Strong vs. Stable: The Impact of European ESG Ratings Momentum and their Volatility on the Equity Cost of Capital*

Ian Berk †

Massimo Guidolin ‡

Monia Magnani§

This draft: April 30, 2023

Abstract

We test the performance of two ESG score-driven quantitative signals on a large, multi-national cross-section of European stock returns. In particular, we ask whether in the cross-section of stock returns, the cost of equity capital to firms 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 1-month, short-term ESG momentum is priced in the cross-section of stock returns and that it lowers the ex-ante cost of capital (at the same time causing realized ex-post average abnormal returns). ESG momentum may represent a novel, priced systematic risk factor, even though such a finding is not robust to changing the definition of the scores or to increasing the estimation period of ESG momentum. There is equally strong evidence that a ESG spread strategy that buys (sells) low (high) ESG score volatility stocks leads to a significant alpha and alters the ex-ante cost of capital, even though as a characteristic volatility is not significantly priced in the cross-section. Both quantitative ESG signals lead to portfolio sorts and long-short strategies that increase the speed of improvement of the aggregate sustainability profile of the resulting portfolios with no (or with negative, risk-adjusted) costs in terms of average ESG scores or their stability.

Keywords: ESG ratings, ESG momentum, ESG score volatility, cross-sectional pricing, systematic risk factor.

JEL codes: G11, G12, C59, G24

*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. Quantitative approaches for a robust assessment and management of risks related to sustainable investing")

†Bocconi University, Milan, Italy. E-mail: ian.berk@bocconialumni.it

‡GREEN, Bocconi University, Milan, Italy (corresponding author). E-mail: massimo.guidolin@unibocconi.it

§University of Liverpool Management School, United Kingdom. E-mail: m.magnani@liverpool.ac.uk

1 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 to moral considerations.[2]

Differently from other strands of research in asset management, ESG strategies of a quantitative nature have struggled to be embraced by investors.[3] Arguably, so far, investors have been relying heavily on the subjective and diverse ESG ratings of companies to measure their ESG exposures, but only on a qualitative basis. This is to some extent 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. Firstly, finding significantly positive alpha for portfolios sorted by ESG rating-related signal (i.e., the ESG scores themselves or some intuitive function thereof), may indicate that such a signal would represent 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 source into account in capital allocation decisions. Secondly, 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.

The goal of our paper is hence to develop and back-test the effectiveness of a set of ESG score-related strategies of a rather 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 the cost of (equity) capital to

[1] According to a recent report by Morgan Stanley, by the end of 2022, the worldwide asset under management of sustainable funds totalled nearly 2.8 trillion, an increase in its proportion of overall AUM to 7% from 4% at the end of 2019. Demand was strongest in Europe, which accounted for 89% of sustainable AUM and almost all of 2022's net inflows to sustainable funds. See https://www.morganstanley.com/ideas/sustainable-funds-performance-demand/

[2] For example, a report from Morningstar published in March 2022 revealed that over 2,700 funds in Europe were already 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 formal quantitative methods applied to investors' preferences for ESG-type quality. In spite of such advantages, they characterize the current 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 arguably gained ground over fundamental analyses of stock returns since the 1980s, so Sorensen et al. (2021) expect quantitative rankings of ESG factors to improve on purely fundamental approaches.

firms more strongly affected by the existence of an upward, virtuous trend in their ESG quality or by their ability to avoid undue fluctuations in such quality, after conditioning on a positively slope of their ESG scores? In other words, conditioning on an improving quality of the ESG scores being an important driver of the cost of equity capital, does it also matter whether such a score improves in a steady way that avoids creating undesirable uncertainty on the ESG quality of the firm? Our study is focused on European stock markets because it is the region in the world with more sustainable investment-related assets under management and where a drive towards ESG-oriented asset management is backed by the strongest support in the legislative agenda (see Alliance (2018)).[4]

As a result of this research strategy, we contribute to the literature in several ways. Firstly, we investigate ESG scores and identify a few biases commonly present in this type of data. Moreover, we propose two ways to deal with these issues and analyse how the raw ESG score data treatment influences the resulting quantitative signals. Secondly, we analyse the optimal ESG momentum computation period, backtesting the strategies' performances 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. We also test the robustness of these results to the choice of estimation sample by testing the significance of the risk-adjusted returns of quintile portfolios sorted by ESG momentum computed with reference to a range of alternative characteristic estimation windows. 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 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 needs to be compensated by means of a positive risk premium.[5]

Our results indicate that quintile portfolios sorted by 1-month ESG momentum earn significantly positive alpha for individual portfolios as well as for a spread portfolio that buys "winners" (the

[4] In the aftermath of the COVID-19 pandemic and with the proposal of a sustainable economic recovery model, in Europe, two major initiatives have come to light: the Just Transition fund which has a budget of up to 40 billion euros and the European Recovery and Resilience Facility which includes 310 billion euros in grants and 250 billion euros in loans. Both of these initiatives prioritize sustainability and a transition to a low-carbon economy.

[5] The connection between finding accurately estimated risk-adjusted 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) type decomposition of current stock prices as the sum of *future* cash flow and discount rate news. Section 5 discusses this implication in depth.

virtuous) and sells "losers" (sinners), as defined in the ESG space. An Andrews (1993) test for a structural break at an unknown date fails to find evidence against the null that the estimated alpha is stable over the entire sample. This is consistent with a conjecture that the 1-month ESG momentum represents a new type of systematic risk. However, the significantly positive alpha earned by the quintile portfolios is not always robust to the choice of data treatment, nor to the choice of momentum portfolio formation period. Both the neutralisation method for dealing with ESG score biases and the period over which the ESG momentum is computed prove to be crucial for the significance of the risk-adjusted mean returns. In fact, a simple ESG momentum tilting strategy, that overweights stocks with high ESG momentum and underweights stocks with low ESG momentum, fails to generate significant alpha. Only our novel method to deal with data biases manages to allow a precisely estimated momentum premium to emerge. On the contrary, we fail to find significantly positive alpha generated by the ESG low volatility factor. However, a long-short spread portfolio that buys stocks with stable ESG scores and and sells stocks with unstable scores earns a precisely estimated risk-adjusted return. Yet, a tilting strategy which overweights stocks with low ESG volatility and underweights stocks with high ESG volatility did not earn significant average excess returns.

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 momentum for several 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 novel risk factor.[6] 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 the firms issuing equity shares might be improving their ESG scores over time, relative to their (possibly, size- and industry-specific) peers. We then carefully investigate whether a trading strategy based on ESG momentum earns significantly positive, risk adjusted returns. To make sure 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 asset pricing tests carefully control for the exposure of ESG-momentum excess returns to

[6] Carhart (1997) is the seminal paper in which a new, priced momentum risk factor was created in an effort to explain the performance of mutual funds.

Carhart's classical momentum factor.

Another literature has noted that low- and high-return volatility stocks might be priced differently by financial markets 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 for returns, also with reference to European data. They present empirical evidence that, contrary to classical asset pricing theory, stocks with low return volatility generate higher risk-adjusted returns than stocks with high volatility. Furthermore, they present evidence that these results are robust to the choice of geographical markets and to each sub-period of 10 years in their sample. Moreover, these effects are not the result of exposures to other risk factors, as these spreads represent genuine Jensen's alpha, i.e., either average abnormal returns and remuneration of some novel risk factor exposure. In this paper, we follow the lead of this literature and investigate the existence of a ESG Low Volatility factor, which posits that the uncertainty on 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. We test this conjecture and assess whether investors should receive a premium for bearing this novel risk. Of course, also with reference to the ESG volatility factor, we make sure that in no way the classical and already well-investigated low volatility factor may explain away and re-absorb the novel ESG volatility factor that we introduce.

There is also a new literature that studies systematic, quantitative strategies based on sorting the universe of 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 outweights companies with the highest increase in the ESG score over a pre-established interval of time). They find that both tilting strategies outperform the MSCI World index in terms of realised, mean returns and of implied, portfolio weighted ESG scores. However, their research 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 returns. In our paper, 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, and assess whether quintile portfolios sorted by this signal may earn significantly positive risk-adjusted returns.

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 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. However, this approach ignores the heteroskedasticity (in ESG scores) that different groups may display. In our paper, we control for these biases and propose two neutralisation methods for ESG score data that take both their mean and variance heterogeneity into account when correcting for size and sector biases. [7]

Using a exposure-matched approach that formally neutralizes 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 turn out to contain the strongest momentum effects and ESG momentum is the strongest for stocks around the medium ESG level. This pattern is consistent with the finding that the ESG momentum effect was strongest among stocks with medium-level ESG scores. Compared to Dor et al. (2022), we extend the analysis to European data and to the empirical uncertainty characterizing the time series of the ESG scores. Even though we carefully neutralize industry and size effects in the ESG scores, we refrain to completely neutralize 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 remainder of this paper is structured as follows. Section 2 describes the data and how to deal with the biases typical of ESG scores. Section 3 explains the ESG momentum signal and analyses the alpha earned by quintile portfolios sorted on it. Section 4 defines the ESG low volatility signal and investigates whether quintile portfolios sorted on this signal earn significantly positive alpha. After a number of robustness checks, Section 5 concludes.

[7] Cantor and Mann (2007) is a related paper that has 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 and we adopt a cross-sectional asset pricing perspective instead of simply discussing the trade-off between accuracy in forecasting credit events and their stability (note it would be hard to provide a definition of sustainability events).

2 Data

The data set comprises a variety of stock market and firm-specific characteristics such as returns, market capitalisation and the Sustainalytics ESG scores. All these data are extracted from Factset.

Our investment universe covers a number of stocks issued in most of the countries classified as European, with the only exception of Iceland and Greece.[8] Furthermore, as an additional quality restriction to select stocks to be analyzed, we exclude all stocks that were not covered by any equity analyst over our sample period. We, also, require a minimum market capitalisation of 400 million euros so to exclude listed companies close to bankruptcy or small capitalization stocks just admitted to listing. These stocks have been reported to be able to affect the performance of many multi-factor asset pricing models and may bias our empirical findings concerning alphas and their statistical significance, see, e.g., Israel and Moskowitz (2013).[9]

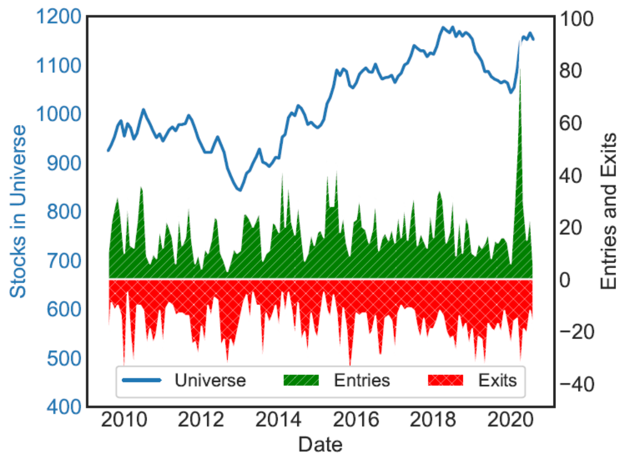
Furthermore, We narrow our sample to stocks with a share price between 0.01 euros and 10,000 euros. The lower bound is introduced to exclude penny stocks because these tend to be less informative due to their extreme returns which are obtained by construction.[10] The upper bound is set to avoid situations in which the optimal portfolio would be otherwise infeasible: if stocks carry too high a price, (nearly) continuous portfolio weight rebalancing cannot be implemented because of the induced discreteness of the portfolio adjustments. Finally, we also 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 minimum turnover filter on the selection of stocks, in the case in which the optimal portfolio strategy might require to buy a stock multiple times its maximum daily volume this would imply a severe market impact, with the risk of practically destroying the signal's profitability so to make our back-testing findings virtual at best.

The sample investigated spans the period July 2009 - April 2020, which corresponds to 130 months. The beginning of the sample corresponds to the starting of Sustainalytics's activities at rating stocks, as covered by Factset. As often remarked by previous literature, ESG scores represent relatively new

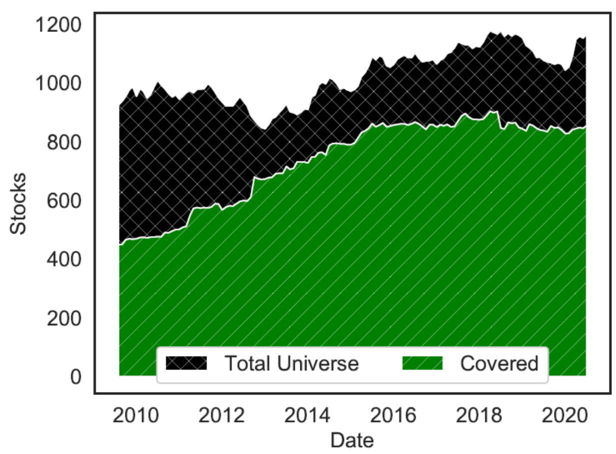
[8] The countries in which stocks were issued include Austria, Belgium, Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Poland, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. The selection of these countries is driven by a set of minimum requirements that ensure one investor would be able to trade in well- functioning stock markets, in terms of investor protection laws and liquidity.

[9] A recent, glamorous example of how a small capitalization score induced by a stock being close to bankruptcy that would defy most rational asset pricing is represented by Hertz Global Holdings, which saw its stock price rise 825% after filing for Chapter 11 bankruptcy in 2021.

[10] However, the minimum market capitalisation constraint turned out to be play hardly any role.



(a) Evolution of the investment universe, the entries, and the exits.



