Proceedings of ICFTBA 2026 Symposium: Strategic Management & Business Analytics: Data-Driven Performance Models
DOI: 10.54254/2754-1169/2026.LD38813

ESG Rating Disagreement and Stock Price Crash Risk: Evidence from Chinese A-Share Listed Companies

Zherui Zhang

School of Mathematics and Statistics, Ningbo University, Ningbo, China
15906501135@163.com

Abstract. Environmental, social and governance (ESG) ratings issued by different agencies frequently point in opposite directions, yet it remains unsettled whether such disagreement helps market participants or confuses them. Drawing on a panel of Chinese A-share listed firms from 2018 to 2024, we measure rating disagreement as the dispersion among three domestic raters—Huazheng, Wind, and SynTao Green Finance—and relate it to subsequent crash risk. Pooled two-way fixed effects estimates yield no significant average relation, indicating that the noise channel and the information channel embedded in disagreement counterbalance each other. Tests based on stock price synchronicity do not detect an information channel at the aggregate level. Splitting the sample by analyst forecast dispersion, however, uncovers a sharp boundary condition: disagreement is followed by significantly lower crash risk only among firms whose information environment is relatively transparent, and the contrast between groups is significant at the 5% level. Disagreement therefore behaves as information where investors are able to verify it and as noise where they are not. Implications for rating-market regulation, disclosure policy, and rating users are discussed.

Keywords: ESG rating disagreement, stock price crash risk, noise effect, information effect, information environment

1. Introduction

China's capital market has attached growing weight to environmental, social and governance (ESG) performance in recent years. A regulatory milestone arrived in September 2018, when the China Securities Regulatory Commission revised the Corporate Governance Code for Listed Companies to require ESG-related disclosure in annual reports. As ESG considerations entered mainstream valuation practice, a diverse rating industry emerged: Huazheng, Wind, SynTao Green Finance, MioTech, and other vendors now score the same universe of listed firms under very different frameworks. The resulting scores often diverge sharply. A striking case is Xingwang Yuda, an A-share company that receives an A under the Wind system but only a CC under the MioTech system —nearly the top of one scale and nearly the bottom of another. Contrasts of this magnitude suggest that disagreement among ESG raters is not occasional measurement error but a persistent feature of the market.

© 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.LD38813

How should market participants read such disagreement? One interpretation treats it as noise: raters apply different standards and possess different capabilities, so conflicting scores simply distort judgment. A rival interpretation treats disagreement as information: each rater probes a different facet of ESG performance, and the spread of scores aggregates these multi-dimensional assessments. The distinction is consequential. Noise misleads rating users and may amplify volatility, whereas information helps market participants understand firms more fully and allocates capital more efficiently. Because crashes occur when long-accumulated bad news finally reaches prices, crash risk offers a natural outcome with which to discriminate between the two readings.

Prior work on disagreement has focused mainly on analyst forecasts and credit ratings, generally concluding that dispersion reflects noise. For ESG ratings, the evidence points in both directions. Berg et al. [1] attribute roughly 56% of rating divergence to measurement choices and 38% to scope, implying that raters are not evaluating the same construct. Serafeim and Yoon [2] report that ratings forecast subsequent ESG news only when raters concur, and Christensen et al. [3] show that richer disclosure is associated with wider, rather than narrower, divergence. On the other side, Shao et al. [4] document lower subsequent crash risk for Chinese firms with greater ESG rating dispersion over 2015–2021. Two gaps remain. First, little evidence covers the most recent years, including the turbulent market conditions of 2024. Second, the conditions under which the information channel operates have received limited attention.

This study addresses both gaps. Using a 2018–2024 panel of A-share firms, we examine whether rating disagreement affects subsequent crash risk and, crucially, where the effect is stronger or weaker. Three contributions follow. First, with the most recent sample available, we show that the noise and information channels cancel out at the aggregate level, a pattern that helps reconcile the mixed conclusions in the literature. Second, we establish a boundary condition: the information channel operates only among firms with a relatively transparent information environment, implying that disagreement information pays off only when investors possess the capacity to verify it. Third, the results carry practical implications for regulators overseeing the rating industry and for investors incorporating ratings into their decisions.

