Social Flourishing (2026), 2(3): 1-24.
https://doi.org/10.63336/SocFlourish.57

Provincial Pollution-Control Investment and Wind–Huazheng Environmental Peer-Position Distance

Xinnan Zhang[1*]

[1] University College London

Corresponding author: Xinnan Zhang (zczvxz2@ucl.ac.uk)

Submitted: 2026-06-17 / Accepted: 2026-09-02 / Published: 2026-09-13

Abstract: ESG-rating disagreement is commonly attributed to persistent differences across providers. We examine variation within a fixed pair by relating province-year pollution-control investment intensity to contemporaneous Wind and Huazheng environmental scores for 4,797 Chinese A-share firms during 2018–2023. We standardize each provider's scores within its industry-year distribution and express them on a common peer scale. The primary outcome is the absolute gap between the two resulting peer positions, which shows how far apart the Wind and Huazheng scores place the same firm within its industry-year peer group. The signed Wind-minus-Huazheng difference preserves provider orientation. A one-standard-deviation increase in investment intensity is associated with 0.019 additional distance units, or 2.800 percent of the outcome standard deviation, with a restricted Webb-bootstrap p value of .005. The same increase is associated with a 0.046-unit decline in the signed difference, orienting the relationship toward a higher Huazheng-implied position. On the common pre-clipping scale, the Huazheng position accounts for most of the signed association. The Wind position changes little. The wider distance also appears within firms and under equal province-year weighting. Disagreement changes across places and years even when the provider pair remains fixed. In province-years with higher pollution-control investment intensity, the peer positions implied by Wind and Huazheng scores lie farther apart, which means that the choice between Wind and Huazheng matters more for comparisons and threshold decisions.

Keywords: environmental ratings, rating disagreement, pollution-control investment, commensuration, China

1. Introduction

Environmental ratings are used to compare firms in investment and lending decisions, but a firm's relative standing can depend on the provider. A firm that appears environmentally strong under one system may appear ordinary or weak under another. Provider choice matters because ESG information is used in capital markets. Sustainability signals are associated with portfolio allocation, expected returns, and market reactions (Avramov et al. 2022; Gibson Brandon, Krueger, and Schmidt 2021; Hartzmark and Sussman 2019; Pástor, Stambaugh, and Taylor 2021; Pedersen, Fitzgibbons, and Pomorski 2021; Krüger 2015; Capelle-Blancard and Petit 2019). Related evidence links rating disagreement and corporate social responsibility to stock-price responses and external finance (Serafeim and Yoon 2023; Cheng, Ioannou, and Serafeim 2014; Qin and Wang

2025). A stable gap between two systems produces a consistent provider-specific ordering. A gap that expands or contracts across province-years changes firm comparisons by different amounts. The unresolved issue is whether the distance between a fixed provider pair remains stable or varies systematically across the places and years in which firms are assessed.

A fixed provider pair can retain durable methodological differences and still exhibit changing distance. Providers make persistent choices about scope, measurement, aggregation, data coverage, and scoring practice, and assessments of the same firm often diverge even on comparable scales (Berg, Kölbel, and Rigobon 2022; Chatterji et al. 2016; Dimson, Marsh, and Staunton 2020; Dorfleitner, Halbritter, and Nguyen 2015; Halbritter and Dorfleitner 2015; Kotsantonis and Serafeim 2019; Billio et al. 2021). Firm size, disclosure, industry exposure, and rating design also shape observed assessments (Drempetic, Klein, and Zwergel 2020; Christensen, Serafeim, and Sikochi 2022; Escrig-Olmedo et al. 2019; Delmas and Doctori Blass 2010; Ilinitch, Soderstrom, and Thomas 1998; Scalet and Kelly 2010; Clementino and Perkins 2021). More comparable records may narrow interpretive differences. A larger or more varied record may give provider-specific measurement and aggregation more room to produce different positions. The positive association between corporate ESG disclosure and rater disagreement illustrates this second possibility (Christensen, Serafeim, and Sikochi 2022). Persistent provider differences can coexist with changing pairwise distance.

China separates formal environmental-policy status from realized pollution-control investment. Low-carbon-city pilots and regional carbon-market coverage classify province-years, but the same policy category can contain very different levels of completed investment relative to industrial activity. Environmental rules alter production, plant location, employment, and abatement investment (Greenstone 2002; Berman and Bui 2001; Walker 2013; Fowlie, Reguant, and Ryan 2016; Ryan 2012; Cai, Chen, and Gong 2016). They also change emissions and the records exposed to monitoring and auditing (Greenstone and Hanna 2014; Shapiro and Walker 2018; Keiser and Shapiro 2019; Tanaka 2015; Pargal and Wheeler 1996; Zou 2021; Duflo et al. 2013). Provincial accounts make a continuous feature of this activity observable. They report annual completed industrial pollution-control investment and secondary-industry value added. Their ratio measures province-year pollution-control investment intensity (National Bureau of Statistics of China 2019–2024). Published studies use the same ratio as a province-level measure of environmental-regulation intensity (Zhang and Chen 2025; Liu et al. 2023; Huang et al. 2025).

Chinese rating studies have used these policy categories to compare broad ESG-rating disagreement across regimes. Disagreement falls after low-carbon-city pilots and regional carbon-market coverage (Yang et al. 2026; Zhang, Qi, and Wu 2025). These discrete policy comparisons show that overall disagreement changes across regimes. They do not reveal whether the environmental positions implied by the same two providers separate differently with realized investment inside a regime. A named pair gives disagreement both magnitude and direction. The absolute distance between the peer positions implied by the two scores measures the size of the separation. The signed difference records which provider's observed score implies the higher position. For Wind and Huazheng, we study how province-year pollution-control investment intensity covaries with both the absolute distance and the signed Wind-minus-Huazheng difference.

We answer this question by matching province-year investment intensity to Wind and Huazheng environmental scores for the same listed firms. The panel contains 22,094 firm-year observations on 4,797 Chinese A-share firms across 31 provinces from 2018 through 2023, with firms assigned the provincial measure by registered province and year. Because the providers report scores on different units, we standardize each score within its own industry-year distribution and construct peer positions on a common scale. The absolute distance between these standardized peer positions measures separation. Their signed Wind-minus-Huazheng difference preserves provider orientation. We relate both outcomes to investment intensity after accounting for firm characteristics and province, industry, and year fixed effects, with inference at the province level. Repeated observations of the same firms and the same provider pair distinguish changes in separation from changes in which provider's score implies the higher peer position.

2. Theoretical Framework and Literature Review

2.1 Commensuration and Fixed-Pair Distance

Environmental ratings commensurate heterogeneous qualities. They select the information included in an assessment, express that information in comparable units, and combine it into an ordering (Espeland and Stevens 1998; Lamont 2012; Fourcade 2011; Mennicken and Espeland 2019). Rankings direct attention and can reorganize the behavior of rated actors and users (Espeland and Sauder 2007; Sauder and Espeland 2009; Dranove and Jin 2010; Jin and Leslie 2003; Zapechelnyuk 2020). The metric and comparison set define a firm's rating position. Emissions, environmental practices, and other records acquire meaning through that comparison. Two rating systems can place the same firm differently because each constructs the comparison in its own way.

Rating research identifies the choices that produce provider differences. Chatterji et al. (2016) distinguish theorization, or which attributes a rater regards as relevant, from commensurability, or whether those attributes are measured consistently. Berg, Kölbel, and Rigobon (2022) separate ESG-rating divergence into scope, measurement, and weight. Scope determines the covered attributes, measurement maps those attributes into indicators, and weight aggregates the indicators into a score. Their decomposition shows that scope and measurement account for most of the divergence across six providers. Together, these distinctions explain persistent differences in the comparable firm positions produced by rating systems.

