Journal of Asset Management (2024) 25:666–699
https://doi.org/10.1057/s41260-024-00377-w

ORIGINAL ARTICLE

Strong vs. stable: the impact of ESG ratings momentum and their volatility on the cost of equity capital

Monia Magnani[1] · **Massimo Guidolin**[2] · **Ian Berk**[3]

Revised: 26 July 2024 / Accepted: 25 August 2024 / Published online: 16 October 2024

Abstract
We test whether in the cross-section of European stocks, the cost of equity capital is more strongly affected by the (upward) "slope" (identified as momentum over a period of time) of their ESG scores or by their "stability" (identified as the volatility of the scores over a period of time), measured around a given slope. We find that short-term ESG momentum is priced in the cross-section of stock returns but that it may increase or decrease the ex-ante cost of capital depending on the specific sample investigated. While short-term ESG momentum may represent a novel, priced systematic risk factor, there is also strong evidence that a ESG spread strategy that buys (sells) low (high) ESG score volatility stocks leads to a significant alpha and lower the ex-ante cost of capital. This suggests that ESG rating stability may carry a more reliable reward than improvements do, in terms of ex-ante equity cost of capital. These results are robust to the use of different sub-samples (over firms and sub-periods) and to forming the two quantitative ESG signals on the basis of alternative rating data.

Keywords ESG ratings · ESG momentum · ESG score volatility · Cross-sectional pricing · Systematic risk factor

JEL Classifications G11 · G12 · C59 · G24

Introduction

The phenomenon of Environmental, Social and Governance (ESG) compliant investing has gained substantial traction over the past decade. In August 2020, CNBC reported that the value of sustainable funds had exceeded $1 trillion for the first time in history.[1] Both institutional and retail investors are increasingly expanding their holdings of ESG-compliant securities due to either social pressure or moral considerations.[2]

Unlike the practical impact of other strands of research in asset management, ESG strategies of a quantitative nature have so far struggled to be embraced by investors.[3] Arguably, so far, investors have been relying heavily on the subjective

✉ Monia Magnani
monia.magnani@unibocconi.it

Massimo Guidolin
massimo.guidolin@unibocconi.it

Ian Berk
ian.berk@bocconialumni.it

1 GREEN and Baffi Carefin Centre, Bocconi University, Milan, Italy

2 Department of Finance, GREEN, and Baffi Carefin Centre, Bocconi University, Milan, Italy

3 Bocconi University, Milan, Italy

1 See https://www.morganstanley.com/ideas/sustainable-funds-performance-demand/.

2 A report by Morningstar published in 2022 revealed that over 2,700 funds in Europe were using ESG-related metrics in their investment decisions, see https://www.morningstar.com/articles/1132007/esg-investing-keeps-pace-with-conventional-investing-in-2022.

3 See Sorensen et al. (2021) for a discussion of quantitative methods applied to investors' ESG preferences. In spite of their advantages, they characterise our times as the "(...) early days of quantitative sustainable investing (...)". Currently, managers typically use fundamental approaches to assess stocks and build ESG portfolios. However, just as the advances in quantitative portfolio theory have gained ground over fundamental analyses of stock returns since the 1980s, so Sorensen et al. (2021) expect quantitative rankings of ESG factors to gain relevance. See also the discussion of the integration of ESG within a smart beta framework in Yasmine and Kooli (2022).

and diverse ESG ratings of companies to measure their ESG exposures, but only on a qualitative basis. To some extent, this may be due to the fact that ESG score-linked investment signals have not yet received the extensive academic coverage that other types of accounting or finance-related signals (e.g. book-to-price ratios) have. Yet, the identification of profitable ESG strategies is important for a number of reasons. First, a significantly positive alpha for portfolios sorted by ESG rating signals (e.g., the ESG scores themselves or some intuitive function thereof), may indicate that the signal represents a new source of systematic risk. Such a finding would be important to academics, to enhance the understanding of the cross-sectional variation in stock returns, and to asset managers, who would want to take this new risk into account in allocation decisions. Second, such evidence of systematic abnormal returns may provide important indications on the links between the cost of equity capital and the dynamics of the ESG phenomenon.[4]

The goal of our paper is to develop and back-test the effectiveness of two ESG score strategies of an intuitive nature. This allows us not only to provide two important examples of how ESG-related quantitative signals may play a leading role in asset pricing and allocation strategies, but also to address an important policy question: is a firm's cost of (equity) capital more strongly affected by the existence of an upward, virtuous trend in their ESG ratings or by their ability to avoid undue fluctuations in such scores? In other words, conditioning on strong ESG scores being an important driver of the cost of capital, does it also matter whether such a score improves in a more or less steady way over time, that avoids creating undesirable uncertainty on the ESG quality of the firm? Our study is focused on the European stock markets because it is the region of the world with more sustainable investment-related assets under management and where a drive towards ESG-oriented portfolios is backed by the strongest support in the legislative agenda (see Alliance for Global Sustainable Investment (2018)). We tackle the issue of the robustness of our answers to these questions by studying two alternative data rating providers, Sustainalytics and MSCI.

We contribute to the literature in several ways. First, we investigate the cross-sectional properties of ESG scores and identify a few biases commonly present in this type of data. We propose two ways to deal with these issues and analyse how the raw ESG score data treatment influences the resulting quantitative signals. To the best of our knowledge, one method based on sector and size rank-neutralisation is novel.[5] Second, we analyse the optimal ESG momentum computation period, back-testing the strategy's performance when winners and losers are defined with reference to a window of 1, 3, 6 and 12 months, respectively. Furthermore, we formally test whether quintile portfolios sorted by optimal ESG momentum, i.e., the ESG momentum computed under the optimal period, may achieve a significantly positive risk-adjusted mean return, obtaining mixed findings across alternative rating providers. Third, we investigate a new quantitative ESG signal related to the volatility of ESG scores, starting from a prior that companies with a stable ESG score may outperform out-of-sample companies with volatile ESG scores because their ratings are subject to less uncertainty. Finally, we also pursue the positive, asset pricing implications of our earlier results and test whether ESG momentum and volatility may represent new sources of systematic risk that need to be compensated by means of a positive risk premium.[6]

Our results indicate that quintile portfolios sorted by (especially) short-term ESG *momentum* earn significant alpha for a spread portfolio that buys "winners" (the virtuous) and sells "losers" (the sinners), as defined in the ESG space. Such an alpha is negative, large, and significant in the case of MSCI ratings data and smaller, positive and statistically significant in the case of Sustainalytics data, as in Berk et al. (2023). Further analysis reveals that such a sign heterogeneity is explained by the strong role played by the average level of ESG scores as a sorting variable in the former case. Andrews' (1993) test for a structural break at an unknown date fails to find evidence against the null hypothesis of a stable alpha over the entire sample. This is consistent with a conjecture that short-term (3 and 6 month and, subject to some limitations on the timeliness of the signals, 1 month)

[4] The connection between finding accurately estimated risk premia in the cross-section of past excess returns and the forecasts of the future equity cost of capital can be traced back to a Campbell and Shiller (1988)'s decomposition of current stock prices as the sum of future cash flow and discount rate news. This is related to the mechanism in Stotz (2022) by which realised and expected future returns may differ.

[5] Kaiser (2020) has investigated ESG integration in the portfolios of smart beta, factor investors. He shows that growth, value, and momentum investors in the EU and US can raise the sustainability level of their portfolio, without sacrificing financial performance. He also acknowledges the presence of size and sector biases in ESG score data and proposes a method to neutralise them by dividing the ESG score of every stock by the average ESG score of the corresponding sector. He further refines this method by scaling the sector-neutralised ESG score by the average ESG score of the corresponding size decile.

[6] Our paper does not test or take a stand on whether a (positive) ESG "level" premium may exist rewarding virtuous versus "sin" stocks as defined in terms of the level of their ESG scores. We focus instead on the changes and the variability of such scores. On the existence and sign of the ESG premium, see e.g. Adler and Kritzman (2008) Bannier et al. (2023), Bolton and Kacperczyk (2021), Bruno et al. (2022), Giese et al. (2021), Halbritter and Dorfleitner (2015), Hong and Kacperczyk (2009), and Stotz (2022). However, in subsection "ESG Momentum", with reference to MSCI ratings, we perform a few tests that reveal the hard-to-estimate nature of the ESG level premium.

ESG momentum represents a new type of systematic risk, a finding confirmed at the intermediate portfolio formation horizons by formal asset pricing tests of the Gibbons et al.'s (1989) type. However, the neutralisation method for dealing with the score biases, the period over which ESG momentum is computed, and the specific data source (i.e., Sustainalytics vs. MSCI) are crucial for the significance of the risk-adjusted mean returns. For instance, a simpler ESG momentum tilting strategy, that over-weights stocks with high ESG momentum and under-weights stocks with low ESG momentum, fails to generate significant alpha. Only our novel approach to deal with data biases leads to a precisely estimated short-term momentum premium. Moreover, also a long-short spread portfolio that buys stocks with stable ESG scores and sells stocks with unstable scores earns a statistically significant risk-adjusted return. Interestingly, this empirical finding is more robust as it holds across different volatility estimation periods and for both Sustainalytics and MSCI data and it is confirmed by standard asset pricing tests that report that under all portfolio formation windows, an ESG volatility factor cannot be rejected and that it simultaneously sets all the alphas of the test portfolios to zero.

There is a long tradition of empirical finance studies concerning momentum that can be traced back to Jegadeesh and Titman (1993) who investigated the performance of winning vs. losing stocks over a range of back-testing intervals and horizons. They reported statistically significant positive returns for zero-investment portfolios based on 3, 6, 9, and 12 month holding periods. Moreover, they found that these results cannot be traced back to other risk factors, thereby grounding the idea that momentum may represent a genuine risk factor as in Carhart (1997). Interestingly, they also show that often minor modifications of the portfolio sorting methodology may lead to reversals of momentum, which is in some sense also reflected by our application. Our paper extends this research in two ways. We apply the concept of momentum to a new setting as we define it not simply with reference to past return performance of winners vs. losers, but instead we coin a new notion of ESG momentum based on the fact that the firms issuing equity might be improving their (size- and industry-specific) ESG scores over time, relative to their peers. We then carefully investigate whether a trading strategy based on ESG momentum earns significant risk-adjusted returns.[7]

Another literature has noted that low- and high-return volatility stocks might be priced differently and that not only the classical CAPM beta exposure would matter in this perspective. For instance, Blitz and Van Vliet (2007) have studied the low volatility effect on stock returns, also with reference to European data. They present empirical evidence that, contrary to classical asset pricing, stocks with low return volatility generate higher risk-adjusted returns than stocks with high volatility. Their results are robust to the choice of geographical markets and of each of the 10 year sub-period in their sample. Moreover, this effect is not the result of exposures to other risk factors, as these spreads represent genuine average abnormal returns possibly compensating a novel risk factor. In this paper, we follow the lead of this literature and investigate the existence of an ESG Low Volatility factor. We posit that the uncertainty about the ESG quality of a stock-issuing firm, defined as the volatility of its ESG score over a certain period, may provide an additional source of systematic risk.[8] We test this conjecture and assess whether investors should receive a premium for bearing this novel risk.[9] While Gibson et al. (2021) have examined the positive relationship between stock returns and ESG rating disagreement, in this paper we focus on a more classical notion of time series variance.

There is also a new strand of literature that studies systematic, quantitative strategies that sort stocks on the basis of their ESG performance. For instance, Nagy et al. (2016) have introduced ESG momentum to the literature. They use ESG scores from MSCI to create tilting strategies and compare their performance to the MSCI World index benchmark. Their signals are ESG score quality (which leads to over-weighting companies with a relatively high ESG score) and ESG score momentum (which over-weights companies with the highest increase in ESG score over a given interval of time). They find that both separate tilting strategies outperform the MSCI World index in terms of mean, realised returns. However, their paper focuses on raw returns, rather than on risk-adjusted returns. Yet, adjusting for a portfolio's risk exposures could wipe out all the realised, average excess returns. We formally test the performance of an ESG momentum strategy to assess whether the risk-adjusted average returns earned by this signal are statistically significant. Additionally, we investigate another ESG signal, ESG score volatility. Bruno et al. (2022) have investigated the recursive, risk- and sector-adjusted performance of several systematic ESG strategies, including 12 month momentum, with

7 To ensure that ESG momentum does not represent a mere re-packaging and re-branding of the classical momentum factor (or at least that it does not strongly correlate with it), our tests control for the exposure of ESG momentum excess returns to Carhart's classical momentum factor.

8 Cantor and Mann (2007) tested the hypothesis that market participants have a strong preference for credit ratings that are not only accurate but also stable as they would like ratings to reflect enduring changes in credit risk. Of course, our analysis concerns ESG and not credit ratings.

9 By formally controlling for the standard, financial volatility premium, we make sure that in no way the already well-investigated low volatility factor may explain away the novel ESG volatility factor.

reference to stocks in the USA and in developed markets. While they fail to find significant alpha, their results are not inconsistent with ours as we report that it is a shorter window ESG momentum that, at least in the case of European equities, turns out to be profitable on a systematic basis.[10]

Using an exposure-matched approach that formally neutralises the ESG level effect applied to US data, Dor et al. (2022) have shown that not only the ESG level but also the ESG rating history may impact stock performance. Companies with improving ESG credentials may outperform those with deteriorating scores, even if their current ESG level is the same. In their empirical results, the E-pillar score turns out to contain the strongest momentum effects and ESG momentum is the highest for stocks around the medium ESG level. We extend this analysis to European data and to the uncertainty characterising the time series of the ESG scores. Even though we carefully neutralise industry and size effects in the ESG scores, we refrain to completely neutralise the level of the ESG scores to be able to ask what are the sustainability policy implications of the ESG-driven strategies under examination.

The paper is structured as follows. Section "Data" describes the data and how to deal with any biases in ESG scores. Section "Main results" explains the ESG momentum and volatility signals and analyses the alpha earned by quintile portfolios sorted on them. Formal asset pricing tests of the new, ESG-driven factors are performed here. Section "Robustness checks" performs robustness checks concerning the definition of the two ESG signals, their stability over time, the span of the universe of firms investigated, and the true out of sample analysis concerning the period 2020–2023. Because ESG signals may become available to investors with a lag compared to what our data set show because of back-filling, we also implement a 1 month ESG momentum strategies after applying a 1 month gap. Section "Discussion and conclusions" concludes.

Data

The data set comprises a variety of stock market and firm-specific characteristics, such as returns, market capitalisation and the Sustainalytics (now Morningstar) and MSCI KLD ESG scores. The reason for this choice of raters resides in Wong et al. (2019), who have recently reported that the scores provided by Sustainalytics and MSCI KLD are among the most highly regarded scores by investors. Apart from their comparable global and impact-related relevance, one reason to supplement the original analysis of Sustainalytics ratings (as in Berg et al. (2022)) with MSCI scores is that Sustainalytics has changed their methodologies over time (also as a result of the company changing ownership structure), and their ESG scores are suspected to have been back-filled using these new methodologies (see the discussion in Berg et al. (2020). MSCI is instead deemed to be free of forward-looking biases. Moreover, while Sustainalytics has been reported to market its ratings in coarsely aggregated, categorical (0–5) scores then aggregated and re-expressed in percentage terms, MSCI determines its rating as a weighted average of key issues scores, all ranging from 0 to 10, considering both positive and negative performance indicators.[11]

Apart from the ESG scores, all company-specific data are extracted from FactSet. In particular, we start by covering a sample of companies that are common to both Sustainalytics and MSCI ratings, but in later robustness checks we extend our analysis of the MSCI-rated stocks to that specific universe (the MSCI Europe Investable Market Index) which turns out to be considerably larger than Sustainalytics's.

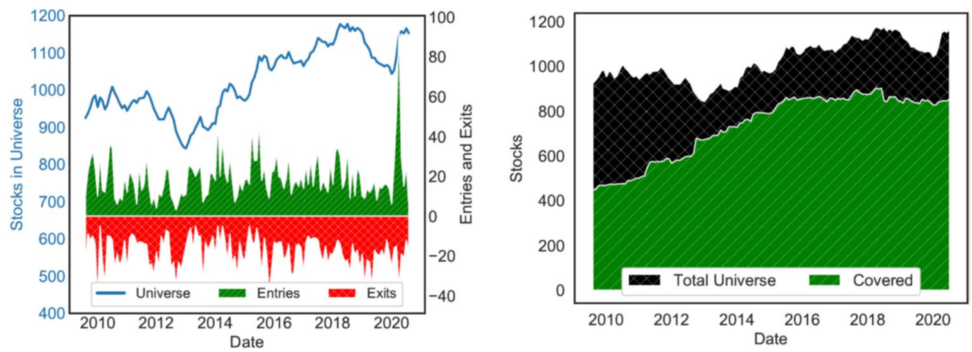
Our investment universe covers a number of stocks issued in the countries of the European Area, hence including the UK, with the exception of Iceland and Greece. The countries are Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the UK. Their selection was driven by a set of minimum requirements that ensure an investor would be able to trade in well-functioning stock markets, in terms of investor-protection laws and liquidity. These filters are imposed on the Sustainalytics' universe of firms.[12] As an additional quality filter, we exclude all stocks that are not covered by any equity analyst over our sample period. We, also, require a minimum market capitalisation of 400 million euros to exclude listed companies close to bankruptcy or small capitalisation stocks just admitted to listing. These stocks have been reported to be able to adversely affect the performance of many multi-factor asset pricing models and may bias our empirical findings

[10] Martellini and Vallée (2021) have studied the performance of a ESG momentum strategy applied to sovereign bonds issued by developed and emerging countries.

[11] For each company, around 4 to 7 most material ESG key issues of the primary industry it belongs to are considered when computing the weighted average, together with any relevant additional ESG topic which may significantly affect the company. The resulting scores are then re-scaled by us to be expressed out of 100. Although also the MSCI's scores range between 0 and 10, they are more granular and such a 11-point scale range may be taken as a closer approximation to a continuous rating. We thank an anonymous referee for drawing our attention on these issues (see https://www.sustainalytics.com/esg-data).

[12] The exclusion of all Greek listed firms derives from such filters, especially the liquidity and analyst coverage ones. However, later robustness checks based on MSCI ratings do cover also 46 Greek and 3 Icelandic stocks over time.

Fig. 1 shows the evolution of the investment universe and the coverage of ESG scores in it. Panel (**a**) refers to the number of stocks, as well as the entries and exits for the whole sample period. Panel (**b**) refers to the number of stocks reporting both Sustainalytics and MSCI ESG scores compared to the total number of stocks under analysis. The sample period is July 2009–December 2019



concerning alphas and their statistical significance in Sections "Main results" and "Robustness checks", see, e.g. Israel and Moskowitz (2013).

We narrow our sample to stocks with a share price below 10,000 euros. We also introduce a lower bound to exclude "penny" stocks because these tend to be less informative due to the extreme returns they may yield by construction. The upper bound is set to avoid situations in which the optimal portfolio would be otherwise infeasible: if stocks carry too high a price, frequent portfolio rebalancing cannot be implemented because of the induced discreteness of the portfolio adjustments. Finally, we impose a minimum liquidity constraint of 1 million euros in terms of trading volume per day to limit the market impact of any of the long-short strategies that we shall be proposing and back-testing. Absent such a minimum turnover filter, in the case in which the optimal portfolio strategy may require to buy a stock multiple times its maximum daily volume, this would imply a severe market impact with the risk of destroying the signal's profitability making our back-testing findings virtual at best.