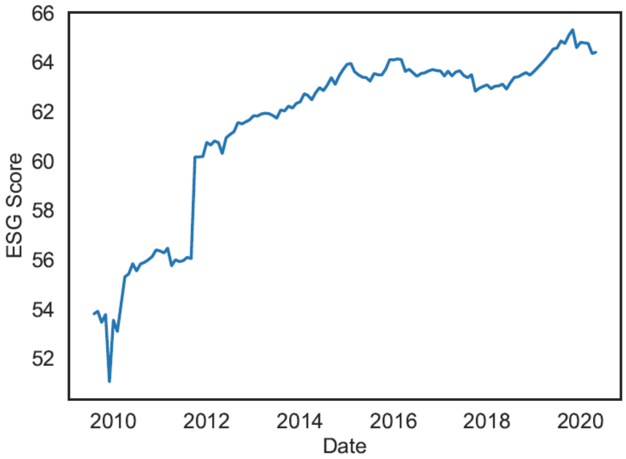
(b) Evolution of the ESG score coverage within the investment universe.

Figure 1 shows the evolution of the investment universe and the coverage of ESG scores in it. Panel (a) shows the evolution of the number of stocks, as well as the entries and exits for the whole sample period. Panel (b) shows the evolution of the number of stocks reporting a Sustainalytics' ESG scores compared to the total number of stocks under analysis. The sample period is July 2009 - April 2020.

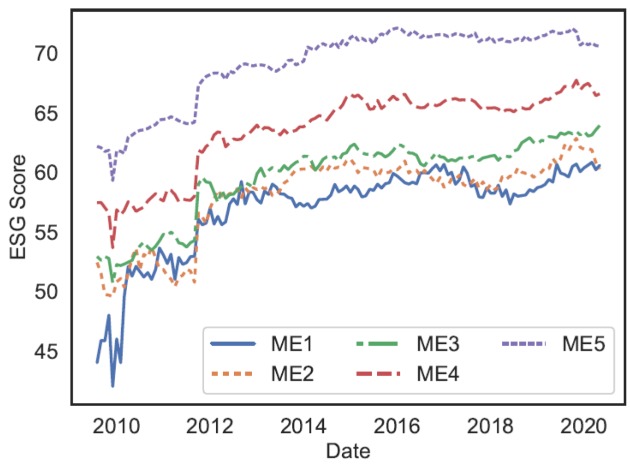
data and are not available before mid-2009. The end date of our sample is carefully chosen to not overlap with the period (2020-2022) dominated by the Covid pandemics to avoid any bias.

Figure 1 shows the evolution of the number of stocks under analysis in our sample. Panel (a) exhibits the total number of stocks in the sample, the new entries, and the exits for every month in our sample. In order to prevent any survivorship bias, we account for all entries and exits at any point in our analysis. Clearly, this procedure causes the resulting panel of data to be unbalanced. However, our backtesting methods are then based on separate cross-sectional spread tests repeated over time, which prevents the variations in the cross-sectional size to bias our results. Panel (b) of Figure 1 shows the evolution of the number of stocks with an ESG scores as a fraction of the total number of stocks under analysis. The difference between these two sets narrows down as the time goes by in our sample because of the growing importance of being rated in the ESG dimension.

Looking at the literature, Kaiser (2020) has noted that the ESG scores strongly depend on the firms' size and differ substantially across different industries. These conclusions are not surprising as there is an abundance of research linking the size of a company with their sustainability policies. For example, Artiach et al. (2010) have documented a stark, positive relationship between the size of a firm and its Corporate Sustainability Performance (CSP). Bos (2017) and Kaiser (2020) show that this relationship translates to higher ESG scores for larger companies. Yet, as far as we are concerned, the sector dependency is embedded in the rating that Sustainalytics gives to each company. They



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



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

Figure 2 shows the evolution of the ESG scores over time. Panel (a) shows the evolution of the mean ESG score across the entire investment universe. Panel (b) shows the evolution of the mean ESG score for each of the five size quintiles used throughout the paper (for example, ME1 includes the companies with the lowest market capitalisation while ME5 represents the quintile containing companies with the highest market capitalisation among the investment universe). The sample period is July 2009-April 2020.

build up a score by rating a company and using this score in 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). Moreover, the same category can attribute different weights for different sectors. This way of evaluating the overall ESG score automatically assigns some sectors higher ratings than others. In order to avoid undesired sector tilts, we address these sector dependencies eliminating this bias through a specific empirical methodology.

Figure 2 shows the evolution of the ESG scores for the companies included in the investment universe. Panel (a) shows the evolution of the average ESG score of all the stocks that have been assigned a score. The figure shows a clear upward trend of the ESG score over the sample period, signalling the increasing urge for companies to have their ESG-directed efforts publicly acknowledged. While the 2008 reduction that can be noted in panel (a) is easily explained by the effects of the Great Financial Crisis, equally visible is the upward surge in average scores that occurs in 2012.[11]. Panel (b) in Figure 2 displays the evolution of the average ESG score across five equally-weighted quintile portfolios sorted by size.[12]. Therefore, ME1 in panel (b) is the quintile including the companies

[11]Boffo and Patalano (2020) emphasized that the assets under management in ESG funds have risen by 60 percent starting 2012, from USD 655 billion in 2012 to USD 1.05 billion in October 2018. This pressure by ESG-screening mutual funds would have pushed listed companies to rush into the ESG-rate universe with a strong acceleration taking place since early 2012

[12]The average market capitalization of each of the quintiles in panel (b) are the same as those reported in the fourth column of Table 1

Table 1: Summary statistics for returns and ESG scores for five size-sorted quintiles

	Returns				ESG Score				
	Mean (%)	σ (%)	SR	Mark. Cap.	Mean	Median	σ	Max	Min
1	0.633	6.454	0.072	670	56.91	58.22	3.86	60.84	42.00
2	0.809	5.536	0.116	1,361	58.06	59.46	3.49	62.85	49.56
3	0.795	5.260	0.119	2,646	59.65	61.08	3.39	63.93	50.70
4	0.777	4.854	0.125	5,591	63.57	65.25	3.40	67.76	53.66
5	0.651	4.803	0.101	29,112	69.11	70.71	3.21	72.13	59.31

The table shows the portfolio summary statistics for returns and mean ESG scores. In the left panel, we report the mean return, the volatility of the returns, the Sharpe Ratio and the market capitalisation in millions of euros. In the right panel, we show for each of the size-sorted quintiles, the average and the median ESG score, the volatility of the ESG score and the maximum and the minimum value of the ESG score.

that have the lowest market capitalisation while ME5 is the quintile portfolio containing the largest stocks.The plots shows that the average ESG score increases with size. ME5 has the highest ESG score for every observation in the sample period, while ME4 has the second largest. However, such ordering does not hold uniformly over time for the average ESG scores of the first three quintiles, even though the general tendency of scores to increase with market capitalization generally holds.

Table 1 shows the relationship between size and the ESG scores on the basis of the same quintile portfolios used in panel (b) of Figure 2. Clearly, as market capitalization increases, the average and median ESG score (together with the minimum and the maximum ESG score) increase monotonically. This suggests that a strong size tilt is present in this ESG data set. Artiach et al. (2010) claims that this relationship between size and ESG score is due to the comparatively higher visibility of large companies. Well capitalized firms, in fact, tend to spend more resources to address ESG related issues, compared to smaller businesses, because of higher pressure from regulators and general stakeholders. At the same time, data show a negative relationship between the average returns and size except for ME1, which is to be expected in the light of the classical "size anomaly". In our analyses, we therefore take into consideration this size effect to avoid any interference with the measurement of the performance of ESG-based signals.

Table 2 makes it clear that a high dispersion of the ESG scores across industries is clearly visible of one compares the energy and utilities sectors. These two industries have roughly the same average market capitalisation (12941 and 12204 million euros, respectively). However, utilities carry an average ESG score that is more than two standard deviations higher than the energy sector's ESG score (using the standard deviation of the latter). The same pattern applies also to the median, the maximum and the minimum ESG scores while the SR of the utilities is only marginally higher than energy's. A similar ranking can be observed for other pairs of industries that display similar average market capitalisations (e.g., the industrials and materials for which the difference in SR however large). Therefore, to avoid as undesirable as implicit industry tilts in our ESG-driven strategies, our methodology shall also take into account sector-dependency.

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

	Returns				ESG Score				
	Mean (%)	σ (%)	SR	Mark. Cap.	Mean	Median	σ	Max	Min
Automobiles	0.504	7.238	0.046	15513	65.75	67.22	3.76	70.15	54.27
Banks	0.835	6.072	0.110	6990	60.14	61.02	3.42	65.29	52.11
Consumer Discr.	0.716	4.908	0.112	7228	62.47	63.17	2.91	66.05	54.35
Consumer Staples	0.807	4.402	0.145	9081	60.86	62.58	3.65	65.22	50.20
Div. Financials	0.642	5.989	0.079	9485	62.60	63.50	2.79	66.53	56.04
Energy	0.869	4.722	0.148	12941	61.17	61.33	1.89	64.64	55.66
Healthcare	0.837	5.137	0.130	4911	61.92	62.81	2.99	65.14	52.35
Industrials	0.805	5.861	0.109	5081	63.87	64.65	2.18	66.68	56.88
Insurance	0.928	5.639	0.135	4269	62.82	63.53	2.19	66.43	55.30
IT	0.374	6.431	0.032	7838	62.89	64.15	3.84	68.35	54.05
Materials	0.787	8.021	0.077	5103	58.72	58.87	1.90	62.63	53.39
Media Entert.	0.814	6.351	0.102	6155	64.75	64.88	2.68	70.30	58.45
Real Estate	0.566	6.098	0.065	5527	63.78	64.69	2.42	67.55	55.97
Telecom	0.627	5.629	0.082	5773	60.39	60.42	3.34	65.02	53.77
Utilities	0.860	4.501	0.154	12204	68.19	68.95	2.83	72.55	59.00

The table reports summary statistics for returns and ESG scores in each industry. On the left-hand panel, we show the mean return, the volatility of the returns, the Sharpe Ratio and the market capitalisation in millions of euros. On the right-hand panel, we have the average and the median ESG score, the volatility of the ESG score, the maximum and the minimum value of the ESG score across the five size quintiles of Table 1. The industries are based on the Global Industry Classification Standard (GICS) classification.

We neutralise the ESG scores from the influence of both size and industry biases and thus remove any unintentional tilts. We propose two approaches to reach this goal. A first method divides the stocks into five size quintiles, where the size of a company is defined by its market capitalisation. We compute the median ESG score for each quintile which is then subtracted from every stock within that quintile. This ensures that the monotonically increasing relationship between size and ESG score is approximately neutralized because we take into consideration only deviations from the median. Afterwards, we compute the ESG score quintiles to which each stock is to be assigned by giving every stock a value between 0 and 1: 0 is given to the stocks with the largest negative deviation from the median ESG score within their appropriate size quintile, and 1 to the stocks with the highest positive deviation from the median ESG score within a size quintile. Stocks falling in between the minimum and the maximum are assigned a score that is proportional to their signed distance from the median, with 0.5 being attached to the median stocks (or to the two stocks straddling the median, when this fails to correspond to a realized ESG score). To correct for any sectoral tilts, we follow a similar procedure and apply it to all sectors, but this time we start off the size-corrected percentile ESG scores constructed following the steps above, rather than the raw ones.

We propose an alternative neutralisation method, which has appeared already in the literature (see, e.g., Heij et al., 2004). It consists of simply standardising the scores across portfolio sorts. Firstly, the stocks are grouped into size quintiles and for each of them, we compute the mean and standard deviation to standardize the ESG scores within the quintile. Next, we apply the same steps to the fifteen industries, but taking the size-standardised ESG scores as the initial score data.[13]

Before proceeding, we introduce terminology that we use throughout the paper. When referring to the un-adjusted, non-neutralised ESG scores we call them "*raw ESG scores*". When we resort to the ESG scores adjusted using the first procedure, we use "*rank-neutralised ESG scores*". Finally, when we refer to the ESG scores adjusted using the second procedure, we will use the term "*standardised ESG scores*" .

[13]We recognize that both neutralisation methods tackle the biases in a sequential manner, by first neutralizing the size tilt and then the industry bias. Although they are common (see, e.g., Kaiser (2020)), such sequential approaches cause the ESG scores obtained after the second, industry-neutralization step to longer be perfectly neutral to size, thus potentially reversing the effect of the first neutralization step. Clearly, one would need to apply simultaneous double sorting to be contemporaneously in control of size and industry biases. However, a double sorting would create certain portfolios including an abysmal number of stocks, while empty portfolios cannot be ruled out. For this reason, we prefer to stick to a sequential, conditional sorting by computing risk-adjusted performances from multi-factor models that contain a long-short factor portfolio that exposes the investor to size risk.

3 Can ESG Momentum Represent a Reliable Risk Factor?

In this section, we first sort our universe of stocks in portfolios; for each sort, we use different ESG momentum computation periods.[14] We then estimate the Fama and MacBeth (1973) cross-sectional regressions to provide preliminary evidence on the effects that ESG momentum has on excess stock returns. Formal asset pricing tests are subsequently deployed to investigate whether ESG momentum generates statistically significant positive abnormal returns. Finally, we perform a robustness analyses concerning the choice of the ESG momentum computation periods.

