SURF

Summer Undergraduate Research Fellowship

SURF CODE：SURF-2025-0307

# The Gap Between Actions and Claims: ESG Divergence and Mutual Fund Returns

APPENDICES

Supervisors: Xiaomeng Shi, Oupin Tang

Participants: Xuan Rong, Guhao Shen, Xintong He, Hanrui Yang, Jingxi Zhao, Shuyi Pan, Siyu Liu

## Abstract

This study examines mutual fund ESG Say-Do gap: the difference between funds' self-reported ESG claims and external ratings, and corresponding fund returns. Using Chinese ESG fund data (2019–2024), we find a significant positive link (p<0.01) between claim - practice gap and returns, but this vanishes with time controls. ESG funds experienced growth post-2022, while tighter regulations simultaneously narrowed performance divergence. Our analysis reveals no systemic greenwashing during this period.

## 01 BACKGROUND

The growing ESG investment sector faces significant measurement challenges, particularly in quantifying the "ESG Say-Do Gap" between funds' prospectus claims and third-party ratings. While existing literature demonstrates low inter-agency ESG rating correlations (<0.6) due to methodological heterogeneity (Berg et al., 2022), current natural language processing applications focus on corporate reports rather than fund documents (Kearney & Liu, 2014), leaving fund-level ESG commitments unquantified. This study addresses these gaps by proposing a Transformer-based model to measure mutual fund ESG Divergence—systematic gaps between funds' self-reported claims and external ratings—and examining its impact on returns. The research urgency is heightened in China, where ESG fund assets surged 400% (2019-2024, Wind Data) yet disconnections between Western pricing models and local practices impede accurate ESG value assessment (Martin et al., 2025). This study contributes a novel methodological framework for quantifying fund-level ESG commitment discrepancies while providing empirical insights into how these divergences affect investment performance in an emerging market context.

## 02 DATA COLLECTION

We compiled a battery of databases from authoritative sources, capturing China's pivotal ESG investment development phase between 2019 Q1 and 2024 Q4. Our comprehensive observations are sourced from two databases, Wind and CSMAR.

We obtained fund prospectus data and ESG ratings for ESG mutual funds from Wind Financial Terminal, utilizing legally mandated disclosure documents that establish ESG objectives, principles, and evaluation methodologies. Quarterly fund returns and the CSI 300 Index benchmark data are sourced from Wind and CSMAR databases respectively. We obtain fund-level control variables, including management fees, manager characteristics, and firm attributes from both Wind and CSMAR.

## 03 METHODOLOGY

We adopt the OLS regression to explore the ESG mutual fund practice-claim divergence and its return. To conduct this empirical analysis, we following the **regression formula** below:

$$Return\_gap_i = \beta_0 + \beta_1 * ESG\_gap_i + \beta_2 * Fund\_share_i + \beta_3 * Fee_i + \beta_4 * Num\_fund_i + \beta_5 * Num\_manager_i + FEs + \epsilon_i \quad (1)$$

Where:

Return_gap: Difference between fund return and benchmark CSI 300 index

ESG_gap: Difference between ESG mutual fund claim score from NLP and practice score from agency

Fund_share: Fund share of the fund

Fee: Management fee of the fund

Num_fund: Total number of funds of a fund company

Num_manager: Number of fund managers

FEs: Year-quarter fixed effect, Manager style fixed effect

Our main interest is $\beta_1$ , with a positive sign indicating that greenwashing ESG mutual funds generate higher returns relative to the stock market benchmark. Other vairables are regarded as control variables in the regression.

### Proxies construction

**Dependent Variable:**

Return_gap = Fund quarterly return – CSI 300 Index quarterly return (2)

**Independent Variable:**

ESG_gap = ESG claim score (from NLP analysis) - ESG practice score (from agency rating) (3)*

**NLP analysis:**

For ESG claim score, we use a Transformer-based NLP model to process ESG-related content from the prospectuses and measure ESG disclosure extent (Gillioz et al., 2020). Specifically, we compute the cosine similarity between each prospectus text and a list of ESG seed words from Huazheng ESG Evaluation System covering environmental, social, and governance dimensions to quantify the ESG level. To ensure robustness, similarity computations are conducted 200 times with averaged results. Our key explanatory variable ESG_gap captures the standardized difference between NLP-derived and agency ESG scores:

*ESG_gap = Standardized (ESG from NLP) – Standardized (ESG practice score) (4)

This standardized difference, ESG_gap, captures the degree of ESG disclosure divergence - the gap between what they say and what they actually behave.hat they actually behave.

