z., & Zhao, X. (2023). ESG investment in China: Doing well by doing good. *Pacific-Basin Finance Journal, 77*, 101907. https://doi.org/10.1016/j.pacfin.2022.101907

[11] Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics, 167*(2), 333–360. https://doi.org/10.1007/s10551-019-04164-1

[12] Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review, 91*(6), 1697–1724. https://doi.org/10.2308/accr-51383

[13] State Council of the People's Republic of China. (2021, October 25). Working guidance for carbon dioxide peaking and carbon neutrality in full and faithful implementation of the new development philosophy. https://english.www.gov.cn/policies/latestreleases/202110/25/content_WS61760047c6d0df5798e3c21.html

[14] Beijing Stock Exchange. (2024, April 12). Continuous supervisory guidelines No. 11 for companies listed on Beijing Stock Exchange—Sustainability report (for trial implementation). https://www.bse.cn/uploads/6/file/public/202404/20240412193432_9yjs1le33w.pdf

[15] Shanghai Stock Exchange. (2024, April 12). Guidelines No. 14 of Shanghai Stock Exchange for self-regulation of listed companies—Sustainability report (trial). https://english.sse.com.cn/news/newsrelease/c/10753174/files/7896ece79ed435684b77591c8d372a4.pdf

[16] Shenzhen Stock Exchange. (2024, April 12). Self-regulatory guidelines No. 17 for companies listed on Shenzhen Stock Exchange—Sustainability report (for trial implementation). https://www.szse.cn/www/English/rules/siteRule/P020240412667555851701.pdf

[17] Chen, Z. H., Kang, J. W., Koedijk, K. G., Gao, X., & Gu, Z. H. (2024). Short-term market reactions to ESG ratings disclosures: An event study in the Chinese stock market. *Journal of Behavioral and Experimental Finance, 43*, 100975. https://doi.org/10.1016/j.jbef.2024.100975

[18] Avramov, D., Cheng, S., Lioui, A., & Tarelli, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics, 145*(2), 642–664. https://doi.org/10.1016/j.jfineco.2021.09.009

[19] Giese, G., Nagy, Z., & Lee, L.-E. (2021). Deconstructing ESG ratings performance: Risk and return for E, S, and G by time horizon, sector, and weighting. *The Journal of Portfolio Management, 47*(3), 94–111. https://doi.org/10.3905/jpm.2020.1.198

[20] Wu, Y., Zhou, R., & Zhang, C. (2024). Size and ESG premiums: Evidence from Chinese A-share market. *The North American Journal of Economics and Finance, 74*, 102246. https://doi.org/10.1016/j.najef.2024.102246

[21] Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics, 116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

[22] Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

[23] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

[24] Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708. https://doi.org/10.2307/1913605

DOI: 10.54254/2754-1169/2026.39165

[25] Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*(3), 607–636. https://doi.org/10.1086/260061.
[26] Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *The Journal of Finance, 47*(2), 427–465. https://doi.org/10.1111/j.1540-6261.1992.tb04398.x
[27] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65–91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x.
[28] Lehmann, B. N. (1990). Fads, martingales, and market efficiency. *The Quarterly Journal of Economics, 105*(1), 1–28. https://doi.org/10.2307/2937816.
[29] Christensen, D. M., Serafeim, G., & Sikochi, A. (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review, 97*(1), 147–175. https://doi.org/10.2308/TAR-2019-0506

DOI: 10.54254/2754-1169/2026.39165

Appendix

Table A1. Correlation between Wind and Sino-Securities ESG Z scores

Predictor Year	Correlation
2018	0.240
2019	0.262
2020	0.301
2021	0.379
2022	0.426
2023	0.560
2024	0.628

Notes: *This table reports Pearson correlations between Wind and Sino-Securities ESG ratings among matched firm-year observations. The increasing correlation suggests gradual convergence between ESG providers, while remaining differences indicate persistent measurement heterogeneity.*