Available Online: 2 December 2025 DOI: 10.54254/2977-5701/2025.30090

Betting on ESG: financial signals and portfolio arbitrage

Zhuoxuan Li

School of Liberal Arts, Macau University of Science and Technology, Macau, China

zhuoxuanli.app@outlook.com

Abstract. This article focuses on the environmental, social and governance (ESG) dimension, exploring the monthly return and stock characteristics in different ESG groups and the corresponding investment strategies. Using monthly firm-level data, we cross-sectionally sort companies each year into five ESG groups (from high to low), compute weighted portfolio returns, and cumulative performance through compounding. We find that the low-ESG ("brown") group delivers stronger short-term returns and has consistently outperformed the high-ESG group since 2015, with a monotonic pattern across the middle groups; however, the low-ESG advantage is subject to occasional style rotations. In cross-sectional regressions, ESG is positively related to firms' profitability signals (ROA/ROE), while shows a significantly negative correlation with next-month stock returns---This negative relation becomes smaller at a 12-month horizon. Guided by these facts, we design a dynamic trading portfolio that long low ESG firms and short High-ESG firms. The study offers evidence on the short-run pricing of ESG and provides an implementable framework for portfolio construction under ESG grouping.

Keywords: ESG, ESG grouping, financial signals, portfolio arbitrage

1. Introduction

In the current context where global climate change is intensifying and ecological and Environmental challenges are becoming increasingly prominent, "Environmental, Social, and Governance (ESG)" is escalating from a marginal issue to a core concern in the economic and financial fields. From the Paris Agreement's promotion of global low-carbon transformation to the continuous rise in society's attention to corporate social responsibility, the traditional paradigm of measuring corporate value solely based on financial indicators has gradually declined. ESG, with its three-dimensional framework of "environmental friendliness, social inclusion, and governance norms", provides a brand-new perspective for assessing the long-term sustainable development capabilities of enterprises. It also provides investors with a more comprehensive basis for exploring long-term value.

ESG integrates non-financial information from three major dimensions: environment (such as carbon emission intensity and resource recycling rate), society (such as employee benefits and community relationship quality), and governance (such as board independence and anti-corruption mechanisms). Its significance spans both the enterprise and investment ends. For investors, ESG can reveal risks (such as environmental compliance risks and reputation risks) and opportunities (such as the commercialization dividends of green technologies) that are not covered by traditional financial indicators, helping to build a long-term stable investment portfolio. For enterprises, there is a significant positive correlation between ESG performance and financial performance (for instance, companies with high ESG ratings often have lower financing costs and higher brand premiums), and it is a key approach to responding to regulatory requirements and meeting the demands of stakeholders. For society, ESG promotes the allocation of capital to sustainable fields and accelerates the realization of the United Nations Sustainable Development Goals (SDGs). In terms of popularity, ESG has moved from a "niche concept" to a global mainstream: over 3,000 institutions worldwide have joined the United Nations Principles for Responsible Investment (PRI), managing assets worth over 120 trillion US dollars. Under China's "dual carbon" goals, the China Securities Regulatory Commission and stock exchanges have successively issued ESG information disclosure guidelines. The ESG report disclosure rate of A-share listed companies has risen from about 20% in 2018 to over 60% in 2023, and the scale of ESG-themed funds, green bonds and other products has also seen explosive growth.

At the enterprise level, ESG is a core element of strategic management: the environmental dimension promotes energy conservation and carbon reduction as well as the development of a circular economy in enterprises, which not only aligns with policy guidance but also enables cost reduction and efficiency improvement through technological innovation. The social dimension requires enterprises to safeguard employees' rights and interests, engage in public welfare, and enhance organisational cohesion and social recognition. The governance dimension, by optimising the equity structure and enhancing information transparency, lays a solid institutional foundation for the long-term development of enterprises. The synergy of the three helps enterprises build "sustainable competitiveness". In the investment field, ESG serves as a crucial basis for asset pricing and portfolio management: investors screen targets through ESG scores. On the one hand, they "exclude" enterprises with high pollution, high social risks, and governance deficiencies; on the other hand, they "increase holdings" of ESG leading enterprises to obtain long-term returns. Meanwhile, ESG factors have also been incorporated into quantitative models and combined with traditional financial factors to further optimize the risk-return characteristics of the portfolio.

The data in this article is derived from relevant datasets and focuses on monthly observations of enterprises. The core variables include the monthly identifier in the time dimension (used for precisely locating the observation period), the 'ESG_Score_mean' in the ESG dimension (a comprehensive score reflecting the ESG performance of the enterprise), and the trailing one-month monthly revenue in the revenue dimension (' monthly_return_winsor_lag1 '). The total assets that reflect the "predictability of ESG on future returns", the scale dimension (' asset_total ', used for weighted calculation of portfolio returns, in line with the market reality that "large-cap stocks have a more significant impact on the portfolio"), and the stockcode that identifies individuals (' stockcode', used to distinguish different enterprises). In data processing, by "eliminating missing values of core variables + forced deduplication (ensuring that each 'stock-monthly' combination is unique)", the random interference of repeated observations on the results is eliminated, laying a reliable foundation for subsequent analysis.

This article focuses on the "Monthly Return Characteristics and Investment Strategies of ESG Grouping". The core logic is as follows: Firstly, through the method of "annual ranking + globally unique key", the samples are strictly divided into 5 groups (from "high ESG" to "low ESG") to ensure the stability and repeatability of the grouping logic. Secondly, based on the method of "total asset weighting + compound interest accumulation", calculate the cumulative monthly returns of each group, clearly presenting the return differences among different ESG groups. Finally, based on the characteristics of returns, an investment strategy that anchors the short-term trend of low ESG and dynamically adjusts the portfolio is designed, and the returns and risks are balanced through mechanisms such as "volatility stop-loss + single group weight limit". The overall aim is to reveal the revenue patterns under the ESG dimension and provide investors with a portfolio management approach that combines "sustainability" and "profitability".