3.1 Portfolios Sorted by ESG Momentum

ESG momentum is defined as the percentage increase of a security's ESG score over a certain period of time. As explained in section 1, Nagy et al. (2016) has used ESG momentum in a tilting strategy to over- and underweight stocks with high and low ESG momentum compared to standard market cap weights. However, to assess whether the change in the ESG score may truly provide an alpha-generating strategy, we need to perform an ESG momentum portfolio sorting following the appropriate literature (Fama and French, 1992, 1993). We therefore start off by computing the x-month ESG momentum, where $x \in \{1, 3, 6, 12\}$ months. The motivation of this range of momentum-measurement periods lays within the literature on momentum. Jegadeesh and Titman (1993) reported rather heterogeneous empirical results depending on a range of momentum computation periods. Since the choice of the period can significantly influence the final results, we further investigate this definition with robustness checks As a first step, we 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 not to any other risk factors (especially, the market portfolio) that ought to cancel out. This zero-cost portfolio is created taking a long position in portfolio 1 and a short position in portfolio 5.

[14]We resort to simpler, univariate sorting instead of bivariate sorting, conditioning on the ESG score level, because such a strategy would prevent us from deriving the endogenous, uncontrolled implications for such average scores of our systematic strategies, based instead on the change and volatility of such scores.

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

	Returns				ESG Score			
	Mean (%)	σ (%)	SR	SE(SR)	Mean	Median	σ	Momentum
	Panel A: 1-Month ESG Momentum							
1	0.851	4.377	0.195	0.088	57.92	58.26	2.93	46.1%
2	0.695	4.397	0.158	0.087	65.44	66.16	3.17	4.1%
3	0.667	4.662	0.143	0.087	69.15	69.82	3.09	-0.1%
4	0.699	4.630	0.151	0.092	64.99	65.67	2.88	-4.1%
5	0.426	4.596	0.093	0.094	57.34	57.73	2.99	-21.0%
1 - 5	0.382	1.543	0.248	0.113				
	Panel B: 3-Month ESG Momentum							
1	0.801	4.508	0.178	0.088	58.90	59.72	2.95	83.5%
2	0.660	4.438	0.149	0.091	65.56	66.57	2.99	6.9%
3	0.551	4.685	0.118	0.091	69.33	70.25	2.88	-0.3%
4	0.764	4.654	0.164	0.086	64.85	65.15	2.50	-7.0%
5	0.565	4.425	0.128	0.093	57.38	57.88	2.62	-29.8%
1 - 5	0.193	1.480	0.130	0.095				
	Panel C: 6-Month ESG Momentum							
1	0.677	4.721	0.144	0.094	60.06	60.97	2.71	123.8%
2	0.663	4.568	0.146	0.093	65.69	66.33	2.69	9.8%
3	0.543	4.525	0.120	0.092	69.77	70.49	2.57	-0.5%
4	0.654	4.547	0.144	0.089	64.69	64.77	2.29	-9.5%
5	0.649	4.395	0.148	0.096	57.50	58.04	2.37	-36.3%
1 - 5	-0.017	1.427	-0.012	0.092				
	Panel D: 12-Month ESG Momentum							
1	0.559	4.781	0.117	0.095	61.42	61.98	2.30	187.5%
2	0.494	4.586	0.108	0.097	66.31	66.86	2.22	15.1%
3	0.587	4.460	0.132	0.091	70.26	70.74	2.43	-0.3%
4	0.731	4.445	0.165	0.093	65.01	65.14	2.04	-12.6%
5	0.632	4.324	0.147	0.096	57.71	58.07	2.12	-43.6%
1 - 5	-0.119	1.420	-0.083	0.092				

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 1 - 5, which takes a long position in portfolio 1 and a short position in 5.

Table 3 shows summary statistics for both the performance and the raw ESG scores for the quintile portfolios sorted by four different ESG momentum definitions. In this table, the rank-neutralised ESG scores are used to produce the sort. In addition to display the summary statistics for the five quintile portfolios, the table shows the summary statistics for the zero cost portfolio 1 - 5. We report the mean return, the standard deviation of the returns, the Sharpe ratio, and the standard error of the Sharpe ratio as derived by Lo (2002) (SE(SR)) for every portfolio. Furthermore, on the right-hand-side of the table we detail the mean ESG score, the median ESG score, the standard deviation of the ESG score, and the average ESG momentum for each of the quintile portfolios.[15] Table 3 displays several notable features. First, the SR of spread portfolio 1 - 5 is positive only 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 steeply affects the results. Of course, this finding naturally shifts our focus in what follows towards ESG momentum defined on a shorter window of 1 month. Second, the table shows that both the mean and median raw ESG score of portfolio 1 are higher than the mean and median ESG score for portfolio 5 under all momentum computation choices, albeit only marginally for the 1-month ESG momentum. This demonstrates that stocks characterized by high ESG momentum are generally also issued by "ESG virtuous" firms, or that are at least quickly becoming so.[16] Third, the mean and median ESG score increase the longer the computation period is. Yet, this may represent an artifact of the fact that over our sample period there is a visible upward trend in ESG scores, as shown in panel (a) of Figure 2.[17] Finally, the table shows a flattening of the mean return and of the SRs across the quintiles as the momentum computation period increases. This suggests that the strength of the signal in the ESG scores declines as the computation period increases, which intuitively makes sense given the informational efficient nature of the European equity market under investigation: if a company's ESG score improves, the information is absorbed quickly and does not take several months to be fully priced in.

[15] We report raw instead of rank-neutralized or standardized scores to be able to propose the policy consideration that follows, even though the raw scores are *not* used in the portfolio sorting process.

[16] However, the mean and median ESG scores of portfolio 1 (the one exhibiting higher ESG momentum) represent an exception as they are generally similar to the mean and medians of quintile 5. However, quintiles 2 and 3 display mean and median scores that exceed those of quintile 5, the portfolio characterized by negative momentum that contains firms that see their ESG positioning decline in relative terms.

[17] The longer the ESG momentum computation period, the more of the earlier months in our sample are ignored because of the need of more initialization data. Because the ESG score have been increasing over time, dropping the earlier, lower scores increases the resulting average and median. This explanation receives additional support by the fact that the standard deviation of the ESG scores decreases when the ESG momentum computation period lengthens.

Table 4: Results for portfolios sorted by 6-month ESG momentum using standardised ESG scores

	Returns				ESG Score			
	Mean (%)	σ (%)	SR	SE(SR)	Mean	Median	σ	Momentum
1	0.727	4.587	0.159	0.094	63.64	64.16	2.72	568.3%
2	0.654	4.367	0.150	0.095	62.63	62.62	2.52	16.6%
3	0.655	4.378	0.150	0.092	63.82	64.23	2.42	-4.6%
4	0.655	4.520	0.145	0.093	63.86	64.28	2.79	-27.5%
5	0.595	4.612	0.129	0.093	63.92	64.82	2.71	-469.8%
1 - 5	0.088	1.015	0.087	0.095				

The five quintiles are sorted by ESG momentum, the percentage increase of the ESG score over a period. The ESG scores are standardised to neutralize both size and sector tilts. In addition, the table provides the summary statistics for a portfolio 1 - 5, which takes a long position in portfolio 1 and a short position in 5.

While Table 3 used the rank neutralised ESG scores to compute the ESG momentum, Table 4 shows the same results but with the standardised ESG scores. To save space, Table 4 only reports the summary statistics for the 6-month momentum.[18] We want to focus our attention on at least two differences in key empirical results with Table 3. First, the long-short portfolio no longer yields a significantly positive SR. This is a first indication that a simple ESG-momentum strategy may only earn fragile risk-adjusted performance, at least as far as European equity data are concerned. However, since the SR obtained implies only a very rough method of risk-adjustment (in essence, by scaling mean returns by total and not systematic risk), we post-pone a thorough assessment of this question to later portions of the paper. Second, standardising the scores is effective in neutralising the sector and size biases but it may be troublesome in light of its policy implications. In fact, under the ESG momentum estimated using standardised scores, portfolio 1 has a lower mean and median raw ESG score than portfolio 5 does. This indicates that the long-short portfolio will not have a positive ESG score spread, albeit such a negative spread turns out to be very small in absolute value. 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 standardised ESG scores, the ESG scores of the quintile

[18]The interested reader can find the full table spanning alternative formation periods in Table A.1 of the Appendix. Here we focus on 6-month ESG momentum because it is the period for which the SR of the long-short portfolio displays the highest t-statistic and hence can be taken as the upper bound of realized momentum performance.

portfolios are very similar in magnitude. This means that an increasing standardised ESG score no longer corresponds to higher raw ESG scores. However, allocating capital to firms enjoying a high mean standardized ESG momentum would not imply transferring capital from virtuous companies to sin ones.

3.2 Regression Analysis of the ESG Momentum Characteristic

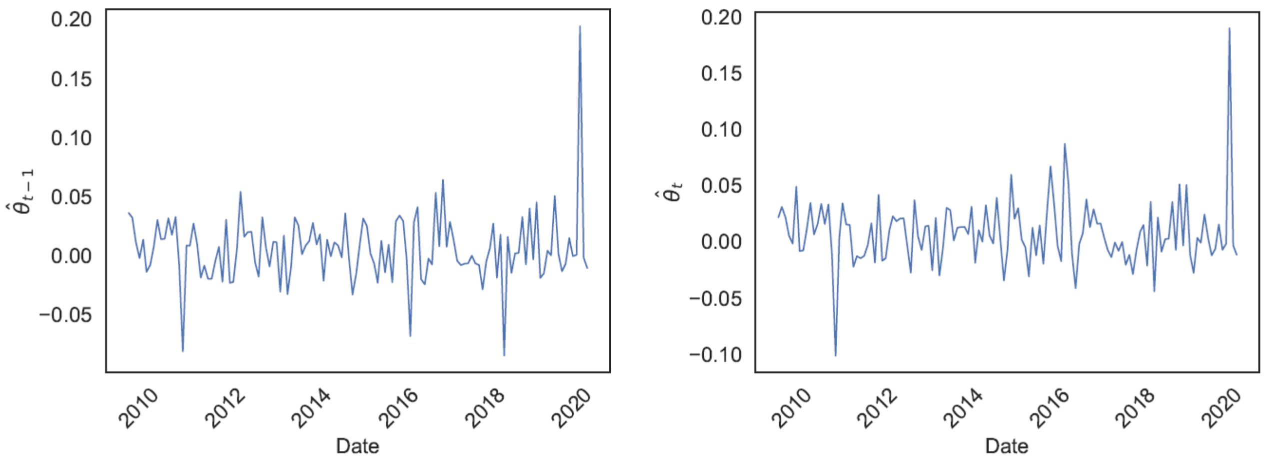
After the identification of the optimal ESG computation period, we turn our attention to the investigation of its effects on the cross-section of realized equity returns. Therefore, we apply standard Fama and MacBeth (1973) analysis, which consists of regressing the excess returns of each stock on the corresponding 1-month ESG momentum and a constant. It must be stressed that the Fama and MacBeth (1973) regression do not provide a formal asset pricing test. Since the ESG momentum is a security-specific characteristic, it could also happen that it does not represent a common risk factor. However, a significant average effect of the 1-month ESG momentum on the excess returns would provide another indication, on top of those derived from Table 3, that the momentum characteristic may indeed be a source of risk.

The regressions 1 and 2 presented below provide the functional forms of the models for the cross-sectional regressions. Here, $r_{i,t}$ is the excess return of security i at time t, $c_{i,t}$ the 1-month ESG momentum of security i at time t, and $\varepsilon_{i,t}$ a random error term. The difference between these two models relates to the influence of past and current levels of the characteristic. Equation 1 focuses on the relationship between the excess returns of the securities and the 1-month ESG momentum of the previous month. The parameters are estimated using cross-sectional regressions, so that estimation is performed at every point in time, to allow for time-varying coefficients (hence the subscript $t-1$ in θ_{t-1}). On the contrary, Equation 2 capture the linear, contemporaneous relationship between stock excess returns and the contemporaneous 1-month ESG momentum. Also in this case, the regression is estimated cross-sectional at every point in time, hence the subscript t applied to θ_t.

$$r_{i,t} = a_t + \theta_{t-1} c_{i,t-1} + \varepsilon_{i,t} \tag{1}$$

$$r_{i,t} = a_t + \theta_t c_{i,t} + \varepsilon_{i,t} \tag{2}$$

The definition of what the subscript i may refer to, has been left so far intentionally vague. The choice of the test securities implies tackling a trade-off between the noise inherent in the use of



(a) Dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1.

(b) Dynamics of the cross-sectional OLS estimate of θ_t in Equation 2.

Figure 3 shows the dynamics of the Fama and MacBeth (1973) cross-sectional OLS estimates of the regression of excess portfolio returns on their 1-month ESG momentum and a constant. The portfolios are sorted by the 1-month ESG momentum, the percentage increase of the ESG score over the last month. Panel (a) shows the dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1. In this equation, $c_{i,t-1}$ is portfolio i's 1-month ESG momentum at time $t-1$ and $r_{i,t}$ the excess return of portfolio i at time t. Panel (b) shows the dynamics of the cross-sectional OLS estimate of θ_t in Equation 2. In this case, $c_{i,t}$ is portfolio i's 1-month ESG momentum at time t and $r_{i,t}$ the excess return of portfolio i at time t.

individual stock excess returns and the loss of information implied by the use of portfolios. In the former case, recalling section 2 in Section 2, we have between 400 and 800 data points available to perform estimation at each month. However, given that the ultimate goal is to create common asset pricing factors within the ESG space, it is desirable to minimise the stock-specific noise. Therefore, in this paper we use the approach by Fama and MacBeth (1973), and sort portfolios by the characteristic investigated to diversify away firm-specific noise.[19]

Figure 3 shows the dynamics of the cross-sectional OLS estimates for the predictive and contemporaneous Fama and MacBeth (1973) regressions. Panel (a) shows the dynamics for the OLS estimate of θ_{t-1} in Equation 1. θ_{t-1} turns out to be generally positive but also rather volatile. Panel (b) shows a similar patterns, but with less spikes pointing downward, to negative estimates. The unreported regression R-squares are on average high, but this is influenced by the fact that the cross-sectional regressions are estimated with few observations.