Our sample spans the period July 2009–December 2019. The beginning of the sample corresponds to the starting of Sustainalytics's activities at European rating stocks on a large scale, as covered by FactSet. As often remarked in the literature, ESG scores represent relatively new data and are not available before mid-2009. The end of our sample is carefully chosen to not overlap with the period (January 2020–March 2023) dominated by the Covid-19 pandemic to avoid any biases due to the presence of extreme outliers in the data.[13]

Figure 1 shows the evolution over time of the number of stocks in our sample. Panel (a) exhibits the total number of stocks, the new entries, and the exits for every month in our sample. To prevent any survivorship bias, we account for all entries and exits at any point.[14] Clearly, this procedure causes the resulting panel data set to be unbalanced. Our back-testing methods rely on repeated cross-sectional spread tests over time. This approach prevents variations in cross-sectional size from influencing our results. Panel (b) of Fig. 1 shows the evolution of the number of stocks with an ESG score assigned by both raters, as a fraction of the total number of stocks under analysis.

In the literature, Kaiser (2020) notes that the ESG scores strongly depend on the firms' size and differ substantially across different industries (see also Bruno et al. (2022)). These findings align with prior research that consistently links a company's size to its sustainability policies. For instance, Artiach et al. (2010) established a strong positive correlation between firm size and Corporate Sustainability Performance (CSP). Bos (2017) and Kaiser (2020) show that this relationship translates into higher ESG scores for larger companies. As far as our data are concerned, any sectoral dependence may be embedded in the rating that the raters assign to each company, as they build their scores by assessing a company according to a variety of different subcategories. Some subcategories only apply to certain sectors (e.g. fleet carbon emission is irrelevant for an IT company, but highly relevant for a transport company). Furthermore, within the same category, varying weights may be assigned to different sectors. To mitigate unintended biases, we address these sectoral dependencies using a novel (to our knowledge) methodology, which we describe in the following.

Figure 2 shows the evolution of the mean ESG scores for the companies included in our sample. Panels (a) (for Sustainalytics) and (c) (for MSCI) show the evolution of

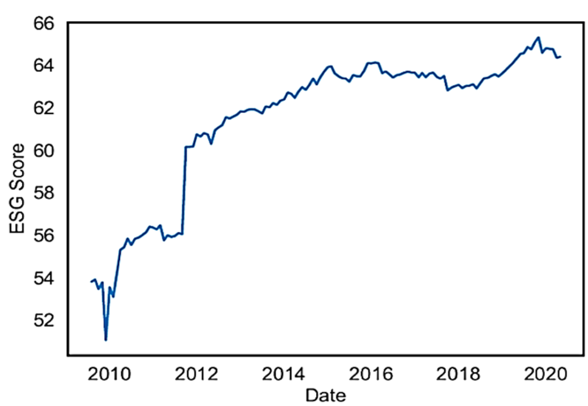
[13] Nonetheless, in subsection "Covid and post-Covid MSCI data", we perform robustness checks based on MSCI ratings that also cover the 2000–2023 pandemics and its aftermath.

[14] On average, our sample include 712 companies per month, ranging from 445 to a maximum of 812. Because all sorting methods are applied at the univariate level and the resulting portfolio include an equal number of stocks, this delivers an average of 142 stocks in each portfolio, with any reminders allocated to the extreme portfolios 1 and 5. Such an average of 142 stocks oscillates from approximately 90 stocks per quintile portfolio early on in our sample to a maximum of approximately 162.

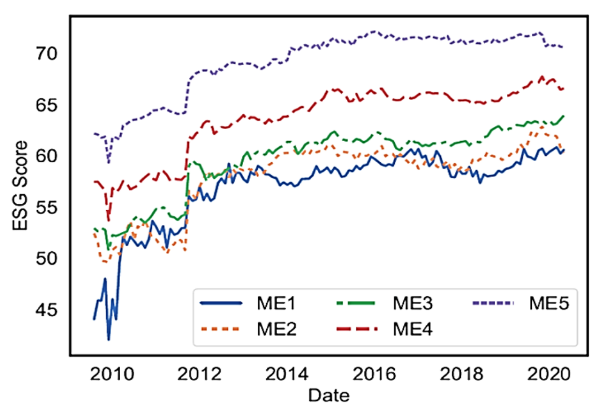
the average of all the stocks that have been assigned an ESG score. Especially panel (a) shows a clear upward trend over the sample period. Such a trend sets off in 2013 in the case of panel (c). This may reflect the increasing urge for companies to have their ESG-directed efforts publicly acknowledged. A reduction in 2009–2010 (2010–2011) can be noted in both panels (a) and (c) but this may be explained by the effects of the European Sovereign Debt crisis, possibly a sign that when firms are pressed to scrape valuable resources in critical situations, they may relent on their sustainability development efforts (see, e.g., Idowu et al. 2017). In spite of some differences, it is also notable the existence of an upward trend in average scores that occurs in 2012–2014. Panels (b) and (d) in Fig. 2 display the evolution of the average ESG score across five equally-weighted quintile portfolios sorted by size. Therefore, ME1 is the quintile portfolio including companies with the lowest market capitalisation while ME5 is the quintile containing the largest stocks. The plots (in particular panel (b)) show that the average ESG score tends to increase with the size. ME5 essentially carries the highest ESG score for every observation in the sample period, while ME1 generally has the lowest. Yet, this ordering does not consistently apply to the average ESG scores of the bottom three quintiles over time, despite the general trend of scores increasing with market capitalisation.

Table 1 shows the relationship between size and the ESG scores on the basis of the same quintile portfolios used in panels (b) and (d) of Fig. 2. As market capitalisation increases, the average and median ESG score (together with the minimum and the maximum score) increase monotonically. This suggests that a strong size tilt is present in the data (see Artiach et al. (2010) for a discussion of the potential motives driving this empirical regularity). Additionally, the data reveal a negative correlation between average cross-section. In our analysis, we consider the size effect to ensure it does not interfere with the measurement of ESG signal performance.

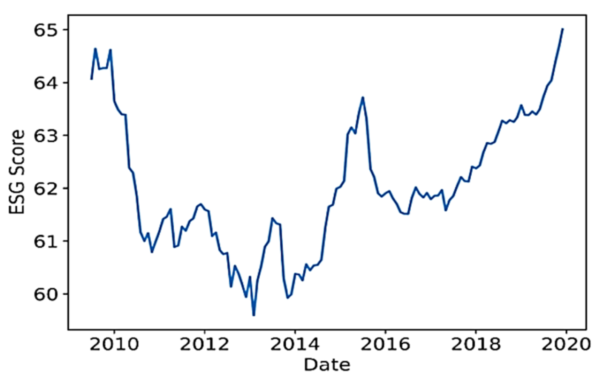
Table 2 shows a high dispersion of mean ESG scores across industries. For instance, this is clear if one compares the industrials and materials sectors. These two industries



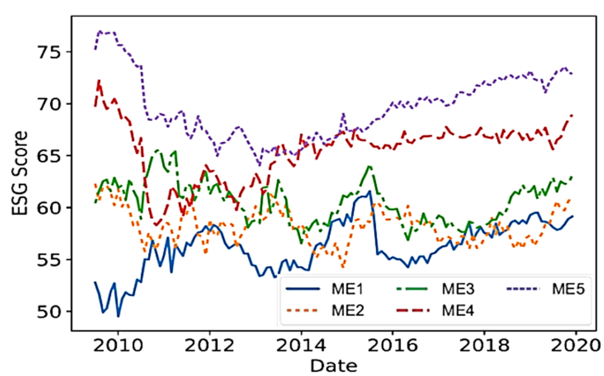
(a) Evolution of the mean ESG score of the investment universe.



(b) Evolution of the mean ESG score per size bucket.



(c) Evolution of the mean ESG score of the investment universe.



(d) Evolution of the mean ESG score per size bucket.

Fig. 2 shows the evolution of the mean ESG scores over time. Panels (**a**) and (**c**) refer to the ESG score across the entire investment universe for Sustainalytics and MSCI, respectively. Panels (**b**) and (**d**) refer to ESG scores for each of the five size quintiles used in the paper, for Sustainalytics and MSCI, respectively. The common sample period is July 2009–December 2019

Table 1 Summary Statistics and ESG scores for five size-sorted quintiles

	Stock returns				Sustainalytics ESG scores					MSCI ESG scores				
	Mean (%)	σ (%)	Sharpe Ratio	Mark. Cap.	Mean	Median	σ	Max	Min	Mean	Median	σ	Max	Min
1	0.633	6.454	0.072	670	56.91	58.22	3.86	60.84	42.00	56.20	56.32	2.41	61.59	49.52
2	0.809	5.536	0.116	1361	58.06	59.46	3.49	62.85	49.56	58.23	58.28	1.68	62.30	54.16
3	0.795	5.260	0.119	2646	59.65	61.08	3.39	63.93	50.70	60.59	60.86	2.04	65.60	56.56
4	0.777	4.854	0.125	5591	63.57	65.25	3.40	67.76	53.66	65.41	66.31	2.78	72.23	58.30
5	0.651	4.803	0.101	29,112	69.11	70.71	3.21	72.13	59.31	69.62	69.44	3.14	77.06	64.05

The table shows the portfolio summary statistics of returns and mean ESG scores by size quintiles. In the left panel, we report the mean return, the volatility of the returns, the Sharpe Ratio (SR) and the market capitalisation in millions of euros. In the two rightmost panels (for Sustainalytics and MSCI, respectively), we show the average and the median ESG score, the volatility of the score, and the maximum and the minimum value of the score for each of the size-sorted quintiles

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

have roughly the same average market capitalisation (5081 and 5103 million euros, respectively). However, industrials carry an average ESG score that is approximately two standard deviations higher than the materials sector's ESG score (using the standard deviation of the former) for both Sustainalytics and MSCI. The same pattern applies also to the median and the minimum ESG scores while the Sharpe ratio (SR) of the industrials is only marginally higher than materials producers. Less disparity can be observed for other pairs of industries that display similar average market capitalisation (e.g. the healthcare and industrials sectors). The rightmost section of Table 2 concerning the MSCI scores proposes other examples of industries with similar average market capitalization and very different mean and volatility of the ESG scores. To avoid any undesirable, implicit industry tilts in our ESG-driven strategies, our methodology takes sector-dependency into account.

We neutralise the ESG scores from the influence of both size and industry biases, so as to remove any unintentional tilts. To this end, we follow two approaches. A first method divides the stocks into five size quintiles, where the size of a company is defined by its market capitalisation. We calculate the median ESG score for each quintile and then subtract this median from the score of every stock within that quintile. This approach effectively neutralises the monotonically increasing relationship between size and ESG score since we focus solely on deviations from the medians. Next, we compute the score quintiles to which each stock belongs by assigning the stock a value between 0 and 1: zero is attributed to the stocks with the largest negative deviation from the median ESG score within their size quintile; one to the stocks with the highest positive deviation from the median score within a quintile. For stocks falling between the minimum and maximum, we assign a score that is proportional to their signed distance from the median, with a value of 0.5 given to the median stocks (or to the two stocks straddling the median if this does not correspond to an actual score). To mitigate sectoral biases, a similar process is applied to all sectors. However, this time, we begin with the size-corrected ESG scores percentiles derived from the earlier steps, rather than the raw scores.

We also entertain an alternative neutralisation method, which consists of simply standardising the scores across portfolio sorts. To begin, stocks are categorised into size quintiles. Within each quintile, we calculate the mean and standard deviation to standardise the ESG scores. Next, the same procedure is applied to a fifteen-industry classification, with the initial data being the size-standardised ESG scores.[15]

Throughout the paper, when referring to the un-adjusted, non-neutralised ESG scores we call them "*raw scores*". When we resort to the ESG scores adjusted using the first procedure, we use the expression "*rank-neutralised ESG scores*". Finally, when we refer to the ESG scores adjusted using the second procedure, we refer to them as "*standardised ESG scores*".

How often do ESG ratings change and why?

It is well known that Sustainalytics and MSCI assign ESG ratings and apply revisions to them according to a rather different

[15] We recognise that both neutralisation methods tackle the biases in a sequential manner, by first neutralising the size tilt and then the industry bias. Although they are common (see, e.g. Kaiser (2020)), sequential approaches like ours may lead to the ESG scores obtained after the second industry-neutralisation step no longer being perfectly size-neutral. This may partially reverse the effect of the first neutralisation step. While employing simultaneous double sorting could control for both size and industry biases at the same time, it might result in portfolios with a limited number of stocks or even completely empty portfolios. For this reason, we stick to a sequential, conditional sorting. We take care of the potential reversal of the original size-neutralisation by computing risk-adjusted performances from multifactor models that contain a factor that exposes to size risk.

Table 2 Summary statistics of returns and ESG scores for fifteen industries

	Returns				Sustainalytics ESG scores					MSCI ESG Scores				
	Mean (%)	σ (%)	Sharpe Ratio	Mark. Cap.	Mean	Median	σ	Max	Min	Mean	Median	σ	Max	Min
Automobiles	0.504	7.238	0.046	15,513	65.750	67.220	3.760	70.150	54.270	56.899	57.460	4.346	63.273	47.857
Banks	0.835	6.072	0.110	6,990	60.140	61.020	3.420	65.290	52.110	57.806	57.871	2.407	63.243	50.243
Consumer Discr.	0.716	4.908	0.112	7,228	62.470	63.170	2.910	66.050	54.350	64.335	64.823	2.352	67.747	58.284
Consumer Staples	0.807	4.402	0.145	9,081	60.860	62.580	3.650	65.220	50.200	64.868	66.866	6.521	73.073	50.495
Financials	0.642	5.989	0.079	9,485	62.600	63.500	2.790	66.530	56.040	54.423	54.288	3.049	60.261	48.121
Energy	0.869	4.722	0.148	12,941	61.170	61.330	1.890	64.640	55.660	65.544	65.400	2.346	69.571	61.557
Healthcare	0.837	5.137	0.130	4,911	61.920	62.810	2.990	65.140	52.350	60.427	59.977	2.504	65.493	54.511
Industrials	0.805	5.861	0.109	5,081	63.870	64.650	2.180	66.680	56.880	63.235	63.649	2.262	67.160	57.360
Insurance	0.928	5.639	0.135	4,269	62.820	63.530	2.190	66.430	55.300	56.210	56.359	2.500	61.497	48.900
IT	0.374	6.431	0.032	7,838	62.890	64.150	3.840	68.350	54.050	59.180	56.206	6.103	73.121	52.920
Materials	0.787	8.021	0.077	5,103	58.720	58.870	1.900	62.630	53.390	60.226	60.022	3.275	67.451	52.715
Media Entert.	0.814	6.351	0.102	6,155	64.750	64.880	2.680	70.300	58.450	67.874	67.154	3.146	74.962	61.409
Real Estate	0.566	6.098	0.065	5,527	63.780	64.690	2.420	67.550	55.970	60.256	59.950	3.224	71.517	52.588
Telecom	0.627	5.629	0.082	5,773	60.390	60.420	3.340	65.020	53.770	63.837	62.444	6.581	76.977	25.997
Utilities	0.860	4.501	0.154	12,204	68.190	68.950	2.830	72.550	59.000	71.590	71.791	2.858	76.526	66.912

The table reports summary statistics of returns and ESG scores in each industry. In the left panel, we show the mean return, the volatility of the returns, the Sharpe Ratio and the market capitalisation in millions of euros. In the right panel, for both providers of ESG scores, we display the average, the median, the volatility, the maximum and the minimum value of the scores. The firms are classified in 15 industries on the basis of the Global Industry Classification Standard (GICS) codes

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

logic. Sustainalytics provides risk ratings that measure a company's exposure to industry-specific, material ESG risks and how well the company is managing those risks. They collect data from a variety of sources, including company disclosures, government databases, and third-party research. They emphasise transparency and aim to provide clear insights into how ratings are derived. Sustainalytics updates its ESG Risk Ratings at least once a year, but they are also subject to review whenever significant new information becomes available. This includes major company events, news, and regulatory changes. MSCI ESG Ratings evaluate a company's exposure to ESG risks and opportunities relative to its industry peers. They use a mix of public data sources, proprietary models, and direct company engagement. To gather information updates its ESG ratings annually but also conducts quarterly reviews to capture any significant changes or events that may impact a company's rating. Such events might include new regulatory compliance issues, major corporate actions, or significant incidents related to environmental, social, or governance factors. MSCI also has mechanisms for more immediate updates when critical incidents occur. At least in principle, this frequent review mechanism ensures that MSCI ratings are relatively up-to-date and reflective of current conditions. Moreover, Sustainalytics is noted for its emphasis on transparency in how its ratings are derived, which can provide clearer insights into the risk levels assigned to companies; yet, such careful approach may be reflected in a lower responsiveness of their ratings to news. Therefore, we expect the variation over time of MSCI ratings to exceed that of Sustainalytics, given the greater attention placed by the former on fast updating.

Figure 3 shows instead the variation over our sample of the cross-sectional mean proportion of the ESG rating updates or initiation as a fraction of the companies included in our data (at each point in time, t) set over a pre-defined time window, identified with 1, 3 and 12 months before t. Because the two plots share the same scales (the right scale is used for the 12 month rating activity index), a comparison of panels (a) and (b) immediately reveals a much stronger rate of activity by MSCI vs. Sustainalytics. For instance, over the 12 month horizon, the mean cross-sectional ESG rating activity indicator is 45.5% (i.e., on average between 2009 and 2019, on every possible 12 month window, almost 46% of the companies in the sample receive a new ESG rating or their rating is changed in some of its components and dimensions) in the case of MSCI vs. 28.6% in the case of Sustainalytics.[16] However, all curves describing the intensity of the rating activity steeply increase during our sample, even though such a trend is more visible in the case of MSCI. For instance, focusing now on 3 month cumulative activity, Sustainalytics goes from a rate of 3–4% during 2009–2010, to an average in excess of 10% of all of our companies in the final part of the sample; MSCI goes from a rate of approximately 5% in 2009–2010 to always exceed 10% (except May and June of 2019) after 2016. Of course, as expected, also the variability of the rating activity differs across panels (a) and (b). Such a spread is maximum in the case of the 3 month cumulative activity (the sample standard deviation of the rating activity indicator is 4.0% in the case of Sustainalytics vs. 5.7% in the case of MSCI), it persists at a 12 month horizon (15.3% in the case of Sustainalytics vs. 18.5% for MSCI). Finally, the most crucial question is whether there could be enough variability in ratings to support ESG score-driven portfolio strategies over short horizons, and in particular 1 month. Focusing now on the rating activity indicator over a simple 1 month window, we find that on average Sustainalytics initiates ratings or revises old ones for 3.26% of the companies in our sample, while MSCI does so for 3.80% of the same. If one applies these rates to our average number of 712 companies, this means that on average there are on every month between 24 (Sustainalytics) and 27 (MSCI) companies whose rating is revised and affected. Of course, by the end of the sample, when the number of companies exceed 800 and the rates of activity average between 4 and 5% for both raters, we would be investigating between 35 and 45 companies on which there is rating activity on every month. When the horizon is 3 months, the number are of course much larger, with on average between 52 (Sustainalytics) and 75 companies (MSCI) affected by rating changes on every 3 month period, and these figures are reaching triple digits for both raters after 2016. Although they are never extreme, these rates of activity seem sufficient to at least create very short-term long-short portfolio that systematically buy stocks whose ESG rating has improved and sell those whose rating has declined in overall terms.[17]

[16] The numerator of the statistics excludes cases in which the ESG ratings are simply re-affirmed to be identical to the previous occasion in which they were published. In this sense, our indicator may easily under-estimate the intensity of the rating activity and this is especially the case of MSCI ratings. However, it is possible that the indicator may reflect the fact that the ratings of the same company are revised multiple times over the intervals of 3 and especially 12 months. Nonetheless the last type of activity would be relevant to both our portfolios.