2. Theoretical analysis and research hypotheses

Two opposing channels link rating disagreement to subsequent crash risk. Crash risk originates in managerial incentives to conceal unfavorable news: when bad news accumulates inside the firm and eventually reaches the market all at once, prices adjust abruptly [5, 6]. Anything that improves the flow of firm-specific information limits such hoarding, whereas anything that clouds the information environment facilitates it.

2.1. The noise channel

Research on analyst and credit-rating dispersion suggests that conflicting third-party assessments can function as noise. In the ESG context, this channel may operate through several connected steps. Contradictory ratings first leave investors uncertain about which score to trust. Such uncertainty then discourages informed trading and weakens the extent to which prices incorporate firm fundamentals. When prices become less informative, managers can conceal unfavorable news more easily, and greenwashing becomes harder for outsiders to detect [1]. Once the concealed news finally surfaces, the abrupt price adjustment appears as a crash. This reasoning leads to the first hypothesis:

DOI: 10.54254/2754-1169/2026.LD38813

H1a: ESG rating disagreement significantly increases future stock price crash risk; in other words, disagreement acts as noise.

2.2. The information channel

The opposite channel emphasizes information production. Each rating agency builds its assessment on a distinct indicator system and data collection process, so dispersion across agencies implies that more facets of a firm's ESG performance have been uncovered and brought to the market [3]. Conflicting scores can also grab attention: investors who encounter disagreement have an incentive to gather further evidence on their own, and this additional search raises corporate transparency [4]. Where transparency improves, concealed bad news is released earlier and more gradually, which lowers crash risk. This reasoning leads to the second hypothesis:

H1b: ESG rating disagreement significantly reduces future stock price crash risk; in other words, disagreement acts as information.

2.3. Investor information search as the mechanism

The information channel just described contains two reinforcing components. First, disagreement stimulates search: Shao et al. [4] show that ESG rating dispersion is followed by increased online retrieval and interactive questioning about the affected firms, as investors redirect their limited attention toward fundamentals. Second, disagreement itself constitutes output: because each rater applies a distinct methodology, a wider spread of scores reflects a broader set of firm-specific ESG information already placed in the public domain [3]. Intensified search and multi-dimensional information release jointly enhance transparency, so that bad news reaches prices before it can accumulate. Formally:

H2: ESG rating disagreement lowers future crash risk by stimulating investor information search and promoting the release of firm-specific information.

2.4. The conditioning role of the information environment

A crucial feature of disagreement information is that it arrives as a semi-finished product. Contradictory scores do not speak for themselves; investors must cross-check them against financial reports, analyst commentary, and media coverage before they become decision-useful. Verification capacity, however, is unevenly distributed across firms. Where the information environment is transparent, investors can triangulate conflicting ratings at low cost, and the informational value of disagreement is realized. Where the environment is opaque, verification is costly or infeasible, and the same disagreement adds to confusion rather than knowledge. The value of disagreement is therefore conditional:

H3: The information effect of ESG rating disagreement is stronger for firms with a better information environment.

3. Research design

3.1. Sample and data sources

Our panel covers Chinese A-share listed firms over the period 2018–2024. ESG scores come from three domestic raters: Huazheng, Wind, and SynTao Green Finance. A transaction-level database that tracks individual stocks from 2018 to 2025 supplies weekly returns, market capitalization, free-

DOI: 10.54254/2754-1169/2026.LD38813

float capitalization, and turnover. Accounting items are collected from corporate financial statements, with 2023–2024 values supplemented through the open AKShare interface (Eastmoney source). Analyst forecast dispersion comes from an analyst consensus forecast database. We apply five screens in sequence: (1) financial firms are dropped; (2) firms under special treatment (ST or *ST) are dropped; (3) firms listed for less than one year are dropped; (4) firm-years with fewer than 30 trading weeks are dropped; and (5) continuous variables are winsorized at the 1st and 99th percentiles. These filters leave 28,056 firm-year observations.