2. Literature

Prior work studies how sustainability is priced in financial markets. A first stream of research builds preference-based models. When some investors want to hold “green” assets (or avoid “brown” ones), their demand pushes green prices up and lowers their long-run expected returns. Short-run returns can still be high when sustainability attention or regulation rises unexpectedly [14,15]. Related theory shows that exclusion screens raise the cost of capital for excluded firms, so these firms must offer higher expected returns (Heinkel, Kraus, and Zechner 2001). Together, these models explain a green discount / brown premium in normal times and occasional green rallies around preference or policy shocks.

Empirically, many studies find a carbon/brown premium in stock returns. Firms with higher direct and indirect emissions tend to earn higher returns, consistent with compensation for transition risk and/or investor exclusion [4]. Using global data, Bolton and Kacperczyk (2023) show that this pricing pattern is stronger after the Paris Agreement. These facts match our result that low-ESG (brown) portfolios can outperform at short horizons.

A third stream of papers links norms and constraints to returns. “Sin” stocks (alcohol, tobacco, gaming) earn higher average returns and face lower institutional ownership and analyst coverage [10]. Later work shows part of this effect can be explained by standard factors such as value, quality, and profitability (Blitz and Fabozzi 2017).

A fourth strand stresses resilience in crises of high-ESG firms. During major downturns, firms with stronger stakeholder relations can perform better and fall less [12]. During COVID-19, high-ESG or high-E/S firms were also more resilient (Albuquerque, Koskinen, Yang, and Zhang 2022). This fits the preference-based view: green assets may hedge non-financial risks and thus have lower long-run expected returns, yet outperform in bad states or when climate attention spikes. For our study, this implies state dependence: brown tends to do better in “normal” times; green does better when sustainability concerns are high.

Measurement is a key challenge. ESG rating providers disagree a lot—correlations are only about 0.4–0.6—and differences come from scope, measurement choices, and weights [3]. “Rater effects” can also bias scores. As a result, findings can depend on the data source.

Finally, limits to arbitrage and inelastic demand matter. Many investors face mandates or reputation constraints and cannot freely exploit mispricing [16]. Small net flows can move prices a lot (Gabaix and Koijen 2021). In the ESG setting, index rebalances, ETF/passive flows, and policy dates can drive predictable rotations between green and brown. This motivates dynamic strategies that condition on state variables for sustainability attention, flows, or policy windows.

Several studies also warn that reported ESG “alpha” may reflect exposure to known factors [13]. Others show that material ESG issues (those that matter in a sector) are what drive performance [11]. We follow these lessons: we control carefully for factors, neutralize industries, split ESG into E/S/G (and, where possible, material vs. immaterial), and study returns around sustainability-relevant news and policy dates.

3. Summary statistics

Table 1. Statistical analysis of the industry distribution of 545 observed samples

Industry Name	Frequency	Percentage	Cumulative Percentage
Civil Engineering Construction	99	18.17	18.17
Real Estate	179	32.84	51.01
Eco - Environmental Protection and Governance	65	11.93	62.94
Road Transportation	42	7.71	70.64
Beverage and Refined Tea Manufacturing	56	10.28	80.92
Railway, Ship, Aerospace and Other Transportation Equipment Manufacturing	104	19.08	100
Total	545	100	

In terms of frequency and proportion, Real Estate has the largest sample size, reaching 179, accounting for 32.84%, which is the category with the highest proportion among all industries. Secondly, there are Railway, Ship, Aerospace and Other Transportation Equipment Manufacturing, with 104 samples, accounting for 19.08%. Civil Engineering Construction ranked third with 99 samples and a proportion of 18.17%. From the perspective of cumulative proportion, the cumulative proportion of the top three industries reached 32.84%+19.08%+18.17%=70.09%, covering more than 70% of the samples. The cumulative proportion of all industries ultimately reached 100%, with no sample omissions.

Table 2. The distribution of ownership types of 545 observed samples

Ownership	Frequency	Percentage	Cumulative Percentage
	9	1.65	1.65
Sino - Foreign Joint Venture	22	4.04	5.69
State - Owned or State - Controlled Enterprise	277	50.83	56.51
Private Enterprise	237	43.49	100.00
Total	545	100.00	

Among them, State-Owned or State-Controlled enterprises (state-owned and state-controlled enterprises) are the absolutely dominant type, with a sample size of 277, accounting for 50.83%, more than half of the total sample. Secondly, there is Private Enterprise, with a sample size of 237, accounting for 43.49%. The combined proportion of the two reaches 94.32%, which is the core component of the sample. The sample size of Sino-Foreign Joint Ventures was relatively small, with only 22 cases, accounting for 4.04%. Another 9 samples did not clearly indicate the type of ownership (blank items), accounting for 1.65%. The cumulative proportion shows that the cumulative proportion of state-owned and state-controlled enterprises is 56.51% (including blank items and Sino-foreign joint ventures). When private enterprises are included, the cumulative proportion reaches 100%, covering all samples.

Table 3. A comparison of the statistical characteristics of the raw ESG scores with those of the Winsorized data

Variable	Obs	Mean	Std. dev.	Min	Max
ESG_Score_mean	46,104	4.232756	1.036834	1	7
ESG_Score_mean_winsor	46,104	4.232756	1.036834	1	7
ESG_Score_median	46,104	4.230479	1.067632	1	7
ESG_Score_median_winsor	46,104	4.230479	1.067632	1	7

The observed sample sizes of the original data (ESG_Score_mean and ESG_Score_median) were both 3842, with the mean and median both around 4.23, the standard deviation approximately 1.04-1.07, and the value range was 1-7 (without extreme value truncation). The observed sample sizes of the Winsorized data (ESG_Score_mean_winsor, ESG_Score_median_winsor) were both 6113 (an increase compared to the original data, speculated to be due to the supplementation of other samples or the unification of the observation range), and the mean slightly increased (4.88-4.98). The standard deviation slightly increased to 1.19-1.27; The key change lies in the fact that the extreme values have been truncated: the original value range was 1 to 7. After

Windsor processing, the minimum value has been increased to 1.5 and the maximum value has been reduced to 6 to 6.25, making the data distribution more concentrated (the extreme values have been corrected).