Persistent provider differences do not imply constant pairwise distance. A fixed procedure governs how one provider evaluates available firm information, but distance depends on whether two procedures move in parallel as that information changes. Different choices over scope, measurement, and aggregation make parallel movement more plausible for some forms of information than for others. Information environments and social networks can also sustain disagreement among evaluators (Cookson and Niessner 2020). Pairwise distance emerges from the interaction between provider procedures and the material being evaluated.

2.2 Pollution-Control Investment and Competing Predictions

Pollution-control investment matters for rating distance through the environmental material presented to rating systems. Environmental regulation changes emissions, behavior around monitoring, and the accuracy of reports (Shapiro and Walker 2018; Zou 2021; Duflo et al. 2013). Research on regulation also documents changes in plant behavior, investment, location, productivity, and pollution (Greenstone 2002; Berman and Bui 2001; Greenstone and Hanna 2014; Walker 2013; Fowlie, Reguant, and Ryan 2016; Ryan 2012; Tanaka 2015; Pargal and Wheeler 1996; Cai, Chen, and Gong 2016; Keiser and Shapiro 2019). These changes alter corporate conduct and the record available to outsiders. Rating research explains how separate systems translate that record into scores.

Province-year pollution-control investment intensity captures a continuous feature of realized environmental activity. It records completed industrial pollution-control investment relative to secondary-industry value added. Firms registered in the same province-year share the measure. Wind and Huazheng provide separate firm-year assessments. Higher intensity may accompany differences in the amount, composition, or documentation of environmental activity. Linking the ratio to the two observed provider positions reveals whether those outputs move in parallel across province-years.

Constant distance requires the positions implied by the two scores to move in parallel as evaluated information changes. Provider-specific choices over scope, measurement, and aggregation make nonparallel movement plausible. One system may cover an attribute that the other omits, represent the same attribute on a different scale, or give it a different role in the aggregate score. The same information can coincide with different movements in the two provider positions. The balance between record comparability and record variety determines whether distance contracts or expands.

More comparable records predict convergence. Environmental records organized through common categories, units, and benchmarks should promote more parallel positions. Chatterji et al. (2016) find greater agreement where social ratings attain greater commensurability. An investment increase dominated by records structured around shared standards would narrow Wind–Huazheng distance.

H1a (convergence). Province-year pollution-control investment intensity is negatively associated with the absolute distance between the environmental peer positions implied by Wind and Huazheng scores.

More varied environmental material predicts divergence. Research on environmental and sustainability reporting shows that the quantity and quality of observable records vary with performance, materiality, market demand, legitimacy incentives, and disclosure requirements (Clarkson et al. 2008; Hummel and Schlick 2016; Matsumura, Prakash, and Vera-Muñoz 2014; Dhaliwal et al. 2011; Grewal, Riedl, and Serafeim 2019; Khan, Serafeim, and Yoon 2016; Marquis, Toffel, and Zhou 2016; Delmas and Burbano 2011; Christensen, Hail, and Leuz 2021). Completed investment can combine projects directed at different pollutants, technologies, facilities, and administrative objectives. Expressing that spending relative to industrial value added creates a common intensity scale across varied underlying activities. Christensen, Serafeim, and Sikochi (2022) document an analogous association between more corporate ESG disclosure and wider rating disagreement. An investment increase dominated by a broader or more varied set of activities and records would give provider-specific evaluation more room to produce greater distance.

H1b (divergence). Province-year pollution-control investment intensity is positively associated with the absolute distance between the environmental peer positions implied by Wind and Huazheng scores.

2.3 Provider Orientation

Absolute distance removes provider order and records disagreement magnitude. Rating disagreement has been represented with standard deviations, ranges, and absolute pairwise gaps. Avramov et al. (2022), for example, construct two-provider uncertainty from the absolute gap between provider percentile ranks. These measures answer how far evaluations stand apart. Equal distances are assigned to Wind-above-Huazheng and Huazheng-above-Wind orderings of the same size.

Provider identity matters to users who select one rating system. Billio et al. (2021) show that mapping several ESG ratings into common classes still produces different investable universes across providers. Two systems that place a firm on opposite sides of a cutoff produce different classifications under the same screening rule, portfolio threshold, or lending criterion. The signed difference between the peer positions implied by the two scores preserves provider order. Positive Wind-minus-Huazheng values place the Wind-implied position above Huazheng, and negative values place the Huazheng-implied position above Wind. Together, absolute distance and signed difference describe magnitude and orientation within one pairwise comparison.

The historical information sets and evaluative rules embedded in Wind and Huazheng determine their ordering. Converting the observed aggregate scores to a common peer scale makes that ordering directly comparable. The signed Wind-minus-Huazheng difference shows which provider's score places a firm higher within its industry-year peer group.

RQ1. How is province-year pollution-control investment intensity associated with the final signed Wind-minus-Huazheng environmental peer-position difference?

3. Method

3.1 Data

The analysis combines four sources at the firm-year level. Wind and Huazheng provide environmental scores. Provincial statistical accounts provide completed industrial pollution-control investment and secondary-industry value added. Listed-firm records provide registered province, industry, and financial and governance covariates (Wind 2026; Sino-Securities Index 2024; National Bureau of Statistics of China 2019–2024). We match the two provider scores by firm and fiscal year and assign the provincial measure by each firm's registered province and year. The resulting record holds the rated company and date constant across providers and links the company to its contemporaneous province-year measure.

The estimation sample requires one score from each provider, a valid provincial match, and complete covariates. It contains 22,094 observations on 4,797 Chinese A-share firms from 2018 through 2023. The firms are registered in 31 provinces, and the panel spans 184 province-years. The rating-source sample used to construct peer positions is larger than the estimation sample and contains 30,288 unique firm-years from 2018 through 2024. Constructing peer positions before the final sample intersection preserves each provider's industry-year comparison set. Appendix A1 reports the sample intersection and the six singleton provider-industry-year cells, none of which enters the estimation sample.

The provincial accounts report pollution-control investment in ten-thousand yuan and secondary-industry value added in hundred-million yuan. We use the archived series from which the estimation panel was built. Appendix A1 reconciles these values with the available China Statistical Yearbook editions and identifies which publication vintages match the archived rows. Repeated firms permit within-company comparisons, and the uneven number of listed firms across province-years motivates a separate equally weighted province-year specification.

3.2 Measurement

Pollution-control investment intensity measures completed industrial pollution-control investment relative to the scale of the provincial industrial economy. Let p index province and t year. Converting the reported units gives

$$P_{pt}^{raw} = \frac{\text{completed industrial pollution-control investment}_{pt}}{10,000 \times \text{secondary-industry value added}_{pt}}.$$

After assigning P_{pt}^{raw} to firms by registered province and year, we clip its firm-year distribution separately in each year at the first and ninety-ninth percentile order-statistic bounds. With $\mathcal{W}_t^{FY}(\cdot)$ denoting that operation, the explanatory variable is

$$P_{pt} = \mathcal{W}_t^{FY}\left(P_{pt}^{raw}\right).$$

Higher values indicate more completed industrial pollution-control investment per unit of secondary-industry value added. Because clipping follows the province-year merge, the percentile bounds reflect the firm-year distribution used in the baseline analysis. The variable characterizes the province-year of a firm's registered location; it is not firm expenditure. Appendix A1 reports the number of affected observations.

The rating outcomes begin with the aggregate environmental score issued by each provider. We define peer groups from the industry code, using its first two characters for manufacturing firms and its first character for other firms. In the full rating-source sample, we standardize each provider's scores separately by derived industry and year. For provider $r \in \{W, H\}$, firm i, industry group j, and year t, the standardized industry-year peer position is

$$Z_{it}^{r} = \frac{R_{it}^{r} - \overline{R}_{jt}^{r}}{s_{jt}^{r}},$$

where R_{it}^{r} is the observed provider score, $\overline{R}_{jt}^{r}$ is its provider-specific industry-year mean, and s_{jt}^{r} is the corresponding sample standard deviation. The transformation preserves each provider's within-group ordering and expresses the two positions in common standard-deviation units. Wind and Huazheng issue the scores. Industry-year standardization converts them into the peer positions used in the analysis.