However, in order to draw conclusions about the possible effect of ESG momentum on the excess portfolio returns, Fama and MacBeth (1973) demonstrate how to perform a statistical test of a zero average effect. This cross-sectional statistic test is a t-test for the null hypothesis that the average

[19]For consistency, we sort all available stocks into equally weighted quintile portfolios. This method provides five data points in the cross-section each point in time.

effect is zero:

$$t = \frac{\bar{\hat{\theta}}}{s(\hat{\theta})/n} \sim t(n-k), \tag{3}$$

where $\bar{\hat{\theta}}$ is the time series average of the estimated coefficient for either the predictive or contemporaneous regressions and $s(\hat{\theta})$ is the standard deviation of the time series of estimates, n the number of months in the sample, and k the number of regressors, i.e., one. Performing this test yields a t-statistic of 2.117 for the time series in panel (a) of Figure 3, which is significant at a 5% size level. Similarly, performing the test for the series in panel (b) yields a t-statistic of 2.942, which is significant at a 1% size. Therefore, the null hypothesis that the average effect of ESG momentum on the excess portfolio returns is zero can be rejected which is consistent with ESG momentum producing a significantly positive effect on European equity excess returns. However, as already mentioned, the Fama and MacBeth (1973) cross-sectional regressions concern security-specific characteristics and cannot represent a formal asset pricing test. The significance of the average effects, however, is a preliminary (in some ways, necessary) evidence that ESG momentum may indeed be a source of systematic risk.

We apply the same procedure to ESG momentum computed from the standardised ESG scores. To save space, the plots of the predictive and contemporaneous estimates of the cross-sectional regression coefficients are shown in Figure A.1 in the Appendix. We discuss here only the t-statistics. The average effect of the predictive coefficient yields a t-statistic of 1.397, which is not statistically significant. The contemporaneous average effect, instead, has a t-statistic of 1.710, which is significant at a 10% size. These weaker results are not surprising given the insignificant SR reported in Table 4. Moreover, these findings provide additional evidence that the returns earned by ESG momentum strategies may not be robust to the choice of treatment of ESG scores, at least with reference to European data.

3.3 Is ESG Momentum Compensated by Abnormal Returns?

Given the preliminary evidence suggesting that ESG momentum computed with the rank neutralised ESG scores may indeed represent a source of (at least, firm-specific) risk, we now turn to investigate whether there is actual alpha that portfolios sorted by 1-month ESG momentum may earn. This requires adjusting the returns by their actual, systematic risk exposure, as one cannot rule out the possibility that the mean excess returns earned by the portfolios sorted by 1-month ESG momentum may be caused by risk exposures that SR is unable to identify. 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 (henceforth SMB), High Minus Low (HML), Robust Minus Weak (RMW), and Conservative Minus Aggressive (CMA) portfolios. These factors approximate market risk, size risk, value risk, profitability risk, and investment risk respectively. Given the objective of our paper, besides the five Fama and French (2015) factors, we also include the Carhart (1997) momentum factor. The data for all these factors are downloaded from the data repository maintained by Kenneth French at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, who provides factor data for the European market as a whole. Finally, the seventh factor is the low volatility factor, see Blitz and Van Vliet (2007).[20] This factor is added to our factor model in the light of the results in Blitz and Van Vliet (2007) who motivate the existence of the anomaly by the existence of sin stocks.[21] By incorporating the low volatility factor, we ensure that the returns from investing in morally questionable assets, to be identified with low or even declining ESG scores (inverse momentum), may not represent a potential source of alpha.

In Table 5, Panel A shows the estimated coefficients, whereas Panel B shows the corresponding t-statistics. First, we note the decreasing estimates of Jensen's α when the ESG momentum of a portfolio decreases. The table shows that the estimate of α for portfolio 1 is significant at a 5% size level and the alpha estimates for portfolios 2 and 4 are significant at a 10% size. Moreover, the alpha estimate for the long-short portfolio, which can be seen as the clean strategy implementation of the ESG momentum signal, is significant at a 1% level. This is an additional, formal evidence of the economic value generated by sorting by 1-month ESG momentum. The portfolios with the strongest signal display both the highest alphas, as well as statistical significance at the highest confidence levels. Therefore, there is value-added in sorting portfolios by ESG momentum since these portfolios do earn risk-adjusted abnormal returns in the European stock market. Panel A also shows that portfolios sorted by ESG momentum carry a substantial exposure to market risk. These five portfolios all imply a positive coefficient with a large t-statistic. However, the effect vanishes for portfolio 1 - 5, indicating that the pure ESG momentum signal is not influenced by market risk.[22] On the opposite, the negative

[20]This latter factor mimicking portfolio is constructed by computing the rolling, 36-month volatility for every stock in the universe. Then, we sort the quintile portfolios in order of ascending volatility and create a long-short portfolio. This zero cost portfolio has a long position in the portfolio containing stocks with the lowest volatility and a short position in the portfolio comprised of stocks with the highest volatility.

[21]Sin stocks, they argue, are systematically under-priced because many investors shun them, not wanting to be associated with these firms.

[22]The exposure to market risk of the quintile portfolios could be caused by the granularity of the portfolios. Since each quintile portfolio contains between 100 and 200 stocks at every point in time, they can be seen as small market portfolios. Therefore, it is intuitive that these portfolios imply a substantial, often aggressive (with beta in excess of 1, for quintiles 1-4) exposure to the market factor in univariate regressions.

exposures to the low volatility factor of the long-only quintile portfolios carries over to the long-short portfolio. This is inconsistent with the initial intuition that stocks marked by increasing ESG ratings could be low volatility stocks. In fact, it appears that high ESG momentum stocks behave more like volatile stocks. This can be explained by the increasing importance that ESG considerations are having to investors.[23] Finally, classical Carhart's return momentum provide scant contribution to the returns earned by ESG momentum. This confirms that ESG momentum and actual momentum are two separate factors that do not exhibit any significant overlap.

Crucially, Table 5 reports that some of the individual ESG momentum-sorted portfolio return series do earn significant alpha. However, in order to assess whether our seven-factor model is unable to explain these excess returns, we must perform a formal test on the joint significance of the pricing errors. It is indeed possible that we are dealing with a system of equations that is only connected via correlations in the error terms and that by taking these into account, all excess returns would become jointly insignificant. The test proposed by Gibbons et al. (1989) (henceforth, GRS test) has this exactly aim. The test statistic is:

$$z = \frac{T-n-k+1}{n(T-k)} \frac{1}{q_{11}} \hat{\boldsymbol{\alpha}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} \sim F_{n,T-n-k+1}, \tag{4}$$

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 from the regressions, T the number of months in the sample, n the number of portfolios, and q_{11} the first element of $(\boldsymbol{X}'\boldsymbol{X})^{-1}$ (where $\boldsymbol{X}$ collects the factors used in the factor model). In the case of Table 5, the test statistic is 2.412, with an associated p-value of 0.040. This leads to conclude that the estimated abnormal excess returns are jointly significant. We therefore conclude that the postulated seven-factor model is unable to adequately price the ESG momentum portfolios and that the Jensen's alphas earned by these portfolios are jointly significant.

[23] Blitz and Van Vliet (2007) argue that the low volatility anomaly can be explained by fund managers all investing in the same stocks and ignoring others. Because there is still only a limited number of high quality ESG stocks in Europe, the rising interest in sustainable investing could explain the negative relationship between portfolios sorted by 1-month ESG momentum and the low volatility factor.

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

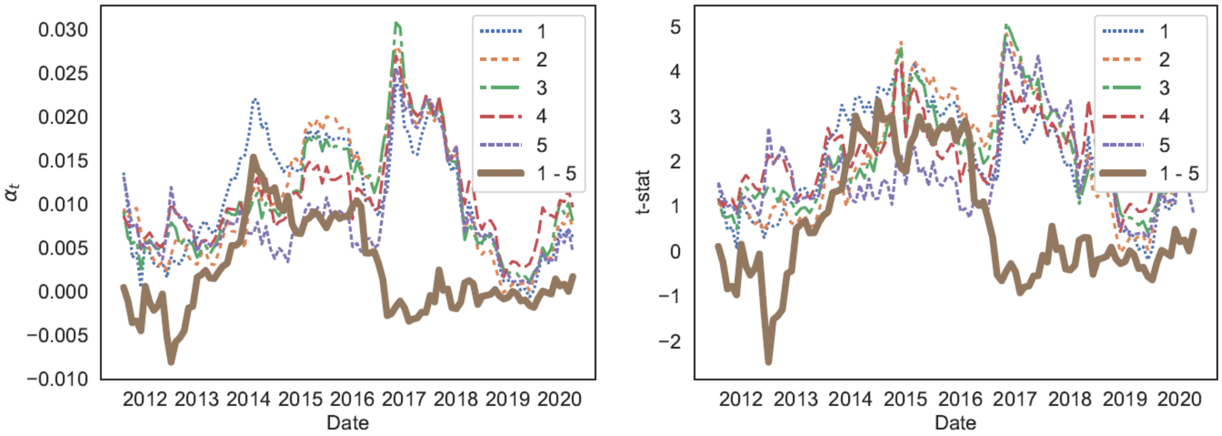
	α	Mkt. - Rf	SMB	HML	RMW	CMA	Mom.	Low Vol.	R^2
	Panel A: Regression Coefficients								
1	0.577	0.575	-0.157	-0.162	0.495	0.048	-0.064	-0.417	0.780
2	0.450	0.616	-0.053	0.095	0.532	0.005	-0.116	-0.262	0.769
3	0.424	0.636	0.049	0.197	0.585	-0.149	-0.145	-0.263	0.777
4	0.431	0.619	-0.078	0.208	0.592	-0.259	-0.086	-0.307	0.782
5	0.124	0.627	0.110	0.326	0.545	-0.416	-0.080	-0.215	0.760
1 - 5	0.453	-0.052	-0.267	-0.488	-0.050	0.464	0.016	-0.202	0.243
	Panel B: t-statistics								
1	2.204	6.950	-0.769	-0.739	1.804	0.232	-0.977	-3.887	
2	1.720	6.893	-0.224	0.336	1.716	0.027	-1.658	-2.319	
3	1.525	6.333	0.183	0.608	1.627	-0.683	-1.965	-1.981	
4	1.725	6.870	-0.300	0.717	1.816	-1.333	-1.330	-2.870	
5	0.457	6.049	0.408	1.080	1.606	-2.002	-1.222	-1.784	
1 - 5	3.102	-1.220	-2.605	-3.942	-0.297	4.834	0.425	-3.485	

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 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 OLS estimates and the R^2 of the regression. Panel B reports the t-statistics of the OLS estimates in panel A. The standard errors are estimated using the heteroskedasticity and autocorrelation consistent estimator (HAC) proposed by Newey and West (1987).

To assess whether these results are due to our specific data treatment, we perform similar regressions for the portfolios sorted by 6-month ESG momentum computed using the standardised ESG scores. To save space, we only comment on the significance of the coefficient estimates and of the resulting alphas.[24] We find that using a 6-month interval to compute ESG momentum, only portfolio 1 and, rather crucially, the long-short portfolio 1 - 5 display significant alphas at a 10% test size or less. The factor model therefore seems to do a better job at pricing the ESG momentum-sorted portfolios but some evidence of abnormal returns remains.[25] However, computing the test inEquation 4 yields

[24]The interested reader can find the full results in Table A.2 in the Appendix.

[25]The quintile portfolios carry, just as for the momentum computed using the rank neutralised ESG scores, a significant



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

Figure 4 shows the dynamics of the OLS estimate of α and the dynamics of its corresponding t-statistic. The estimates are computed using a rolling window of 24 months of data. The portfolios are sorted by the 1-month ESG momentum, the percentage increase of the ESG score over one month. The risk factors incorporated in the OLS regression are the five factors Mkt - Rf, SMB, HML, RMW, and CMA as proposed by Fama and French (2015), the Momentum factor proposed by Carhart (1997), and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007).

a statistic of 0.958, which implies a p-value of 0.447, so that the null of a jointly zero alpha across alternative portfolios cannot be rejected.

Because the GRS test applied to Table 5 has shown that the ESG momentum computed using the rank neutralised ESG scores earns jointly significant alpha, we now need 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 has then faded as time progresses. Figure 4 shows the behaviour of the estimated pricing errors over time, when the same seven factors are included in the risk factor model as in Table 5. We use a rolling window of 24 months of data to estimate the alphas. Panel (a) shows that especially during the period 2014-2016, the dispersion in the alpha estimates was large. However, this translates into high and significant alpha estimates for the long-short portfolio 1 - 5. However, after this period, the alphas of the long leg and of the short leg seem to converge and to decline substantially. This is a troublesome development for the ESG momentum signal, because it appears that the signal may have been fading in recent years, at least as far as European stocks are concerned. Nonetheless, panel (b) of Figure 4 shows that over the entire sample, only one estimate of the long-short portfolio's alpha turned out to be significantly negative (at the end of 2012), which stresses that the ESG momentum signals tends to produced a consistent sign of alpha over time.

relationship with the market and the low volatility factor portfolios.