[17] One residual check concerns the occurrence of adequate downgrading activity to make a portfolio strategy that shorts negative ESG momentum implementable. It turns out that on average 47.5% of the rating activity undertaken by Sustainalytics consists of (aggregate) downgrades; the corresponding statistic is 43.3% in the case of MSCI.

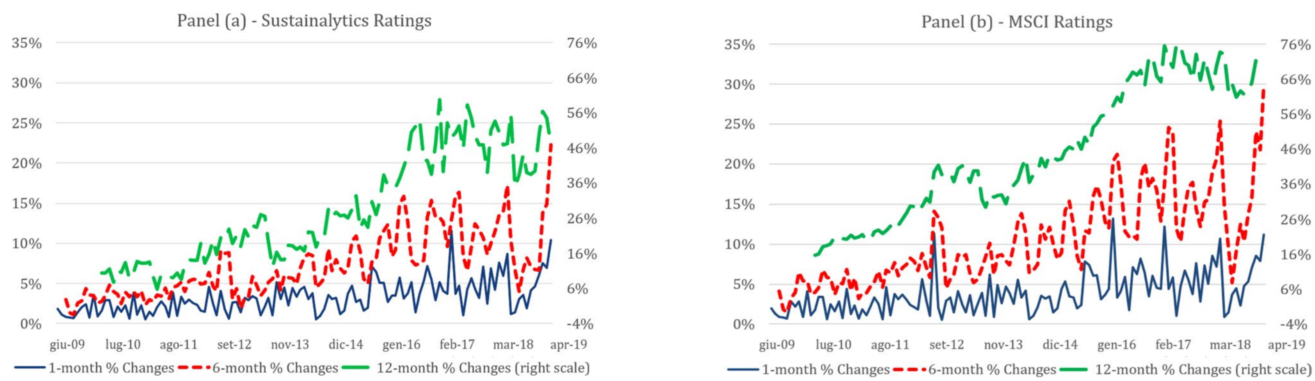


Fig. 3 shows the evolution of the cross-sectional mean proportion of the ESG rating updates or initiation as a fraction of the companies included in our data set over a pre-set period of time, identified with 1, 3 and 12 months. Panel (**a**) refers to Sustainalytics; panel (**b**) to MSCI data. The common sample period is July 2009–December 2019

Main results

In this section, we first sort our universe of stocks into portfolios; for each sort, we use different ESG momentum or volatility computation periods.[18] We then estimate the Fama and MacBeth (1973) cross-sectional regressions to provide preliminary evidence on the effects that ESG momentum and volatility each have on stock returns. Formal asset pricing tests are subsequently implemented to investigate whether the two ESG characteristics generate statistically significant abnormal returns and whether they provide explanatory power for the cross-section.

ESG Momentum

Portfolio sorting

ESG momentum is defined as the percentage increase of a security's ESG score over a certain period of time. As discussed in the Introduction, Nagy et al. (2016) has used ESG momentum in devising a strategy to adjust the weights of stocks with high and low ESG momentum relative to standard market capitalization weights. Yet, to evaluate whether changes in the ESG score can genuinely support an alpha-generating strategy, it is essential to conduct ESG momentum portfolio sorting according to the conventions of the asset pricing literature. We start by computing the x-month ESG momentum, where $x \in \{1, 3, 6, 12\}$ months. The motivation for this range of momentum-measurement periods reflects the well-documented finding that the results of empirical tests often vary depending on the range of momentum computation periods (see Jegadeesh and Titman (1993)). Crucially, the cases of $x \geq 3$ are important because ESG rating changes are normally communicated with a delay. In fact, entertaining a range of horizons over which momentum is computed is supposed to neutralise the fact that there may be long and variable lags in the reporting of ratings and of their changes: Readers worried of the timeliness with which rating changes are made available to market participants may want to pay attention to the case of $x = 3$ months or higher. Because the selection of the computation period may significantly influence the final results, we investigate this selection in Table 3, in which the two adjacent panels deal with Sustainalytics and MSCI ratings. We first sort stocks by their x-month ESG momentum in a descending manner. Next, we create equally-weighted quintile portfolios based on the descending ESG momentum scores, such that portfolio 1 contains stocks with the highest ESG momentum and portfolio 5 stocks with the lowest ESG momentum. Additionally, we create a long-short portfolio that only has exposure to ESG momentum and minimises its sensitivity to other risk factors (when different portfolios carry exposures with similar signs), especially, to the market portfolio. This zero-cost portfolio is conventionally created taking contemporaneous long and short positions in the extreme quintile portfolios of stocks (i.e., 5–1).

Table 3 shows summary statistics for both the performance metrics and the raw ESG scores of the quintile portfolios sorted by four different ESG momentum definitions. In this table, we use the rank-neutralised ESG scores to produce the sort. The table also shows the summary statistics for the zero-investment portfolio 5–1. We report the mean return, the standard deviation of the returns, the Sharpe ratio, and the standard error of the SR (Std. Err(SR)) as derived by Lo

[18] We resort to simple, univariate sorting instead of bivariate sorting that would condition on the ESG score level. However (see, for instance, Table 4 to follow) we also investigate what the effects of reversing the order of ESG score level and momentum sorting would be.

Table 3 Results for portfolio sorted by ESG momentum using four computation periods

	Sustainalytics								MSCI							
	Returns				Sustainalytics raw ESG scores				Returns				MSCI raw ESG scores			
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Momentum	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Momentum (%)
Panel A: 1-month ESG Momentum																
1	**0.851**	4.377	**0.195**	0.088	57.920	58.260	0.927	46.1%	−0.208	4.037	−0.047	0.089	59.050	58.750	2.290	77.0
2	0.695	4.397	0.158	0.087	65.440	66.160	1.002	4.1%	**0.932**	4.171	**0.228**	0.086	60.659	61.551	3.162	0.0
3	0.667	4.662	0.143	0.087	69.150	69.820	0.977	−0.1%	**0.750**	3.961	**0.194**	0.082	63.192	62.813	2.407	0.0
4	0.699	4.630	0.151	0.092	64.990	65.670	0.911	−4.1%	**0.796**	3.930	**0.207**	0.084	64.896	64.987	2.141	0.0
5	0.426	4.596	0.093	0.094	57.340	57.730	0.946	−21.0%	**2.270**	3.821	**0.599**	0.089	62.607	62.711	3.045	−6.9
5 −1	**-0.425**	1.543	**−0.248**	0.113					**2.479**	1.414	**1.782**	0.157				
Panel B: 3-month ESG Momentum																
1	0.801	4.508	**0.178**	0.088	58.900	59.720	0.933	83.5%	−0.164	4.033	−0.036	0.088	58.075	57.917	2.617	120.6
2	0.660	4.438	0.149	0.091	65.560	66.570	0.946	6.9%	**0.777**	3.973	**0.200**	0.082	60.868	61.164	2.914	0.0
3	0.551	4.685	0.118	0.091	69.330	70.250	0.911	−0.3%	**0.587**	3.977	0.152	0.081	63.667	63.199	2.439	0.0
4	0.764	4.654	0.164	0.086	64.850	65.150	0.791	−7.0%	**0.847**	3.907	**0.221**	0.080	64.655	64.911	2.555	−0.1
5	0.565	4.425	0.128	0.093	57.380	57.880	0.829	−29.8%	**1.797**	3.533	**0.514**	0.081	63.747	63.938	2.478	−17.7
5 −1	-0.236	1.480	−0.130	0.095					**1.961**	1.436	**1.394**	0.126				
Panel C: 6-month ESG Momentum																
1	0.677	4.721	0.144	0.094	60.060	60.970	0.857	123.8%	0.245	3.989	0.066	0.085	54.082	53.845	3.751	181.8
2	0.663	4.568	0.146	0.093	65.690	66.330	0.851	9.8%	**0.628**	4.021	0.161	0.085	63.776	63.488	3.028	2.2
3	0.543	4.525	0.120	0.092	69.770	70.490	0.813	−0.5%	**0.631**	4.096	0.158	0.083	64.251	64.233	2.290	0.0
4	0.654	4.547	0.144	0.089	64.690	64.770	0.724	−9.5%	**1.083**	3.813	**0.289**	0.082	67.275	67.530	2.961	−1.3
5	0.649	4.395	0.148	0.096	57.500	58.040	0.749	−36.3%	**1.258**	3.589	**0.355**	0.084	62.633	62.184	3.108	−31.5
5 −1	-0.028	1.427	0.012	0.092					**1.013**	1.319	**0.800**	0.114				
Panel D: 12-month ESG Momentum																
1	0.559	4.781	0.117	0.095	61.420	61.980	0.727	187.5%	0.463	4.065	0.118	0.088	46.460	46.912	4.384	286.4
2	0.494	4.586	0.108	0.097	66.310	66.860	0.702	15.1%	**0.564**	3.780	0.153	0.088	67.174	67.092	3.132	21.8
3	0.587	4.460	0.132	0.091	70.260	70.740	0.768	−0.3%	**0.669**	3.888	**0.176**	0.087	71.826	72.363	3.110	1.2
4	0.731	4.445	0.165	0.093	65.010	65.140	0.645	−12.6%	**0.967**	3.630	**0.271**	0.088	70.188	70.029	2.608	−12.6
5	0.632	4.324	0.147	0.096	57.710	58.070	0.670	−43.6%	**0.971**	3.697	**0.267**	0.087	58.790	58.561	3.303	−47.4
5 −1	0.073	1.420	0.084	0.092					**0.508**	1.210	**0.456**	0.091				

The five quintiles are sorted by ESG momentum, the percentage increase of the ESG score over a period. The table also provides the summary statistics for a portfolio 5–1, which takes a long position in portfolio 5 and a short position in 1. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

Table 4 Results for portfolios sorted by 1-month MSCI ESG score levels

	Returns				Raw ESG scores				Rank neutralised scores		
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	1-month Mom.	Mean	Median	σ
1-Month level strategy portfolio returns											
1	0.645	4.724	0.124	0.090	90.65	89.41	2.728	5.4%	0.914	0.912	0.022
2	0.407	4.954	0.070	0.092	74.66	73.87	2.729	2.7%	0.703	0.701	0.028
3	0.562	4.754	0.105	0.090	62.76	62.73	1.842	5.7%	0.495	0.494	0.031
4	0.350	4.838	0.060	0.090	50.98	51.24	1.861	7.7%	0.292	0.298	0.032
5	0.317	4.955	0.052	0.094	31.96	31.26	2.992	4.0%	0.096	0.095	0.023
1–5	**0.328**	1.702	0.157	0.099							

	α	Mkt–Rf	SMB	HML	RMW	CMA	Momentum	Low Vol.	R-square
1-Month level strategy 7-factor regression estimates									
1	**0.386**	**0.425**	**−0.398**	**−0.545**	−0.211	−0.172	0.120	**−0.453**	**0.788**
	(2.117)	(8.022)	−(2.987)	−(3.538)	−(1.342)	−(0.683)	(1.860)	−(8.152)	(0.000)
2	0.266	**0.502**	**−0.404**	**−0.498**	−0.091	−0.217	0.002	**−0.398**	**0.833**
	(1.614)	(11.349)	−(3.199)	−(3.336)	−(0.588)	−(1.174)	(0.030)	−(6.587)	(0.000)
3	**0.448**	**0.484**	**−0.388**	**−0.449**	−0.111	−0.194	0.005	**−0.409**	**0.825**
	(2.576)	(9.279)	−(2.902)	−(2.703)	−(0.721)	−(0.899)	(0.086)	−(6.582)	(0.000)
4	0.218	**0.492**	−0.255	**−0.455**	−0.058	−0.113	−0.032	**−0.378**	**0.804**
	(1.249)	(11.378)	−(1.976)	−(3.026)	−(0.335)	−(0.488)	−(0.491)	−(5.650)	(0.000)
5	0.306	**0.462**	**−0.381**	**−0.467**	−0.161	−0.172	−0.003	**−0.409**	**0.771**
	(1.639)	(8.006)	−(2.579)	−(2.868)	−(0.858)	−(0.653)	−(0.034)	−(6.368)	(0.000)
1–5	0.081	−0.037	−0.016	−0.078	−0.050	0.000	0.122	−0.043	0.100
	(0.559)	−(1.083)	−(0.241)	−(0.715)	−(0.393)	(0.005)	(1.863)	−(1.051)	(0.305)

The five quintiles are sorted by the mean of rank-neutralised MSCI ESG scores. The upper panel of the table also provides the summary statistics for a portfolio 1–5, which takes a long position in portfolio 1 and a short position in 5. The bottom panel shows the OLS estimates and t-statistics for the regression of excess portfolio returns on seven risk factors and a constant. The risk factors are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). The standard errors are estimated using a HAC consistent estimator proposed by Newey and West (1987). The values that appear in the parentheses in the R-square column are the p-values associated to the F statistics that test the overall regression significance. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

(2002) for every portfolio. For each rater, on the right-hand-side of the table we detail the mean and median raw ESG scores, the standard deviation of the raw ESG score, and the average rank-neutralised ESG momentum for each of the quintile portfolios.[19]

Table 3 has a number of implications that crucially differ for Sustainalytics vs. MSCI ratings. First, ESG momentum offers potential explanatory power for the cross-section of European stock returns. In the case of the Sustainalytics scores, the SR of the spread portfolio 5–1 is negative for the 1- and 3 month ESG momentum windows. Moreover, only 1 month ESG momentum yields a long-short portfolio with a statistically significant SR. This confirms the general wisdom that the length of the momentum formation period affects the results (see also the robustness checks in Section "Leaving a one-month gap between ESG momentum measurement and portfolio formation"). When momentum and the resulting portfolio strategies are based on MSCI data instead, both the mean returns and the SR of the spread portfolios 5–1 are sizeable and statistically significant for all momentum windows, even though the performance statistics quickly decline as the window expands beyond 6 months. This finding naturally shifts our focus in what follows towards ESG momentum defined on relatively short windows of 1 and 3 months, as only in this case we find evidence of significant performance by an ESG-momentum driven strategy for both raters under investigation. Second, at 1- and 3 month formation periods, the rate of change of the ESG ratings is modest and as a result the intermediate quintiles (2-3-4) display low rates of momentum (e.g., for 3 month momentum, 6.9%, -0.3%, and -7% in the case of

[19] We report raw instead of rank-neutralised scores to be able to express sensible policy considerations, even though the raw scores are not used in the portfolio sorting process.

Sustainalytics and 0%, 0%, and -0.1% in the case of MSCI). Of course, such momentum scores increase over longer portfolio formation horizons and even become extreme (especially in the case of the MSCI ratings) in the 12 month case.[20] Third, while in the case of Sustainalytics, the mean 5–1 spread returns and the corresponding Sharpe ratios are negative, i.e., high ESG momentum commands a higher (risk-adjusted) performance than low momentum does, in the MSCI case the opposite occurs. For instance, using the former ratings, the mean spread and Sharpe ratio at 1 month are -0.425% and -0.248; using MSCI data, these are 2.479% and 1.782. This means that when MSCI rating momentum is used, there is a positive differential in returns from investing in low or negative ratings momentum, which may at first appear surprising.

Fourth, Table 3 shows a striking difference across Sustainalytics vs. MSCI in the way both the means and medians raw ESG scores are distributed across momentum-sorted portfolios. In the MSCI panel, mean and median ESG scores follow an asymmetric, tent-shaped pattern in which the extreme quantiles, in particular the high momentum one, have lower ESG scores than quantiles 2, 3, and 4. Therefore the investment strategy results indicate that stocks who had low ESG scores in the first place and still have low ESG scores, although they were able to improve their ESG performance in comparison with the other groups receive declining valuations in spite of their efforts at improving their ESG standing. This establishes a sort of primacy of the (initial) mean and median ESG scores over their subsequent changes (momentum), at least over very short periods of time. On the contrary, in the Sustainalytics panel, mean and median ESG scores follow a similar asymmetric tent-shaped pattern in which however the extreme quantiles, in particular the low momentum one, yield lower ESG scores than quantiles 2, 3, and 4. Therefore the strategy performance suggests that firms of poor ESG quality would be punished by declining valuations when efforts at improving their standing fail, as shown by low or negative momentum. On the opposite, firms that implement ESG understand the benefits of ESG integration in their businesses and hence invest in further ESG momentum and rewarded by investors for that. Yet, also in this case, the initial mean and median ESG scores matter more than momentum does *per se*.

To corroborate this point, Table 4 shows the performances obtained by a (rank-neutral) ESG score level-based strategy. Visibly, the ordering by rank-neutralised score (which are indeed on a 0-1 scale very different from the raw ones) implies also a very tight ordering by raw score; furthermore, the sorting based on the mean carries over to the median. Such an ordering seems to matter at least in the sense of generating a statistically significant average return spread of 0.33% per month but fails to generate a statistically significant estimated Sharpe ratio or alpha.[21] Crucially, a level ordering fails to organise in any meaningful way the cross section of rank neutralised momentum ESG scores, which range from +5.4% for the highly scoring stocks to +4% for the poorly scoring ones.[22]

Finally, in Table 3 the mean and median ESG scores are higher the longer the portfolio formation period is. Yet, this may represent an artifact of the visible upward trend in ESG scores over our sample period, as shown in panels (a) and (c)

[20] Because ESG ratings change rather infrequently, we need to carefully handle the cases in which, especially over short horizons, a majority of our stocks is characterised by zero momentum derived from no changes in their ESG scores. This explains why in Table 3, for Sustainalytics at the 1-month (3-month) horizon, portfolio 3 is on average collecting 142 stocks and most of the time, all or almost carry a zero-momentum score, with the result that the mean momentum score is -0.1% (-0.3%). Especially at the shortest, 1-month horizon, portfolios 2 and 4 tend to include a substantial fraction of stocks with no ESG score change (precisely, the fractions are on average 62.4% in the case of portfolio 2 and 58.8% in the case of portfolio 4), which explains their modest and symmetric momentum scores (4.1% and -4.1%). Even though our paper is based on the comparisons of the average returns of portfolios 1 and 5 according to the ESG score momentum-induced ranking, we allocate the zero momentum stocks across portfolios 2, 3, and 4 using the following algorithm: at every point in time, we randomly draw which stocks ought to be allocated across the three portfolios according to an equal probability distribution and compute the equivalent of Table 3 500 times. Table 3 represents an average of these 500 tables. The randomisation seems to have modest effects as in comparisons, for instance, the first, the last, and the average of such Tables imply no substantial differences for the results concerning portfolios 2 and 4. For example, the Sharpe ratio of the 1-month momentum sorted portfolio 2 (4) in the "averaged" Table 3 is 0.158 (0.151) vs. 0.157 (0.151) in the first copy of Table 3 obtained at random and 0.158 (0.152) in the last, 500th copy of the same. The randomisation is applied at all horizons but as the horizon grows, it becomes less impactful as there are more frequent rating changes over time. In the case of MSCI, we have followed the same logic because at the 1-month horizon we have now that portfolios 2, 3, and 4 always and only include stocks with zero rating momentum but such stocks minimally spill over into portfolios 1 and 5 (precisely, the fractions of zero ESG momentum stocks are on average 6.1% in the case of portfolio 1 and 12.7% in the case of portfolio 5). Also in this case, at every point in time, we randomly draw which zero-momentum stocks ought to be allocated across all portfolios according to an equal probability discrete distribution and compute Table 3, 500 times. Yet, we have noted that comparing the first, the last, and the average of such Tables makes no differences for the results concerning portfolios 1 and 5 and for the spread 5-1 portfolio returns.