## 04 DATA ANALYSIS AND RESULTS

**Figure 1** displays the quarterly distribution of ESG mutual funds from 2019 to 2024. According to the figure, the number of ESG funds is rare across the starting-up period and begins to explode since 2022. This growth momentum continues through 2023Q4, indicating a growing market interest in ESG investment vehicles during this period. The number of ESG funds becomes stable from 2024, but still remained at relatively high level.The chart clearly shows a long - term quarterly growth trend in the number of ESG funds from 2019 to 2024, reflecting the increasing recognition and adoption of ESG investment strategies in the financial market over this period.

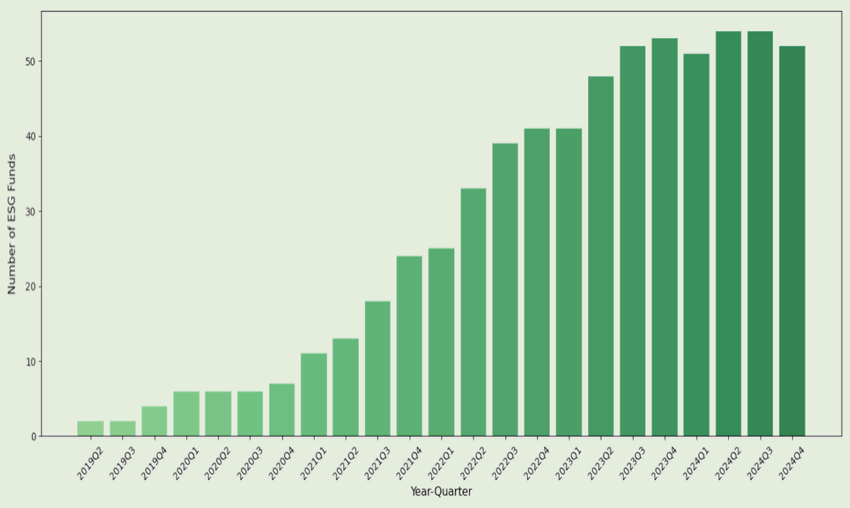


Figure 1. Number of ESG Funds per Quarter (2019-2024)

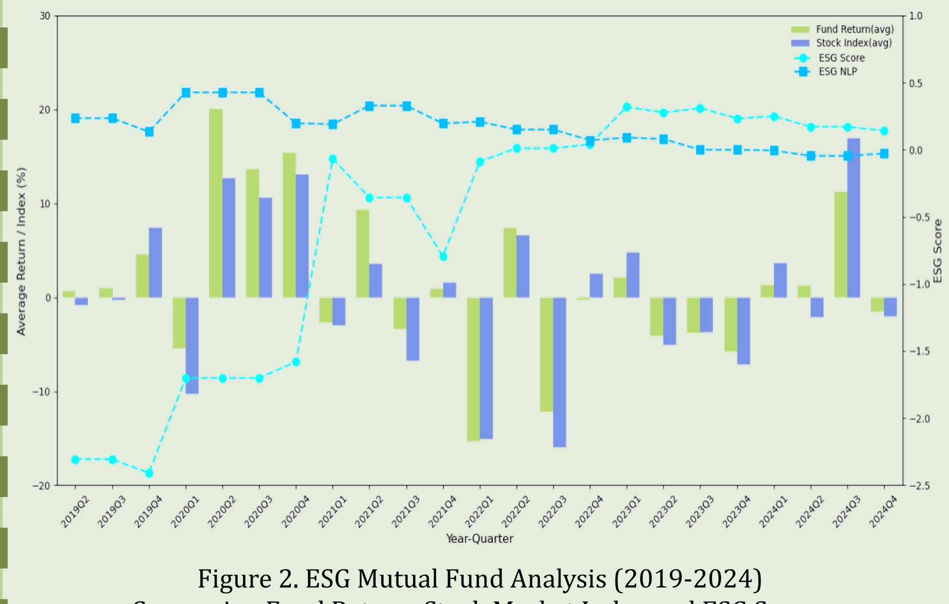


Figure 2. ESG Mutual Fund Analysis (2019-2024)
Comparing Fund Return, Stock Market Index and ESG Scores

**Figure 2** presents the "Say vs. Do" Gap in ESG mutual funds, their returns relative to stock market indices, and distinct discrepancies between fund claims and external ESG ratings. Notably, before 2022, a significant gap existed between ESG NLP (communicated claims) and ESG scores (actual performance), highlighting widespread greenwashing where firms overstated their commitments in disclosures while underperforming in implementation. Since 2022, this gap has narrowed significantly, reflecting strengthened market oversight and regulation that drive companies to align ESG statements with actions—bolstering the integrity and completeness of corporate ESG performance. ESG funds demonstrated strong defensive characteristics during market turbulence in 2019Q4 and 2020Q1, while also capitalizing on recovery opportunities—notably achieving over 20% returns in 2020Q2 compared to the index's 10%. From 2023Q3 onward, ESG fund performance has stabilized with an upward trajectory, reflecting their maturation and increased alignment with broader market dynamics.