Table 6: Andrews (1993) test of a structural break at unknown date for six ESG momentum portfolios

	Andrews Test Stat	Critical Value		
		10%	5%	1%
1	4.61	19.82	22.13	27.25
2	6.52	19.82	22.13	27.25
3	7.92	19.82	22.13	27.25
4	5.58	19.82	22.13	27.25
5	5.87	19.82	22.13	27.25
1 - 5	1.73	19.82	22.13	27.25

This table shows the Andrews (1993) test for a structural break at an unknown break date for five quintile portfolios sorted by descending 1-month ESG momentum. The null hypothesis is the absence of a structural break. Additionally, the table provides the Andrews (1993) test statistic for the long-short portfolio 1 - 5, which takes a long position in portfolio 1 and a short position in portfolio 5. 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. As typical in the applications of this test, the supremum of the Chow break test statistics is computed over the period January 2011 - August 2018, such that the first and last 15% of dates are not considered as structural break candidates. The critical values of the associated non-standard distribution are reported for the 10%, 5% and 1% size test levels.

Table 6 shows the test results for 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 is computed using the rank neutralised ESG scores. 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, since none of the test statistics exceeds the critical values. This means that the earlier suspicion of a structural break can be put at rest, even though the evidence of considerable variation in estimated alphas uncovered in Figure 4 are a reason for some concern.

In light of this evidence of stability over time, we now assess the potential of ESG momentum to represent a new systematic factor. The ESG momentum factor is proxied by the long-short portfolio 1 - 5, where the 1-month ESG momentum is computed with the rank neutralised ESG scores. Table 7 shows the Pearson correlation matrix of the ESG momentum factor and the seven other risk factors employed so far as well as the t-statistics of these correlations. The lower triangle of the matrix represented in the table displays the sample correlations while the upper triangle (printed in italics) shows the corresponding t-statistics (see, e.g., Sheskin, 2000). The table shows a number of interesting features. First, the ESG momentum factor is essentially uncorrelated with the low volatility factor. This is surprising at first 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 Low Vol and ESG momentum factors are indeed implicitly, indirectly correlated.[26]

Table 7 also shows that the ESG momentum factor is negatively correlated with the HML factor and positively with the RMW. These correlations suggest that high ESG momentum companies are profitable, growth companies. An intuitive argument would be that these companies are new companies since they are still growing. Since many established companies tend to be typically heavily invested already while increasing their ESG score often requires additional, significant investments, profitable growth stocks do fall at the intersection between wanting to become more sustainable and having the means to accomplish it.

Table 7: Correlation matrix of the ESG Momentum factor vs. seven other systematic risk factors

	ESG Momentum	Mkt - Rf	SMB	HML	RMW	CMA	Mom.	Low Vol.
ESG Momentum	1.00	*-1.31*	*-1.81*	*-3.39*	*2.01*	*-0.21*	*1.88*	*0.71*
Mkt - Rf	-0.12	1.00	*0.09*	*6.28*	*-4.28*	*0.66*	*-4.47*	*-11.41*
SMB	-0.16	0.01	1.00	*0.42*	*-0.59*	*-1.12*	*-0.23*	*-4.26*
HML	-0.29	0.49	0.04	1.00	*-14.19*	*8.74*	*-6.35*	*-7.96*
RMW	0.18	-0.36	-0.05	-0.78	1.00	*-5.70*	*5.25*	*7.07*
CMA	-0.02	0.06	-0.10	0.61	-0.45	1.00	*-1.40*	*-0.86*
Momentum	0.16	-0.37	-0.02	-0.49	0.42	-0.12	1.00	*5.47*
Low Vol.	0.06	-0.71	-0.35	-0.58	0.53	-0.08	0.44	1.00

This table reports the correlation matrix of the ESG Momentum factor with seven other risk factors. The lower triangle shows the 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 highest 1-month ESG momentum, and taking a short position in the quintile portfolio containing stocks with the lowest 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).

[26] This result showcases the importance of using a standard, state-of-the-art factor pricing model to learn the true, underlying relationships instead of simply resorting to pairwise correlations, as disregarding other risk factors may lead to incorrect inferences on a pairwise relationship.

3.4 An ESG Momentum Tilting Strategy

The long-short portfolio built in subsection 3.3 has as the disadvantage that, by ignoring the information in quintiles 2-4, it ignores 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. Classic portfolio theory states that this is very unlikely to lead to efficient portfolios and 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 simple and straightforward tilting strategy which consists of computing the ESG momentum for every stock in our European equity universe and ranking them by ESG momentum. However, the stocks will now be ranked in ascending order, such that the stock with the highest ESG momentum has the highest rank and vice versa for the stocks with the lowest ESG momentum. 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 highest ESG momentum are over-weighted and the stocks with the lowest ESG momentum are under-weighted, relative to the equally weighted portfolio.

Table 8: Returns and ESG scores of tilting strategies based on 1, 3, 6 and 12-month ESG momentum

	Returns						ESG Score	
	Mean (%)	σ (%)	SR	SE(SR)	α	t-stat	Mean	Median
1	0.498	4.528	0.101	0.088	0.003	0.616	63.23	63.96
3	0.494	4.551	0.099	0.089	0.003	0.601	63.76	64.46
6	0.482	4.558	0.096	0.092	0.002	0.472	64.29	64.98
12	0.551	4.617	0.109	0.093	0.003	0.689	65.12	65.54

This table shows summary statistics of returns and ESG scores for a tilting strategy based on 1-, 3-, 6-, and 12-month ESG momentum. The ESG momentum is computed using the rank neutralised ESG scores. The tilting strategy consists of ranking every stock in the investment universe based on their x-month ESG momentum, where $x \in \{1, 3, 6, 12\}$. The x-month ESG momentum is the percentage change of the stock's ESG score over the last x months. 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 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).

Table 8 shows two important results. First, all average abnormal returns are insignificant. In fact, none of the strategies manages to generate a significant SR. Therefore, the tilting strategies fail to earn significant, risk-adjusted returns. Second, the 12-month ESG momentum implies now the highest mean return, the highest SR, and the highest ESG score mean and median. This result differs from our earlier finding that 1-month momentum would guarantee the highest realized performances.[27] Finally, the mean and median ESG scores 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.[28]

3.5 Robustness Checks

We performed afresh the exercise in subsection 3.3 with reference to alternative definitions of ESG momentum, i.e., when the number of months to estimate momentum exceed one. Table 9 shows the estimated alphas and Gibbons et al. (1989) test statistics for portfolios sorted by 1-, 3-, 6-, and 12-ESG momentum respectively. In this case, ESG momentum is computed using the rank neutralised ESG scores.

Table 9: Robustness to momentum estimation period length under rank neutralised ESG scores

	1-Month		3-Month		6-Month		12-Month	
	α	t-stat	α	t-stat	α	t-stat	α	t-stat
1	0.592	2.225	0.558	2.057	0.0450	1.588	0.359	1.393
2	0.461	1.789	0.416	1.635	0.416	1.703	0.244	0.864
3	0.418	1.482	0.333	1.317	0.328	1.376	0.327	1.312
4	0.411	1.626	0.437	1.614	0.372	1.542	0.432	1.798
5	0.112	0.408	0.278	1.095	0.344	1.334	0.237	0.860
1 - 5	0.484	3.343	0.270	2.126	0.109	0.885	0.120	1.094
GRS test	2.412 (0.040)		1.015 (0.412)		0.539 (0.746)		0.911 (0.476)	

This table performs a robustness analysis with respect to the choice of the ESG momentum estimation period. It reports the estimated alphas for the ESG momentum portfolios sorted by 1-, 3-, 6-, and 12-month ESG momentum respectively. Momentum is computed using rank neutralised 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 momentum 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 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).

[27] Nonetheless, no selection of the length of the estimation period delivers significant alphas. 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.

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

Table 9 shows that the pricing error for portfolio 1, the portfolio with the highest ESG momentum, decreases monotonically both in level and in terms of the implied t-statistic when the computation period increases. The same applies to the long-short portfolio 1 - 5. This confirms our earlier suspicion that the choice of momentum computation period is a very delicate one as the statistical significance of average abnormal returns on the long-short portfolio disappears entirely for formation periods longer than 3 months.[29]

Table 10 shows the same robustness analysis applied to ESG momentum computed from the standardised ESG scores. The table reveals that, as one would expect, none of the ESG momentum estimation periods lead to a significant GRS statistics. Therefore, the null hypothesis of $\boldsymbol{\alpha} = 0$ cannot be rejected and we conclude that none of these alternative momentum estimation windows eventually manages to earn jointly significant risk-adjusted returns. This is unsurprising when one considers that the individual series already displayed insignificant alpha.

Table 10: Robustness to momentum estimation period length under standardised ESG scores

	1-Month		3-Month		6-Month		12-Month	
	α	t-stat	α	t-stat	α	t-stat	α	t-stat
1	0.285	1.391	0.414	1.609	0.499	1.859	0.384	1.270
2	0.411	1.420	0.287	1.031	0.300	1.449	0.317	1.249
3	0.290	1.184	0.534	1.916	0.404	1.606	0.388	1.367
4	0.541	1.909	0.418	1.506	0.377	1.442	0.280	1.265
5	0.524	1.697	0.514	1.787	0.388	1.350	0.287	1.029
1 - 5	-0.097	-1.180	-0.082	-0.919	0.158	1.850	0.078	1.066
GRS test	0.958 (0.447)		1.028 (0.405)		0.797 (0.554)		0.470 (0.798)	

This table performs a robustness analysis with respect to the choice of the ESG momentum estimation period. It reports the estimated alphas for the ESG momentum portfolios sorted by 1-, 3-, 6-, and 12-month ESG momentum respectively. Momentum 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 momentum 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 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).

[29] This is in line with what reported in Table 3, where the SR displayed a similar patter and even turned negative for the 6-month and 12-month ESG momentum definitions.

4 ESG Volatility as a New Systematic Risk Factor

In this section, we ask whether instead of the trend in ESG scores, it is possible to capitalize on their volatility to build a new, systematic risk factor priced in the European cross section of stocks. As with the ESG momentum, portfolios are sorted by a variety of ESG score volatility computation periods. Then, we perform the Fama and MacBeth (1973) cross-sectional regressions, estimate the alphas earned by the quintile portfolios sorted by ESG volatility and analyse their behaviour over time. The section concludes with the proposal of a tilting strategy based on the ESG volatility signal.

4.1 Portfolios Sorted by ESG Volatility

ESG volatility is defined as the volatility of the ESG score over a time window. When assessing x-month ESG volatility we consider a set of volatility windows where $x \in \{18, 24, 30, 36\}$ to follow Blitz and Van Vliet (2007). The x-month ESG volatility is computed as the sample standard deviation of the ESG score over the x months. Stocks are sorted by ascending ESG volatility so that the ones sorted in the highest ranked quintile display the lowest ESG volatility. Then, we form equally weighted quintile portfolios where portfolio 1 contains stocks with the most stable ESG scores while in portfolio 5 there are stocks with the most volatile scores over the previous x months.

Table 11 shows that for two estimation windows (18 and 24 months) the long-short portfolio is characterized by a significant SR. Therefore, we focus the bulk of our attention on results concerning portfolios sorted on 24-month ESG score volatility since it reports the highest SR. Also compared to the table concerning ESG momentum (Table 3), the SR turns out to be higher (as much as 0.254 per month) and imply higher t-statistics. This is a favourable early sign for a possible abnormal performance by ESG score volatility-sorted portfolio. Another interesting result is that the standardised ESG score volatility does not highly correlate with raw ESG volatility. For instance, the portfolio with the lowest standardised ESG score volatility has the highest overall volatility of the raw ESG scores. In fact, all average and median raw ESG scores across quintiles tend to increase with the ESG volatility within the portfolios. This is problematic for sustainability-driven investors, as this implies that investing in stable ESG scores characterized by low ESG volatility leads to lower investors' portfolio ESG scores.[30]

[30]The difference in the mean ESG score between portfolio 1 and 5 is even greater than it was for the ESG momentum computed with the standardised ESG scores from the previous section. A sustainability-driven investor aims at overweighting stocks characterized by relatively high ESG (raw) scores or to exclude stocks with low ESG (raw) scores.

Table 11: Results for portfolios sorted by ESG volatility using four computation periods

	Returns				ESG Score			
	Mean (%)	σ (%)	SR	SE(SR)	Mean	Median	σ	Mean Vol.
	Panel A: 18-month ESG Volatility							
1	0.600	4.369	0.138	0.102	63.53	63.23	2.50	0.103
2	0.654	4.275	0.153	0.096	63.87	63.99	1.88	0.155
3	0.582	4.528	0.129	0.098	64.81	65.31	1.73	0.202
4	0.569	4.503	0.127	0.098	65.92	66.47	1.90	0.267
5	0.325	4.747	0.069	0.097	66.27	66.53	2.05	0.427
1 - 5	0.227	1.165	0.196	0.074				
	Panel B: 24-month ESG Volatility							
1	0.697	4.375	0.160	0.107	63.87	63.48	2.21	0.122
2	0.636	4.602	0.138	0.100	64.61	64.53	1.39	0.179
3	0.521	4.414	0.118	0.104	65.82	65.94	1.48	0.230
4	0.606	4.653	0.131	0.105	66.35	66.61	1.53	0.299
5	0.343	4.825	0.071	0.104	66.89	67.02	1.74	0.465
1 - 5	0.303	1.196	0.254	0.087				
	Panel C: 30-month ESG Volatility							
1	0.835	4.448	0.188	0.107	64.13	63.35	1.90	0.138
2	0.904	4.310	0.210	0.104	65.47	65.36	1.04	0.199
3	0.763	4.221	0.181	0.105	66.31	66.20	1.15	0.253
4	0.571	4.651	0.123	0.107	66.84	66.95	1.11	0.326
5	0.666	4.572	0.146	0.110	67.46	67.80	1.39	0.497
1 - 5	0.116	1.171	0.099	0.104				
	Panel D: 36-month ESG Volatility							
1	0.761	4.488	0.170	0.103	64.44	63.76	1.75	0.151
2	0.864	4.242	0.204	0.100	65.83	65.67	1.18	0.216
3	0.670	4.281	0.157	0.105	66.87	66.84	0.95	0.273
4	0.599	4.525	0.133	0.104	67.19	67.40	1.11	0.350
5	0.622	4.394	0.142	0.108	67.90	68.06	1.32	0.524
1 - 5	0.083	0.996	0.084	0.106				

The five quintiles are sorted by ESG score volatility, computed using the standardised 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.