The two outcomes come from one ordered comparison. The raw signed difference subtracts the Huazheng peer position from the Wind peer position,

$$\Delta_{it}^{WH,raw} = Z_{it}^{W} - Z_{it}^{H}.$$

We then clip the raw signed difference within year at the first and ninety-ninth percentile order-statistic bounds,

$$\Delta_{it}^{WH} = \mathcal{W}_t^{FY}\left(\Delta_{it}^{WH,raw}\right).$$

The resulting Δ_{it}^{WH} is the secondary outcome. Positive values place the Wind-implied peer position above Huazheng, and negative values place Huazheng above Wind. The primary outcome removes provider order after clipping,

$$D_{it}^{WH} = \left|\Delta_{it}^{WH}\right|.$$

The distance D_{it}^{WH} records the magnitude of the final separation. The signed difference Δ_{it}^{WH} records its orientation. Appendix A1 reports the number of signed observations affected by clipping.

All models use the same thirteen firm-year controls. They cover cash flow, fixed assets, revenue growth, loss status, size, leverage, return on equity, board size, board independence, CEO-chair duality, state ownership, Big Four auditor status, and audit opinion. Table 1 defines their observed scales and reports descriptive statistics. Keeping the adjustment set fixed makes outcome construction the only design difference between the baseline distance and signed equations.

3.3 Analytical Strategy

We estimate the two outcomes with the same baseline equation. For $Y_{it} \in \{D_{it}^{WH}, \Delta_{it}^{WH}\}$,

$$Y_{it} = \beta P_{pt} + X_{it}^{\top}\gamma + \alpha_p + \eta_j + \tau_t + \varepsilon_{it},$$

where X_{it} contains the thirteen controls, α_p denotes province fixed effects, η_j derived-industry fixed effects, and τ_t year fixed effects. In the distance equation, a negative β accords with convergence hypothesis H1a and a positive β with divergence hypothesis H1b. In the signed equation, the same coefficient answers RQ1 by estimating the corresponding change in Wind-minus-Huazheng orientation. Province effects absorb time-invariant differences across registered locations, industry effects absorb persistent differences across peer groups, and year effects absorb common annual movements. The baseline coefficient is an adjusted average firm-year association estimated from within-province changes over time. Province-years with more listed firms receive greater weight.

The explanatory variable varies by province-year, so standard errors are clustered by province. This level follows the source of common variation in the regressor and the resulting dependence in regression errors (Abadie et al. 2023; Cameron and Miller 2015). Conventional tests use a t_{30} reference distribution for the 31 clusters. We also report restricted wild-cluster bootstrap-t p values based on Webb weights and apply Holm adjustment to the two baseline outcomes (Cameron, Gelbach, and Miller 2008; MacKinnon and Webb 2017; MacKinnon, Nielsen, and Webb 2023a, 2023b; Roodman et al. 2019). Coefficient magnitudes are reported in outcome units for a one-standard-deviation increase in P, keeping substantive scale distinct from a threshold label for statistical significance (Zhao et al. 2024; Li, Zhang, and Ju 2025). Appendix A3 gives the finite-cluster implementation.

We use three additional specifications to change the comparison represented by β. Firm and year fixed effects estimate the association within repeatedly observed companies. Equal province-year weights give every observed province-year the same total influence and retain the baseline controls and fixed effects. Adding province-specific linear trends estimates covariance in deviations around a separate fitted path for each province. These specifications compare the average firm-year, repeated firms, the average province-year, and variation around province-specific linear paths.

To locate the signed association in the two provider positions, we estimate the baseline equation separately for the pre-clipping Wind and Huazheng positions. Because these equations use one sample and design matrix, their coefficients satisfy

$$\widehat{\beta}_{\Delta,raw} = \widehat{\beta}_W - \widehat{\beta}_H.$$

This identity applies only to the raw signed outcome; annual clipping and the absolute-value operation make the final signed and distance outcomes separate estimands. Appendix A2 reports leave-one-year and lagged-intensity estimates, and Table 5 compares alternative clipping orders and peer scales.

4. Results

4.1 Descriptive Evidence

Pollution-control investment intensity varies substantially across the firm-year sample. Expressed in thousandths, it averages 0.922, with a standard deviation of 0.826 and an interquartile range from 0.399 to 1.243. All firms registered in the same province-year receive the same value. The distribution reflects both changes in the provincial ratio and the concentration of listed firms across locations. Equal province-year weighting shows how that concentration affects the estimated relationship.

Wind and Huazheng positions also exhibit substantial pairwise separation. Absolute distance averages 0.838 peer-position units and has a standard deviation of 0.694. The final signed Wind-minus-Huazheng difference averages -0.005, even though its standard deviation is 1.088. Opposite provider orderings largely cancel in the signed mean. Absolute distance retains their magnitude. The two outcomes separate how far the positions stand apart from which provider's score implies the higher position. Table 1 reports the complete distributions.

Table 1 Variables and descriptive statistics

Variable	N	Mean	SD	P25	Median	P75	Min	Max
Pollution-control investment intensity (thousandths)	22,094	0.922	0.826	0.399	0.695	1.243	0.047	7.653
Absolute Wind–Huazheng distance	22,094	0.838	0.694	0.308	0.654	1.196	< 0.001	3.443
Final signed Wind–Huazheng difference	22,094	-0.005	1.088	-0.684	-0.033	0.627	-2.689	3.443
Raw Wind peer position	22,094	0.010	1.000	-0.707	-0.314	0.423	-1.485	6.941
Raw Huazheng peer position	22,094	0.014	1.006	-0.717	-0.147	0.602	-4.478	5.670
Cash-flow ratio	22,094	0.051	0.067	0.013	0.049	0.089	-0.156	0.265
Fixed-asset share	22,094	0.195	0.147	0.080	0.165	0.279	0.002	0.691
Operating-revenue growth	22,094	0.118	0.336	-0.054	0.079	0.228	-0.654	2.062
Loss indicator	22,094	0.165	0.372	0.000	0.000	0.000	0.000	1.000
Logged board size	22,094	2.214	0.172	2.079	2.303	2.303	1.792	2.708
Independent-director share	22,094	0.379	0.053	0.333	0.364	0.429	0.300	0.571
CEO-chair duality	22,094	0.329	0.470	0.000	0.000	1.000	0.000	1.000
Big Four auditor indicator	22,094	0.063	0.243	0.000	0.000	0.000	0.000	1.000
Audit-opinion indicator	22,094	0.965	0.183	1.000	1.000	1.000	0.000	1.000
State-owned-enterprise indicator	22,094	0.287	0.452	0.000	0.000	1.000	0.000	1.000
Firm size	22,094	22.355	1.291	21.421	22.147	23.072	19.997	26.490
Leverage	22,094	0.417	0.200	0.259	0.410	0.561	0.050	0.916
Return on equity	22,094	0.045	0.158	0.020	0.065	0.116	-0.975	0.415

Notes. Pollution-control investment intensity is multiplied by 1,000 for display in this table. All decimal statistics are displayed to three places; an absolute distance below 0.001 is displayed as < 0.001. Absolute distance is the absolute value of the final annually clipped signed Wind-minus-Huazheng difference. Raw provider positions are standardized separately within provider-specific derived-industry-year distributions in the full rating-source sample. Board size is reported on its observed logged scale, and firm size follows the source description as log total assets. Every variable has 22,094 nonmissing observations.

4.2 Pollution-Control Investment Intensity and Rating Distance

Pollution-control investment intensity is positively associated with absolute Wind–Huazheng distance. The positive coefficient accords with divergence hypothesis H1b; convergence hypothesis H1a predicts a negative coefficient. The baseline coefficient is 23.461. With 31 province clusters, the restricted Webb-bootstrap p value is .005 and the Holm-adjusted p value for the paired baseline family is .009. Table 2 reports the clustered standard error, confidence interval, and remaining model statistics.