3.2. Variable definitions

Crash risk in year t+1 serves as the outcome of interest. For every firm-year we estimate, with weekly returns, the extended market model of Chen et al. [7] and Hutton et al. [6]:

$$r(i,w) = \alpha + \beta 1 \cdot r(m,w-2) + \beta 2 \cdot r(m,w-1) + \beta 3 \cdot r(m,w) \quad (1)$$

$$+\beta 4 \cdot r(m,w+1) + \beta 5 \cdot r(m,w+2) + \varepsilon(i,w)$$

Here r(i,w) denotes the weekly return of firm i with cash dividends reinvested, and r(m,w) denotes the weekly market return weighted by free-float capitalization. The residual ε enters the firm-specific weekly return W = ln(1 + ε), from which we build two standard crash measures: negative conditional skewness (NCSKEW) and the down-to-up volatility ratio (DUVOL) [7, 8]. Higher NCSKEW or DUVOL values signal greater crash risk. For robustness we also construct CRASH, an indicator equal to one when at least one weekly W in year t+1 lies more than 3.09 standard deviations below its own annual mean.

The explanatory variable is ESG rating disagreement (Disagree). Letter grades are first converted to numbers (AAA = 9 down to C = 1 for Huazheng and Wind; A+ = 10 down to D = 1 for SynTao Green Finance), then z-standardized within each agency-year cell so that scale differences do not contaminate the measure. Disagree equals the standard deviation of the standardized scores across agencies and requires scores from at least two raters. The cell-mean of the standardized scores (ESG_mean) absorbs the rating level. A rank-based variant, in which letter grades are replaced by within-year percentiles before dispersion is computed, appears in the robustness section.

Controls comprise firm size (Size, the natural logarithm of total market capitalization), the book-to-market ratio (BM), average weekly turnover (Turnover), leverage (Lev), return on assets (ROA), and the volatility of firm-specific weekly returns (Sigma). The mechanism variable is stock price synchronicity (Synch), obtained as the logit transformation of the R^2 from a simple market-model regression; a lower Synch means that prices carry more firm-specific information. In the heterogeneity tests we proxy the information environment with analyst forecast dispersion (FDISP)—higher FDISP corresponding to a poorer environment—and use Synch and Size as alternatives.

3.3. Model specification

H1a and H1b are evaluated with the following two-way fixed effects specification:

$$Crash(i,t+1) = \alpha + \beta \cdot Disagree(i,t) + \gamma \cdot Controls(i,t) + \mu(i) + \lambda(t) + \varepsilon(i,t) \quad (2)$$

In equation (2), Crash stands for NCSKEW or DUVOL measured in year t+1, μ(i) absorbs firm fixed effects, λ(t) absorbs year fixed effects, and standard errors are clustered by firm. A significantly positive β would support H1a (noise), while a significantly negative β would support

DOI: 10.54254/2754-1169/2026.LD38813

H1b (information). The mechanism test for H2 proceeds in two steps: Synch is regressed on Disagree, and Synch is then entered alongside Disagree in the crash regression. For H3, observations are split at the sample median of FDISP into a poor-environment and a good-environment group, and the interaction Disagree × Info_poor supplies a formal test of whether the two group coefficients differ.

4. Empirical results

4.1. Descriptive statistics

Table 1. Descriptive statistics

Variable	N	Mean	SD	Min	Max
NCSKEW	28,056	-0.344	0.753	-2.627	1.577
DUVOL	28,056	-0.201	0.482	-1.386	1.002
Disagree	28,056	0.612	0.365	0.043	1.709
ESG_mean	28,056	0.023	0.771	-1.583	2.224
Size	28,054	22.609	1.016	20.980	25.849
BM	27,660	0.530	0.377	0.053	1.957
Turnover	27,724	2.830	2.546	0.253	13.886
Lev	27,419	0.414	0.201	0.057	0.898
ROA	27,419	0.028	0.069	-0.278	0.197
Sigma	28,056	0.051	0.020	0.018	0.116

Summary statistics appear in Table 1. NCSKEW and DUVOL average -0.344 and -0.201, values in line with those documented in prior crash-risk studies [9], which reassures us that the measures are constructed correctly. Disagree averages 0.612 with a standard deviation of 0.365, so dispersion across raters is both common and economically variable.