Table 12: Summary statistics for quintile portfolios sorted by 24-month ESG score volatility

	Returns				ESG Score			
	Mean (%)	σ (%)	SR	SE(SR)	Mean	Median	σ	Mean Vol.
1	0.633	4.490	0.141	0.101	65.62	65.81	1.29	0.025
2	0.557	4.405	0.127	0.101	64.84	64.40	1.94	0.047
3	0.618	4.620	0.134	0.105	65.12	65.27	1.20	0.065
4	0.525	4.775	0.110	0.104	65.06	65.23	1.31	0.090
5	0.424	4.774	0.089	0.105	66.70	67.06	2.11	0.147
1 - 5	0.158	1.066	0.149	0.071				

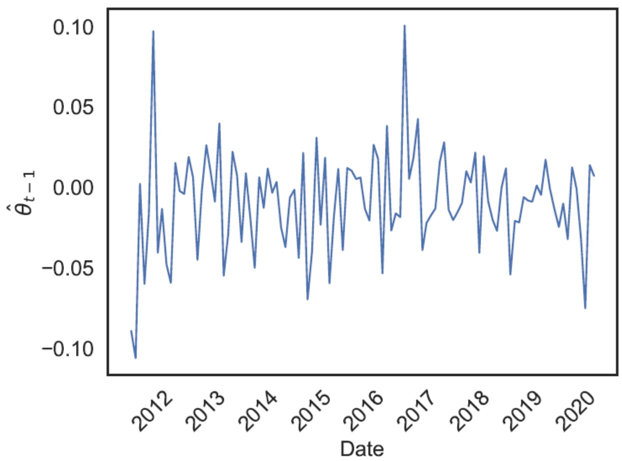
The five quintiles are sorted by ESG score volatility. The ESG scores are standardised to neutralize both size and sector tilts. The ESG volatility is computed using the rank neutralised ESG scores. The table also provides the summary statistics for a portfolio 1 - 5, which takes a long position in portfolio 1 and a short position in portfolio 5.

Table 12 shows the same summary statistics as Table 11 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.[31] We focus on an estimation window of 24 months because this is the only period for which the long-short portfolio earns a significant SR. In Table 12 as in Table 11 we find the same peculiarity of a lower mean and median ESG scores for portfolio 1 compared to portfolio 5. This result has some 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.

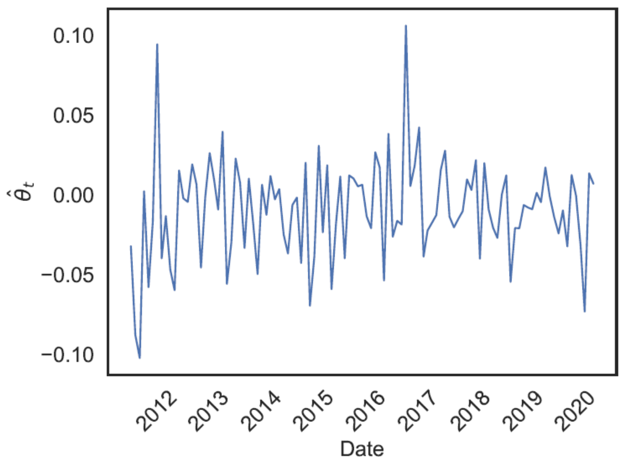
4.2 Regression Analysis of the ESG Volatility Characteristic

The Fama and MacBeth (1973) cross-sectional regressions consist of regressing the excess portfolio returns of the portfolios sorted by ascending ESG volatility on a constant and the corresponding ESG volatility. Equation 5 and Equation 6 regress $r_{i,t}$, the excess return of portfolio i at time t, on $c_{i,t}$ the 24-month ESG volatility of security i at time t. The difference between the two models stems from the fact that Equation 5 uses the 24-month ESG volatility of the previous month as a predictor. Also in this case, the parameters of the model are estimated with cross-sectional regressions at every point in time to allow for changing effects throughout the sample, hence the subscript $t-1$ in θ_{t-1}.

[31] The interested reader can find the full set of results in Table A.3 in the Appendix.



(a) Dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1.



(b) Dynamics of the cross-sectional OLS estimate of θ_t in Equation 2.

Figure 5 shows the dynamics of the Fama and MacBeth (1973) cross-sectional OLS estimates of the regression of excess portfolio returns on their 24-month ESG volatility and a constant. The portfolios are sorted by the ascending 24-month ESG volatility. Panel (a) shows the dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1. In this equation $c_{i,t-1}$ is portfolio i's 24-month ESG volatility at time $t-1$ and $r_{i,t}$ is the excess return of portfolio i at time t. Panel (b) shows the dynamics of the cross-sectional OLS estimate of θ_t in Equation 2.

$$r_{i,t} = a_t + \theta_{t-1} c_{i,t-1} + \varepsilon_{i,t} \tag{5}$$

$$r_{i,t} = a_t + \theta_t c_{i,t} + \varepsilon_{i,t} \tag{6}$$

Figure 5 shows the evolution of the estimated θ coefficients from the above equations. The two panels present similar patterns, which is unsurprising given the stickiness (caused by stability accompanied by large sample autocorrelations) of the ESG scores. A visual inspection of Figure 5 reveals that the estimated effect is negative on average, even though it is noisy. This is in line with the expectation based on the results in Table 11, in which the higher the ESG volatility is, the lower the realized average portfolio return. This suggests that a low ESG volatility anomaly may indeed exist. When we compute the t-statistic for the estimated average θ_{t-1} in Equation 5, we obtain a value of -2.991. Furthermore, the t-statistic for the estimated average θ_t in Equation 6 yields a value of -3.063. Both these t-statistics are significantly different from zero at a 1% test size or lower. This is consistent with the existence of a negative relationship between ESG score volatility and excess portfolio returns. This is preliminary evidence that ESG low volatility may represent a source of systematic risk.

Fama and MacBeth (1973) regressions are performed afresh when ESG volatility is computed for the standardised ESG scores. To save space, the plots for the series of estimated coefficients are reported in the Appendix. In this case, the t-statistic for the estimated average θ_{t-1} is -2.118; the t-statistic for the estimated average θ_t is -2.099. These two t-statistics are lower than their counterparts for portfolios sorted by ESG volatility computed using the rank neutralised ESG scores. However, both average effects are still significantly different from zero at a test size of 5%. These results, combined with the fact that, regardless of which score neutralisation approach we use, the optimal ESG volatility estimation period selected is always the same, suggest that the ESG volatility may be a more robust source of abnormal performance compared to ESG momentum.

4.3 ESG Volatility and Abnormal Portfolio Performance

As usual, it remains vital to distinguish whether a strategy earns return because of its implied exposure to one ore more risk factors, or whether it could earn true risk-adjusted abnormal returns. To this end, we adopt the same seven-factor linear model as in section 3, which includes tradeable proxies for market, size, value, profitability, investment, momentum, and volatility risks. Table 13 shows in panel A the OLS estimation results for the factor loadings for each of the quintile portfolios built in subsection 4.1, while the associated t-statistics are reported in panel B.

The table shows that the average pricing errors for the quintile portfolios are insignificant at a size test level of 5%. Therefore, the null hypothesis of a zero alpha cannot be rejected for any of the quintile portfolios. However, the alpha for the long-short portfolio is highly significant, even at a 1% size: the purest form of the ESG low volatility signal earns significant alpha. Moreover, consistently with the results obtained in section 3, Table 13 implies high and significant estimates of the market factor exposures for all the quintile portfolios. However, the effect disappears for the long-short portfolio, illustrating that the pure ESG low volatility signal is indeed not significantly affected by market risk for any reasonable confidence level.[32]

Additionally, Table 13 shows a negative and significant estimate for the Low Volatility factor for every quintile portfolio. However, this significantly negative relationship disappears for the long-short portfolio 1 - 5, As a result, the ESG low volatility factor is crucially not significantly explained by or correlated with the classical low volatility anomaly and the resulting factor. This emphasizes that, similarly to how ESG momentum and classical momentum were completely unrelated, the low volatility and ESG low volatility factors are separate factors. This reveals one of the strengths of

[32] In section 3, we speculated that the significant exposure of the quintile portfolios sorted by ESG momentum to market beta risk may be caused by the granularity of the quintile portfolios. The fact that the quintile portfolios sorted by ESG volatility display the same exposure strengthens this reasoning, since these portfolios exhibit the same granularity as the portfolios sorted by ESG momentum.

ESG-linked strategies: they appear to be uncorrelated with traditional factors and therefore may generate substantial diversification gains.

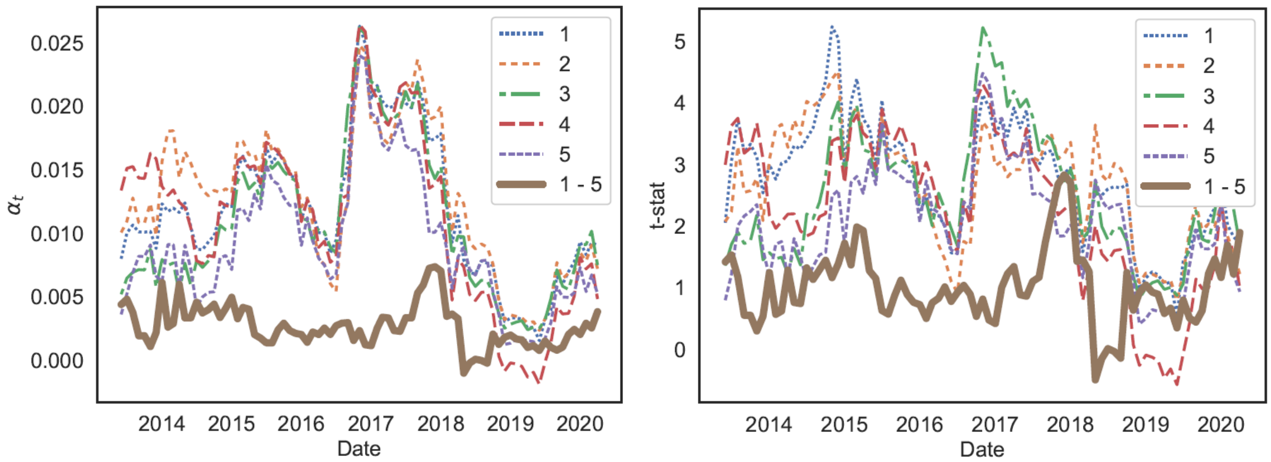
Table 13: Estimates of regressions of ESG volatility-sorted portfolio excess returns on 7 risk factors

	α	Mkt. - Rf	SMB	HML	RMW	CMA	Mom.	Low Vol.	R^2
				Panel A: Regression Coefficients					
1	0.441	0.677	-0.063	0.063	0.724	0.014	-0.119	-0.261	0.780
2	0.436	0.725	0.045	0.074	0.595	0.192	-0.122	-0.248	0.793
3	0.364	0.700	-0.119	0.076	0.570	0.193	-0.124	-0.248	0.808
4	0.430	0.707	-0.163	0.053	0.555	0.110	-0.087	-0.320	0.803
5	0.180	0.734	-0.025	0.122	0.594	0.138	-0.110	-0.317	0.829
1 - 5	0.261	-0.058	-0.038	-0.059	0.130	-0.123	-0.009	0.056	0.383
				Panel B: t-statistics					
1	1.855	7.318	-0.258	0.216	1.852	0.059	-1.694	-2.268	
2	1.785	7.838	0.185	0.246	1.551	0.812	-1.714	-2.026	
3	1.673	9.049	-0.556	0.262	1.693	0.880	-1.824	-2.354	
4	1.733	7.215	-0.740	0.217	1.648	0.406	-1.088	-2.693	
5	0.695	7.287	-0.100	0.380	1.548	0.721	-1.907	-2.889	
1 - 5	3.036	-1.312	-0.595	-0.630	1.109	-0.952	-0.220	1.108	

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 the ascending 24-month ESG volatility, the volatility of the ESG score over the last 24 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 OLS estimates and the R^2 of the regressions. Panel B reports the t-statistics of the OLS estimates in panel A. In the table, the standard errors are estimated using the heteroskedasticity and autocorrelation consistent estimator (HAC) proposed by Newey and West (1987).

Finally, Table 13 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, i.e., Jensen's alpha, which is highly significant (the t-stat is 3.04) and measures a rather massive risk-adjusted excess return earned by the ESG low volatility long-short portfolio.

Because none of the individual quintile portfolios earns a significant alpha at a 5% size, it seems prudent to also conduct the GRS test in Equation 4 as an encompassing way to ask whether the



(a) Dynamics of the rolling OLS estimate of α in Equation 9.

(b) Dynamics of the t-statistic of the rolling OLS estimate of α in Equation 9.