[21] In the bottom panel of Table 4, we report an economically small and imprecisely estimated alpha of 0.08% per month. In fact, this result provides justification to our efforts to isolate non-level ESG signals as a basis to quantitative investment strategies.

[22] We have repeated this level-based cross-sectional analysis applied to Sustainalytics data but failed to find any significant results for either performances of momentum scores. To save space, we omit these outputs.

of Fig. 2.[23] Furthermore, the average ESG momentum tends to increase as the length of the formation period increases from 1 to 12 months. Especially over long formation windows, the average portfolio momentum may become extreme as over several months it is not uncommon for rank-neutralised scores below 50 in the interval [0,100] to even double. Because of the general positive trends in ESG scores, positive momentum prevails over negative momentum, and the positive and negative are more extreme (e.g., +124% for quintile 1 vs. -36% for quintile 5 in the case of Sustainalytics and +182% vs. -32% for MSCI, both estimated at 6 months). The table also shows a decrease in the mean return and Sharpe Ratios (SRs) as the portfolio formation period extends. This observation suggests that the ESG signal's strength diminishes as the computation period lengthens, aligning with the expected behaviour in the context of the informational efficiency of the European equity markets under examination.

Abnormal portfolio performance

Given the earlier evidence suggesting that ESG momentum may represent a source of (at least, firm-specific) risk, we investigate whether there is actual alpha that portfolios sorted by 1 month ESG momentum may consistently earn. This necessitates the adjustment of returns based on their (estimated) systematic risk exposure. It is crucial to consider the possibility that the mean excess returns obtained by the portfolios sorted by 1 month ESG momentum might be attributed to risk exposures that cannot be identified by the Sharpe Ratio, as discussed in previous research (see, e.g. Bruno et al. (2022)). In particular, we select a linear factor model that includes the five Fama and French (2015) factors, i.e., excess market returns and the returns on the Small Minus Big (SMB), High Minus Low (HML), Robust Minus Weak (RMW), and Conservative Minus Aggressive (CMA) portfolios. These factors approximate market, size, value, profitability, and investment risks, respectively. Given the objective of our paper, besides the five Fama and French (2015) European factors, we also include the Carhart (1997) momentum factor.[24] Finally, a seventh factor is the low volatility factor added to our model in light of the results in Blitz and Van Vliet (2007) who motivate the existence of the anomaly by the listing of sin stocks.[25]

Panel A in Table 5 shows the estimated coefficients and t-statistics for strategies driven by Sustainalytics rating momentum, while panel B does the same for MSCI rating momentum. First, there is evidence that the estimates of Jensen's α vary when the ESG momentum of a portfolio decreases, moving from top to bottom. In the case of Sustainalytics data, the alphas are weakly decreasing, while in the case of MSCI these are strongly increasing. In both cases, this is consistent with the findings in Table 3. Panel A shows that the estimate of α for portfolio 1 is significant at a 5% size level while for portfolios 2 and 4 they are significant only at a 10% size. Nonetheless, the alpha estimate for the long-short portfolio, which can be seen as a clean strategy implementation of the ESG momentum signal, is significant at a 1% level. The estimated alphas are also non-negligible, ranging from 0.12 to 0.58 per cent per month. In panel B, the estimates of the αs are much larger as they range between -0.74 and 1.69 per cent per month, but also in this case only the extremes are statistically significant. The alpha for the long-short portfolio which buys negative momentum, high mean ESG rating stocks to sell positive momentum, low rated stocks is significant at a 1% level. This is additional, formal evidence of the economic value that can be generated through sorting European stocks by 1 month ESG momentum. The table also shows that portfolios sorted by ESG momentum carry a substantial, precisely estimated exposure to market risk, generally inferior to 1 as this is a sample of relatively large-size firms and the market betas are anyway obtained within a multi-factor model. For each rating provider, the five portfolios all imply a positive coefficient with a large t-statistic. However, the effect vanishes for portfolio 5–1, indicating that the pure ESG momentum signal is not influenced by market risk. On the contrary, the negative exposures to the low volatility factor of the long-only quintile portfolios fail to strongly carry over to the long-short portfolios, as visible especially in panel B. The classical Carhart's return momentum provides a scant contribution to the returns earned by ESG momentum, especially insofar as the long-short portfolios go. This confirms that ESG momentum and actual momentum are two separate factors that do not exhibit much overlap. Finally, while the Jensen's alpha derived from the Sustainalytics implementation of the

[23] The longer the ESG momentum computation period, the more of the earlier months in our sample are ignored because of the need for a longer initialisation sample. Because the ESG scores have been increasing over time, dropping the earlier, lower scores increases the resulting average and median in a mechanical way.

[24] The data on the factors are downloaded from the data repository maintained by Kenneth French at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Blitz and Fabozzi (2017), Yasmine and Kooli (2022) and Sagbakken and Zhang (2022) with explicit reference to European data, have stressed the importance of including all of the Fama and French's five factors when assessing the average abnormal performance of sin vs. virtuous stocks.

[25] This latter factor mimicking portfolio is constructed by computing the rolling, 36-month volatility for every stock in our investment universe. Next, we sort the quintile portfolios in order of ascending, past realised volatility and create a long-short portfolio. This zero-cost portfolio implies a long position in the portfolio containing stocks with the lowest volatility and a short position in the portfolio collecting stocks with the highest volatility.

Table 5 Estimates of regressions of ESG momentum-sorted portfolio excess returns on 7 risk factors

	α	Mkt.–Rf	SMB	HML	RMW	CMA	Momentum	Low Vol.	R-square
Panel A: Sustainalytics rating-based Momentum strategy									
1	**0.581**	**0.569**	– 0.155	– 0.162	0.492	0.048	– 0.065	**– 0.417**	**0.776**
	(2.215)	(7.010)	– (0.768)	– (0.731)	(1.804)	(0.230)	– (0.987)	– (3.893)	(0.000)
2	0.446	**0.617**	– 0.052	0.095	0.527	0.005	– 0.116	**– 0.264**	**0.766**
	(1.724)	(6.911)	– (0.226)	(0.333)	(1.736)	(0.027)	– (1.670)	– (2.311)	(0.000)
3	0.428	**0.639**	0.049	0.196	0.588	– 0.149	– 0.144	**– 0.265**	**0.771**
	(1.527)	(6.301)	(0.184)	(0.609)	(1.619)	– (0.684)	– (1.950)	– (2.000)	(0.000)
4	0.426	**0.613**	– 0.078	0.210	0.595	– 0.262	– 0.086	**– 0.310**	**0.789**
	(1.718)	(6.916)	– (0.299)	(0.724)	(1.816)	– (1.330)	– (1.340)	– (2.858)	(0.000)
5	0.123	**0.623**	0.110	0.323	0.539	**– 0.413**	– 0.079	– 0.217	**0.758**
	(0.458)	(6.104)	(0.413)	(1.085)	(1.587)	– (2.012)	– (1.215)	– (1.767)	(0.000)
5–1	**– 0.457**	0.052	**0.267**	**0.494**	0.050	**– 0.462**	– 0.016	**0.204**	**0.243**
	– (3.108)	– (1.218)	– (2.619)	– (3.939)	– (0.298)	(4.787)	(0.427)	– (3.448)	(0.000)
Panel B: MSCI rating–based Momentum strategy									
1	**– 0.736**	**0.470**	– 0.258	**– 0.860**	– 0.394	– 0.251	0.038	**– 0.529**	**0.720**
	– (2.773)	(6.764)	– (1.650)	– (3.561)	– (1.169)	– (0.768)	(0.315)	– (5.996)	(0.000)
2	0.287	**0.577**	**– 0.491**	**– 0.480**	– 0.054	– 0.145	**0.186**	**– 0.442**	**0.775**
	(1.341)	(8.679)	– (3.143)	– (2.347)	– (0.259)	– (0.665)	(2.123)	– (5.664)	(0.000)
3	0.296	**0.550**	**– 0.399**	**– 0.433**	– 0.044	– 0.207	**0.170**	**– 0.451**	**0.825**
	(1.819)	(10.090)	– (2.945)	– (2.622)	– (0.263)	– (1.165)	(2.145)	– (6.290)	(0.000)
4	0.306	**0.537**	**– 0.315**	**– 0.540**	– 0.031	– 0.201	0.067	**– 0.471**	**0.813**
	(1.536)	(10.175)	– (2.274)	– (3.146)	– (0.149)	– (0.924)	(0.754)	– (6.342)	(0.000)
5	**1.694**	**0.565**	– 0.260	**– 0.641**	– 0.244	– 0.284	0.172	**– 0.425**	**0.785**
	(8.249)	(7.723)	– (1.815)	– (3.715)	– (1.267)	– (1.211)	(1.963)	– (5.232)	(0.000)
5–1	**2.430**	0.095	– 0.003	0.219	0.150	– 0.032	0.134	0.103	**0.126**
	(14.436)	(1.654)	– (0.029)	(1.254)	(0.675)	– (0.152)	(1.946)	(1.809)	(0.034)

This table shows the OLS estimates and t-statistics for the regression of excess portfolio returns on seven risk factors and a constant. The portfolios are sorted by their 1-month ESG momentum, the percentage increase of the ESG score over one month, in descending order. The 5–1 portfolio is a long-short portfolio which takes a long position in portfolio 5, and a short position in portfolio 1. The incorporated risk factors are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). Panel A reports the results for the returns of strategies driven by Sustainalytics rating momentum, while panel B for the case of MSCI ratings. The standard errors are estimated using a HAC consistent estimator proposed by Newey and West (1987). The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

strategy seems to be partially explained by some correlation with other factors (in particular, size, value, and the investment factors, besides low volatility, as discussed above, to achieve an overall R-square of 24%), in the case of MSCI the 5-1 excess returns represent the realisations of an almost pure factor as these are not explained by any other priced risk component and yield a modest R-square of 13% only.

Crucially, Table 5 indicates that some of the individual ESG momentum-sorted portfolios do earn significant alpha. However, to assess whether our seven-factor model is able to explain these excess returns, we must perform a formal test on the joint significance of the pricing errors.[26] The test proposed by Gibbons et al. (1989) (GRS) has this exact aim. The GRS test statistic is

$$z = \frac{T-n-k+1}{n(T-k)} \frac{1}{q_{11}} \hat{\boldsymbol{\alpha}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} \sim F_{n,T-n-k+1} \quad (1)$$

where $\hat{\boldsymbol{\alpha}}$ is the vector of estimated alphas from the factor model, $\hat{\boldsymbol{\Sigma}}$ is the estimated covariance matrix of the residuals

[26] It is possible that we may be dealing with a system of equations that are only connected via correlations in the error terms and that by taking these into account, all abnormal excess returns may become jointly insignificant.

Table 6 Correlation matrix (lower triangle) and associated t-ratios (upper triangle) of the ESG Momentum factor vs. seven other systematic risk factors

	ESG Momentum	Mkt–Rf	SMB	HML	RMW	CMA	Momentum	Low Vol.
Panel A: Sustainalytics ratings								
ESG Momentum	1	*1.31*	*1.81*	*3.41*	*−2.00*	*0.21*	*−1.87*	*−0.72*
Mkt–Rf	0.12	1	*−0.09*	*−6.22*	*4.26*	*−0.67*	*4.44*	*11.52*
SMB	0.16	−0.01	1	*−0.42*	*0.60*	*1.13*	*0.23*	*4.24*
HML	**0.29**	**−0.49**	−0.04	1	*14.03*	*−8.81*	*6.41*	*8.00*
RMW	**−0.18**	**0.36**	0.05	**0.77**	1	*5.70*	*−5.24*	*−7.02*
CMA	0.02	−0.06	0.10	**−0.61**	**0.45**	1	*1.39*	*0.85*
Momentum	−0.16	**0.37**	0.02	**0.49**	**−0.42**	0.12	1	*−5.49*
Low Vol.	−0.06	**0.71**	**0.35**	**0.59**	**−0.53**	0.08	**−0.44**	1
Panel B: MSCI ratings								
ESG Momentum	1	*0.33*	*−0.67*	*0.00*	*0.47*	*0.33*	*1.98*	*1.59*
Mkt–Rf	0.03	1	*1.47*	*−3.78*	*2.87*	*0.19*	*4.49*	*8.59*
SMB	−0.07	0.15	1	*0.23*	*0.68*	*1.47*	*−0.95*	*1.09*
HML	0.00	**−0.37**	0.02	1	*14.83*	*−7.98*	*4.51*	*7.03*
RMW	0.05	**0.29**	0.07	**0.84**	1	*5.69*	*−3.43*	*−5.66*
CMA	0.04	0.02	0.15	**−0.65**	**0.52**	1	*0.99*	*1.46*
Momentum	0.23	**0.43**	−0.10	**0.43**	**−0.34**	0.10	1	*−7.61*
Low Vol.	0.17	**0.67**	0.11	**0.60**	**−0.51**	0.15	**−0.63**	1

This table reports the correlation matrix of the ESG Momentum factor and of seven other risk factors. The lower triangle shows the pairwise correlations and the upper triangle shows the corresponding t-statistics (in italics). The ESG Momentum factor is constructed by taking a long position in the quintile portfolio containing stocks with the lowest 1-month ESG momentum and taking a short position in the quintile portfolio containing stocks with the highest 1-month ESG momentum. The risk factors considered are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA as proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

from the regressions, T the number of months in the sample, n the number of portfolios and q_{11} is the first element of $(X'X)^{-1}$ (where X collects the factors used in the factor model). In the case of Table 5, in the panel for the Sustainalytics momentum strategy, the test statistic is 2.417, with an associated p-value of 0.040; in the case of MSCI momentum, the test statistic is 39.37, with an associated p-value of 0.000. This means that in both applications the estimated abnormal excess returns are jointly significant and cause the set of factors to be rejected as in-sample explanations of the long-short realised returns. Hence, the postulated seven-factor model is unable to adequately price the ESG momentum portfolios and the resulting average abnormal returns may be a real effect.[27]

We now assess the potential of ESG momentum to represent a new systematic risk factor. The ESG momentum factor is proxied by the long-short portfolio 5–1, where the 1 month ESG momentum is computed using the rank-neutralised ESG scores. Table 6 shows the Pearson correlation matrix of the ESG momentum factor and the seven other risk factors as well as the t-statistics of these correlations.

Panel A concerns the case of Sustainalytics while panel B refers to MSCI. The lower triangle of the matrix displays the sample correlations while the upper triangle (printed in italics) shows the corresponding t-statistics. Both panels show that the ESG momentum factor is only weakly correlated with value and profitability when is extracted using Sustainalytics ratings and it is otherwise linearly unrelated to all other traditional factors.[28] In particular, ESG momentum is essentially uncorrelated with the classical momentum and low volatility factors.[29]

To supplement this information and even though for every formation horizon we have built only five portfolios of

[27] However, we need to be mindful of the fact that the sign of the premium of such an alleged factor derived from the same universe of European firms is different across Sustainalytics vs. MSCI rating data.

[28] These correlations suggest that high ESG momentum companies are profitable, growth companies. An intuitive argument would be that these companies are new since they are still growing. Many established companies, in fact, tend to be typically heavily invested already.

[29] In the case of Sustainalytics, this may be surprising given the fact that Table 5 showed a significant estimate for the low volatility factor. However, the table also shows that HML and the Low Volatility factor are significantly negatively correlated, indicating that the

Table 7 F-tests of the pricing performance of an ESG momentum factor

	Momentum-sorted Ptf.	Sustainalytics ratings			MSCI Ratings		
		Full Model Estimated a_i	Partial F-tests	Joint F-test	Full Model Estimated a_i	Partial F-tests	Joint F-test
1-month Strategy	1	0.114 (0.063)			0.146 (0.030)		
	2	0.053 (0.148)	Exclude ESG Mom Factor	388.004	0.056 (0.197)	Exclude ESG Mom Factor	441.636
	3	0.061 (0.143)	*88.816 (0.000) REJECT*		0.043 (0.232)	90.960 (0.000) *REJECT*	
	4	0.156 (0.001)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)	0.114 (0.002)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)
	5	0.011 (0.052)	11.481 (0.043) REJECT		0.145 (0.030)	30.298 (0.000) REJECT	
3-month Strategy	1	0.073 (0.108)			0.089 (0.090)		
	2	0.010 (0.048)	Exclude ESG Mom Factor	304.992	0.063 (0.095)	Exclude ESG Mom Factor	279.684
	3	0.085 (0.063)	*74.557 (0.000) REJECT*		0.059 (0.094)	11.230 (0.047) *REJECT*	
	4	0.068 (0.245)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)	0.072 (0.338)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)
	5	0.007 (0.158)	*10.101 (0.072) NOT REJECT*		0.090 (0.092)	9.936 (0.077) *NOT REJECT*	
6-month Strategy	1	0.041 (0.034)			0.050 (0.055)		
	2	0.025 (0.305)	Exclude ESG Mom Factor	211.187	0.023 (0.456)	Exclude ESG Mom Factor	171.058
	3	0.067 (0.070)	*94.581 (0.000) REJECT*		0.045 (0.122)	91.000 (0.000) *REJECT*	
	4	0.070 (0.043)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)	0.052 (0.061)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)
	5	0.003 (0.883)	*10.005 (0.075) NOT REJECT*		0.044 (0.058)	8.746 (0.120) *NOT REJECT*	
12-month Strategy	1	0.034 (0.079)			0.045 (0.044)		
	2	0.024 (0.227)	Exclude ESG Mom Factor	132.504	0.016 (0.502)	Exclude ESG Mom Factor	119.276
	3	0.187 (0.002)	*99.020 (0.000) REJECT*		0.020 (0.400)	91.001 (0.000) *REJECT*	
	4	0.078 (0.010)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)	0.054 (0.021)	$\boldsymbol{\alpha} = \mathbf{0}$ (GRS test)	(0.000)
	5	0.004 (0.903)	22.270 (0.000) REJECT		0.046 (0.042)	17.170 (0.004) REJECT	

This table reports the estimated alphas from an unconstrained system regression model that includes an ESG Momentum factor and seven other risk factors. The ESG Momentum factor is constructed by taking a long position in the quintile portfolio containing stocks with the highest x-month ESG momentum and taking a short position in the quintile portfolio containing stocks with the lowest x-month ESG momentum, where x = 1, 3, 6 and 12 months. The risk factors considered are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA as proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). The table also performs the three F-tests of asset pricing restrictions concerning the role played by the ESG Momentum factor. We have italicised cases in which the rejection or non-rejection of a null conforms to the hypothesis of ESG Momentum serving as a factor.