4.2. Baseline regression

Table 2. Baseline regression of ESG rating disagreement on crash risk

	(1) NCSKEW	(2) NCSKEW	(3) DUVOL	(4) DUVOL
Disagree	0.018 (1.14)	0.010 (0.66)	0.013 (1.32)	0.010 (0.98)
Controls	No	Yes	No	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	28,056	27,233	28,056	27,233
R^2 (within)	0.000	0.082	0.000	0.068

DOI: 10.54254/2754-1169/2026.LD38813

Table 2 presents the baseline estimates. Parentheses contain t-statistics computed from firm-clustered robust standard errors, and *, **, *** mark significance at the 10%, 5%, and 1% levels throughout. The Disagree coefficient is positive yet statistically indistinguishable from zero in all four columns. Pooled across the market, therefore, disagreement neither raises nor lowers crash risk in a detectable way, and neither H1a nor H1b survives at the full-sample level. This null average is exactly what the competing-hypotheses design anticipates: disagreement carries both noise and information, and the two forces offset one another. Replacing firm effects with industry and year effects leaves the conclusion intact (0.016 with t = 1.29 for NCSKEW; 0.012 with t = 1.52 for DUVOL).

4.3. Robustness tests

Table 3. Robustness tests

Specification	NCSKEW coef. (t)	DUVOL coef. (t)
Excluding year 2024	0.009 (0.52)	0.007 (0.63)
Firms covered by all three agencies	0.023 (1.19)	0.017 (1.31)
Rank-based disagreement	0.056 (1.05)	0.056 (1.66)*
Crash dummy (linear probability model)	0.010 (1.23)	—

Table 3 subjects the baseline to four perturbations: dropping 2024, the year of extreme market swings; keeping only firms rated by all three agencies; rebuilding disagreement from within-year ranks rather than standardized scores; and replacing the continuous outcomes with the CRASH indicator in a linear probability model. In each case the Disagree coefficient stays positive and insignificant, so the null average relation is not an artifact of any particular choice.

4.4. Mechanism test

Table 4. Mechanism test based on stock price synchronicity

	(1) Synch	(2) NCSKEW	(3) DUVOL
Disagree	-0.001 (-0.22)	0.013 (0.82)	0.011 (1.12)
Synch	—	0.865 (21.18)***	0.471 (17.97)***
Controls	Yes	Yes	Yes
Firm/Year FE	Yes	Yes	Yes
N	27,233	27,233	27,233
R^2 (within)	0.129	0.044	0.035

Column (1) of Table 4 regresses synchronicity on disagreement. The coefficient, -0.001, is far from significance, so the chain from disagreement to richer firm-specific price information to lower synchronicity finds no support on average; H2 is rejected. One plausible reason is that synchronicity mixes information with noise and is therefore an imperfect gauge of transparency, a limitation emphasized in the literature [5]. Columns (2) and (3) show that Synch itself enters the crash regressions with large, highly significant coefficients, which indicates that information-related

DOI: 10.54254/2754-1169/2026.LD38813

quantities do predict crashes; the weak first stage thus more likely reflects the proxy than the theory. Sharper evidence emerges once the sample is conditioned on the information environment.

4.5. Heterogeneity: the conditioning role of the information environment

We partition the sample at the median of analyst forecast dispersion: firm-years above the median form the poor-environment group and those below form the good-environment group. Because FDISP exists only for firms followed by analysts, this exercise relies on 12,821 firm-year observations, split roughly evenly between the two groups.