Table 2 Baseline associations and province-trend comparison

Statistic	Absolute distance baseline	Signed difference baseline	Absolute distance with province trends	Signed difference with province trends
Pollution-control investment intensity coefficient	23.461	-56.009	7.790	-19.203
Province-clustered CR1 SE	6.539	22.646	6.985	15.070
t statistic	3.588	-2.473	1.115	-1.274
t_{30} p value	.001	.019	.274	.212
95 percent t_{30} CI	[10.106, 36.816]	[-102.259, -9.759]	[-6.475, 22.056]	[-49.981, 11.574]
Restricted Webb-bootstrap p value	.005	.010	.314	.243
Webb exceedances in 99,999 draws	449	1,000	31,408	24,339
Holm-adjusted Webb p value	.009	.010	—	—
Adjusted R^2	.034	.013	.034	.016
Observations	22,094	22,094	22,094	22,094
Province clusters	31	31	31	31
Firm controls	Yes	Yes	Yes	Yes
Province fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Province-specific linear trends	No	No	Yes	Yes

Notes. Each column estimates the adjusted association between pollution-control investment intensity and the named outcome. All equations include the thirteen controls in Table 1 and province, derived-industry, and year fixed effects. CR1 standard errors are clustered by province, and conventional p values and confidence intervals use a reference distribution. Restricted Webb p values use null-imposed, CR1-studentized wild-cluster bootstrap-t inference with six-point Webb weights, 99,999 draws, and seed 20260810. Holm adjustment covers the two baseline outcomes. The final two columns add province-specific linear time trends.

A one-standard-deviation increase in pollution-control investment intensity corresponds to 0.019 additional distance units, or 2.800 percent of the standard deviation of absolute distance. Average pairwise distance is 0.838 units. The estimate represents incremental widening within an already substantial provider separation.

The baseline estimate uses departures from persistent provincial and derived-industry levels after removing annual movements shared across the sample. Higher completed investment relative to industrial scale coincides with greater separation between the two peer positions in this comparison. Province-years containing more listed firms contribute more because the coefficient averages across firm-years. The alternative specifications change the comparison population, the weighting of province-years, or the temporal variation used for estimation.

4.3 Provider Orientation

The final signed equation identifies provider orientation. Its investment-intensity coefficient is -56.009, with a restricted Webb-bootstrap p value of .010. A one-standard-deviation increase in intensity corresponds to a 0.046-unit decline in the signed outcome, or 4.300 percent of its standard deviation. The outcome subtracts Huazheng from Wind, so the negative association points toward a higher Huazheng-implied peer position relative to Wind.

The two coefficients add direction to the widening distance. Higher investment intensity is associated with greater separation and a lower Wind-minus-Huazheng difference. On average, the wider separation is oriented toward a higher Huazheng-implied position.

Matched pre-clipping equations locate the raw signed relationship in the two observed provider positions. The investment-intensity coefficient is -4.307 for Wind and 50.424 for Huazheng. Their difference equals the raw signed coefficient of -54.731 on the common sample and design matrix, subject to displayed rounding. Most of the negative raw Wind-minus-Huazheng coefficient lies in the positive Huazheng-position coefficient. The Wind coefficient is comparatively small. Table 4 reports the separate inference for all three pre-clipping equations.

4.4 Comparison Targets and Temporal Variation

Shifting the comparison from the average firm-year to repeated firms or the average province-year leaves the positive distance association nearly unchanged. The coefficient is 23.461 in the baseline, 22.254 within firms, and 22.852 under equal province-year weighting. The latter estimates have Webb-bootstrap p values of .006 and .007.

Provider orientation changes more with the comparison population. The signed coefficient is -46.067 within firms and -20.719 under equal province-year weighting, compared with -56.009 in the baseline. Repeated observations preserve the Huazheng orientation. Equal weighting reduces its magnitude. Table 3 reports the paired outcome family under both alternative targets.

Table 3 Alternative comparison targets

Statistic	Absolute distance firm FE	Signed difference firm FE	Absolute distance equal province-year weights	Signed difference equal province-year weights
Pollution-control investment intensity coefficient	22.254	-46.067	22.852	-20.719
Province-clustered CR1 SE	6.129	23.963	6.488	15.412
t statistic	3.631	-1.922	3.522	-1.344
t_{30} p value	.001	.064	.001	.189
95 percent t_{30} CI	[9.739, 34.770]	[-95.006, 2.873]	[9.603, 36.102]	[-52.193, 10.756]
Restricted Webb-bootstrap p value	.006	.018	.007	.208
Webb exceedances in 99,999 draws	567	1,761	680	20,789
Observations	22,094	22,094	22,094	22,094
Province clusters	31	31	31	31
Firm controls	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	No	No
Province fixed effects	No	No	Yes	Yes
Industry fixed effects	No	No	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Equal province-year weights	No	No	Yes	Yes

Notes. Columns 1 and 2 estimate within-firm associations with firm and year fixed effects. Their province-clustered bootstrap inference includes the absorbed-score correction for the 71 firms observed in two provinces across 384 firm-years. Columns 3 and 4 assign every observed province-year total analytic weight one and retain the baseline controls and province, derived-industry, and year fixed effects. Coefficients are expressed per one-unit increase in the investment ratio.

Table 4 Provider-position location of the raw signed association

Statistic	Wind peer position	Huazheng peer position	Raw signed Wind-minus-Huazheng difference
Pollution-control investment intensity coefficient	-4.307	50.424	-54.731
Province-clustered CR1 SE	8.152	23.410	22.454
t statistic	-0.528	2.154	-2.438
t_{30} p value	.601	.039	.021
95 percent t_{30} CI	[-20.954, 12.341]	[2.614, 98.234]	[-100.587, -8.874]
Restricted Webb-bootstrap p value	.634	.012	.012
Webb exceedances in 99,999 draws	63,427	1,224	1,225
Adjusted R^2	.147	.129	.013
Observations	22,094	22,094	22,094
Province clusters	31	31	31
Firm controls	Yes	Yes	Yes
Province, industry, and year fixed effects	Yes	Yes	Yes

Notes. The three outcomes are pre-clipping provider positions estimated on one sample and design matrix. Linearity gives , so , subject to displayed rounding. The identity locates the raw signed coefficient only. Final signed difference and absolute distance are separately constructed outcomes reported in Tables 2, 3, and 5.

The province-trend specification uses departures from a separate fitted linear path for each province. In those departures, the absolute-distance coefficient is 7.790 and the signed coefficient is -19.203. The one-standard-deviation distance contrast is 0.006 units, compared with 0.019 in the broader within-province baseline. The smaller estimate places part of the baseline relationship in gradual provincial trajectories. Table 2 reports the corresponding uncertainty.

The direction of both relationships persists across each leave-one-year window, the lagged-intensity specification, and the alternative outcome constructions. Absolute-distance coefficients remain positive, and signed coefficients remain negative. Appendix A2 and Table 5 report the corresponding magnitudes and inference.