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

stocks, in Table 7 we perform a typical asset pricing test. As a first step, for each of the four horizons, we use the spread 5–1 portfolio returns as a candidate risk factor that we shall be calling *ESG_mom(x)*, where $x = 1$, 3, 6 and 12 months. Next, we set up and compare (using F-tests applied on cross-equation restrictions) four systems of time series, factor-based portfolio regressions, in which the estimation occurs by SUR (seemingly unrelated) GLS methods. The first system is unconstrained and contains the seven classical, finance factors already described before augmented to include *ESG_mom(x)*; the second system is constrained in

Footnote 29 (continued)

Low Vol and ESG momentum factors are indeed implicitly, indirectly correlated. This outcome underscores the significance of employing a state-of-the-art factor pricing model to unveil the genuine underlying relationships. Relying solely on pairwise correlations, without accounting for other relevant risk factors, can lead to erroneous conclusions regarding the pairwise relationship.

that all the portfolio exposures vs. *ESG_mom(x)* are constrained to be zero, so that an F-test built as an R-squared comparison between the first and second system allows to test whether *ESG_mom(x)* may be a priced risk factor; the third system contains *ESG_mom(x)* as an explanatory factor but it constraints all the alphas of the five regressions in the system to be zero, so that an GRS/F-test built as an R-squared comparison between the first and third system allows to test whether *ESG_mom(x)* sets the alphas to zero; the fourth system simultaneously implements all the constrains of the second and third models and essentially tests whether *ESG_mom(x)* belongs to these linear pricing models and it can set to zero (in the sense of not being statistically significant) all the abnormal, ESG momentum-sorted portfolio returns.[30] Therefore the fourth model estimated tests whether the seven factor model would be able to exclude *ESG_mom(x)* and at the same time to "correctly price" the five ESG-momentum sorted portfolios.

Table 7 shows the results of such system of regressions tests. Notably, in spite of the rather different behaviour of the signs of the returns of the ESG momentum spread portfolio across the two raters (especially at a one-month horizon), the two corresponding panels reveals results which are somewhat homogeneous. Moreover, in all the eight cases featured (four horizons in the case of Sustainalytics and four in the case of MSCI), the joint F-test that compares the first unconstrained model with the fourth, constrained one, lead to bitter rejections of the restrictions, thus giving evidence (consistent with Table 5) that a classical, seven-factor model cannot price the ESG-sorted portfolios. It is more interesting to note that in all the eight cases, the null of the ESG momentum factor not belonging to the set of factors is rejected with p-values of 0.000, while in only 4 cases (and always when $x = 3$ and 6 months) the null of all the individual portfolio regression alphas being zero cannot be rejected (hence it leads to p-values that exceed 0.05, even though in one case, MSCI and $x = 6$ months, the resulting p-value exceeds 0.1). All in all, Table 7 provides evidence that the ESG-momentum factor may be real, in the sense of always helping to explain the spread returns built on the basis of momentum and also contributing (at least at the intermediate horizons) to set the model pricing performance "to be right", at least in a statistical sense.

[30] Apart from when it is explicitly indicated, all systems contain estimable intercepts in each regression, arguably to be interpreted as Jensen's alphas. The GRS/F-tests are all implemented using the estimated covariance matrix of the pricing errors obtained from the first, unconstrained model.

ESG Volatility

Portfolio sorting

We define x -month ESG volatility as the sample standard deviation of (size- and industry-adjusted) ESG scores over a fixed time window of x months, where $x \in \{18, 24, 30, 36\}$. Stocks are sorted by ascending ESG volatility and we form equally-weighted quintile portfolios where portfolio 1 contains stocks with the most stable ESG scores while in portfolio 5 we collect the stocks with the most volatile scores over the previous x months. Also in this case, we repeat the exercise separately for Sustainalytics and MSCI data in different panels of the Tables.

Table 8 shows that for two estimation windows (18 and 24 months), the long-short (quintile 1–5) portfolio is characterised by a significant realised SR for both rating providers. Therefore, we focus our attention on the results concerning portfolios sorted on 18-month ESG volatility since it leads to high SR across our data sets. Moreover, when compared to Table 3, the mean spread returns and Sharpe ratios (here from the 1–5 portfolio), now carry the same sign across the two raters, i.e., it is profitable to go long in low-variance ESG rating stocks and to short the ones with volatile ratings. Furthermore, the rank-neutralised ESG score volatility does not highly correlate with raw ESG volatility. For instance, the portfolio with the lowest rank-neutralised ESG volatility has by far the highest overall standard deviation of the raw ESG scores. In fact, all average and median raw ESG scores across quintiles tend to increase with the ESG volatility across portfolios. This may be problematic for sustainability-driven investors, as it implies that investing in firms with stable ESG scores characterised by low ESG volatility leads to lower portfolio ESG mean and median scores. This strengthens the finding in Table 5 that it is long positions in negative momentum, intermediate-ESG score firms that tend to generate the highest average abnormal returns.[31]

Abnormal portfolio performances

It remains vital to establish whether a strategy earns a return because of its implied exposure to one or more risk factors or because it earns true risk- adjusted abnormal returns. To distinguish between the two cases, we adopt the same seven-factor linear model as in subsection "ESG Momentum", which includes tradeable proxies for market, size, value, profitability, investment, momentum, and volatility risks. Panel A of Table 9 shows the OLS estimates (and the

[31] A sustainability-driven investor aims at over-weighting stocks characterised by relatively high raw ESG scores or to exclude stocks with low raw ESG scores irrespective of their predicted market performance and reward-to-risk ratio.

Table 8 Results for portfolio sorted by ESG volatility using four computation periods

	Sustainalytics								MSCI							
	Returns				Sustainalytics raw ESG scores				Returns				MSCI raw ESG scores			
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Volatility	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Volatility
Panel B: 18-month ESG Volatility																
1	0.598	4.378	0.137	0.103	63.911	62.562	0.791	3.3%	**0.844**	3.864	**0.207**	0.089	60.523	60.044	3.667	0.9%
2	0.653	4.302	0.152	0.097	64.554	64.662	0.591	4.9%	0.570	3.801	0.138	0.089	61.910	61.868	2.074	3.3%
3	0.583	4.503	0.129	0.098	64.360	65.941	0.552	6.3%	**0.807**	3.911	**0.195**	0.092	63.964	63.826	2.176	5.8%
4	0.570	4.454	0.128	0.098	65.557	66.465	0.604	8.4%	0.548	3.621	0.139	0.090	65.026	65.330	2.163	9.2%
5	0.323	4.740	0.068	0.098	66.131	66.370	0.651	13.6%	0.618	3.985	0.155	0.088	64.951	65.012	1.946	16.5%
1–5	**0.275**	1.153	**0.239**	0.074					**0.226**	1.151	**0.196**	0.090				
Panel B: 24-month ESG Volatility																
1	0.691	4.425	0.156	0.108	63.608	63.152	0.697	3.9%	**1.732**	1.448	**0.183**	0.089	61.259	61.043	2.653	1.7%
2	0.643	4.620	0.139	0.100	64.812	64.317	0.437	5.7%	1.545	1.390	0.156	0.095	63.153	62.636	2.284	4.6%
3	0.516	4.393	0.118	0.105	65.451	65.819	0.466	7.4%	1.303	1.121	0.150	0.096	65.185	65.243	2.484	7.3%
4	0.603	4.614	0.131	0.106	66.009	67.014	0.488	9.5%	1.449	1.231	0.136	0.095	64.892	65.328	2.084	10.6%
5	0.347	4.830	0.072	0.104	67.535	67.554	0.545	14.8%	0.797	0.690	0.147	0.094	65.667	65.777	1.807	17.9%
1–5	**0.344**	1.188	**0.256**	0.087					**0.936**	2.948	**0.317**	0.071				
Panel B: 30-month ESG Volatility																
1	**0.841**	4.430	0.190	0.106	63.936	63.372	0.603	4.3%	**2.104**	1.792	**0.194**	0.093	61.389	61.493	2.903	2.4%
2	**0.907**	4.275	**0.212**	0.105	65.958	65.330	0.331	6.2%	**2.352**	2.016	**0.207**	0.093	63.252	62.593	2.625	5.6%
3	**0.771**	4.193	0.184	0.104	66.555	65.879	0.366	7.9%	**2.040**	1.767	**0.207**	0.100	65.785	66.250	2.136	8.4%
4	0.574	4.647	0.124	0.108	66.732	67.477	0.355	10.3%	**1.370**	1.146	**0.200**	0.093	65.306	65.306	1.796	11.8%
5	0.661	4.534	0.146	0.109	67.500	67.907	0.438	15.6%	**1.617**	1.338	0.180	0.094	66.328	65.835	1.999	18.9%
1–5	0.179	1.173	0.153	0.103					**0.487**	0.962	**0.507**	0.094				
Panel B: 36-month ESG Volatility																
1	0.757	4.538	0.167	0.102	64.990	63.536	0.551	4.8%	**1.851**	1.630	0.179	0.091	62.012	62.123	2.562	3.0%
2	**0.870**	4.292	**0.203**	0.099	65.147	66.138	0.373	6.9%	**2.249**	2.049	0.173	0.095	64.476	63.891	3.265	6.4%
3	0.666	4.312	0.155	0.105	67.461	66.410	0.299	8.5%	**1.714**	1.465	**0.226**	0.096	65.755	66.229	2.397	9.2%
4	0.594	4.528	0.131	0.104	66.744	67.209	0.349	11.1%	**1.455**	1.261	0.175	0.097	66.315	66.092	1.694	12.7%
5	0.620	4.374	0.142	0.109	68.481	67.397	0.418	16.8%	1.572	1.300	0.161	0.095	66.581	66.658	2.442	19.7%
1–5	0.137	0.987	0.085	0.105					**0.279**	0.807	**0.346**	0.099				

The five quintiles are sorted by ESG score volatility, computed using rank-neutralised ESG scores. The table also provides the summary statistics for a portfolio 1–5, which takes a long position in portfolio 1 and a short position in portfolio 5. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

associated t-ratios) of the factor loadings for each of the quintile portfolios built in subsubsection "Portfolio sorting" with reference to Sustainalytics data, while panel B performs the same task with reference to MSCI.

The table shows that, with one exception, the average pricing errors for the quintile portfolios are not statistically significant at standard confidence levels and the null hypothesis of a zero alpha cannot be rejected for any of them.[32] However, the alpha for the long-short portfolio is highly significant in both panels: the pure form of the ESG volatility signal earns significant average abnormal returns. Interestingly, the low minus high ESG rating volatility premium is of similar magnitude regardless of whether Sustainalytics or MSCI scores are used, 0.261 and 0.297 per cent per month, respectively. Moreover, consistent with the results obtained in subsection "ESG Momentum", Table 9 implies high and significant estimates of the market factor exposures for all the quintile portfolios. Yet, the effect disappears for the long-short strategies, illustrating that the ESG volatility signal is indeed not significantly affected by market risk for any reasonable confidence level.

Additionally, Table 9 shows a negative and significant estimate of the exposure to the equity Low Volatility factor for every quintile portfolio and both rating providers. However, this negative relationship completely disappears for the long-short portfolio 1–5 and in both panels. As a result, the ESG volatility factor is not explained by or significantly correlated with the classical low volatility anomaly and the resulting factor. This emphasises that, similarly to how ESG momentum and classical momentum were largely unrelated, the low volatility and ESG volatility are separate factors. This reveals one of the strengths of ESG-linked strategies: they appear to be uncorrelated with traditional factors and therefore may generate substantial diversification gains. Finally, Table 9 shows that the long-short portfolio carries no significant risk exposure to any of the seven risk factors. The only significant estimate is the average abnormal return which is highly significant (the t-stat is 3.04 in panel A and 3.05 in panel B) and measures a rather substantial risk-adjusted excess return of over 3 per cent per year (3.13% in the case of Sustainalytics and 3.57% in the case of MSCI) earned by the ESG volatility long-short portfolio.

Table 10 performs the same formal asset pricing tests on systems of factor-based return regressions that we have already commented with reference to Table 7 to test the performance of the ESG volatility factor taken to represent the spread portfolio that goes long into the lowest ESG volatility portfolio and shorts the highest ESG volatility one. Rather impressively, all the tests proposed work out in seven out of the eight combinations of portfolio formation horizons and rating providers:[33] the F-tests reject the null hypothesis of the ESG volatility factor implying zero exposures by the five portfolios examined; the GRS tests cannot reject the null hypothesis that when the ESG volatility factor is included the resulting test portfolio alphas are jointly zero, with p-values ranging from 0.095 to 0.631; a joint F-test that the seven classical factors can jointly set all the alphas to zero and exclude the ESG volatility from the pricing equations always lead to rejections with a p-value that is essentially zero.

Robustness checks

Alternative definitions of ESG momentum and volatility

We repeat the empirical tests in subsubsection 3.1.3 using alternative definitions of ESG momentum, e.g., when the number of months used to estimate momentum exceed one. Table 18 in the Supplementary Information shows the estimated alphas and the GRS test statistics for long-short portfolios sorted by 1-, 3-, 6-, and 12- ESG momentum, respectively. ESG momentum is computed using the rank-neutralised ESG scores. We find that as the computation period increases, the estimated pricing error of the long-short portfolio 5–1 converge monotonically towards zero both in levels and in terms of the implied t-statistic. This confirms that the choice of the momentum computation period is an important one as the statistical significance of average abnormal returns on the long-short portfolio fades substantially (in panel A, with reference to Sustainalytics data, disappears entirely) in the case of formation periods longer than 3 months.

We perform similar robustness tests for ESG volatility. Table 11 shows the average abnormal returns and the corresponding t-statistics for the five quintile portfolios and the long-short portfolio for four ESG volatility-computation

[32] Given that none of the quintile portfolios individually exhibits a statistically significant alpha at the 5% significance level, it is advisable to further assess the efficacy of the postulated seven-factor model in pricing the returns of portfolios categorised by ESG volatility. To achieve this comprehensive evaluation, the GRS test can be employed. Computing this test for the quintile portfolios sorted by ESG volatility yields a statistic of 1.510 in the case of the Sustainalytics data; the probability that this value is exceeded under the null hypothesis is 0.194. The GRS test is instead 1.144 in the case of the MSCI data, with an even higher p-value of 0.344.

[33] The only exception occurs with reference to the 24-month horizon for Sustainalytics ratings, where we find that the null that by adding the ESG volatility factor all the alphas are jointly set to zero is rejected with a p-value of 0.047. Nonetheless, the test of the ESG volatility factor belonging to the system of regressions and the joint F test always reject the null and establish evidence in favor of the ESG volatility factor.

Table 9 Estimates of regressions of ESG volatility-sorted portfolio excess returns on 7 risk factors

	α	Mkt.–Rf	SMB	HML	RMW	CMA	Momentum	Low Vol.	R-square
Panel A: Sustainalytics rating-based Volatility strategy									
1	0.437	**0.672**	−0.063	0.063	0.730	0.014	−0.119	**−0.262**	**0.781**
	(1.877)	(7.360)	−(0.259)	(0.214)	(1.857)	(0.059)	−(1.676)	−(2.281)	(0.000)
2	0.440	**0.725**	0.045	0.075	0.597	0.193	−0.122	**−0.246**	**0.784**
	(1.789)	(7.915)	(0.183)	(0.244)	(1.557)	(0.820)	−(1.707)	−(2.042)	(0.000)
3	0.364	**0.707**	−0.120	0.076	0.571	0.193	−0.123	**−0.245**	**0.799**
	(1.657)	(8.985)	−(0.562)	(0.261)	(1.711)	(0.884)	−(1.817)	−(2.338)	(0.000)
4	0.432	**0.701**	−0.162	0.052	0.551	0.111	−0.088	**−0.323**	**0.793**
	(1.742)	(7.235)	−(0.742)	(0.216)	(1.667)	(0.402)	−(1.076)	−(2.703)	(0.000)
5	0.180	**0.742**	−0.025	0.121	0.598	0.137	−0.110	**−0.313**	**0.839**
	(0.702)	(7.241)	−(0.099)	(0.377)	(1.558)	(0.720)	−(1.898)	−(2.886)	(0.000)
1 -5	**0.264**	−0.058	−0.038	−0.058	0.131	−0.124	−0.009	0.056	**0.386**
	(3.071)	−(1.305)	−(0.600)	−(0.631)	(1.112)	−(0.941)	−(0.220)	(1.100)	(0.000)
Panel B: MSCI rating−based Volatility strategy									
1	0.607	**0.566**	**−0.371**	**−0.682**	−0.098	0.026	0.105	**−0.458**	**0.788**
	(1.807)	(9.010)	−(2.346)	−(3.770)	−(0.428)	(0.110)	(1.185)	−(5.969)	(0.000)
2	0.290	**0.503**	**−0.471**	**−0.550**	−0.133	−0.268	0.108	**−0.510**	**0.804**
	(1.480)	(8.397)	−(3.315)	−(2.929)	−(0.604)	−(1.225)	(1.210)	−(6.278)	(0.000)
3	**0.450**	**0.538**	**−0.395**	**−0.636**	−0.151	0.026	0.130	**−0.492**	**0.787**
	(2.199)	(9.032)	−(2.676)	−(3.328)	−(0.700)	(0.124)	(1.439)	−(5.960)	(0.000)
4	0.221	**0.478**	**−0.444**	**−0.696**	−0.254	−0.101	0.130	**−0.521**	**0.822**
	(1.182)	(8.911)	−(3.120)	−(3.874)	−(1.215)	−(0.484)	(1.487)	−(7.501)	(0.000)
5	0.310	**0.573**	**−0.414**	**−0.547**	−0.232	−0.288	0.093	**−0.448**	**0.791**
	(1.583)	(8.832)	−(2.680)	−(2.825)	−(0.902)	−(1.037)	(0.894)	−(5.211)	(0.000)
1 -5	**0.297**	−0.007	0.042	−0.135	0.135	**0.314**	0.011	−0.010	**0.139**
	(3.053)	−(0.180)	(0.577)	−(1.117)	(0.950)	(2.260)	(0.217)	−(0.226)	(0.038)

This table shows the OLS estimates and t-statistics from the regression of excess portfolio returns on seven risk factors and a constant. The portfolios are sorted by the ascending 18-month ESG volatility, the volatility of the ESG score over the last 18 months. The 1–5 portfolio is a long-short portfolio which takes a long position in portfolio 1 and a short position in portfolio 5. The incorporated risk factors are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). Panel A reports the results for the returns of strategies driven by Sustainalytics rating volatility, while panel B for the case of MSCI ratings. The standard errors are estimated using a HAC consistent estimator proposed by Newey and West (1987). The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

periods. It also displays the GRS statistics for testing the joint significance of the quintile portfolios' average abnormal returns. The table shows consistent results across the two panels. First, the 1–5 spread strategy yields a significant alpha only at horizons of 18 and 24 months. Interestingly, the resulting average abnormal returns (arguably, the low ESG volatility factor premia) are of similar magnitude, between 0.26 and 0.30% per month, across the two data providers. Second, even though the spread strategy yields statistically significant alphas at horizons inferior to 24 months, globally, a GRS test fails to reject the null that all the alphas are jointly equal to zero. Finally, for both data providers, it is the low(er) volatility portfolios 1 and 2 that are driving most of the differences. Their alphas are positive and statistically significant at the 30- and 36-month horizons.