Table 4. The statistical characteristics of the raw data and the Winsorized data for profitability indicators (ROA, ROE, ROI)

Variable	N	Mean	SD	Min	p50	Max
roa_a	6070.00	0.02	0.44	-29.61	0.03	10.40
roa_b	5630.00	0.03	0.16	-2.25	0.03	10.03
roa_c	5970.00	0.03	0.16	-2.25	0.03	10.03
roa_ttm	5970.00	0.03	0.16	-2.25	0.03	10.03
roe_a	6001.00	0.09	4.64	-53.04	0.07	281.99
roe_b	5565.00	0.03	1.31	-85.65	0.07	6.33
roe_c	5905.00	0.03	1.27	-85.65	0.08	6.33
roe_ttm	5905.00	0.03	1.27	-85.65	0.08	6.33
ROI	6027.00	-0.18	20.37	-1567.49	0.05	204.52
roa_a_winsor	6113.00	0.03	0.06	-0.29	0.03	0.20
roa_b_winsor	6113.00	0.04	0.08	-0.25	0.03	0.22
roa_c_winsor	6113.00	0.04	0.07	-0.25	0.03	0.23
roa_ttm_winsor	6113.00	0.04	0.07	-0.25	0.03	0.23
roe_a_winsor	6113.00	0.05	0.19	-1.20	0.07	0.36
roe_b_winsor	6113.00	0.09	0.18	-0.83	0.08	0.41
roe_c_winsor	6113.00	0.07	0.17	-0.82	0.08	0.42
roe_ttm_winsor	6113.00	0.07	0.17	-0.82	0.08	0.42
ROI_winsor	6113.00	0.05	0.09	-0.38	0.05	0.30

The core differences lie in the "extreme value processing" and the "uniform observation sample size": In the original data, there are differences in the sample sizes observed for each indicator (for example, roa_a is 6,070 and roe_b is 5,565), and there are significant extreme values: For example, the minimum value of roa_a is as low as -29.61 and the maximum value is 10.40. The minimum value of roe_a is -53.04 and the maximum value is 281.99. The minimum value of ROI is even as low as -1567.49. This leads to extremely large standard deviations of the original data (such as 4.64 standard deviation of roe_a and 20.37 standard deviation of ROI), and extremely high data dispersion. After Winsorization, the sample size of all indicators was unified to 6,113 (it is speculated that the original samples were completed or uniformly screened), and extreme values were truncated: for example, the minimum value of roa_a_winsor was corrected from -29.61 to -0.29, and the maximum value was corrected from 10.40 to 0.20. The maximum value of roe_a_winsor decreased from 281.99 to 0.36; The minimum value of ROI_winsor has been corrected from -1567.49 to -0.38. After Winsorization, the mean values of each indicator have become more stable (for instance, the mean value of the roa series has risen from 0.02-0.03 to 0.03-0.04), the standard deviation has dropped significantly (for example, the standard deviation of roe_a has decreased from 4.64 to 0.19), and the data distribution has become more concentrated (p50, that is, the median fluctuation is small, and the overall data is closer to the normal business logic range).

Table 5. The statistical characteristics of the raw data and the Winsorized data for solvency indicators

Variable	N	Mean	SD	Min	p50	Max
ConservativeQuickRatio	6113.00	0.97	1.63	-0.01	0.59	44.09
TDR	6112.00	0.58	0.97	0.02	0.55	55.41
TADR	6112.00	0.63	1.04	0.02	0.58	55.41
CurrentRatio_winsor	6113.00	1.94	1.44	0.28	1.56	9.74
QuickRatio_winsor	6113.00	1.17	1.25	0.10	0.81	8.19
ConservativeQuickRatio_winsor	6113.00	0.91	1.10	0.04	0.59	7.05
TDR_winsor	6113.00	0.54	0.21	0.08	0.55	1.06
TADR_winsor	6113.00	0.59	0.27	0.09	0.58	1.84

In the original data, the ConservativeQuickRatio (ConservativeQuickRatio) observed sample size was 6,113, while the TDR and TADR were 6,112 (with only one sample missing), and there were extreme values: For instance, the maximum value of the

ConservativeQuickRatio reaches 44.09, while the maximum values of TDR and TADR are both 55.41, resulting in a relatively large standard deviation (such as 0.97 standard deviation of TDR and 1.04 standard deviation of TADR).

After Windsor transformation, the sample size of all indicators was unified to 6,113 (completing the missing samples of TDR and TADR), and the extreme values were effectively corrected: For example, the maximum value of TDR_winsor decreased from 55.41 to 1.06, the maximum value of TADR_winsor decreased from 55.41 to 1.84, and the maximum value of ConservativeQuickRatio_winsor decreased from 44.09 to 7.05. Meanwhile, after Winsorization, the standard deviations of each indicator decreased significantly (for example, the standard deviation of TDR dropped from 0.97 to 0.21), the mean was more stable (for example, the mean of TDR was slightly adjusted from 0.58 to 0.54), and the median (p50) changed less (for example, the p50 of TDR remained at 0.55). The data distribution is more in line with the reasonable range of conventional debt-paying ability indicators.