Table 5 Outcome-construction and peer-scale comparisons

Outcome construction	Coefficient	Province CR1 SE	t	t_{30} p	95 percent t_{30} CI	Webb p	Adjusted R^2
Panel A Absolute after clipping signed difference	23.461	6.539	3.588	.001	[10.106, 36.816]	.005	.034
Panel A Absolute raw signed difference	24.210	7.068	3.425	.002	[9.774, 38.645]	.005	.031
Panel A Annual clipping applied directly to raw absolute difference	23.189	6.615	3.506	.001	[9.680, 36.699]	—	.033
Panel B Absolute year-only z-position difference	28.300	7.550	3.748	<.001	[12.881, 43.718]	—	.042
Panel B Signed year-only z-position difference	-45.115	24.835	-1.817	.079	[-95.835, 5.604]	—	.058
Panel C Absolute industry-year percentile-position difference	5.800	1.875	3.093	.004	[1.970, 9.630]	—	.020
Panel C Signed industry-year percentile-position difference	-17.971	7.568	-2.375	.024	[-33.427, -2.516]	—	.016

Notes. Every equation includes 22,094 observations, 31 province clusters, thirteen firm controls, and province, derived-industry, and year fixed effects. Panel A changes the order of annual clipping and the absolute-value operation. Panel B standardizes each provider in the full rating-source sample within year. Panel C forms average-tie percentile positions within provider, derived industry, and year. Coefficient magnitudes are comparable within each panel because the outcome units differ across panels. Webb p values are reported where the numerical authority contains the corresponding bootstrap record; an em dash denotes an unreported value.

5. Discussion

Most explanations of rating disagreement begin with differences across providers. Our results identify systematic variation within a fixed pair. Wind and Huazheng remain the same providers throughout the sample, yet the distance between the environmental peer positions implied by their scores covaries with province-year pollution-control investment intensity. Provider methods account for persistent differences between the two orderings. Province-year environmental material is associated with how far apart those orderings stand at a given time. Observed disagreement reflects both the measurement systems and the environmental material they assess.

The positive distance association fits a commensuration account in which varied environmental activity creates more opportunities for two systems to place the same firm differently. Each system constructs an ordering by selecting, representing, and aggregating environmental qualities (Espeland and Stevens 1998). Differences in scope, measurement, and weight sustain provider-specific mappings (Chatterji et al. 2016; Berg, Kölbel, and Rigobon 2022). Commonly structured records can keep those mappings moving together. More varied activities and records can expose the points at which they diverge. The Wind–Huazheng estimates align with the second possibility.

Evidence from China's environmental policies distinguishes formal policy exposure from realized investment. Low-carbon-city pilots and regional carbon-market coverage are discrete regime changes associated with lower broad ESG-rating disagreement (Yang et al. 2026; Zhang, Qi, and Wu 2025). Completed pollution-control investment relative to industrial activity records continuous realized variation within and across such regimes. Its positive association with environmental peer-position distance shows that policy status and realized investment describe different empirical features. Following one named pair also preserves an ordering that broad cross-provider disagreement cannot display.

For users who rely on one rating system, the direction of a gap matters alongside its size. Equal absolute distances can place Wind above Huazheng or Huazheng above Wind, producing different rankings under the chosen system. The negative final Wind-minus-Huazheng association points on average toward a higher Huazheng-implied position. On the matched pre-clipping scale, the positive Huazheng-position coefficient accounts for most of the negative raw signed coefficient, and the Wind coefficient is small. Absolute distance measures the extent of separation. Signed position identifies the provider ordering.

The positive distance association appears in the average firm-year, repeatedly observed firms, and the average province-year, making magnitude the more stable feature of the fixed-pair relationship. Provider orientation varies more with weighting. The Huazheng orientation remains within firms and becomes smaller under equal province-year weighting, indicating that heavily represented province-years contribute more to the signed baseline relationship. Both associations are smaller around province-specific linear paths, placing part of the baseline pattern in gradual provincial change.

Rating users convert these measurements into portfolio, lending, and classification decisions (Billio et al. 2021; Avramov et al. 2022; Cheng, Ioannou, and Serafeim 2014; Qin and Wang 2025). Related evidence connects rating disagreement and public ESG information to market reactions, firm value, and corporate greenwashing (Serafeim and Yoon 2023; Aouadi and Marsat 2018; Hu et al. 2023). A common peer scale shows how much measured position changes with provider choice and whether the two positions straddle a decision threshold. Wider fixed-pair distance places a larger interval of possible thresholds between the positions. In practical decisions, a wider gap means that a larger set of portfolio, lending, or classification thresholds would classify the same firm differently under Wind and Huazheng.

6. Limitations

The estimates identify adjusted associations, not causal effects of pollution-control investment. Investment intensity varies by province-year and is assigned through a firm's registered location, so contemporaneous enforcement, industrial composition, reporting, environmental conditions, and the composition of listed firms can move with both the ratio and provider positions. Province, derived-industry, and year fixed effects remove persistent location and peer-group differences and common annual movements. Firm fixed effects, equal province-year weights, and province-specific trends alter the firms, weights, or temporal variation represented by the coefficient. They do not create exogenous variation in completed investment. The evidence shows where the association appears and how it changes across estimands.

The provincial ratio measures completed expenditure relative to industrial activity. It does not reveal which projects were completed, which pollutants or facilities they addressed, or which inspections, emissions, enforcement actions, and corporate reports accompanied them. Those records would be needed to distinguish the two theoretical interpretations of the positive distance coefficient. The archived series also spans several publication vintages. Available yearbooks verify the 2019–2023 investment entries and the 2023 value-added denominator against same-year rows. The 2019 yearbook reports 2017 investment by province, and the inspected 2018–2022 value-added rows come from different vintages. Appendix A1 reports the province-year reconciliation and keeps source documentation distinct from the statistical estimates.

The rating evidence operates at the level of observed aggregate Wind and Huazheng environmental scores. Historical indicator inputs, category maps, aggregation rules, analyst decisions, and methodology versions would be required to attribute the output relationship to scope, measurement, weight, or a particular rating procedure. The matched pre-clipping equations locate the raw signed coefficient in the two observed provider positions, not in their internal production processes. The study covers one provider pair, one rating pillar, Chinese A-share firms, and the 2018–2023 period. Applying the paired-outcome design to other providers, pillars, markets, and periods would show how widely province-year variation in fixed-pair distance recurs.

7. Conclusion

Environmental-rating disagreement combines persistent provider differences with variation inside a fixed pair. We examine that variation by linking the industry-year peer positions implied by Wind and Huazheng environmental scores to province-year pollution-control investment intensity.

Across 22,094 firm-years, higher investment intensity accompanies wider Wind–Huazheng separation and a lower signed Wind-minus-Huazheng position. The wider distance appears within firms and under equal province-year weighting. On the common pre-clipping scale, the signed relationship is concentrated in the Huazheng position.

These findings change how disagreement should be read. The size of a rating gap and the provider on each side carry distinct information. For investors, lenders, and other users who compare firms or apply thresholds, provider choice can alter both a firm's relative position and its classification.

Competing Interests

The author declares no conflicts of interest.

References

[1] Abadie, A., S. Athey, G. W. Imbens, and J. M. Wooldridge. 2023. "When Should You Adjust Standard Errors for Clustering?" Quarterly Journal of Economics 138 (1), 1–35. https://doi.org/10.1093/qje/qjac038.

[2] Aouadi, A., and S. Marsat. 2018. "Do ESG Controversies Matter for Firm Value? Evidence from International Data." Journal of Business Ethics 151 (4), 1027–1047. https://doi.org/10.1007/s10551-016-3213-8.

[3] Avramov, D., S. Cheng, A. Lioui, and A. Tarelli. 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.

[4] Berg, F., J. F. Kölbel, and R. Rigobon. 2022. "Aggregate Confusion: The Divergence of ESG Ratings." Review of Finance 26 (6), 1315–1344. https://doi.org/10.1093/rof/rfac033.

[5] Berman, E., and L. T. M. Bui. 2001. "Environmental Regulation and Productivity: Evidence from Oil Refineries." Review of Economics and Statistics 83 (3), 498–510. https://doi.org/10.1162/00346530152480144.

[6] Billio, M., M. Costola, I. Hristova, C. Latino, and L. Pelizzon. 2021. "Inside the ESG Ratings: (Dis)agreement and Performance." Corporate Social Responsibility and Environmental Management 28 (5), 1426–1445. https://doi.org/10.1002/csr.2177.