Standardised ESG momentum and volatility

Table 19 in the Supplementary Information reports the results of robustness analyses applied to ESG momentum computed from the standardised ESG scores. The table reveals that none of the ESG momentum estimation periods lead to statistically significant alphas from the long-short 5–1 strategies for either raters under investigation or to significant GRS statistics. Therefore, the null hypothesis of $\alpha = 0$ cannot be rejected and we conclude that none of the

Table 10 F-tests of the pricing performance of an ESG volatility factor

	Volatility-sorted Ptf.	Sustainalytics ratings			MSCI ratings		
		Full model estimated α_i	Partial F-tests	Joint F-test	Full model estimated α_i	Partial F-tests	Joint F-test
18-month Strategy	1	0.019 (0.124)			0.037 (0.076)		
	2	0.028 (0.150)	Exclude ESG Vol Factor	98.824	0.019 (0366)	Exclude ESG Vol Factor	101.188
	3	0.042 (0.034)	99.854 (0.000) *REJECT*		0.047 (0.030)	91.542 (0.000) *REJECT*	
	4	0.043 (0.040)	**α = 0** (GRS test)	(0.000)	0.023 (0.230)	**α = 0** (GRS test)	(0.000)
	5	0.008 (0.549)	10.001 (0.095) *NOT REJECT*		0.031 (0.084)	9.832 (0.095) *NOT REJECT*	
24-month Strategy	1	0.015 (0.304)			0.031 (0.125)		
	2	0.060 (0.027)	Exclude ESG Vol Factor	94.105	0.040 (0.059)	Exclude ESG Vol Factor	96.636
	3	0.016 (0.284)	82.440 (0.000) *REJECT*		0.019 (0.360)	87.770 (0.000) *REJECT*	
	4	0.048 (0.094)	**α = 0** (GRS test)	(0.000)	0.026 (0.207)	**α = 0** (GRS test)	(0.000)
	5	0.008 (0.164)	11.221 (0.047) REJECT		0.029 (0.144)	5.572 (0.350) *NOT REJECT*	
30-month Strategy	1	0.015 (0.354)			0.032 (0132)		
	2	0.043 (0.052)	Exclude ESG Vol Factor	79.984	0.029 (0.162)	Exclude ESG Vol Factor	94.780
	3	0.022 (0.194)	79.005 (0.000) *REJECT*		0.024 (0.249)	91.000 (0.000) *REJECT*	
	4	0.044 (0.089)	**α = 0** (GRS test)	(0.000)	0.025 (0.246)	**α = 0** (GRS test)	(0.000)
	5	0.009 (0.679)	8.004 (0.156) *NOT REJECT*		0.032 (0.132)	3.122 (0.681) *NOT REJECT*	
36-month Strategy	1	0.016 (0.324)			0.032 (0.145)		
	2	0.027 (0.132)	Exclude ESG Vol Factor	100.086	0.019 (0.315)	Exclude ESG Vol Factor	92.118
	3	0.029 (0.166)	95.583 (0.000) *REJECT*		0.033 (0.114)	84.881 (0.000) REJECT	
	4	0.040 (0.093)	**α = 0** (GRS test)	(0.000)	0.023 (0.292)	**α = 0** (GRS test)	(0.000)
	5	0.008 (0.640)	3.515 (0.621) *NOT REJECT*		0.029 (0.140)	3.502 (0.623) *NOT REJECT*	

This table reports the estimated alphas from an unconstrained system regression model that includes an ESG Volatility factor and seven other risk factors. The ESG Volatility factor is constructed by taking a long position in the quintile portfolio containing stocks with the lowest x-month ESG Volatility and taking a short position in the quintile portfolio containing stocks with the highest x-month ESG Volatility, where x = 18, 24, 30 and 36 months. The risk factors considered are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA as proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007). The table also performs the three F-tests of asset pricing restrictions concerning the role played by the ESG Volatility factor. We have italicised cases in which the rejection or non-rejection of a null conforms to the hypothesis of ESG Volatility serving as a factor

The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

alternative momentum estimation windows earn jointly significant risk-adjusted returns.

However, for most horizons the spread strategies alphas are signed in the same way, i.e., negative momentum portfolios yield higher returns than positive momentum ones. Additional, unreported checks reveal that also in Table 12, and now also in the case of Sustainalytics, the negative momentum firms are characterised by a higher starting standardised rating than high momentum firm are. Moreover, a combined assessment of Table 5 and Table 19 indicates that it is the rank neutralisation of the scores that acts on Sustainalytics to flip to negative the sign of the mean abnormal returns of the 5–1 quintile. Hence, the precise method through which one can neutralise the size and industry biases plays a key role in defining an ESG momentum factor.

While Table 3 used the rank-neutralised ESG scores to compute the ESG momentum, Table 12 shows the same results but with the standardised ESG scores. To save space,

Table 11 Robustness to alternative volatility estimation period lengths

	18-Month		24-Month		30-Month		36-Month	
	α	t-stat	α	t-stat	α	t-stat	α	t-stat
Panel A: Sustainalytics rating-based Volatility strategy								
1	0.297	1.380	0.436	1.876	0.407	1.378	0.371	1.198
2	0.427	1.614	0.377	1.771	**0.491**	2.076	**0.611**	2.336
3	0.405	1.452	0.379	1.666	0.420	1.821	0.441	1.511
4	0.402	1.873	0.425	1.742	0.098	0.461	0.307	0.951
5	0.087	0.406	0.226	0.693	0.282	1.215	0.340	1.021
1–5	**0.259**	3.062	**0.273**	3.099	0.034	0.450	0.103	0.805
GRS test	1.528		1.2815		2.0049		1.5492	
	(0.194)		(0.279)		(0.089)		(0.181)	
Panel B: MSCI rating-based Volatility strategy								
1	0.607	1.807	0.596	1.361	**0.456**	2.068	**0.429**	2.054
2	0.290	1.480	0.398	1.981	0.284	1.444	0.162	0.810
3	**0.450**	2.199	0.187	0.939	0.221	0.987	0.316	1.531
4	0.221	1.182	0.256	1.236	0.248	1.284	0.212	0.973
5	0.310	1.583	0.326	1.707	0.360	1.846	0.383	1.950
1–5	**0.297**	3.053	**0.271**	2.821	0.096	0.876	0.046	0.380
GRS test	1.144		0.633		0.424		0.651	
	(0.344)		(0.748)		(0.903)		(0.733)	

This table reports the results of a robustness analysis with respect to the choice of the ESG score volatility estimation period. It reports the estimated alphas for the ESG volatility portfolios sorted by 18-, 24-, 30-, and 36-month ESG volatility, respectively. The ESG volatility is computed using the standardised ESG scores. The table also reports the p-values (in parentheses) from Gibbons et al. (1989) test of the joint significance of the alphas for a range of volatility estimation periods. α is the estimated intercept in a regression of excess portfolio returns on a constant and seven risk factors. The risk factors are the five risk factors Mkt–Rf, SMB, HML, RMW, and CMA proposed by Fama and French (2015), the Carhart (1997) Momentum factor, and the Low Volatility factor proposed by Blitz and Van Vliet (2007). The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

Table 12 only reports the summary statistics for 6-month momentum.[34] We focus our attention on two differences in key empirical results relative to Table 3. First, the long-short portfolios no longer yield a significantly positive SR. This is a first indication that a simple ESG-momentum strategy based on standardised score may only earn fragile risk-adjusted performance, at least in the European equity market. Second, standardising the scores is effective in removing sector and size biases but it may be also troublesome in light of its policy implications. Under the ESG momentum estimated using standardised scores, portfolio 5 has a lower or identical mean and median raw ESG score than portfolio 1 does. This indicates that the long-short portfolio will not have a positive raw ESG score spread, albeit such a negative spread turns out to be small in absolute value.[35] Therefore, *per se*, building ESG momentum on standardised scores does not contribute to yield more sustainable or ESG-compliant portfolios. More generally, the table shows that when the portfolios are sorted by the ESG momentum with classical, standardised ESG scores, the raw ESG scores of the quintile portfolios are very similar in magnitude. This means that an increasing standardised ESG score no longer corresponds to higher raw ESG scores. Hence allocating capital to firms on the basis of standardised ESG momentum does not imply transferring capital from virtuous companies to sin ones.

Table 12 shows the same summary statistics as Table 7 when ESG volatility is computed using the standardised ESG scores. To save space, Table 12 only shows the summary statistics for stock portfolios sorted using 24-month ESG volatility.[36] We focus on an estimation window of 24 months because in Table 8, this is one of the two horizons for which the long-short strategies earns a significant alpha in the case of both Sustainalytics and MSCI. In Table 13, as

[34] We focus on 6-month ESG momentum because it strikes a balance between the size of the positive alpha obtained from MSCI data and obtaining non-negative alphas from the Sustainalytics ratings.

[35] Table 12 also shows a much larger dispersion in the momentum characterising the five quintiles, with the outer portfolio scores taking extreme values of +568% (+317% in the case of MSCI) and of -470% (-109% in the case of MSCI), both obtained from a 6-month momentum estimation window.

[36] The full set of results is in Table 20 of the Supplementary Information.

Table 12 Effects of an alternative method of industry and size adjustment on ESG momentum

	Returns				Raw ESG scores			
	Mean (%)	σ(%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Momentum
Panel A: Sustainalytics								
1	0.727	4.576	0.160	0.094	64.071	64.031	2.736	566.3%
2	0.654	4.333	0.151	0.095	62.682	63.075	2.549	16.6%
3	0.654	4.409	0.149	0.093	63.483	64.939	2.417	− 4.6%
4	0.653	4.495	0.143	0.094	63.314	64.408	2.787	− 27.5%
5	0.597	4.635	0.127	0.093	63.831	64.259	2.715	− 474.2%
5 -1	− 0.130	1.026	− 0.086	0.096				
Panel B: MSCI								
1	**0.802**	3.844	**0.198**	0.084	61.168	61.375	3.595	317.0%
2	**0.885**	3.984	**0.212**	0.084	64.140	64.686	2.325	8.6%
3	**0.675**	4.008	0.158	0.082	62.640	63.044	3.335	0.0%
4	**0.874**	3.882	**0.214**	0.083	60.764	60.171	2.614	− 0.9%
5	**0.665**	3.714	0.168	0.084	61.913	62.179	2.451	− 109.0%
5 -1	0.137	1.086	0.088	0.090				

The five quintiles are sorted by 6-month ESG momentum, the percentage increase of the ESG score over a period. The table also provides the summary statistics for a portfolio 5–1, which takes a long position in portfolio 5 and a short position in 1. Panel A reports the results for the returns of strategies driven by Sustainalytics rating momentum, while panel B for the case of MSCI ratings. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

in Table 9, we find the same lower mean and median ESG scores for portfolio 1 compared to portfolio 5, especially when Sustainalytics data are used. This result has implications from a sustainable investing point of view: investing in a spread ESG volatility factor may not lead to rewarding green stocks at the expense of the brown ones.

We also investigate the alphas of the portfolios sorted by ESG momentum and volatility computed using the standardised scores and under different estimation periods applied to the definition of ESG volatility. Tables 19 and 20 in the Supplementary Information show the estimated alphas and corresponding t-statistics. The results are much weaker than those in Table 5 and Table 11, especially as far as the MSCI ratings are concerned. None of the GRS tests comes close to reveal the alphas being jointly significant and the average long-short returns turn to be not statistically significant. This occurs everywhere in Table 19 and in Table 20 in the case of the MSCI ratings, in panel B. Nonetheless, especially for the longest horizons, the low ESG score volatility strategies based on Sustainalytics gain some traction and are estimated precisely, even though these are not larger than the average spreads in Table 11.[37] All in all, the evidence of significance of the long-short alphas reported earlier is therefore not completely robust to the data neutralisation method used.

Estimated ESG momentum and volatility alphas over time

Because the GRS tests applied to Table 5 (and reported in Table 18 in the Supplementary Information) have shown that the ESG (reverse) momentum computed using the rank-neutralised ESG scores earns jointly significant alpha, it is appropriate to study the behaviour of the estimated pricing errors over time. One reason for concern is that it would be possible for the majority of the risk-adjusted returns to be earned at the beginning of the sample and that the signal might then have faded as time progresses. Figure 5 in the Appendix shows the behaviour of the estimated pricing errors over the sample period, when the same seven factors are included in the risk factor model as in Table 5.[38] We use a rolling window of 24 months of data to estimate the alphas. Panels (a) and (c) (the former with reference

[37] Unreported results on the regressions underlying Table 19 in Supplementary Information reveal that the momentum-sorted portfolios carry, just as for momentum computed using the rank-neutralised ESG scores, a significant relationship with the market and the low volatility factor portfolios. The regressions underlying Table 20 disclose that for the quintile portfolios the estimated coefficients loading on the market and the classical volatility factor turn out to be significant. Both sets of results suggest some mixing between ESG score momentum and volatility strategies and classical momentum and volatility factors, which may be troublesome and indicates that standardising the scores has limitations.

[38] For clarity and to favor direct comparisons to panels (c) and (d), in panels (a) and (b) that refer to Sustainalytics ratings, we are plotting the negative of the 1 - 5 alpha spread, i.e., the 5 - 1 average abnormal return.

Table 13 Summary statistics of quintile portfolios sorted by 24-month ESG score volatility

	Returns				Raw ESG scores			
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Volatility
Panel A: Sustainalytics								
1	0.687	4.460	0.163	0.107	62.498	62.431	0.697	3.8%
2	0.637	4.655	0.137	0.100	64.999	64.756	0.442	5.7%
3	0.516	4.483	0.117	0.105	66.037	66.377	0.466	7.2%
4	0.607	4.654	0.131	0.105	66.830	66.678	0.476	9.5%
5	0.350	4.752	0.072	0.104	67.208	67.617	0.542	14.8%
1–5	**0.336**	1.207	**0.252**	0.088				
Panel B: MSCI								
1	**0.695**	3.853	0.145	0.088	63.587	64.322	3.018	7.4%
2	0.631	3.914	0.149	0.095	64.436	64.251	2.154	16.9%
3	0.585	3.882	0.139	0.099	63.677	63.719	2.483	25.5%
4	0.637	3.730	0.158	0.091	63.439	63.331	2.142	35.9%
5	0.571	3.902	0.157	0.095	62.134	62.466	2.484	59.4%
1–5	0.124	1.040	-0.100	0.100				

The five quintiles are sorted by 24-month ESG volatility. The table also provides the summary statistics for a portfolio 1- 5, which takes a long position in portfolio 1 and a short position in 5. Panel A reports the results for the returns of strategies driven by Sustainalytics rating volatility, while panel B for the case of MSCI ratings. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

to Sustainalytics and the latter to MSCI data) show that especially during the period 2013–2016, the dispersion in the alpha estimates was large. This translates into significant alpha estimates for the long-short portfolios. Over this period, the structural features of the strategies built on the ratings issued by the two alternative providers are different, something discussed earlier and featured, e.g., in Berg et al. (2022). However, after this period, the momentum alphas of the long leg and of the short leg obtained from Sustainalytics in panel (a) seem to converge and decline substantially to levels that deliver a negative but imprecisely estimated spread, whose sign is consistent with results for MSCI in panel (c). Moreover, Fig. 5 in Supplementary Information shows that over the entire sample, the estimates of the long-short portfolio's alpha turned out to be significantly negative for a very limited period (at the end of 2012) in the case of Sustainalytics, while the precision of the alpha estimates is never problematic in the case of MSCI.

Similarly, to investigate whether the alphas of the quintile portfolios are structural to the entire sample or whether ESG low volatility just recently became important, we analyse the dynamics of the pricing errors of the portfolios over time. As the long-short portfolio, derived from the spread in ESG score volatility using rank-neutralised ESG scores, demonstrated the most robust results in terms of t-statistics, we opt for a 24-month volatility estimation period. Figure 4 provides an insight into the evolving 24-month alpha estimates for both the quintile portfolios and the long-short portfolio. Following an initial period of instability (2012–2015), the estimates tend to converge towards a common value as we approach the end of the sample period. In the case of the Sustainalytics data, especially in the first half of the sample period, the alpha estimate for the individual quintile portfolios are highly significant, even though the long-short portfolio's estimate is not. This suggests that a structural break may have occurred in the performance of the model used for pricing the quintile portfolios. Yet, the estimated alpha for the long-short portfolio turns out to be very stable, explaining the highly significant full-sample estimate reported in Table 11. The inferred coefficients are positive over most of our sample and whenever the estimates turn out to be negative, the t-statistic is very small in absolute value. Even though the 1–5 spread alphas are also generally positive and statistically significant between 2011 and 2014, for MSCI data, the dynamics is a bit different as most individual quintile portfolios imply non-positive alphas but eventually the average abnormal returns are estimated to favour the low rating volatility portfolios over the high volatility ones. Yet, also in the case of MSCI data, there is some visual evidence in Fig. 4 of a change in the mean of the alpha after 2013.

Testing for structural breaks in the ESG momentum and volatility portfolios

Table 14 shows the results of the Andrews (1993) test of a structural break at an unknown break date for the five quintile portfolios and the long-short portfolio. The ESG momentum and volatility returns are computed using the rank-neutralised ESG scores using horizons of 1- and 24 month, respectively. The test uses the seven factors and excludes the first and last 15% of dates as possible break dates. The table shows that the null hypothesis of the absence of a structural break cannot be rejected for any of the portfolios and the raters investigated, since none of the test statistics exceeds the critical values. This means that the concern about a possible structural break can be put to rest, in spite of the considerable variation in estimated alphas shown in Figs. 4 and 5 in Supplementary Information. Therefore, there is no statistically significant evidence indicating that the asset pricing model incorrectly prices the assets because of structural instability.

ESG momentum and volatility tilting strategies

The long-short portfolios built in subsection "ESG Momentum" and subsection "ESG Volatility" have the disadvantage that, by ignoring the information in quintiles 2-4, they ignore 60% of the stocks and their ESG signals. In this sense, the long-short portfolio is formed with reference to a strongly, artificially bounded investment universe. Standard principles imply that this is unlikely to lead to optimised performances. In line with the earlier literature on ESG momentum (see, e.g. Nagy et al. (2016)), this suggests testing the robustness of the ESG momentum signal also using lighter, tilting-based strategies.

To this end, we shall use a straightforward tilting strategy which consists of computing the ESG momentum and volatility for every stock in our (common) European equity universe and ranking them by the resulting estimates. To guarantee consistence with subsection "ESG Momentum", the stocks are ranked in descending (ascending) order, such that the stocks with the lowest ESG momentum (volatility) receive the highest ordering and vice versa for the stocks with the highest ESG momentum (volatility). Next, all the ranks are summed and every stock's rank is divided by the sum. This way, we ensure that the stocks with the lowest ESG momentum (volatility) are over-weighted and the stocks with the highest ESG momentum (volatility) are under-weighted, relative to the equally-weighted portfolio.