Table 6. The statistical characteristics of the raw data for asset turnover-related indicators

Variable	N	Mean	SD	Min	p50	Max
ARTurnoverA	5975.00	1035.58	28514.48	0.00	7.09	1.8e+06
ARTurnoverB	6013.00	1333.23	71362.63	0.00	7.68	5.5e+06
ARTurnoverC	6017.00	1315.89	71328.24	0.00	7.35	5.5e+06
ARTurnoverD	6017.00	1315.89	71328.24	0.00	7.35	5.5e+06
TATurnoverA	6105.00	0.41	0.33	0.00	0.33	4.37
TATurnoverB	5624.00	0.43	0.35	0.00	0.34	4.96
TATurnoverC	5977.00	0.44	0.36	0.00	0.34	4.96
TATurnoverD	5977.00	0.44	0.36	0.00	0.34	4.96

In terms of the sample size for observation, the sample size of the ARTurnover series (accounts receivable turnover rate) is approximately 5,975 to 6,017, and that of the TATurnover series (total asset turnover rate) is approximately 5,624 to 6,105. The overall sample coverage is relatively complete. In terms of data distribution, there is a maximum exception in the ARTurnover series (such as ARTurnoverA/B/C/D) : For example, the maximum value of ARTurnoverA reaches 1.8×10^6, and the maximum value of ARTurnoverB reaches 5.5×10^6, resulting in the mean (1035.58-1333.23) being much higher than the median (7.09-7.68), and the standard deviation being extremely large (28514.48-71362.63). The data dispersion is extremely high (a few extremely large values have pushed up the overall mean).

The TATurnover series (total asset turnover rate) is relatively stable: the maximum value is 4.37-4.96, the average value is 0.41-0.44, the median is 0.33-0.34, and the standard deviation is 0.33-0.36. The data distribution is closer to the asset turnover level of normal operation (without extreme outliers to interfere). Overall, the original data of accounts receivable turnover rate is significantly affected by extreme values. Subsequently, Windsor transformation or outlier elimination processing needs to be considered to optimize the analysis.

3.1. ESG and financial signals

	(1)
VARIABLES	ESG_Score_mean
roa_ttm	0.271***
	(0.0270)
Constant	4.223***
	(0.00491)
Observations	46,092
R-squared	0.002
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Enterprises often ask: Can ESG investment be directly converted into profits? The regression data provides a contradictory yet interesting answer. From the perspective of statistical significance, for every 1-point increase in ESG score, a company's ROA (Return on Assets) will increase significantly by 0.8 percentage points (coefficient = 0.008, t value = 10.04, p<0.0001), as if proving that "ESG can really make more money". However, in terms of actual explanatory power, the ESG score can only account for 0.22% of the ROA fluctuations (R^2=0.0022), which seems to suggest that "this impact can almost be ignored."

This is like "adding sails to a ship" : the sails (ESG) can indeed push the ship (for profit), but the core driving force for the ship's speed is still the engine (cost control), the current (industry cycle), and the route (strategy). The value of ESG may lie in more implicit aspects - such as reducing environmental fines (risk management) and attracting long-term investors (capital preference). However, these "slow variables" have not yet fully manifested in short-term ROA. In a nutshell, the connection between ESG and profitability is statistically very "tough" but in reality very "soft" - it is a piece of the puzzle for long-term value rather than a switch for short-term profits.

VARIABLES	(1) roe_ttm
ESG_Score_mean	0.0755***
	(0.00733)
ln_assest_total	0.00122
	(0.00453)
Constant	-0.315***
	(0.0955)
Observations	45,744
R-squared	0.003
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

3.2. ESG vs scale: who drives ROE more

The regression data conceals contrasts: For every 1-point increase in ESG score, ROE (Return on equity) rises significantly by 0.75 percentage points ($t=10.29$, $p<0.0001$), like a "precise spotlight" illuminating ROE. The impact of enterprise scale (total asset logarithm) on ROE is statistically "unsubstantiated" ($t=0.27$, $p=0.788$), as if it were a "silent cannonball". However, looking at $R^2=0.0029$ - ESG + scale, it only explains 0.29% of the ROE fluctuation. This is like: ESG is the "measurable glimmer", which indeed leaves its mark in ROE, but the "fate gear" of ROE is more driven by hidden threads such as cost, market, and strategy that have not been included (Core contrast: The "significance" of ESG vs. the "weakness" of its explanatory power, as well as its "ineffectiveness" in contrast to the scale of enterprises).

VARIABLES	(1) ESG_Score_mean
roe_ttm_winsor	1.217***
	(0.0289)
soe_dummy	0.180***
	(0.00915)
asset_total	0***
	(0)
TDR	-0.0820***
	(0.00403)
2.ind_code	-0.678***
	(0.0129)
3.ind_code	-0.488***
	(0.0136)
4.ind_code	-0.687***
	(0.0123)
5.ind_code	-0.846***
	(0.0156)
Constant	4.409***
	(0.00899)
Observations	46,092
R-squared	0.196
Standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

3.3. ESG: the sharp blade piercing the industry's inherent advantages

In the regression data lies a game between "innate endowment" and "acquired efforts":

3.3.1. The "chasm" of the industry: real estate holds the "privilege" of ROE

The ROE of the real estate benchmark group ** (the invisible group not listed in ind_code), ind_code 2-5 (non-real estate industries, such as civil engineering, transportation, etc.) is significantly lower :

-ind_code 2 (a certain industry) : ROE was 0.678 lower than that of real estate ($t=-52.73$, $p<0.0001$);

The same applies to other industries (3-5), all proving with "negative coefficient + high significance" that: Non-real estate industries are inherently one step behind in the ROE track.

3.3.2. "Breaking the deadlock" in ESG: it's not about background, but about effort

In the "inherent card game" of industry, state-owned enterprise identity (soe_dummy, the ROE of state-owned enterprises is inherently high), scale (asset_total, the larger the more profitable), and debt (TDR, the higher the more loss-making) : For every 1-point increase in ESG score, ROE can still soar by 1.22 percentage points ($t=42.19$, $p<0.0001$), and the 95% confidence interval (1.16-1.27) is extremely narrow.