[7] Cai, H., Y. Chen, and Q. Gong. 2016. "Polluting Thy Neighbor: Unintended Consequences of China's Pollution Reduction Mandates." Journal of Environmental Economics and Management 76, 86–104. https://doi.org/10.1016/j.jeem.2015.01.002.

[8] Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." Review of Economics and Statistics 90 (3), 414–427. https://doi.org/10.1162/rest.90.3.414.

[9] Cameron, A. C., and D. L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." Journal of Human Resources 50 (2), 317–372. https://doi.org/10.3368/jhr.50.2.317.

[10] Capelle-Blancard, G., and A. Petit. 2019. "Every Little Helps? ESG News and Stock Market Reaction." Journal of Business Ethics 157 (2), 543–565. https://doi.org/10.1007/s10551-017-3667-3.

[11] Chatterji, A. K., R. Durand, D. I. Levine, and S. Touboul. 2016. "Do Ratings of Firms Converge? Implications for Managers, Investors and Strategy Researchers." Strategic Management Journal 37 (8), 1597–1614. https://doi.org/10.1002/smj.2407.

[12] Cheng, B., I. Ioannou, and G. Serafeim. 2014. "Corporate Social Responsibility and Access to Finance." Strategic Management Journal 35 (1), 1–23. https://doi.org/10.1002/smj.2131.

[13] Christensen, D. M., G. Serafeim, and A. Sikochi. 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.

[14] Christensen, H. B., L. Hail, and C. Leuz. 2021. "Mandatory CSR and Sustainability Reporting: Economic Analysis and Literature Review." Review of Accounting Studies 26, 1176–1248. https://doi.org/10.1007/s11142-021-09609-5.

[15] Clarkson, P. M., Y. Li, G. D. Richardson, and F. P. Vasvari. 2008. "Revisiting the Relation between Environmental Performance and Environmental Disclosure: An Empirical Analysis." Accounting, Organizations and Society 33 (4–5), 303–327. https://doi.org/10.1016/j.aos.2007.05.003.

[16] Clementino, E., and R. Perkins. 2021. "How Do Companies Respond to Environmental, Social and Governance (ESG) Ratings? Evidence from Italy." Journal of Business Ethics 171, 379–397. https://doi.org/10.1007/s10551-020-04441-4.

[17] Cookson, J. A., and M. Niessner. 2020. "Why Don't We Agree? Evidence from a Social Network of Investors." Journal of Finance 75 (1), 173–228. https://doi.org/10.1111/jofi.12852.

[18] Delmas, M. A., and V. C. Burbano. 2011. "The Drivers of Greenwashing." California Management Review 54 (1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64.

[19] Delmas, M., and V. Doctori Blass. 2010. "Measuring Corporate Environmental Performance: The Trade-Offs of Sustainability Ratings." Business Strategy and the Environment 19 (4), 245–260. https://doi.org/10.1002/bse.676.

[20] Dhaliwal, D. S., O. Z. Li, A. Tsang, and Y. G. Yang. 2011. "Voluntary Nonfinancial Disclosure and the Cost of Equity Capital: The Initiation of Corporate Social Responsibility Reporting." The Accounting Review 86 (1), 59–100. https://doi.org/10.2308/accr.00000005.

[21] Dimson, E., P. Marsh, and M. Staunton. 2020. "Divergent ESG Ratings." Journal of Portfolio Management 47 (1), 75–87. https://doi.org/10.3905/jpm.2020.1.175.

[22] Dorfleitner, G., J. Halbritter, and M. Nguyen. 2015. "Measuring the Level and Risk of Corporate Responsibility: An Empirical Comparison of Different ESG Rating Approaches." Journal of Asset Management 16 (7), 450–466. https://doi.org/10.1057/jam.2015.31.

[23] Drempetic, S., C. Klein, and B. Zwergel. 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.

[24] Dranove, D., and G. Z. Jin. 2010. "Quality Disclosure and Certification: Theory and Practice." Journal of Economic Literature 48 (4), 935–963. https://doi.org/10.1257/jel.48.4.935.

[25] Duflo, E., M. Greenstone, R. Pande, and N. Ryan. 2013. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." Quarterly Journal of Economics 128 (4), 1499–1545. https://doi.org/10.1093/qje/qjt024.

[26] Escrig-Olmedo, E., M. J. Muñoz-Torres, M. A. Fernández-Izquierdo, and J. M. Rivera-Lirio. 2019. "Rating the Raters: Evaluating How ESG Rating Agencies Integrate Sustainability Principles." Sustainability 11 (3), 915. https://doi.org/10.3390/su11030915.

[27] Espeland, W. N., and M. Sauder. 2007. "Rankings and Reactivity: How Public Measures Recreate Social Worlds." American Journal of Sociology 113 (1), 1–40. https://doi.org/10.1086/517897.

[28] Espeland, W. N., and M. L. Stevens. 1998. "Commensuration as a Social Process." Annual Review of Sociology 24, 313–343. https://doi.org/10.1146/annurev.soc.24.1.313.

[29] Fourcade, M. 2011. "Cents and Sensibility: Economic Valuation and the Nature of 'Nature.'" American Journal of Sociology 116 (6), 1721–1777. https://doi.org/10.1086/659640.

[30] Fowlie, M., M. Reguant, and S. P. Ryan. 2016. "Market-Based Emissions Regulation and Industry Dynamics." Journal of Political Economy 124 (1), 249–302. https://doi.org/10.1086/684484.

[31] Gibson Brandon, R., P. Krueger, and P. S. Schmidt. 2021. "ESG Rating Disagreement and Stock Returns." Financial Analysts Journal 77 (4), 104–127. https://doi.org/10.1080/0015198X.2021.1963186.

[32] Greenstone, M. 2002. "The Impacts of Environmental Regulations on Industrial Activity: Evidence from the 1970 and 1977 Clean Air Act Amendments and the Census of Manufactures." Journal of Political Economy 110 (6), 1175–1219. https://doi.org/10.1086/342808.

[33] Greenstone, M., and R. Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." American Economic Review 104 (10), 3038–3072. https://doi.org/10.1257/aer.104.10.3038.

[34] Grewal, J., E. J. Riedl, and G. Serafeim. 2019. "Market Reaction to Mandatory Nonfinancial Disclosure." Management Science 65 (7), 3061–3084. https://doi.org/10.1287/mnsc.2018.3099.

[35] Halbritter, G., and G. Dorfleitner. 2015. "The Wages of Social Responsibility—Where Are They? A Critical Review of ESG Investing." Review of Financial Economics 26, 25–35. https://doi.org/10.1016/j.rfe.2015.03.004.

[36] Hartzmark, S. M., and A. B. Sussman. 2019. "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows." Journal of Finance 74 (6), 2789–2837. https://doi.org/10.1111/jofi.12841.

[37] Hu, X., R. Hua, Q. Liu, and C. Wang. 2023. "The Green Fog: Environmental Rating Disagreement and Corporate Greenwashing." Pacific-Basin Finance Journal 78, 101952. https://doi.org/10.1016/j.pacfin.2023.101952.

[38] Huang, K., J. Chi, J. Liao, and M. K. Yuen. 2025. "In the Heat of the Moment, Secrets Will Out: Oil Price Uncertainty and Firm Green Innovation Disclosure." International Review of Economics & Finance 98, 103934. https://doi.org/10.1016/j.iref.2025.103934.

[39] Hummel, K., and C. Schlick. 2016. "The Relationship between Sustainability Performance and Sustainability Disclosure—Reconciling Voluntary Disclosure Theory and Legitimacy Theory." Journal of Accounting and Public Policy 35 (5), 455–476. https://doi.org/10.1016/j.jaccpubpol.2016.06.001.

[40] Ilinitch, A. Y., N. S. Soderstrom, and T. E. Thomas. 1998. "Measuring Corporate Environmental Performance." Journal of Accounting and Public Policy 17 (4–5), 383–408. https://doi.org/10.1016/S0278-4254(98)10012-1.