Table 21 in the Supplementary Information shows two important results. First, all average abnormal returns are insignificant for both ESG raters. In fact, none of the strategies manage to generate a significant SR and all fail to earn significant, risk-adjusted returns, which also happen to be economically rather small (or negative) in the case of MSCI. Second, the 12-month ESG momentum now implies the highest positive mean return when MSCI ratings are used, the highest SR, and the highest ESG score means and medians. This result differs from our earlier finding that 1-month momentum would guarantee the highest realised performances.[39] Finally, the mean and median ESG scores generally increase when the momentum computation period increases. These results are robust to applying a tilting strategy based on the standardised ESG scores, even though in that cases the 1-month ESG momentum tilting strategy displays a significant SR.[40]

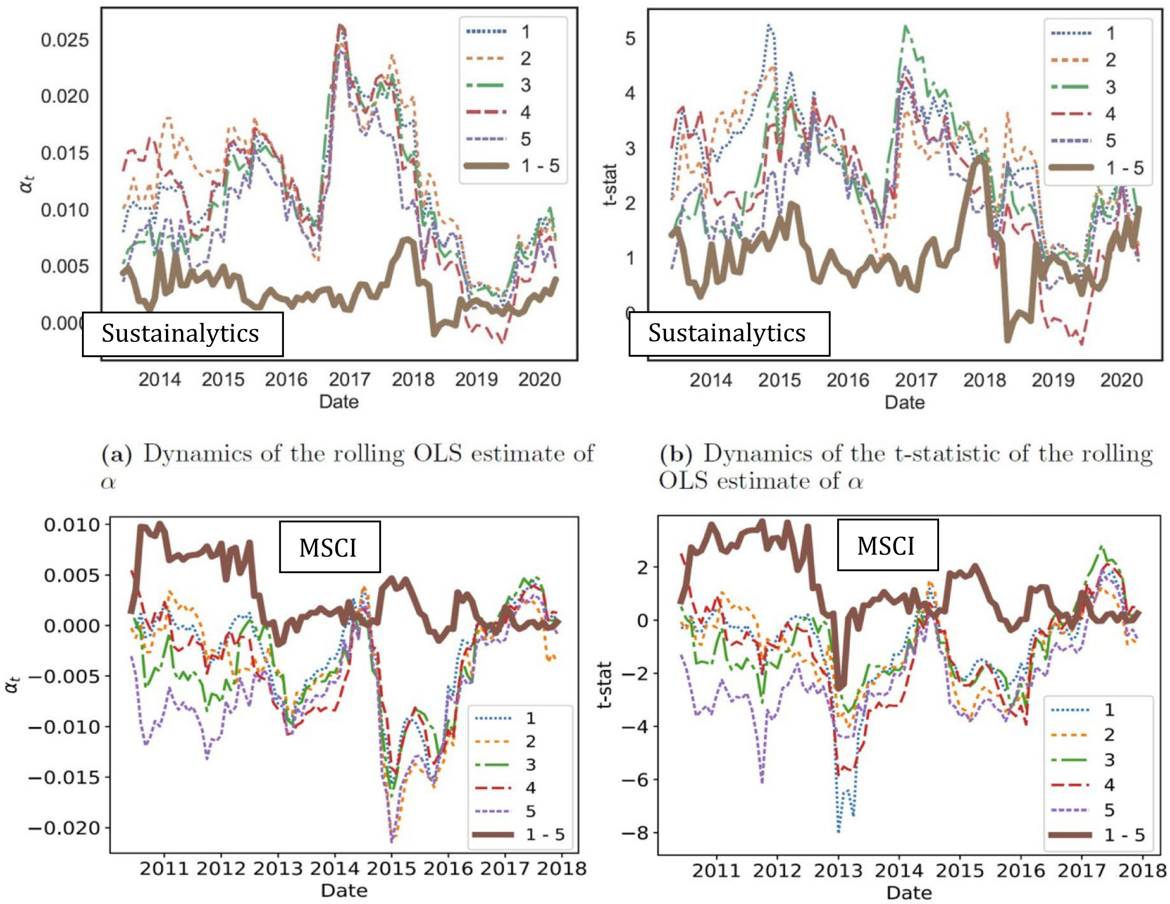
Similarly, we apply the tilting strategy to the ESG volatility signal. Table 22 in the Supplementary Information displays summary statistics for strategy returns for different computation periods. The table shows that none of the tilting strategies earns a statistically significant alpha. Yet, this finding is not totally unexpected after the earlier results on the GRS tests in Table 12. In the top panel concerning Sustainalytics, the 24-month ESG volatility strategy fails to provide the highest SR. Both the 30- and 36-month ESG volatility tilting strategies yield a higher SR, higher alphas, and higher ESG scores. This observation is promising within the context of sustainable investing. Specifically, when considering an ESG volatility tilting strategy, the conventional trade-off between mean realised returns and un-adjusted ESG scores does not materialise. The longer the ESG volatility estimation period, the higher the SR and the higher the mean and median ESG score, despite the absence of statistical significance in the estimated SR and alpha. It is worth noting that this is in contrast with the outcomes for the quintile portfolios. However, Table 9 displays significant average abnormal returns for the long-short portfolio, whereas this is not the case for the tilting strategy. It therefore seems that the tilting strategy is superior to the long-short strategy from a sustainable investing perspective, even though the sustainable investing implied by the tilting strategy dilutes returns. Nonetheless, in the bottom panel, concerning the MSCI-driven strategies, realised mean returns are comparable to those in Section "Main results" and yet both the resulting Sharpe ratios and alphas are either not precisely estimated (the former) or even negative (but not significant, in the latter case). Therefore, any evidence in favour of the effectiveness of low ESG volatility strategies is limited to Sustainalytics data.

A larger universe of Firms for MSCI ratings

As discussed in Section "Data", to increase comparability, we have performed the earlier analyses on the intersection of the firm universes for which we have both Sustainalytics and

[39] Nonetheless, no choice of the length of the estimation period delivers significant alphas. Also Nagy et al. (2016) used the 12-month ESG momentum to build their tilting strategy, which indeed seems to be the superior one among their alternative strategies.

[40] Complete, tabulated estimated are available from the Authors upon request.



(a) Dynamics of the rolling OLS estimate of α

(b) Dynamics of the t-statistic of the rolling OLS estimate of α

(c) Dynamics of the rolling OLS estimate of α

(d) Dynamics of the t-statistic of the rolling OLS estimate of α

Fig. 4 shows the dynamics of the OLS estimates of α and the dynamics of its corresponding t-statistics. The estimates are computed using a rolling window of 24 months of data. The portfolios are sorted by the 24-month ESG volatility. The risk factors incorporated in the OLS regressions are the five factors Mkt–Rf, SMB, HML, RMW, and CMA proposed by Fama and French (2015), the Momentum factor proposed by Carhart (1997), and the Low Volatility factor proposed by Blitz and Van Vliet (2007). Panels (**a**) and (**b**) concern Sustainalytics data, while (**c**) and (**d**) refer to MSCI ratings

MSCI ratings data. In practice, this translates in a substantial sacrifice of the number of firms investigated, as this declines from a potential (maximum) number of 1301 to a (maximum) of 882 firms used earlier.[41] Therefore, as an additional robustness check, we have performed afresh the key tests in Section "Main results" applied to the full universe of MSCI firms. Notably, in this case, in spite of the application of the same filters used before, also 47 Greek and Icelandic firms are included in the analysis. To save space, we focus on 1-month momentum and 18-month low volatility strategies which have also been our main focus earlier. To allow comparisons with earlier results, also in these analyses we focus on a July 2007–December 2019 sample.

Table 15 shows that the results concerning MSCI reported in Section "Main results" get stronger if one allows an even wider universe of firms, as selected by MSCI with reference to the European area. First, while the 5-1 spread portfolio findings for 1-month momentum remain rather strong (the statistically significant mean average spread returns grows from 2.48% to 2.98% per month and the SR goes from 1.78 to 1.54 when going from Tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15, the finding that negative momentum portfolios yield higher realised returns than high momentum ones remains puzzling and likely explained by the structurally higher mean and median (raw and rank-neutralised) ratings that characterised the former quintiles vs. the latter. In other words, also on a wider sample of firms, it remains true that

[41] The corresponding average numbers of firms are 751 in the case of MSCI ratings and 711 in the case of Sustainalytics.

Table 14 Andrew's test of a structural break at unknown date for ESG momentum and volatility

	1-month Momentum		24-month Volatility				
	Sustainalytics	MSCI	Sustainalytics	MSCI			
	Andrews test statistic				Statistic critical values		
					10%	5%	1%
1	4.610	3.628	6.026	3.851	19.82	22.13	27.25
2	6.521	2.448	4.741	5.543	19.82	22.13	27.25
3	7.917	4.330	5.318	3.013	19.82	22.13	27.25
4	5.582	4.906	5.802	3.490	19.82	22.13	27.25
5	5.869	3.148	6.207	4.381	19.82	22.13	27.25
5-1/1-5	1.728	3.680	2.303	4.023	19.82	22.13	27.25

This table shows the Andrews (1993) test for a structural break at an unknown break date for the regression coefficients of five quintile portfolios sorted by ascending 1-month ESG momentum or descending 24-month volatility. The null hypothesis is the absence of a structural break. Additionally, the table provides the Andrews (1993) test statistic for long- short portfolios (5–1 in the case of momentum and 1–5 in the case of volatility). The test statistic is the supremum of the Chow break test statistics over a fraction of the sample period. The test statistic follows a non-standard distribution, which depends on the number of regressors in the model, and the relative size of the fraction of the sample period. The regressors are the constant, the five Fama and French (2015) factors, the Carhart (1997) Momentum factor, and the Blitz and Van Vliet (2007) Low Volatility factor. The supremum of the Chow break test statistics is computed over the period January 2011–August 2018, so that the first and last 15% of dates are not considered as structural break candidates. The critical valueThe estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

s of the associated non-standard distribution are reported for the 10%, 5% and 1% size test levels

in the case of MSCI ratings, the average (and median) level of the ESG score prevails and makes momentum able to express higher performance for stocks with high but stable or declining score vs. the ones with an initially low but increasing score. As a point in case, Table 23 in the Supplementary Information shows that also for the MSCI extended universe of firms, the spread return between the high ESG (both raw and rank-neutralised) score portfolio 1 and the low ESG score portfolio 5 is relatively large (0.30% per month) and statistically significant.

Yet, as in Section "Main results", the corresponding SR (0.176) fails to be significantly positive and the estimated 7-factor model alpha (0.15% per month) is economically small and statistically insignificant. In fact, in Table 23, the abnormal returns derived from a seven-factor linear model lack any linear predictability structure. Once more these result show that while in the case of MSCI rating a simple rating-level strategy fails to be viable as an additional risk factor, the prominence of the level unravels the viability of the momentum strategy. Visibly, in Table 23, 1-month rank-neutralised score momentum fails to differ substantially across portfolios with different mean and median score levels.

Second, Table 15 shows that the ESG volatility-driven strategy gets unconditionally stronger. The average spread return from the 1–5 long-short portfolio increases from 0.28% per month in Table 7 to 0.39% per month; the corresponding SR from 0.24 to 0.26 and these are both statistically significant. The estimated linear factor model alpha stays roughly constant (from 0.30 to 0.26% and both are precisely estimated). Especially this second set of results emphasises that, at least as far as MSCI data are concerned, expanding the sample to include all of the available data would not affect our results and, at least to some extent, actually strengthen them.

Covid and post-Covid MSCI data

A natural question is whether our results on ESG momentum and volatility strategies may be resilient with reference to the 2020–2021 Covid pandemics-dominated sample and also to the inclusion of additional 2022 and (inasmuch as possible) 2023 ratings data.[42] The exact sample period is January 2000–March 2023. Because of data availability limitations, we could extend the analysis only to MSCI ratings and hence we do that with reference to the same universe as in subsection "A larger universe of Firms for MSCI ratings". To achieve consistency with subsection "A larger universe of Firms for MSCI ratings", Table 16 reports on the performance of 1-month momentum and

[42] See Arat et al. (2023) and Rizvi et al. (2020) for analyses of the impact of the pandemics on European financial markets and the asset management industry.

Table 15 Performance of ESG momentum and volatility strategies for the full MSCI universe

	Returns				Raw ESG scores				Neutralised ESG scores		
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	1-month Mom.	Mean	Median	σ
1-month Momentum strategy											
1	**-1.037**	5.338	-0.205	0.111	56.141	55.879	2.291	36.5%	0.465	0.463	0.028
2	0.429	4.844	0.077	0.093	58.294	59.133	2.999	0.0%	0.489	0.496	0.038
3	0.288	4.524	0.051	0.089	61.580	61.230	2.772	0.0%	0.510	0.508	0.035
4	0.332	4.479	0.061	0.089	61.239	61.217	1.943	0.0%	0.513	0.510	0.023
5	**1.940**	4.846	**0.388**	0.109	61.521	61.656	2.439	-6.9%	0.540	0.543	0.028
5 -1	**2.977**	1.970	**1.541**	0.286							
18-month Volatility strategy											
1	**1.096**	4.106	**0.209**	0.090	59.699	59.135	4.278	0.8%	0.522	0.525	0.037
2	**0.868**	4.292	**0.193**	0.097	59.631	59.193	2.255	3.0%	0.522	0.523	0.022
3	**0.872**	4.560	0.183	0.099	61.949	61.933	1.848	5.5%	0.518	0.517	0.025
4	0.687	4.159	0.156	0.089	62.754	62.869	1.654	8.8%	0.494	0.491	0.024
5	0.707	4.514	0.192	0.096	62.228	62.623	2.405	16.6%	0.509	0.507	0.051
1 -5	**0.389**	1.356	**0.259**	0.090							

In the top panel, the five quintiles are sorted by ESG momentum, the percentage increase of the rank-neutralised ESG score over a 1-month period. The table also provides the summary statistics for a portfolio 5–1, which takes a long position in portfolio 5 and a short position in 1. In the bottom panel, the five quintiles are sorted by ESG volatility, the percentage increase of the rank-neutralised ESG score over an 18-month period. The table also provides the summary statistics for the long-short portfolio. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

18-month volatility strategies. It shows no substantial change in strength and significance of the ESG-driven strategy performances starting in 2020. In practice, the long-short abnormal return of the 5–1 momentum strategy turns out to increase somewhat in correspondence to the pandemics, to generate a rather massive SR in excess of 1 and an alpha that exceeds the one in Table 14. In the bottom panel of Table 16, the ESG volatility strategy also displays some resilience to an extension to a recent sample. On the one hand, the quintiles portfolio returns and the resulting alpha are statistically insignificant; on the other hand, the SR increases vs. subsection "A larger universe of Firms for MSCI ratings" (from 0.26 to 0.28 per month) and it is statistically significant. However, the resulting alpha (complete results are available upon request) that nets out the effect of linear factor exposures turns out to be positive but not significant.

Leaving a one-month gap between ESG momentum measurement and portfolio formation

In spite the evidence of non-negligible ESG rating activity on a monthly basis in the cross-section of European stocks under investigation discussed in Section "How often do ESG ratings change and why?", a sensible concern is that, irrespective of what we may see ex-post in the Sustainalytics and MSCI data set, over a short horizon, the information contained in the changes in the ESG ratings (hence, also in their volatility) may not be easily actionable in terms of trading stocks to build the portfolios tracked in Section "Main results". As a result, we decided to propose results for momentum and volatility strategies based not only on short windows of 1 and 18 months, but also of 3, 6 and 12 months in the case of momentum (and 24, 30 and 36 in the case of volatility). Such medium-long windows are supposed to neutralise the fact that there may be long and variable lags in the reporting of ratings and of their changes. In addition, we also perform a robustness check and test the performance of a "t – 2-gap horizon" momentum strategy, in the sense that at time t, the long-short trades that our set up implies are implemented with reference to ESG momentum sorting statistics based not on momentum measured over month t – 1, but instead on month t – 2. This allows one additional "gap" month for the ESG score changes and/or initiation to be made public and become actionable in terms of trades and portfolio formation decisions.

Table 17 reports the results from the 1-month gap momentum strategy (panel (a)) with those of the standard 1-month already reported in Table 3 (here reproduced for clarity in panel (b)). The two sets of performance results are qualitatively similar and simply characterised by some decay in mean portfolio performances, that however hardly extends to the Sharpe ratios. In the case of Sustainalytics, the long-short portfolio that buys high-minus-low ESG momentum keeps reporting a statistically significant spread of 0.27% per month, which implies a rather large and significant Sharpe

Table 16 Performance of the ESG momentum and volatility strategies in 2020–2023

	Returns				Raw ESG scores				Neutralised ESG scores		
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	1-m Mom.	Mean	Median	σ
1-month Momentum strategy											
1	-1.180	6.643	-0.193	0.138	58.934	59.827	3.859	24.5%	0.449	0.459	0.053
2	0.462	5.027	0.072	0.109	61.965	61.810	2.375	0.1%	0.501	0.503	0.039
3	0.590	5.115	0.095	0.108	63.875	64.490	3.358	0.0%	0.506	0.518	0.052
4	0.560	5.048	0.091	0.108	64.833	65.749	3.160	-0.1%	0.509	0.522	0.043
5	**2.630**	6.281	**0.402**	0.133	66.451	66.066	3.434	-6.3%	0.559	0.554	0.049
5 -1	**3.810**	2.191	**1.785**	0.495							
18-month Volatility strategy											
1	0.819	5.729	0.126	0.124	64.086	65.515	5.091	0.7%	0.525	0.539	0.068
2	0.871	6.119	0.127	0.124	61.865	62.604	3.368	2.8%	0.485	0.491	0.047
3	0.843	6.325	0.118	0.120	64.858	65.009	2.222	4.9%	0.515	0.513	0.031
4	0.839	6.193	0.120	0.117	66.656	66.211	2.001	7.8%	0.536	0.527	0.033
5	0.606	6.435	0.079	0.123	69.810	69.692	2.014	14.2%	0.595	0.582	0.037
1 -5	0.213	1.137	**0.284**	0.112							

In the top panel, the five quintiles are sorted by ESG momentum, the percentage increase of the rank-neutralised ESG score over a 1-month period. The table also provides the summary statistics for a portfolio 5–1, which takes a long position in portfolio 5 and a short position in 1. In the bottom panel, the five quintiles are sorted by ESG volatility, the percentage increase of the rank-neutralised ESG score over an 18-month period. The table also provides the summary statistics for the appropriate long-short portfolio. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font

ratio of 0.23. Even though, the mean spread return is visibly lower than the 0.43% reported in Table 3, the Sharpe ratios are similar (in fact, under the assumption of independent portfolio returns, the two Sharpe ratios are not significantly different). In the case of MSCI ratings, it is again the case that it is a strategy that shorts high momentum and goes long in negative momentum stocks that yields a significant mean return spread; yet, both the mean spread and the corresponding Sharpe ratios are lower, which is to be expected because in unreported results we have verified that the inverse relationship between realised ESG momentum and the initial average level of the ESG scores weakens somewhat.

Table 24 in the Supplementary Information show that also the linear factor regressions previously applied in Table 5 deliver results that are qualitatively similar to those reported in Section "ESG Momentum". In the case of Sustainalytics, even though the p-value of the of the returns on the long-short momentum portfolio is exactly 0.05, such an alpha is negative and there is at best faint evidence that the standard seven factors entertained throughout can explain the ESG momentum spread returns. The corresponding R-square is only 26%, only slightly higher than the 24% reported in Table 5. At the same time, the same factors do explain quite accurately the excess returns on the remaining ESG momentum-sorted portfolios, they imply rather high R-square and leave no evidence of mean abnormal returns behind. In the case of the MSCI-based strategy, even though the excess returns of portfolio 1 (the one characterised by the highest momentum) are on average larger (positive) than in Table 5, the alpha of the 5–1 spread portfolio remains positive and statistically significant and (with an R-square of only 14%) characterised by an alpha of 0.47% per month that implies a p-value of essentially zero. Also in this case it is hard to detect major differences between the second panel of Tables 24 and Table 5 which makes us conclude that our assumptions are essentially robust to allowing a 1-month period to ESG rating flows to become available to stock market participants.