Table 5. Group regressions by information environment

	Poor env. (1) NCSKEW	Good env. (2) NCSKEW	Poor env. (3) DUVOL	Good env. (4) DUVOL
Disagree	0.023 (0.53)	-0.074 (-1.92)*	0.010 (0.33)	-0.056 (-2.17)**
Controls	Yes	Yes	Yes	Yes
Firm/Year FE	Yes	Yes	Yes	Yes
N	6,410	6,411	6,410	6,411

Table 6. Interaction test of cross-group differences

	(1) NCSKEW	(2) DUVOL
Disagree	-0.060 (-2.02)**	-0.046 (-2.30)**
Disagree × Info_poor	0.097 (2.37)**	0.070 (2.52)**
Info_poor	Yes	Yes
Controls	Yes	Yes
Firm/Year FE	Yes	Yes
N	12,821	12,821

Table 5 reveals a stark contrast. In the good-environment subsample the Disagree coefficients are -0.074 for NCSKEW and -0.056 for DUVOL, both significantly negative; in the poor-environment subsample they remain positive and insignificant. Table 6 tests the difference directly. With the good-environment group as the reference, the stand-alone Disagree coefficient is significantly negative, while the interaction with Info_poor is significantly positive (0.097 and 0.070, each significant at the 5% level). Economically, a one-standard-deviation rise in disagreement lowers NCSKEW by roughly 0.022 standard deviations for good-environment firms. H3 is therefore supported: disagreement behaves as information only where investors possess the channels and capacity to verify it. The pattern also rationalizes the null baseline—opposite-signed group effects cancel when pooled—and mirrors the investor-capability heterogeneity documented by Shao et al. [4].

DOI: 10.54254/2754-1169/2026.LD38813

5. Conclusions and policy implications

This paper asks whether divergence among ESG raters informs or confuses the stock market, and answers with a 2018–2024 panel of Chinese A-share firms. Three findings stand out. First, the average relation between rating disagreement and subsequent crash risk is null, because the noise and information channels counteract each other. Second, stock price synchronicity does not validate the information channel at the aggregate level. Third, a clear boundary condition emerges: disagreement is followed by significantly lower crash risk only among firms with a relatively transparent information environment, and the cross-group contrast is statistically significant. Whether disagreement is noise or information is thus not a universal property; it hinges on the environment in which investors receive and process the ratings.

Several policy implications follow. For regulators of the rating industry, the priority is transparency of methodologies, indicator weights, and underlying data: because disagreement degenerates into noise where the environment is opaque, making the rating process more transparent attacks noise at its origin, and a disclosure-quality assessment mechanism for rating agencies would reinforce this. For exchanges and disclosure regulators, the lesson is that the informational payoff of disagreement depends on the surrounding information infrastructure; more standardized and comparable ESG disclosure lowers the cost of cross-checking ratings and lets disagreement convert into usable market information. For investors, the practical advice is to treat any single agency score with caution, to track how much raters agree or diverge, and—especially for poorly covered firms such as small caps with few analysts—to verify ratings against annual reports and independent research before acting, thereby guarding against greenwashing.

The study is not without limitations. The mechanism evidence rests on synchronicity alone; direct gauges of investor attention, such as search-engine indices or investor-platform inquiries, would allow a sharper test of the information-search channel. Because FDISP is defined only for analyst-covered firms, extending the heterogeneity conclusion to uncovered firms requires caution. Finally, rating disagreement may be endogenous; the entry or exit of rating agencies could furnish future work with instruments or quasi-experimental variation.

References

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

[2] Serafeim, G., & Yoon, A. (2023). Stock price reactions to ESG news: The role of ESG ratings and disagreement. *Review of Accounting Studies, 28*(3), 1500-1530.

[3] 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.

[4] Shao, Y., Zhang, G. D., & Zhang, Y. Y. (2025). ESG rating disagreement and stock price crash risk: Noise effect or information effect? *Audit & Economy Research, 40*(1), 84-94.

[5] Jin, L., & Myers, S. C. (2006). R^2 around the world: New theory and new tests. *Journal of Financial Economics, 79*(2), 257-292.

[6] Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics, 94*(1), 67-86.

[7] Chen, J., Hong, H., & Stein, J. C. (2001). Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics, 61*(3), 345-381.

[8] Kim, J.-B., Li, Y., & Zhang, L. (2011). Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics, 100*(3), 639-662.

[9] Xu, N. H., Jiang, X. Y., Yi, Z. H., & Xu, X. Z. (2012). Analysts' conflicts of interest, optimistic bias, and stock price crash risk. *Economic Research Journal, 47*(7), 127-140.