**Table 1** provides descriptive statistics for the variables used in our regression analysis. Return_gap averages 0.26% with a median of 0.39% and extreme values spanning -17.33% to 22.45%, reflecting market volatility. The interest variable, ESG_gap, shows a mean divergence of 0.07, indicating the pervasive misalignment between ESG claims and external assessments.Notably, Fee clusters at the 1.20% cap, while Fund_share exhibits severe right-skewness.

| | count | mean | sd | min | max | p50 |
|---|---|---|---|---|---|---|
| Return gap | 642 | 0.26 | 5.56 | -17.33 | 22.45 | 0.39 |
| ESG gap | 620 | 0.07 | 1.55 | -4.08 | 4.99 | 0.03 |
| Fee | 634 | 0.95 | 0.32 | 0.30 | 1.20 | 1.20 |
| Num fund | 634 | 207.96 | 153.17 | 28.00 | 476.00 | 151.00 |
| Fund share | 616 | 210.60 | 360.38 | 0.00 | 2735.64 | 68.30 |
| Num manager | 634 | 57.65 | 37.06 | 11.00 | 127.00 | 46.00 |
| Style | 486 | 2.77 | 1.31 | 1.00 | 4.00 | 3.00 |

Table 1

| DV: Return_gap | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| ESG_gap | 0.362** | 0.464** | 0.612*** | 0.235 |
| | (0.162) | (0.182) | (0.170) | (0.148) |
| Fund_share | | -0.000 | 0.000 | -0.001 |
| | | (0.001) | (0.001) | (0.001) |
| Fee | | -1.757*** | -1.089** | -1.579** |
| | | (0.413) | (0.528) | (0.720) |
| Num_fund | | 0.002 | -0.004 | -0.007** |
| | | (0.006) | (0.007) | (0.003) |
| Num_manager | | -0.006 | 0.021 | 0.035** |
| | | (0.023) | (0.026) | (0.015) |
| Constant | 0.049 | 1.814*** | 0.384 | 1.135 |
| | (0.171) | (0.477) | (0.599) | (0.824) |
| Style FEs | No | No | Yes | Yes |
| Year_quarter FEs | No | No | No | Yes |
| Observations | 585 | 577 | 431 | 430 |
| Adj. $R^2$ | 0.010 | 0.013 | 0.038 | 0.397 |

Table 2

**Table 2** reports OLS estimates examining the relationship between ESG_gap and Return_gap. Column (1) shows a baseline positive and significant effect of ESG_gap on returns. This effect strengthens with control variables in Column (2) and further intensifies under Style fixed effects in Column (3), suggesting that initial estimates are downwards biased by omitted investment heterogeneity. However, introducing Year-quarter fixed effect in Column (4) renders ESG_gap statistically insignificant, indicating the apparent "greenwashing premium" is highly likely attributed to time-varying market factors. Moreover, Fee consistently erodes returns across all models, aligning with expense sensitivity. Nevertheless, our analysis is constrained by sample reduction induced by Style, which may bias coefficients. (Footnotes:To enhance the readability of the data in the tables, we have multiplied Return_gap by 100.)

## 05 DISCUSSION AND CONCLUSION

Our findings challenge conventional wisdom about ESG greenwashing premiums. While baseline models suggest funds with larger claim-practice gaps earn higher returns ($\beta_1$ = 0.612, p<0.01), this relationship disappears when controlling for time-varying conditions. The dramatic narrowing of ESG gaps post-2022 coincides with China's regulatory tightening, suggesting effective policy intervention. Before 2022, widespread divergences between prospectus claims and external ratings indicated prevalent overstatement. The convergence demonstrates improved market discipline and authentic ESG integration.

Our NLP methodology advances ESG authenticity measurement by capturing disclosure-practice misalignment that traditional rating-based approaches miss. However, sample reduction from style controls (577 to 431 observations) may introduce selection bias, limiting generalizability.

The explosive growth in ESG funds (400% asset surge 2019-2024) alongside evolving claim-practice dynamics highlights ongoing supervision needs. Future work should explore heterogeneous effects across market cycles, investor types, and regulatory regimes to better understand ESG investment authenticity mechanisms.

*Key References:*

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

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.

Gillioz, A., Casas, J., Mugellini, E., & Khaled, O. A. (2020). Overview of the Transformer-based Models for NLP Tasks. In *2020 15th Conference on Computer Science and Information Systems (FedCSIS)* (pp. 179-183). IEEE.

Kearney, C., & Liu, S. (2014). Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, *33*, 171-185.

Martin, H., Zhou, Y., & Raman, R. (2025). Financial Metrics and Environment, Social, Governance (ESG) Performance: A Cross Border Comparison of China and the UK Construction Industries. *Buildings*, *15*(8), 1236.

WORLD ENVIRONMENT DAY