Figure 5 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 volatility of the ESG score over the past 24 months. The risk factors incorporated in the OLS regressions are the five factors Mkt - Rf, SMB, HML, RMW, and CMA as proposed by Fama and French (2015), the Momentum factor proposed by Carhart (1997), and the Low Volatility factor, as proposed by Blitz and Van Vliet (2007).

postulated seven-factor model is able to price the returns on the portfolios sorted by ESG volatility. Computing this test for the quintile portfolios sorted by ESG volatility yields a statistic of 1.286. The probability that this value is exceeded under the null hypothesis is 0.277. Since this is higher than the critical value at a 5% test size, the null hypothesis of $\boldsymbol{\alpha} = 0$ cannot be rejected. Therefore, the hypothesis that the asset pricing model is capable of pricing the assets is not rejected.

We have also analyzed abnormal returns when the sorting of stocks is performed using ESG volatility computed from the standardised ESG scores. Table A.4 in the Appendix reports insignificant alpha estimates for the quintile portfolios at a 5% size. However, it exhibits a significant alpha estimate for the long-short portfolio even if insignificant estimates for the factor loadings.[33] More generally, the results for the variance of the standardised ESG scores are fully consistent with those obtained for the rank neutralised scores. The GRS test for the joint significance of the alphas yields a statistics of 1.518 to which it is associated a p-value of 0.192, also in this case the null that the vector of alphas is equal to the zero vector cannot be rejected.

To investigate whether the insignificant alphas of the quintile portfolios are structural to the entire sample or whether ESG low volatility just recently became important, we now analyze the dynamics of the portfolio pricing errors of the portfolios over time. Because the long-short portfolio formed by

[33] Also in this case, estimated coefficients for the quintile portfolios are precisely estimated in the case of the market and the classical volatility factors.

spread of the ESG score volatility computed with the rank neutralised ESG scores led to the highest t-statistics, we proceed under a 24-month volatility estimation period. Figure 5 shows the dynamics of the rolling 24 month alpha estimates for each of the quintile portfolios and the long-short portfolio. After some initial (2012-2015) period of volatility, the estimates tend to converge to common values towards the end of the sample. Especially in the first half of the sample, the alpha's estimates of 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 estimate reported in Table 13 . 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.

The findings in Figure 5 justify performing Andrews (1993) tests for the presence of structural breaks at an unknown time in the model. Similarly to the tests conducted in section 3, after accounting for the presence of a constant, we use eight regressors and truncate a total of 15% of the data on both sides of the sample to conduct this test. The null hypothesis of this test is that there is no structural break in the model. Table 14 reports the test statistics and the critical values, obtained from Andrews (1993), for 10%, 5%, and 1% size levels. None of the test statistics exceeds the critical value for any of the confidence levels. Therefore, the null hypothesis that there is no structural break in the model is not rejected. Therefore, there is no statistically significant evidence indicating that the asset pricing model used incorrectly prices the assets because of structural instability.

Table 14: Andrews (1993) test for a structural break at an unknown break date for six portfolios

	Andrews Test Statistic	Critical Value		
		10%	5%	1%
1	6.03	19.82	22.13	27.25
2	4.74	19.82	22.13	27.25
3	5.32	19.82	22.13	27.25
4	5.80	19.82	22.13	27.25
5	6.21	19.82	22.13	27.25
1 - 5	2.30	19.82	22.13	27.25

This table shows the Andrews (1993) test for a structural break at an unknown break date for five quintile portfolios sorted by ascending 24-month ESG volatility. The null hypothesis is the absence of a structural break. Additionally, the table provides the Andrews (1993) test statistic for the long-short portfolio 1 - 5, which takes a long position in portfolio 1 and a short position in portfolio 5. 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 values of the associated non-standard distribution are reported for the 10%, 5% and 1% size test levels.

4.4 ESG Volatility Tilting Strategies and Robustness Checks

As in subsection 3.4, testing the ESG low volatility signal by sorting quintile portfolios according to ascending ESG volatility has the disadvantage that the signals of 60% of the universe are ignored. We therefore proceed investigating the ESG volatility signal through a simple tilting strategy identical in spirit to the one in subsection 3.4.[34] The tilting strategy can be interpreted as a "signal weighted" strategy.

Table 15 displays summary statistics for the ESG volatility tilting strategy for different ESG volatility computation periods. The table shows that none of the tilting strategies earns a precisely estimated alpha. Yet, this finding is not totally unexpected after the earlier results of the GRS test. The GRS test investigated the joint significance of the mean abnormal returns of all the portfolios

[34] It consists of computing the x-month ESG volatility for every stock in the universe and then in ranking all stocks in descending order, such that the stocks with the lowest ESG volatility have the highest numerical value. Subsequently, we compute the percentile to which the stock belongs. This gives every stock a value between 0 and 1, where 0 is assigned to the stock with the highest ESG volatility and 1 to the stock with the lowest ESG volatility. The percentiles are then summed and every stock's percentile is divided by the total, ensuring that the portfolio weights sum to one. Adopting this strategy, we enforce that the lowest ESG volatility stocks are over-weighted in the portfolio and that the highest volatility stocks are under-weighed compared to the equally weighted portfolio.

sorted by ESG volatility. The test thus considered the signals of all stocks.

Table 15: Returns and ESG scores of tilting strategies based on 18, 24, 30, and 36-month ESG volatility

	Returns						ESG Score	
	Mean (%)	σ (%)	SR	SE(SR)	α	t-stat	Mean	Median
18	0.699	4.361	0.149	0.098	0.005	1.122	64.29	64.40
24	0.641	4.440	0.133	0.100	0.004	0.893	64.89	64.75
30	0.860	4.308	0.187	0.105	0.006	1.573	65.41	65.15
36	0.854	4.305	0.185	0.102	0.006	1.240	65.78	65.53

This table shows summary statistics of returns and ESG scores for a tilting strategy based on 18-, 24-, 30-, and 36-month ESG volatility. The tilting strategy consists of ranking every stock in the investment universe based on their x-month ESG volatility, where $x \in \{18, 24, 30, 36\}$. The x-month ESG volatility is the volatility of the stock's ESG score over the last x months. The ESG volatility is computed using the standardised ESG scores. 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 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 shows that when it comes to tilting strategies, the 24-month ESG volatility 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 is an encouraging signal in a sustainable investing perspective. In fact, from the point of view of an ESG volatility tilting strategy, the traditional trade-off between mean realized returns and ESG scores fails to emerge. The longer the ESG volatility estimation period, the higher the SR and the higher the mean and median ESG score, even though the estimated SR and alpha are not statistically significant. Interestingly, this was not the case for the quintile portfolios.[35] However, Table 11 also 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.

We have also repeated the tests in Table 15, using the ESG volatility computed from the rank neutralised ESG scores. An unreported table (available upon request from the authors) shows the same main findings as in Table 15. However, the 36-month ESG score volatility tilting strategy yields the highest risk-adjusted return, even though with not statistically significant alphas or SR.

[35]For instance, in Table 11, portfolio 1 always had a lower ESG mean and median than portfolio 5.

Table 16: Robustness to volatility estimation period length under standardised ESG scores

	18-Month		24-Month		30-Month		36-Month	
	α	t-stat	α	t-stat	α	t-stat	α	t-stat
1	0.298	1.383	0.432	1.855	0.408	1.391	0.368	1.198
2	0.432	1.614	0.377	1.785	0.487	2.081	0.614	2.310
3	0.404	1.469	0.378	1.673	0.425	1.828	0.441	1.520
4	0.399	1.853	0.425	1.733	0.097	0.463	0.309	0.945
5	0.087	0.410	0.224	0.695	0.284	1.228	0.340	1.016
1 - 5	0.211	2.468	0274	3.036	0.034	0.447	0.104	0.804
GRS test	1.510 (0.194)		1.286 (0.277)		1.983 (0.089)		1.564 (0.180)	

This table performs 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 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)

Given our earlier results, it is also crucial to check whether the imprecise estimation of the portfolio alphas may be caused by the choice of the volatility estimation window, similarly to what found in section 3. Table 16 shows the average pricing errors and corresponding t-statistics for the five quintile portfolios and the long-short portfolio for four ESG volatility computation periods. Moreover, it displays the GRS statistics for testing the joint significance of the quintile portfolios' average abnormal returns. The table shows rather heterogeneous results. The most prominent feature is that for portfolios sorted by ascending 30-month ESG volatility, the alphas earned by the portfolios are jointly significant, albeit only at a 10% size. This confirms that the significance of the SR may not represent the optimal selection criterion for the identification of the new ESG risk factor. Moreover, Table 16 exhibits that in the cases of 30- and 36-month ESG volatility, the first three portfolios earn in the aggregate higher mean returns than they do under the other estimation periods (18- and 24-month). This explains the good performance of the 30- and 36-month ESG volatility tilting strategies observed before. Since the first three portfolios are generally being over-weighted and the last two under-weighted, the returns of the former portfolios dominate total performance.

We also investigate the alphas of the portfolios sorted by ESG volatility computed with the rank neutralised ESG scores. A table in the Appendix shows the estimated alphas and corresponding t-statistics. The results are closely aligned to those in Table 13: the 30-month ESG volatility yields the highest GRS test statistic, although this fails to be statistically significant using a 10% size test. The marginal evidence of joint significance of the alphas shown earlier is therefore not robust to the data neutralisation method used. The table also emphasizes that the alpha earned by the long-short portfolio for the 24-month ESG volatility signal is not the only long-short portfolio that earns a precisely estimated alpha. The 30-month and 36-month ESG volatility sorts also yield long-short portfolios that earn significant alpha. This is another example of the failure of the Sharpe ratio to properly capture risk exposure.

5 Concluding Remarks

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 growing and stable sustainable investing company profiles, as captured by the Sustainalytics ESG scores, can be profitable and generate risk-adjusted returns that ought to attract resources towards firms that are steadily improving their ESG ratings.

We start out by analysing and dealing with ESG data issues. We propose two novel ways to neutralise the size and sector biases typically present in ESG score data. The first method group all the stocks in size quintiles. We then compute the median ESG score within each size quintile and subtract this median score. Next, we compute the percentile to which every stock belongs, giving every stock a value between 0 and 1. We apply a similar procedure to size-neutralised ESG scores, 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. Then, we compute the average ESG score and the standard deviation of the ESG score within every size quintile. The stocks are then standardised using the mean and standard deviation corresponding to the size quintile that the stock belongs to. After standardising the ESG scores by firm size, we standardise the size-standardised ESG scores by sectors using a similar procedure. These methods ensure that the size and sector biases present in the data set are neutralised and no longer clumped with the ESG signals.

After correcting for size and sector dependencies in the data in the ways mentioned above, 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 decreasing ESG momentum with reference to a range of computation periods. We estimate cross-sectional regressions to investigate the average effect of ESG momentum on the excess returns of portfolios sorted by ESG momentum. Next, we analyse the abnormal returns of portfolios sorted by decreasing ESG momentum under a standard linear factor model that includes the five Fama and French (2015) factors, the Carhart (1997) momentum factor, and the Blitz and Van Vliet (2007) 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, only correlated significantly with the classical, Blitz and Van Vliet (2007) low volatility equity factor.

However, at other estimation horizons, ESG momentum appears to be fragile because it fails to yield significant alphas. Therefore, we extend our analysis to a related but distinct ESG score low volatility signal. We simply define ESG volatility as the sample standard deviation of the ESG score of a stock over a certain 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 then 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, but the test of Gibbons et al. (1989) shows that the quintile portfolios do not earn jointly significant risk-adjusted mean returns. Therefore, ESG low volatility may not represent a source of systematic risk. Moreover, structural break tests fail to reject the null hypothesis of the absence of a break-point date in the estimated alphas. These results appear to be robust to the selection of the ESG volatility estimation periods.

More generally, we find that the method by which the ESG data is neutralised, significantly influences the returns earned by portfolios sorted by an ESG signal. For instance, we report that 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 3-month, 6-month or 12-month ESG momentum. Moreover, the significantly positive mean excess returns are not earned by portfolios sorted by ESG momentum if the ESG momentum is computed using the standardised ESG scores.

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

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

where $\rho = 1/[1 + \exp(\overline{d-p})]$, ($\overline{d-p}$ 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 the required, expected stock returns. Equivalently, current positive alphas forecast higher expected profitability and/or a lower equity cost of capital. Hence, our findings of positive and precisely estimated alphas from a number of implementations of ESG score momentum- and volatility-driven systematic strategies possibly map into an inference of declining cost of capital for firms characterized by strong ESG momentum and low ESG score volatility. This finding must be added to the commonly reported empirical result that high ESG scores would be associated with a lower expected cost of capital and hence higher, current realized stock returns (see, e.g., Pástor et al. (2022)). In a policy perspective, fostering a virtuous cycle of steadily (never excessively volatile) through rising ESG scores may yield the future expected social payoffs deriving from a low of cost of equity capital. On the other hand, policymakers may also derive a less optimistic implication from our findings. If firms were to grasp that immediate benefits (in terms of higher stock prices per unit of fundamentals) are available if any progress in the ESG score arena is kept muted and slowly occurring, this may quickly deprive the ESG scores from their very signalling nature as firms may learn how to "game" the system. Barring the possibility that regulations concerning ESG scoring by the corresponding, specialized agencies may be introduced soon, this represents a looming concern for investors as well as the market regulatory authorities (see the discussion in Redondo Alamillos and de Mariz (2022)).

The sensitivity of the results to the approaches for correcting size and sector biases illustrates the necessity of further research into this topic. In general, the ESG signals are fragile.[36] Therefore,

[36]Compared to standard financial data, there are no global regulations governing or requiring reporting on ESG

a statistically solid treatment of the problems in ESG data is most needed in the current quest to discover profitable ESG quantitative signals. Of course, it may be fruitful to investigate other ESG signals. For instance, a researcher could look at the percentage of revenue spent on ESG related issues and policies by each individual firm and create a resulting portfolio sorting. This might be interpreted as a more objective measure of a company's efforts to address ESG related risks.