Discussion and conclusions

In this paper, we have performed a systematic investigation of equity trading strategies based on ESG signals, with particular reference to the European equity space. We report that stable, sustainable investing company profiles, as captured by ESG ratings issued by MSCI and Sustainalytics, generate trading signals that are profitable and yield positive risk-adjusted returns. Interestingly, provided that the ESG ratings are stable over time, it is possible that also firms with weak but negative momentum in their ESG scores may benefit from a comparatively lower cost of equity capital.

We propose two ways to neutralise the size and sector biases in ESG scores. The first method, new to the best of our knowledge, groups all stocks in size quintiles. We then compute the median ESG score within each quintile and from the raw values we subtract this median score. Next, we compute the quintile to which every stock belongs, giving every stock a value between 0 and 1. Finally, we apply a similar procedure to size-neutralise the resulting ESG scores

Table 17 Results for portfolio sorted by ESG momentum using a 1-month (t – 2)-gap horizon scheme

	Sustainalytics								MSCI							
	Returns				Sustainalytics Raw ESG Scores				Returns				MSCI Raw ESG Scores			
	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Momentum	Mean (%)	σ (%)	Sharpe Ratio	Std Err (SR)	Mean	Median	σ	Momentum (%)
Panel A: 1-month ESG Momentum with a 1-month gap in implementation																
1	0.348	4.006	0.079	0.083	57.801	61.147	0.907	46.5%	0.012	3.694	−0.005	0.083	62.604	62.838	3.088	78.4
2	0.518	4.133	0.118	0.081	62.636	66.940	1.007	4.2%	0.761	3.920	**0.186**	0.081	64.917	65.178	2.119	0.0
3	0.593	4.626	0.121	0.086	71.107	66.694	0.981	−0.1%	0.751	3.931	**0.183**	0.081	63.282	62.810	2.421	0.0
4	0.614	4.938	0.118	0.087	64.702	67.798	0.942	−3.9%	**0.879**	4.191	**0.202**	0.080	60.757	61.630	3.187	0.0
5	0.080	4.738	0.010	0.085	56.488	58.664	0.916	−21.8%	**1.052**	3.939	**0.259**	0.081	59.137	58.691	2.279	−6.8
5 -1	**-0.268**	1.307	**−0.230**	0.054					**1.040**	1.289	**0.782**	0.074				
Panel B: 1-month ESG Momentum (as in Table 3)																
1	**0.851**	4.377	**0.195**	0.088	57.920	58.260	0.927	46.1%	−0.208	4.037	−0.047	0.089	59.050	58.750	2.290	77.0
2	0.695	4.397	0.158	0.087	65.440	66.160	1.002	4.1%	**0.932**	4.171	**0.228**	0.086	60.659	61.551	3.162	0.0
3	0.667	4.662	0.143	0.087	69.150	69.820	0.977	−0.1%	**0.750**	3.961	**0.194**	0.082	63.192	62.813	2.407	0.0
4	0.699	4.630	0.151	0.092	64.990	65.670	0.911	−4.1%	**0.796**	3.930	**0.207**	0.084	64.896	64.987	2.141	0.0
5	0.426	4.596	0.093	0.094	57.340	57.730	0.946	−21.0%	**2.270**	3.821	**0.599**	0.089	62.607	62.711	3.045	−6.9
5 -1	**-0.425**	1.543	**−0.248**	0.113					**2.479**	1.414	**1.782**	0.157				

The five quintiles are sorted by ESG momentum, the percentage increase of the ESG score over a period. The table also provides the summary statistics for a portfolio 5–1, which takes a long position in portfolio 5 and a short position in 1. The estimated coefficients significant at a test size of 5% or lower are emphasized using bold face font from the first step, but this time using sectors to group the stocks. We refer to this method as rank-neutralised ESG scores. The second method consists of first grouping all the stocks in five size quintiles and then computing the average ESG score and the standard deviation of the ESG score within every quintile. The stocks are then standardised using the mean and standard deviation corresponding to the size quintile that the stock belongs to.

After correcting for size and sector dependencies in the data, we analyse two ESG score-linked strategies. The first strategy is based on ESG momentum, the percentage change in the ESG score of a stock over a given time interval. We sort quintile portfolios by their ESG momentum with reference to a range of computation periods. We estimate cross-sectional regressions to investigate the average effect of ESG momentum on portfolio excess returns. Next, we analyse the abnormal returns of portfolios sorted by decreasing ESG momentum using a standard linear factor model that includes the five Fama and French (2015) factors, Carhart (1997)'s momentum, and Blitz and Van Vliet (2007)'s low volatility factor. The portfolios sorted by 1-month ESG momentum earn significantly positive alpha, implying that 1-month ESG momentum can be interpreted as a new source of risk. Moreover, the ESG momentum factor, which is proxied by a long-short portfolio, generally fails to correlate with the classical equity factors.

However, at longer estimation horizons, ESG momentum appears to be fragile because it does not always yield significant alphas. Moreover, especially with reference to MSCI data, there is evidence that the initial average level of the ESG scores may "prevail" over momentum, leading to the result that a positive alpha could be estimated for a portfolio long in low momentum and short in high ESG momentum stocks. Therefore, we extend our analysis to a similar but distinct ESG volatility signal. We define ESG volatility as the sample standard deviation of the ESG score of a stock over a period of time. We sort portfolios by increasing ESG volatility over four different estimation periods. Cross-sectional regressions reveal a significantly negative average relationship between ESG volatility and excess portfolio returns. We explore the extent and estimation precision of the implied mean abnormal returns for the ESG volatility-sorted portfolios within the same seven-factor model examined previously. The long-short portfolio earns statistically significant mean excess returns even though Gibbons et al. (1989)'s test shows that the quintile portfolios do not earn jointly significant risk-adjusted mean returns. Structural break tests fail to reject the null hypothesis of the absence of a break-point date in the estimated alphas. These results are robust to the selection of the ESG volatility estimation periods, to using alternative sample periods and apply with equal strength to both Sustainalytics and MSCI ratings. We also find that the method by which the ESG data is neutralised from biases, significantly influences the returns earned by portfolios sorted by the ESG signals.[43]

Our findings carry a few interesting policy implications. On the one hand, by resorting to the Campbell and Shiller's stock return decomposition (see Campbell and Shiller (1988)),

$$r_{t+1} - E_t\left[r_{t+1}\right] = \alpha + r_{t+1}^{k-factor} + \epsilon_{t+1} - E_t\left[r_{t+1}^{k-factor}\right]$$
$$= \left(\alpha + \epsilon_{t+1}\right) + r_{t+1}^{k-factor} - E_t\left[r_{t+1}^{k-factor}\right] = \sum_{s=0}^{\infty} \rho^s (E_{t+1}\left[\Delta d_{t+1+s}\right]$$
$$- E_t\left[\Delta d_{t+1+s}\right]) - \sum_{s=1}^{\infty} \rho^s (E_{t+1}\left[r_{t+1+s}^{k-factor}\right] - E_t\left[r_{t+1+s}^{k-factor}\right])$$

where $\rho = 1/[1 + \exp\left(d \bar{-} p\right)]$, ($d \bar{-} p$ is the average log dividend-price ratio) and $r_{t+1+s}^{k-factor}$ is the return implied by some multi-factor model, we can see that a current high alpha can only derive from either future positive upward revisions in the expected rate of growth in cash flows (here represented by the natural log of dividends) or from future downward revisions in expected stock returns. In light of the recent literature (see, e.g. El Ghoul et al. (2018) and Stotz (2022)) that has excluded that high realised returns on ESG stocks may derive from higher future expected profitability, current positive alphas forecast lower equity cost of capital. Hence, our findings of positive and statistically significant alphas from a number of ESG score momentum- and volatility-driven systematic strategies may be consistent with a declining cost of capital for firms characterised by ESG momentum (at least, when the scores given by specific raters are employed) and, consistently, low ESG score volatility. However, our very analysis shows that because it is well known that the ESG ratings of different providers rate the same companies differently (see, e.g., Berg et al. 2022), the precise connection between a company's ESG strength and the equity cost of capital it faces, should always be mediated by the quality or in fact, the very deep meaning that different raters assign to their scores. Our findings also supplement the growing empirical evidence that while it may no longer be so obvious that high ESG scores would be associated with a lower expected cost of capital and hence higher, current realised stock returns (as originally reported by, e.g., Pástor et al. (2022) and as also confirmed by our results in Table 4), see e.g., the aggregate evidence in Friede et al. (2015), hard-to-ignore but never trivial connections exist between sustainability ratings and firms' financial decisions.[44] From a policy

[43] For instance, portfolios sorted by 1-month ESG momentum computed with the rank-neutralised ESG scores earn jointly significant alpha. However, these significantly positive mean abnormal returns are not earned by portfolios sorted by ESG momentum if momentum is computed using the classical, standardised ESG scores.

[44] Such a relationship may result from investors' readiness to pay more for these firms, either because these are less risky in the long- run or as an expression of preferences for long-run non-pecuniary benefits, as in Fama and French (2007). As discussed by Stotz

perspective, fostering a virtuous cycle of stable and reliable ESG scores may thus yield the social payoffs deriving from a future, expected low of cost of equity capital. Moreover, at least the results obtained in the case of Sustainalytics, reveal that the European capital markets may have embraced a hyper-efficient assessment of ESG quality in which high momentum and improving ESG (size and industry adjusted) scores lead to a lower cost of equity capital and hence free up additional resources for further ESG investments. This could be beneficial in light of a looming climate crisis.[45] That notwithstanding, policymakers may also derive a less optimistic perspective from our findings. If firms were to realise that immediate benefits (in terms of higher stock prices per unit of fundamentals) are available if any progress in their ESG score is kept muted and occurs gradually, this may quickly deprive the ESG scores of their very signalling nature, as firms may learn how to "game" the system. Barring the possibility that stringent regulations concerning ESG scoring by the relevant, specialised agencies may be introduced soon, this represents a looming concern for investors and the market regulatory authorities (see the discussion in Redondo Alamillos and de Mariz (2022)). Finally, the very hyper-efficient assessment of ESG quality discussed earlier may hide the seeds of a dangerous ESG-rating driven bubble that policy-makers are likely to deem undesirable for its fall-out in terms of system risks. We leave this intriguing possibility to future research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1057/s41260-024-00377-w.

Acknowledgement Massimo Guidolin acknowledges the financial support of Ministero dell'Università e della Ricerca PRIN 2020 PROT. 2020B2AKFW ("Fin4Green–Finance for a Sustainable, Green and Resilient Society"). We thank two anonymous referees for constructive feedbacks and Martijn Boermans, Dirk Broeders, Philipp Krueger, Brent Lindquist, Svetlozar Rachev, Yasmine van der Straten for valuable suggestions. We also thank participants at the workshop "Climate Change, Local Development and Financial Markets" held at Bocconi University in April 2023, at the "2nd Conference on Sustainable Banking & Finance CSBF" held at Parthenope University in June 2024 and at the "International Risk Management Conference" held at Bocconi University, June 2024. Last but not least we also thank seminar participants at De Nederlandsche Bank and Texas Tech University for useful comments.

Funding Open access funding provided by Università Commerciale Luigi Bocconi within the CRUI-CARE Agreement.

Footnote 44 (continued)

(2022), in efficient capital markets, return realisations should equal their expectations in the long-run; yet, over limited sample periods, a deviation of realised returns from expected returns (i.e. an unexpected return) can be explained by unexpected news or by a variation in tastes and preferences that may favor ESG-driven characteristics.

[45] We thank an anonymous referee for suggesting this perspective on our momentum findings.

References

Adler, T., and M. Kritzman. 2008. The cost of socially responsible investing. *Journal of Portfolio Management* 35 (1): 52.

Alliance for Global Sustainable Investment (2018). *Sustainable Investment Review*.

Andrews, D.W. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* 61 (4): 821–856.

Arat, Emre, Britta Hachenberg, Florian Kiesel, and Dirk Schiereck. 2023. Greenium, credit rating, and the Covid-19 pandemic. *Journal of Asset Management* 24 (7): 547–557. https://doi.org/10.1057/s41260-023-00320-5.

Artiach, T., D. Lee, D. Nelson, and J. Walker. 2010. The determinants of corporate sustainability performance. *Accounting and Finance* 50 (1): 31–51.

Bannier, Christina E., Yannik Bofinger, and Björn. Rock. 2023. The risk-return tradeoff: are sustainable investors compensated adequately?. *Journal of Asset Management* 24 (3): 165–172. https://doi.org/10.1057/s41260-023-00303-6.

Berg, F., K. Fabisik, and Z. Sautner 2020. Is history repeating itself? the (un) predictable past of ESG ratings. European Corporate Governance Institute, Finance Working Paper No. 708.

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

Berk, I., M. Guidolin, and M. Magnani. 2023. New ESG rating drivers in the cross-section of European stock returns. *Journal of Financial Research*, 46, S133-S162.

Blitz, D., and F.J. Fabozzi. 2017. Sin stocks revisited: Resolving the sin stock anomaly. *Journal of Portfolio Management* 44 (1): 105.

Blitz, D.C., and P. Van Vliet. 2007. The volatility effect. *Journal of Portfolio Management* 34 (1): 102–113.

Bolton, P., and M. Kacperczyk. 2021. Do investors care about carbon risk?. *Journal of Financial Economics* 142 (2): 517–549.

Bos, J. 2017. The materiality of ESG factors for equity investment decisions: Academic evidence. Technical report, Nationale Nederlanden Investment Partners.

Bruno, Giovanni, Mikheil Esakia, and Felix Goltz. 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.

Campbell, J.Y., and R.J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1 (3): 195–228.

Cantor, R., and C. Mann. 2007. Analyzing the tradeoff between ratings accuracy and stability. *Journal of Fixed Income* 16 (4): 60–67.

Carhart, M.M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52 (1): 57–82.

Dor, A.B., J. Guan, and Y. Sun. 2022. Is incorporating ESG considerations costly?. *Journal of Portfolio Management* 48 (7): 75–87.

El Ghoul, S., O. Guedhami, H. Kim, and K. Park. 2018. Corporate environmental responsibility and the cost of capital: International evidence. *Journal of Business Ethics* 149: 335–361.

Fama, E.F., and K.R. French. 2007. Disagreement, tastes, and asset prices. *Journal of Financial Economics* 83 (3): 667–689.

Fama, E.F., and K.R. French. 2015. Dissecting anomalies with a five-factor model. *Review of Financial Studies* 29 (1): 69–103.

Fama, E.F., and J.D. MacBeth. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81 (3): 607–636.

Friede, G., T. Busch, and A. Bassen. 2015. ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment* 5 (4): 210–233.

Gibbons, M.R., S.A. Ross, and J. Shanken. 1989. A test of the efficiency of a given portfolio. *Econometrica* 57 (5): 1121–1152.

Gibson, B.R., P. Krueger, and P.S. Schmidt. 2021. ESG rating disagreement and stock returns. *Financial Analysts Journal* 77 (4): 104–127.

Giese, G., Z. Nagy, and L.-E. Lee. 2021. Deconstructing ESG ratings performance: Risk and return for e, s, and g by time horizon, sector, and weighting. *Journal of Portfolio Management* 47 (3): 94–111.

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.

Hong, H., and M. Kacperczyk. 2009. The price of sin: The effects of social norms on markets. *Journal of Financial Economics* 93 (1): 15–36.

Idowu, S. O., S. Vertigans, and A. Burlea Schiopoiu. 2017. Corporate Social Responsibility in Times of Crisis. Springer: Cham

Israel, R., and T.J. Moskowitz. 2013. The role of shorting, firm size, and time on market anomalies. *Journal of Financial Economics* 108 (2): 275–301.

Jegadeesh, N., and S. Titman. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance* 48 (1): 65–91.

Kaiser, L. 2020. ESG integration: Value, growth and momentum. *Journal of Asset Management* 21 (1): 32–51.

Lo, A.W. 2002. The statistics of Sharpe Ratios. *Financial Analysts Journal* 58 (4): 36–52.

Martellini, L., and L.-S. Vallee. 2021. Measuring and managing ESG risks in sovereign bond portfolios and implications for sovereign debt investing. *Journal of Portfolio Management* 47 (9): 198–223.

Nagy, Z., A. Kassam, and L.-E. Lee. 2016. Can ESG add alpha? An analysis of ESG tilt and momentum strategies. *Journal of Investing* 25 (2): 113–124.

Newey, W.K., and K.D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3): 703–708.

Pastor, L., R.F. Stambaugh, and L.A. Taylor. 2022. Dissecting green returns. *Journal of Financial Economics* 146 (2): 403–424.

Redondo Alamillos, R., and F. de Mariz. 2022. How can European regulation on ESG impact business globally?. *Journal of Risk and Financial Management* 15 (7): 291.

Rizvi, S.K.A., N. Mirza, B. Naqvi, and B. Rahat. 2020. Covid-19 and asset management in EU: A preliminary assessment of performance and investment styles. *Journal of Asset Management* 21: 281–291.

Sagbakken, S.T., and D. Zhang. 2022. European sin stocks. *Journal of Asset Management* 23 (1): 1–18.

Sorensen, E., M. Chen, and G. Mussalli. 2021. The quantitative approach for sustainable investing. *Journal of Portfolio Management* 47 (8): 38–49.

Stotz, O. 2022. Expected and realized returns on stocks with high-and low-ESG exposure. *Journal of Asset Management* 22: 133–150.

Wong, C., A. Brackley, and E. Petroy. 2019. Rate the raters 2019: Expert views on ESG ratings. Sustainability https://www.sustainability.com/thinking/rate-raters-2019.

Yasmine, B., and M. Kooli. 2022. Smart beta ESG disclosure. *Journal of Asset Management* 21 (1): 567–580.

Monia Magnani is a postdoctoral researcher at Bocconi University. During her PhD studies, she also held a position as a Visiting Researcher at De Nederlandsche Bank (DNB, Dutch Central Bank). Previously, after graduating from Bocconi University with an MSc. in Finance, she spent the early part of her career in banking, in the Risk Management divisions of the two major Italian banking groups. Her research focuses on ESG investing and the finance of climate change. She is currently also researching the process of formation of rational bubbles and their impact on macroeconomic policies and asset prices.

Massimo Guidolin is a professor of finance at Bocconi University Milan, Italy, known for his expertise in financial econometrics, portfolio management, and asset pricing. He has held academic positions at prestigious institutions, including the University of Viriginia and the Federal Reserve Bank of St. Louis. His research has been published in top-tier journals such as the American Economic Review, the Journal of Financial Economics, Journal of Econometrics, and the Review of Financial Studies. He sits in the editorial board of a number of journals straddling empirical finance, forecasting, and data analysis, like the Journal of Financial Econometrics, the International Journal of Forecasting, and the Journal of Economic Dynamics and Control. Guidolin's work often explores topics like asset allocation, risk management, and the behavior of financial markets under uncertainty.

Ian Berk is currently an investment professional, after starting his career in investment banking. He holds an MSc from Bocconi University, where he focused on ESG signals and related factors in asset pricing applications. He also holds an MSc from Erasmus University Rotterdam, where he focused on option pricing methods.