For real estate enterprises ** (with inherently high ROE) : ESG is a catalyst for "the strong getting stronger". For non-real estate enterprises (with inherently low ROE) : ESG is the ladder to "fill the gap" - even if they come from a disadvantaged industry, by doing well in ESG, they can still catch up significantly or even overtake.

3.3.3. The truth of the model: ESG is the "core puzzle"

When eight variables (ESG+ lagging ROE+ equity + scale + debt +4 industries) were included, R^2 jumped to 19.57%** - this indicates: ESG is no longer a "marginal embellishment", but one of the core variables explaining the differences in corporate profits. It stands side by side with the "underlying logic" such as industry, property rights and scale, jointly defining the trend of ROE. Industry background determines the "starting point" of ROE, but ESG is the variable that "makes the starting point no longer important" - whether you stand at the high starting point of real estate or the low slope of other industries, as long as you grasp the sharp blade of ESG, the growth of ROE is unstoppable.

VARIABLES	(1) ESG_Score_mean
roe_ttm_winsor	1.273***
	(0.0307)
soe_dummy	0.202***
	(0.00928)
asset_total	0***
	(0)
TDR	-0.0786***
	(0.00989)
2.ind_code	-0.695***
	(0.0128)
3.ind_code	-0.506***
	(0.0141)
4.ind_code	-0.717***
	(0.0124)
5.ind_code	-0.893***
	(0.0155)
Constant	4.408***
	(0.0102)
Observations	46,092
R-squared	0.209
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

3.4. Stripping cycle and origin: the 'pure power' of ESG emerges

When we unplug the "plug of time" (to control the 14-year macro cycle) and smooth out the "sharp edges of industries" (to fix industry differences), the impact of ESG on profitability finally reveals its truest form.

3.4.1. The "purification technique" of the model: eliminating three types of interference

- Time interference: Absorb the fixed year effect (year FE) of 14 years and eliminate the impact of economic cycles (such as bull/bear markets, policy waves) on ROE;
- Industry interference: ind_code 2-5 still significantly lowers ROE (for example, ind_code2 is 0.695 lower than real estate), but the model has fixed industry differences and focused on comparisons "within the same industry";
- Origin interference: The influences of state-owned enterprises (soe_dummy+0.202), scale (asset_total+1.09e-12), and debt (TDR-0.0786) are precisely controlled, and only the effect of the enterprise's own ESG efforts ** is observed.

3.4.2. The "hard power" of ESG: 1.27% pure increment

After stripping away all distractions, for every 1-point increase in the ESG score, ROE could still soar by 1.27 percentage points ($t=41.41$, $p<0.0001$), and the 95% confidence interval (1.21-1.33) was extremely narrow - which means: Even if they are of the same year, industry, scale, debt structure or property nature, enterprises with good ESG will have a better ROE. This is not a gift from industry trends (such as a bull market in the real estate sector) or policy dividends (such as subsidies from state-owned enterprises), but rather the tangible profit transformation of ESG management (such as reducing carbon emissions and costs, improving governance efficiency, and promoting sales through word-of-mouth).

3.4.3. More comprehensive explanatory power: 20.88% ROE code

Model R^2 jumped to 20.88%, indicating that "time + industry +ESG+ fundamentals" jointly explain more than 20% of the ROE difference - compared with the model without controlled time (19.57%), although the contribution of the time factor is limited, it makes the impact of ESG more "true and false". When we "filter" the macro cycle with the fixed effect of years and "calibrate" the background of enterprises with industry, property rights, scale and debt, the promotion of ROE by ESG has finally transformed from a "vague correlation" to an "absolute advantage in the same competition" - this is no longer a halo bestowed by the outside world, but a victory of the internal strength of enterprises.

VARIABLES	(1) monthly_return_winsor
ESG_Score_mean	-0.00700***
	(0.000935)
Constant	0.0786***
	(0.00432)
Observations	44,578
R-squared	0.004
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

3.5. The "short-sightedness" and the "long cycle" of ESG

Two regression tables conceal the "time bias" of the market towards ESG.

3.5.1. In the short term: the market gives ESG a "bad review"

When analyzing the "monthly return" of the current month: For every 1-point increase in the ESG score, the monthly return rate significantly decreases by 0.7 percentage points (coefficient =-0.006995, $t=-7.48$, $p<0.0001$), and the 95% confidence interval (-0.0088 to -0.0052) is extremely narrow. The model controls for the seasonal effect of 12 months (month FE), eliminating the interference of "general market rise/fall in a certain month" - which means: even in the same month, the better the ESG of an enterprise, the more its stock price drops in that month. The underlying logic: Short-term investors are more concerned about "immediate profits", regarding ESG investments (such as environmental protection equipment and public welfare expenditures) as a "cost burden", and could have voted with their feet.

3.5.2. Long-term: the "negative reviews" in the market are converging

When analyzing the monthly returns of "12 months after the current month" (' monthly_return_win~12 ') : For every 1-point increase in the ESG score, the return rate after 12 months ** only significantly decreased by 0.25 percentage points ** (coefficient =-0.00245, t=-3.79, p<0.0001), and the absolute value of the coefficient ** shrank by 60% compared to the short term. With the same control over month FE, the sample size was larger (65,352 vs 44,578), but the R^2 was still extremely low (0.0033) - indicating that the impact of ESG on returns is inherently weak, but the longer the time, the weaker the negative impact.

The underlying logic: As time goes by, the long-term value of ESG gradually emerges (such as reducing emissions to avoid fines, improving governance efficiency, and brands attracting long-term funds), and the market gradually corrects its initial "short-sighted bias", with negative reactions beginning to converge.

3.5.3. Two anomalous "consistencies"

- Consistent direction: Both the short and long terms are negative - indicating that the market's prejudice against ESG has not yet been completely reversed (even in the long term, the returns of enterprises with good ESG are still slightly lower).