[41] Jin, G. Z., and P. Leslie. 2003. "The Effect of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards." Quarterly Journal of Economics 118 (2), 409–451. https://doi.org/10.1162/003355303321675428.

[42] Keiser, D. A., and J. S. Shapiro. 2019. "Consequences of the Clean Water Act and the Demand for Water Quality." Quarterly Journal of Economics 134 (1), 349–396. https://doi.org/10.1093/qje/qjy019.

[43] Khan, M., G. Serafeim, and A. Yoon. 2016. "Corporate Sustainability: First Evidence on Materiality." The Accounting Review 91 (6), 1697–1724. https://doi.org/10.2308/accr-51383.

[44] Kotsantonis, S., and G. Serafeim. 2019. "Four Things No One Will Tell You about ESG Data." Journal of Applied Corporate Finance 31 (2), 50–58. https://doi.org/10.1111/jacf.12346.

[45] Krüger, P. 2015. "Corporate Goodness and Shareholder Wealth." Journal of Financial Economics 115 (2), 304–329. https://doi.org/10.1016/j.jfineco.2014.09.008.

[46] Lamont, M. 2012. "Toward a Comparative Sociology of Valuation and Evaluation." Annual Review of Sociology 38, 201–221. https://doi.org/10.1146/annurev-soc-070308-120022.

[47] Li, D. M., H. L. Zhang, and Q. R. Ju. 2025. "Statistical Significance, Narrative, and the Scholastic Fallacy: How Ritualized Statistics Exaggerate Social Science Theories." Transformative Society 1 (2), 39–62. https://doi.org/10.63336/TransSoc.28.

[48] Liu, L., Y. Zhao, Y. Yang, S. Liu, X. Gong, and P. Jiang. 2023. "Study on the Measurement of Industrial Eco-Efficiency, Spatial Distribution and Influencing Factors in Yangtze River Economic Belt." PLOS ONE 18 (4), e0283964. https://doi.org/10.1371/journal.pone.0283964.

[49] MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb. 2023a. "Cluster-Robust Inference: A Guide to Empirical Practice." Journal of Econometrics 232 (2), 272–299. https://doi.org/10.1016/j.jeconom.2022.04.001.

[50] MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb. 2023b. "Fast and Reliable Jackknife and Bootstrap Methods for Cluster-Robust Inference." Journal of Applied Econometrics 38 (5), 671–694. https://doi.org/10.1002/jae.2969.

[51] MacKinnon, J. G., and M. D. Webb. 2017. "Wild Bootstrap Inference for Wildly Different Cluster Sizes." Journal of Applied Econometrics 32 (2), 233–254. https://doi.org/10.1002/jae.2508.

[52] Marquis, C., M. W. Toffel, and Y. Zhou. 2016. “Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing.” Organization Science 27 (2), 483–504. https://doi.org/10.1287/orsc.2015.1039.

[53] Matsumura, E. M., R. Prakash, and S. C. Vera-Muñoz. 2014. “Firm-Value Effects of Carbon Emissions and Carbon Disclosures.” The Accounting Review 89 (2), 695–724. https://doi.org/10.2308/accr-50629.

[54] Mennicken, A., and W. N. Espeland. 2019. “What’s New with Numbers? Sociological Approaches to the Study of Quantification.” Annual Review of Sociology 45, 223–245. https://doi.org/10.1146/annurev-soc-073117-041343.

[55] National Bureau of Statistics of China. 2019–2024. China Statistical Yearbook. Beijing, China Statistics Press.

[56] Pargal, S., and D. Wheeler. 1996. “Informal Regulation of Industrial Pollution in Developing Countries: Evidence from Indonesia.” Journal of Political Economy 104 (6), 1314–1327. https://doi.org/10.1086/262061.

[57] Pástor, Ľ., R. F. Stambaugh, and L. A. Taylor. 2021. “Sustainable Investing in Equilibrium.” Journal of Financial Economics 142 (2), 550–571. https://doi.org/10.1016/j.jfineco.2020.12.011.

[58] Pedersen, L. H., S. Fitzgibbons, and L. Pomorski. 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.

[59] Qin, J., and M. Wang. 2025. “ESG Rating Disagreement and Bank Loan Availability: Evidence from China.” PLOS ONE 20 (1), e0317191. https://doi.org/10.1371/journal.pone.0317191.

[60] Roodman, D., J. G. MacKinnon, M. Ø. Nielsen, and M. D. Webb. 2019. “Fast and Wild: Bootstrap Inference in Stata Using boottest.” Stata Journal 19 (1), 4–60. https://doi.org/10.1177/1536867X19830877.

[61] Ryan, S. P. 2012. “The Costs of Environmental Regulation in a Concentrated Industry.” Econometrica 80 (3), 1019–1061. https://doi.org/10.3982/ECTA6750.

[62] Sauder, M., and W. N. Espeland. 2009. “The Discipline of Rankings: Tight Coupling and Organizational Change.” American Sociological Review 74 (1), 63–82. https://doi.org/10.1177/000312240907400104.

[63] Scalet, S., and T. F. Kelly. 2010. “CSR Rating Agencies: What Is Their Global Impact?” Journal of Business Ethics 94 (1), 69–88. https://doi.org/10.1007/s10551-009-0250-6.

[64] Serafeim, G., and A. Yoon. 2023. “Stock Price Reactions to ESG News: The Role of ESG Ratings and Disagreement.” Review of Accounting Studies 28 (3), 1500–1530. https://doi.org/10.1007/s11142-022-09675-3.

[65] Shapiro, J. S., and R. Walker. 2018. “Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade.” American Economic Review 108 (12), 3814–3854. https://doi.org/10.1257/aer.20151272.

[66] Sino-Securities Index. 2024. ESG Ratings Methodology V2.1. May 2024. https://www.chindices.com/esg-ratings.html.

[67] Tanaka, S. 2015. “Environmental Regulations on Air Pollution in China and Their Impact on Infant Mortality.” Journal of Health Economics 42, 90–103. https://doi.org/10.1016/j.jhealeco.2015.02.004.

[68] Walker, W. R. 2013. “The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce.” Quarterly Journal of Economics 128 (4), 1787–1835. https://doi.org/10.1093/qje/qjt022.

[69] Wind. 2026. “Wind ESG Rating.” Accessed August 10, 2026. https://www.wind.com.cn/mobile/ESG/esgRating/zh.html.

[70] Yang, Y., Y. Deng, M. Guo, and S. Li. 2026. “Command-and-Control Environmental Regulation and ESG Rating Disagreement: Evidence from Low-Carbon City Pilot Policy in China.” Applied Economics, published online January 5, 1–16. https://doi.org/10.1080/00036846.2025.2612267.

[71] Zapechelnyuk, A. 2020. “Optimal Quality Certification.” American Economic Review: Insights 2 (2), 161–176. https://doi.org/10.1257/aeri.20190387.

[72] Zhang, J., F. Qi, and M. Wu. 2025. “Market-Oriented Environmental Regulation and ESG Rating Divergence.” Research in International Business and Finance 79, 103074. https://doi.org/10.1016/j.ribaf.2025.103074.

[73] Zhang, Q., and Y. Chen. 2025. “Environmental Regulation, Coal De-Capacity, and PM2.5 in China.” Scientific Reports 15, 7785. https://doi.org/10.1038/s41598-025-91134-5.

[74] Zhao, X., D. M. Li, Z. Z. Lai, P. L. Liu, S. H. Ao, and F. You. 2024. “Percentage Coefficient (bp)—Effect Size Analysis (Theory Paper 1).” arXiv:2404.19495. https://doi.org/10.48550/arXiv.2404.19495.

[75] Zou, E. Y. 2021. “Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality.” American Economic Review 111 (7), 2101–2126. https://doi.org/10.1257/aer.20181346.