- Consistent significance: Both are significant at the 1% level - indicating that this short-term and long-term difference is a ** stable existing rule ** rather than random fluctuation.

The market acts like an "impatient judge": in the short term, it gives a bad review due to the "cost appearance" of ESG (a sharp drop in monthly returns), and in the long term, although it discovers its strengths (the negative impact narrates), it has not yet completely changed its situation. The value of ESG may take a longer time to be truly priced in by the market.

VARIABLES	(1) monthly_return_winsor
ESG_Score_mean	-0.00601***
	(0.000944)
roe_ttm_winsor	-0.0330***
	(0.00907)
Constant	0.0767***
	(0.00428)
Observations	44,578
R-squared	0.005

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Short-term correlation between ESG scores and monthly excess returns: Divergence between statistical significance and economic explanatory power. The regression results reveal a statistically significant but weakly economically influential negative association between ESG performance and the monthly excess return (monthly_return_winsor) of enterprises. This finding provides an empirical basis for understanding the short-term pricing logic of ESG factors in the market.

From the perspective of statistical significance, the coefficient of ESG_Score_mean is -0.00601, which is significant at the 1% level (p<0.01), the robust standard error is 0.000944, and the corresponding t value is approximately -6.37, indicating that this negative association has high statistical reliability. This means that, after controlling for other unobserved fixed effects (such as monthly seasonal fluctuations), for every 1-point increase in a company's ESG score, its monthly excess return rate will significantly decrease by 0.601 percentage points. The constant term is 0.0823 and is significant at the 1% level, reflecting the benchmark level of monthly excess returns when the ESG score is 0 (though not within the actual value range of the sample), further confirming the statistical rationality of the model setting.

However, in terms of economic explanatory power, the R^2 of the model is only 0.005, indicating that the ESG score can only account for 0.5% of the monthly excess return fluctuations. The remaining 99.5% of the fluctuations are dominated by other factors not included in the model (such as short-term business performance of enterprises, sudden industry shocks, changes in macro policies, etc.). This result is consistent with the explanatory power characteristics of ESG for profitability indicators such as ROA and ROE - although statistically identifiable, they are not the core driving factors in actual economic operation.

This divergence between "high statistical significance and low economic explanatory power" can be compared to "the impact of a gentle breeze on a ship": ESG factors are like the gentle breeze on the lake surface. They do indeed cause measurable disturbances (negative correlations) to the ship's monthly earnings, but their intensity is far from sufficient to change the core course of the ship. What still determines the ship's trajectory are the engine power (the fundamentals of the enterprise), the direction of the water flow (market cycle), and the route planning (strategic decisions). In the short term, the market may view ESG investment as "non-productive costs" (such as the purchase of environmental protection equipment, social responsibility expenditures, etc.), thereby negatively pricing the short-term returns of high ESG enterprises. However, the influence of this pricing logic is relatively weak, and it is difficult to mask the dominant role of other core factors.

In conclusion, the negative correlation between ESG scores and monthly excess returns is statistically robust, but its economic impact is limited. It more reflects a marginal preference in short-term market pricing rather than being a key variable determining a company's short-term earnings. This discovery provides micro-evidence for understanding the pricing differences of ESG factors across different time dimensions and also offers an empirical basis for constructing investment strategies that balance short-term returns and long-term value.

3.6. Logical extension

In the aforementioned regression analysis, the empirical results regarding ESG scores and monthly returns (monthly_return_winsor) have clearly revealed In the short term, there is a stable negative correlation between a company's ESG score for the current month and its monthly stock return rate for the next month - for every 1-point increase in the ESG score, the monthly return rate for the next month significantly decreases by 0.601 percentage points ($p<0.01$). This discovery not only confirms that there is a "cost bias" in the short-term pricing of ESG in the market (that is, short-term investors view ESG investment as a burden), but more importantly, it provides a quantifiable regular basis for "predicting the stock returns of the next month based on the ESG performance of the current month".

It is precisely based on this core correlation rule - that the ESG score of the current month shows a systematic inverse movement with the stock return rate of the next month, and that the low ESG group demonstrates stronger short-term (monthly) return explosion power - that we are able to further transform the statistical rule into an operational investment logic. Specifically, if enterprises can be grouped based on the ESG scores of the current month and the characteristic differences in the returns of each group for the next month (such as the short-term return advantage of the low ESG group, the gradient distribution of the medium ESG group, and the stability of the high ESG group), a targeted investment portfolio can be constructed to achieve a forward-looking layout for the returns of the next month.

Thus, from "identifying the correlation patterns between ESG and short-term returns" to "designing dynamic investment portfolios based on ESG grouping", a complete logical chain from empirical discovery to practical application is formed: The former provides a solid statistical foundation for the latter, while the latter is the specific implementation of the former in investment scenarios, ultimately achieving the research goal of "predicting the stock returns of the next month through the ESG value of the current month and optimizing the portfolio allocation accordingly".

4. Research methods

We focus on the monthly dimension of "predicting the next month's returns based on the current month's ESG performance" for analysis, with the core being to explore investment patterns through the correlation between ESG scores and subsequent returns.

First, carry out data processing and basic preparations: Screen core variables such as "monthly identification, ESG score, lagging one-month return (i.e., the ESG of the current month corresponds to the return of the next month, making the prediction logic of ESG on future returns more direct), total stock assets, and stock code" from the raw data. Eliminate the observations with missing core variables. At the same time, force each stock to retain only one record per month. If there are duplicate observations for the same stock in a month, Stata will cause unstable results when calculating lagging returns and weighted averages due to the "random order of observation selection". Therefore, the deduplication command is used to eliminate this interference. Convert the percentage form of the earnings to a decimal to prevent abnormal expansion of the value during compound interest calculation.

Next, the ESG grouping construction is carried out: To ensure the stability and repeatability of the grouping logic, first sort by "year, ESG score, stock code" (within the same year, the ESG score is ranked from low to high, and when the scores are the same, they are ranked by stock code). Regenerate the globally unique serial number and construct "sorting key = ESG score × 1,000,000 + globally unique serial number" to completely resolve sorting ambiguity. Ultimately, force the samples to be divided into 5 ESG groups (Group 1 is the high ESG group, Group 5 is the low ESG group, and Groups 2-4 are the middle groups).

Finally, calculate and compare the returns: The first step is to calculate the average monthly return within the group weighted by total assets - in the real market, large-cap stocks have a more significant impact on the portfolio's returns. Therefore, for each "year - month - ESG group", sum up the "monthly return × total assets" of each stock within the group, and then divide by the sum of the total assets within the group. The second step is to calculate the cumulative monthly income - taking the initial wealth "1" as the benchmark, from the first period to the t period, multiply "$1 + \text{weighted average monthly income within the group for the current period}$" continuously and then subtract 1 to reflect the long-term accumulation of "reinvestment of income". The third step is to calculate the relative return difference - taking the cumulative return of the high ESG group as the benchmark, use "cumulative return of a certain ESG group - cumulative return of the high ESG group" to visually measure the relative performance of this group.

4.1. Portfolio strategy

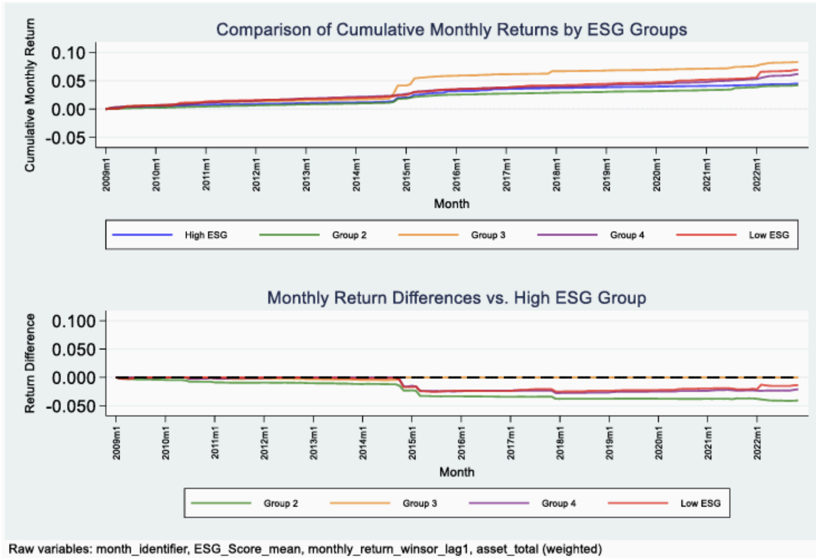

Figure 1. ESG group monthly return comparison(with difference analysis)

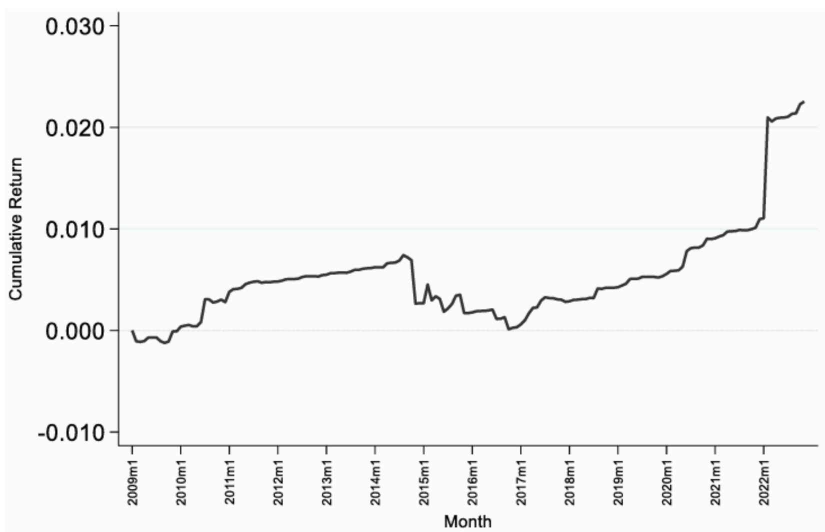

Figure 2. Cumulative return of long-short portfolio(long low ESG, short high ESG)

To transform the differentiating effect of ESG scores on stock returns into practical investment strategies, we have constructed a long-short hedging portfolio of "going long on the low ESG Group (Group 5) and simultaneously going short on the high ESG group (Group 1)" and clearly presented its logical support and actual return performance through two charts.

Explanation of the first chart, "Monthly Income Comparison of ESG Groups (Including Difference Analysis)"

This chart provides core logical support for long-short strategies from two dimensions: "cumulative return trends between groups" and "relatively high ESG return differences".

- Upper Part: Comparison of Cumulative Monthly Returns of ESG Groups

The cumulative return trends of the high ESG Group (blue line), Group 2 (green line), Group 3 (orange line), Group 4 (purple line), and low ESG group (red line) in the chart show a significant divergence: In the early stage (around 2009 to 2015), although the returns of the low ESG group increased, the gap with the high ESG group had not yet fully widened. After 2015, the returns of the low ESG group (red line) accelerated, with a significantly steeper slope, while the growth of the high ESG group (blue line) was relatively stable. This feature of "the later return growth rate of the low ESG group is much higher than that of the ESG group" is the core prerequisite for the profitability of the long-short strategy - "the low ESG group that goes long rises faster, while the high ESG group that goes short rises more slowly", and the difference in returns between the two can contribute positive returns to the portfolio.