Supplementary Appendix

Data Provenance and Reproducibility

A1. Sample Construction and Statistical Vintages

The empirical sample combines four linked secondary-data sources. A firm-year panel supplies company identifiers, financial controls, registered province, and derived industry. Provider records supply contemporaneous Wind and Huazheng environmental scores. Provincial accounts supply completed industrial pollution-control investment and secondary-industry value added. Firm-location records connect each company-year to its registered province. Requiring a unique firm-year match and complete analysis variables yields 22,094 observations on 4,797 firms during 2018–2023. Annual sample counts rise from 3,080 in 2018 to 4,538 in 2023.

Provider-specific peer positions are constructed before the final sample intersection. The full rating-source sample contains 30,288 unique firm-years. For each provider, the environmental score is standardized within its derived-industry-year distribution, preserving separate provider orderings on a common scale. Six singleton provider-industry-year cells occur in the source sample, and all six fall outside the estimation sample. Every retained score belongs to a peer distribution with within-cell variation.

Table A1 Data sources and empirical roles

Source or input	Unit	Empirical role
Listed-firm panel	Firm-year	Supplies identifiers, financial controls, registered province, and derived industry
Wind and Huazheng environmental scores	Provider-firm-year	Supplies the two observed aggregate ratings used to construct provider-specific peer positions
Provincial environmental accounts	Province-year	Supplies completed industrial pollution-control investment and secondary-industry value added
Firm-location records	Firm-year	Links each listed firm to its registered province in each year
China Statistical Yearbook, 2019–2024 editions	Province-year published rows	Establishes the publication-vintage coverage summarized in Table A2

Notes. All inputs are secondary data. Provider-issued scores remain separate from the peer positions obtained through industry-year standardization throughout the analysis.

Raw pollution-control investment intensity is completed industrial pollution-control investment divided by secondary-industry value added after converting the reported units to a common ratio. The provincial archive contains 184 complete province-years used by the analysis. Merging these values to firms and clipping the resulting firm-year distribution within each year at the first and ninety-ninth percentiles changes 201 observations and produces the final intensity measure.

The yearbook comparison identifies the publication vintage available for each input. The 2019–2023 investment entries match all 155 corresponding province rows in the 2020–2024 yearbooks. The provincial investment table in the 2019 yearbook reports 2017, so it does not provide a same-year check for the archived 2018 numerator. For secondary-industry value added, all 31 archived 2023 entries match the 2024 yearbook. The inspected editions for 2018–2022 contain values from different statistical vintages. Table A2 records these differences without replacing the archived inputs used in estimation.

Table A2 China Statistical Yearbook row reconciliation by data year

Data year	Yearbook edition	Investment table reference year	Investment row match	Value-added row match at one decimal	Statistical-vintage status
2018	2019	2017	N/C	0/31	Investment and value added retained at archive vintage
2019	2020	2019	31/31	3/31	Investment verified; value added retained at archive vintage
2020	2021	2020	31/31	0/31	Investment verified; value added retained at archive vintage
2021	2022	2021	31/31	2/31	Investment verified; value added retained at archive vintage
2022	2023	2022	31/31	0/31	Investment verified; value added retained at archive vintage
2023	2024	2023	31/31	31/31	Both inputs verified

Notes. N/C denotes a noncomparable reference year. The province rows in the 2019 yearbook investment table describe 2017. The archived province-year panel contains complete value added for 186 rows and complete investment for 184 rows. The two blank investment cells are Tibet in 2019 and 2022, and neither enters the final firm-year sample.

A2. Leave-One-Year and Lagged-Intensity Estimates

Leaving one calendar year out at a time traces the paired outcomes across every five-year window in the sample. Absolute-distance coefficients remain positive, ranging from 19.605 when 2018 is omitted to 27.506

when 2019 is omitted. Signed Wind-minus-Huazheng coefficients remain negative, ranging from -61.798 to -49.632. Every leave-one-year sample preserves both signs. Sampling precision varies across windows.

The lagged specification associates current rating outcomes with the preceding observed value of the already annually clipped province-year intensity within the 2018–2023 analysis window. It begins after 2018 and retains 18,978 firm-years. The coefficient is 14.510 for absolute distance and -32.132 for the signed difference. Table A3 reports province-clustered CR1 standard errors and t_{30} p values for the leave-one-year and lagged specifications.

Table A3 Leave-one-year and lagged-intensity estimates

Specification	N	Absolute coefficient	Province CR1 SE	t_{30} p	Signed coefficient	Province CR1 SE	t_{30} p
Omit 2018	19,014	19.605	8.766	.033	-61.798	27.830	.034
Omit 2019	18,951	27.506	6.239	<.001	-59.834	22.126	.011
Omit 2020	18,809	26.912	7.345	<.001	-57.338	23.034	.019
Omit 2021	18,261	21.496	6.960	.004	-49.632	20.065	.019
Omit 2022	17,879	20.601	7.969	.015	-52.430	27.036	.062
Omit 2023	17,556	22.837	6.424	.001	-57.035	22.390	.016
Lagged intensity	18,978	14.510	7.071	.049	-32.132	20.725	.132

Notes. Every equation includes the thirteen firm controls and province, derived-industry, and year fixed effects. Standard errors are CR1-clustered by province, and p values use 30 reference degrees of freedom. The lagged equation uses the preceding observed province-year intensity within the analysis window. Each coefficient is an adjusted association in the displayed sample.

A3. Province-Clustered Inference and Absorbed Firm Fixed Effects

The common sample contains 31 province clusters. Cluster size ranges from 53 to 3,707 firm-years, the largest province accounts for 16.778 percent of observations, and the coefficient of variation in cluster size is 1.297. For G clusters, N observations, and design rank K, the CR1 covariance estimator applies the correction

$$c_{CR1} = \frac{G}{G-1}\frac{N-1}{N-K}.$$

The baseline design has $N = 22,094$, $K = 70$, and $c_{CR1} = 1.037$. Conventional province-clustered inference uses a $t_{G-1} = t_{30}$ reference distribution.

Finite-cluster inference uses a restricted, null-imposed, CR1-studentized wild-cluster bootstrap-t procedure with six-point Webb weights and 99,999 draws (Roodman et al. 2019; MacKinnon, Nielsen, and Webb 2023b). For observed statistic t_{obs} and bootstrap statistics t_b^*, the symmetric two-sided p value is

$$p_{WCB} = \frac{1 + \#\{|t_b^*| \geq |t_{obs}|\}}{99,999 + 1}.$$

Holm adjustment treats the two baseline outcomes as one family. The resulting p values are reported with the conventional province-clustered statistics so that the paired outcome family is evaluated under the same finite-cluster procedure.

The firm-fixed-effect equations absorb 4,797 firm means before estimating the slope coefficients. Their equivalent full-design rank is 4,816, which gives $c_{CR1} = 1.321$. Seventy-one firms, covering 384 firm-years, are observed in two provinces. The bootstrap score construction projects each pseudo-error draw off the absorbed firm indicators before aggregating scores by province. Table A4 summarizes finite-cluster inference for the baseline and firm-fixed-effect equations.

Table A4 Restricted Webb-bootstrap inference diagnostics

Specification	Finite-sample rank	CR1 correction	Observed t	Exceedances	Webb p	Holm-adjusted Webb p
Baseline absolute distance	70	1.037	3.588	449	.005	.009
Baseline signed difference	70	1.037	-2.473	1,000	.010	.010
Firm-FE absolute distance	4,816	1.321	3.631	567	.006	—
Firm-FE signed difference	4,816	1.321	-1.922	1,761	.018	—

Notes. Every bootstrap test is restricted, null-imposed, CR1-studentized, symmetric, and two-sided. Holm values are defined for the paired baseline family. Firm-fixed-effect results incorporate the absorbed-score construction described above.