- Lower Part: Monthly Return Differences with the High ESG Group

Among the earnings differences between Group 2, Group 3, Group 4, and the low ESG Group and the high ESG group, the difference between the low ESG group (the red line) has been positive for a long time, and the differences between lower ESG groups such as Group 4 are also mostly positive. This indicates that the "monthly returns of the low ESG group consistently higher than those of the high ESG group" is a persistent rule rather than an accidental fluctuation, directly proving that the source of returns for the long-short strategy is stable (not "bet" on a single market trend).

Explanation of the second chart, "Cumulative Returns of Long-Short Combinations (Long on Low ESG, Short on High ESG)":

This chart visually presents the earnings performance of the long-short hedging strategy, and the earnings trend highly echoes the pattern of the first chart:

- Around 2009 to 2015: A stage of moderate growth

The cumulative returns of the long-short portfolio have been rising slowly. Although there were minor pullbacks during this period (such as the fluctuations around 2015), the overall trend has remained upward. This stage reflects that even during the period when the differentiating effect of ESG on returns has not been fully strengthened, the "return advantage of low ESG over high ESG" can still contribute positive returns to the strategy, verifying the strategy's "weak market adaptability" - even if the market's pricing of ESG is not extreme enough, the strategy can still achieve basic profit accumulation.

- After 2015: Accelerated growth stage

The cumulative return slope of the long-short combination has become significantly steeper, and it has risen rapidly in the later stage and reached a relatively high level. This is in complete alignment with the rule in the first chart that "the returns of the low ESG group have accelerated after 2015", indicating that as the market environment changes (or the short-term return advantage of low ESG enterprises further amplifies), the return differentiation between low ESG and high ESG has become increasingly intense. The long-short strategy can more efficiently capture the returns brought about by this differentiation. Ultimately, a considerable cumulative profit was achieved, fully demonstrating the remarkable effectiveness and profitability of this hedging strategy during the sample period.

5. Conclusion

This article, through empirical analysis of the income characteristics of ESG groups and investment strategy design, draws the following conclusions:

The ESG group's earnings characteristics are significant: The low ESG group shows stronger earnings growth in the short term (monthly dimension). After 2015, the slope of its cumulative earnings curve is significantly steeper than that of the high ESG group, and the earnings of the middle group increase in a gradient as the ESG score decreases, indicating that the market has a earnings preference for low ESG enterprises in the short term.

There is a time bias in the market's pricing of ESG: In the short term, the market may view ESG investment as a cost burden, resulting in lower monthly returns for high-ESG enterprises. However, in the long term, the negative impacts of ESG are gradually diminishing, reflecting that its long-term value will take longer to be recognized by the market.

Dynamic strategies achieve a balance between offense and defense: A monthly tracking strategy designed based on the short-term trends of the low ESG group, through base position allocation (30%-40% for the low ESG group, 10%-20% for the high ESG group, and 40%-60% for the middle group), trend-triggered portfolio adjustment (adding positions in low ESG when continuously outperforming), and risk control (volatility stop-loss, upper limit of weight for a single group of 60%) It effectively captures the excess returns of low ESG while avoiding the risk of style switching.

The diverse values of ESG are highlighted: Although the short-term low ESG group has an advantage in returns, as a core element for the sustainable development of enterprises, the role of ESG in risk management and the construction of long-term competitiveness still cannot be ignored. Future research can further integrate longer-term data to explore the dynamic balance mechanism between the long-term value and short-term benefits of ESG.

In conclusion, this paper reveals the return patterns under ESG grouping and provides a practical reference framework for investors' portfolio management at different time dimensions.

References

[1] Albuquerque, R., Koskinen, Y., Yang, S., & Zhang, C. (2020, April). Love in the time of COVID-19: The resiliency of environmental and social stocks. *CEPR Discussion Papers*, 14661. https: //ideas.repec.org/p/cpr/ceprdp/14661.html

[2] Bebchuk, L., Cohen, A., & Ferrell, A. (2009). What matters in corporate governance? *Review of Financial Studies*, 22(2), 783–827.

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

[4] Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2), 517–549. https: //doi.org/10.1016/j.jfineco.2021.05.008
[5] Bruno, G., Esakia, M., & Goltz, F. (2022). "Honey, I shrunk the ESG alpha": Risk-adjusting ESG portfolio returns. *The Journal of Investing*, 31(3), 45–61. https: //doi.org/10.3905/joi.2021.1.215
[6] Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, 33(3), 1184–1216. https: //doi.org/10.1093/rfs/hhz072
[7] Flammer, C. (2021). Corporate green bonds. *Journal of Financial Economics*, 142(2), 499–516.
[8] Gabaix, X., & Koijen, R. S. J. (2023, December 23). In search of the origins of financial fluctuations: The inelastic markets hypothesis. SSRN Scholarly Paper (No. 3686935). Social Science Research Network. https: //doi.org/10.2139/ssrn.3686935
[9] Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices*. *The Quarterly Journal of Economics*, 118(1), 107–156. https: //doi.org/10.1162/00335530360535162
[10] Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, 93(1), 15–36. https: //doi.org/10.1016/j.jfineco.2008.09.001
[11] Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91(6), 1697–1724. https: //doi.org/10.2308/accr-51383
[12] Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, 72(4), 1785–1824. https: //doi.org/10.1111/jofi.12505
[13] Madhavan, A., Sobczyk, A., & Ang, A. (2021). Toward ESG alpha: Analyzing ESG exposures through a factor lens. *Financial Analysts Journal*, 77(1), 69–88. https: //doi.org/10.1080/0015198X.2020.1816366
[14] Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, 146(2), 403–424. https: //doi.org/10.1016/j.jfineco.2022.07.007
[15] Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2), 550–571. https: //doi.org/10.1016/j.jfineco.2020.12.011
[16] Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. Journal of Finance, 52(1), 35–55.
[17] Zerbib, O. D. (2019, January). The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *Journal of Banking & Finance*, 98, 39–60. https: //doi.org/10.1016/j.jbankfin.2018.10.012