References

Alliance, G. S. I. (2018). Global sustainable investment review.

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

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

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

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

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

Cantor, R. and Mann, C. (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., Guan, J., and Sun, Y. (2022). Is incorporating esg considerations costly? *Journal of Portfolio Management*, 48(7):75–87.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56.

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

data. ESG data typically are deficient in the following areas: quantity (i.e., not enough data are available); consistency (companies often report different ESG data items and it remains rare to find an ESG data item that is disclosed uniformly across a reasonably large investment universe); quality (the methodologies used to derive or compute the reported ESG data may vary across companies). See the discussion in Kotsantonis and Serafeim (2019)

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

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

Heij, C., Heij, C., de Boer, P., Franses, P. H., Kloek, T., van Dijk, H. K., et al. (2004). *Econometric Methods with Applications in Business and Economics*. Oxford University Press.

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

Jegadeesh, N. and Titman, S. (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.

Kotsantonis, S. and Serafeim, G. (2019). Four things no one will tell you about esg data. *Journal of Applied Corporate Finance*, 31(2):50–58.

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

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

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

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

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

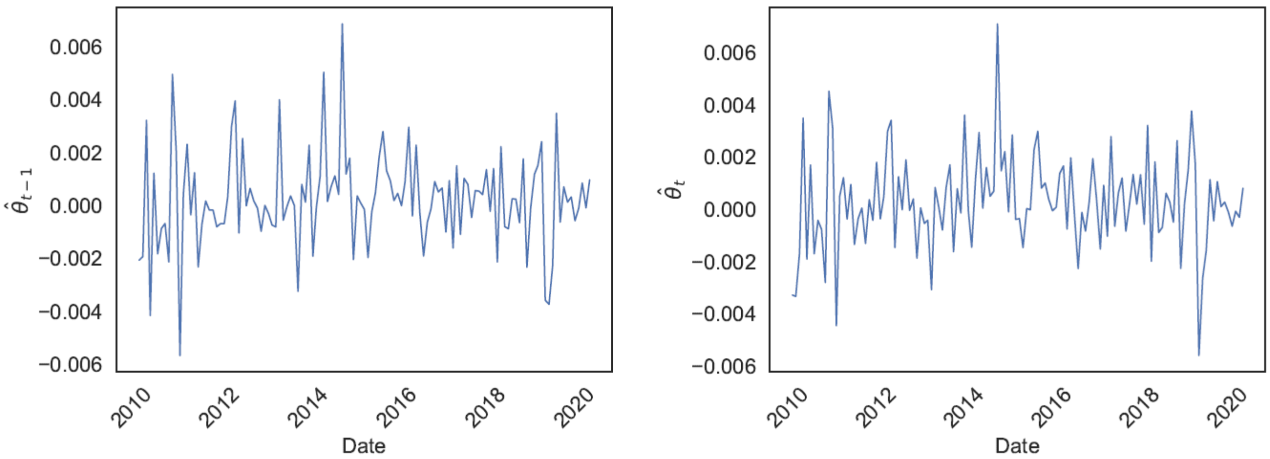
Sheskin, D. J. (2000). Parametric and nonparametric statistical procedures. *Chapman and Hall/CRC: Boca Raton, FL.*

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

Appendix

Table 17: Summary statistics for portfolios sorted by standardised ESG scores momentum

	Returns				ESG Score			
	μ (%)	σ (%)	SR	SE(SR)	μ	Med.	σ	mean Mom.
	Panel A: 1-Month ESG Momentum							
1	0.605	4.402	0.138	0.091	62.57	63.13	2.92	1.898
2	0.667	4.330	0.154	0.091	62.85	62.79	3.09	0.075
3	0.691	4.494	0.154	0.090	63.63	64.17	3.15	-0.009
4	0.754	4.511	0.168	0.087	63.20	63.19	3.47	-0.099
5	0.700	4.680	0.150	0.088	62.76	63.53	3.07	-1.948
1 - 5	-0.137	1.052	-0.129	0.086				
	Panel B: 3-Month ESG Momentum							
1	0.651	4.510	0.145	0.091	63.09	63.66	2.77	3.566
2	0.618	4.361	0.142	0.092	62.66	63.17	3.13	0.125
3	0.762	4.377	0.174	0.093	63.57	63.72	2.41	-0.024
4	0.703	4.552	0.155	0.085	63.43	63.75	3.33	-0.185
5	0.710	4.635	0.153	0.091	63.43	64.39	2.94	-3.487
1 - 5	-0.102	1.009	-0.101	0.095				
	Panel C: 6-Month ESG Momentum							
1	0.727	4.587	0.159	0.094	63.64	64.16	2.72	5.683
2	0.654	4.367	0.150	0.095	62.63	62.62	2.52	0.166
3	0.655	4.378	0.150	0.092	63.82	64.23	2.42	-0.046
4	0.655	4.520	0.145	0.093	63.86	64.28	2.79	-0.275
5	0.595	4.612	0.129	0.093	63.92	64.82	2.71	-4.698
1 - 5	0.088	1.015	0.087	0.095				
	Panel D: 12-Month ESG Momentum							
1	0.664	4.510	0.147	0.093	64.40	64.63	2.55	7.196
2	0.658	4.387	0.150	0.094	63.10	63.33	2.18	0.219
3	0.602	4.477	0.135	0.094	64.31	64.92	2.17	-0.088
4	0.610	4.415	0.138	0.095	64.36	64.85	2.20	-0.407
5	0.564	4.573	0.123	0.094	64.70	65.36	2.26	-6.508
1 - 5	0.054	1.079	0.050	0.074				



(a) Dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1.

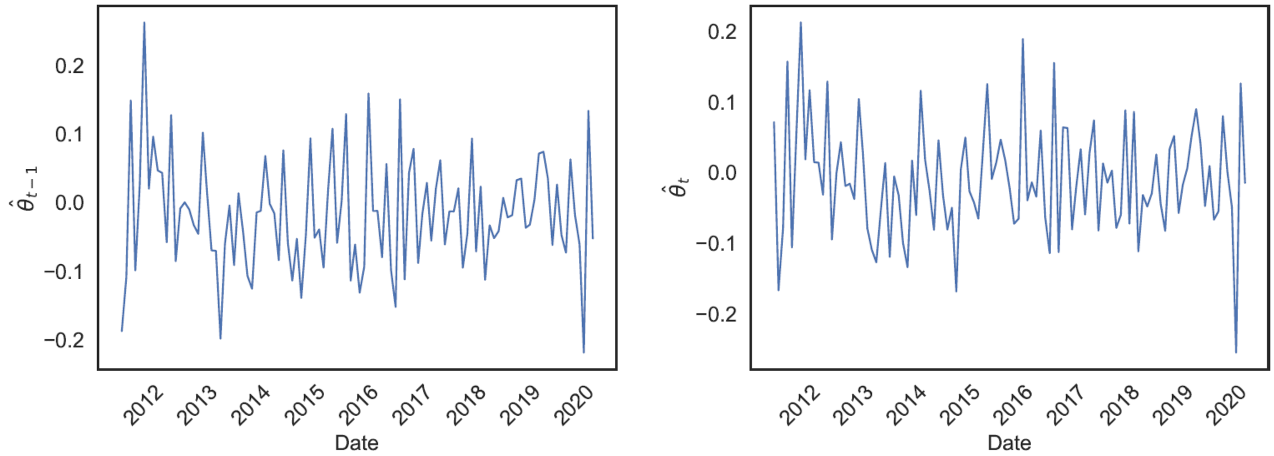
(b) Dynamics of the cross-sectional OLS estimate of θ_t in Equation 2.

Figure A.1 shows dynamics of the Fama and MacBeth (1973) cross-sectional OLS estimates of the regression of excess portfolio returns on their 1-month ESG momentum and a constant. The portfolios are sorted by the 1-month momentum computed on the basis of standardised ESG scores as the percentage increase of the ESG score over the last month. Panel (a) shows the dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1. In this equation, $c_{i,t-1}$ is portfolio i's 1-month ESG momentum at time $t-1$ and $r_{i,t}$ the excess return of portfolio i at time t. Panel (b) shows the dynamics of the cross-sectional OLS estimate of θ_t in Equation 2. In this case, $c_{i,t}$ is portfolio i's 1-month ESG momentum at time t and $r_{i,t}$ the excess return of portfolio i at time t.

Table 18: Estimates of regression of ESG momentum-sorted portfolio excess returns on 7 risk factors

	α	Mkt. - Rf	SMB	HML	RMW	CMA	Mom.	Low Vol.	R^2
	Panel A: Regression Coefficients								
1	0.513	0.626	-0.052	0.071	0.538	-0.008	-0.141	-0.302	0.778
2	0.346	0.607	0.004	0.109	0.586	-0.068	-0.063	-0.302	0.798
3	0.373	0.606	0.014	0.184	0.668	-0.092	-0.111	-0.267	0.786
4	0.381	0.630	-0.038	0.106	0.460	-0.173	-0.074	-0.289	0.792
5	0.357	0.621	-0.062	0.156	0.549	-0.125	-0.107	-0.298	0.775
1 - 5	0.156	0.005	0.010	-0.085	-0.011	0.118	-0.035	-0.004	0.015
	Panel B: t-statistics								
1	1.859	6.433	-0.195	0.241	1.530	-0.035	-1.961	-2.643	
2	1.449	6.925	0.018	0.419	1.839	-0.320	-0.997	-2.830	
3	1.606	6.751	0.055	0.671	2.070	-0.429	-1.581	-2.380	
4	1.442	6.446	-0.148	0.381	1.379	-0.779	-1.046	-2.484	
5	1.350	6.341	-0.217	0.519	1.501	-0.552	-1.606	-2.308	
1 - 5	1.850	0.189	0.169	-1.310	-0.125	1.294	-0.749	-0.113	

The table shows the OLS estimates and t-statistics for the regression of the excess portfolio returns on seven risk factors and a constant. The portfolios are sorted by their 6-month ESG momentum, the percentage increase of the ESG score over six months. The standardised ESG scores are used for the ESG momentum. 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 OLS estimates and the R^2 *of the regression. Panel B reports the t-statistics of the OLS estimates in panel A.*



(a) Dynamics of the cross-sectional OLS estimate of θ_{t-1} in Equation 1.

(b) Dynamics of the cross-sectional OLS estimate of θ_t in Equation 2.

Figure A.2 shows dynamics of the Fama and MacBeth (1973) cross-sectional OLS estimates of the regression of excess portfolio returns on their 24-month ESG volatility and a constant. The ESG volatility is computed using the rank neutralised ESG scores. Panel (a) shows the dynamics of the cross-sectional OLS estimate of θ_{t-1} *in Equation 1. In this equation* $c_{i,t-1}$ *is portfolio i's 24-month ESG volatility at time* $t-1$ *and* $r_{i,t}$ *the excess return of portfolio i at time t. Panel (b) shows the dynamics of the cross-sectional OLS estimate of* θ_t *in Equation 2. Here,* $c_{i,t}$ *is portfolio i's 24-month ESG volatility at time t and* $r_{i,t}$ *the excess return of portfolio i at time t.*

Table 19: Portfolios sorted by ESG rank-neutralized volatility using four computation periods

	Returns				ESG Score			
	μ (%)	σ (%)	SR	SE(SR)	μ	Med.	σ	mean Vol.
	Panel A: 18-Month ESG Volatility							
1	0.566	4.348	0.130	0.097	65.20	65.42	1.76	0.021
2	0.589	4.301	0.137	0.096	64.44	64.56	2.08	0.041
3	0.571	4.674	0.122	0.100	64.29	64.35	1.90	0.057
4	0.464	4.609	0.101	0.097	64.33	64.74	1.65	0.080
5	0.481	4.699	0.103	0.096	65.92	66.42	2.48	0.134
1 - 5	0.036	1.192	0.031	0.067				
	Panel B: 24-Month ESG Volatility							
1	0.633	4.490	0.141	0.101	65.62	65.81	1.29	0.025
2	0.557	4.405	0.127	0.101	64.84	64.40	1.94	0.047
3	0.618	4.620	0.134	0.105	65.12	65.27	1.20	0.065
4	0.525	4.775	0.110	0.104	65.06	65.23	1.31	0.090
5	0.424	4.774	0.089	0.105	66.70	67.06	2.11	0.147
1 - 5	0.158	1.066	0.149	0.071				
	Panel C: 30-Month ESG Volatility							
1	0.768	4.352	0.177	0.107	65.93	66.05	0.96	0.028
2	0.883	4.339	0.204	0.100	65.42	64.84	1.90	0.052
3	0.694	4.559	0.153	0.109	65.54	65.82	0.96	0.072
4	0.766	4.485	0.171	0.104	65.87	66.00	0.95	0.099
5	0.619	4.607	0.135	0.110	67.26	67.52	1.69	0.157
1 - 5	0.095	1.029	0.093	0.099				
	Panel D: 36-Month ESG Volatility							
1	0.832	4.322	0.193	0.103	65.91	65.91	0.78	0.032
2	0.678	4.363	0.156	0.100	66.07	65.80	1.66	0.057
3	0.789	4.415	0.179	0.107	65.67	66.03	1.08	0.078
4	0.583	4.453	0.131	0.103	66.68	66.88	1.01	0.106
5	0.637	4.483	0.142	0.106	67.70	67.67	1.58	0.166
1 - 5	0.139	0.921	0.152	0.101				

Table 20: Regression estimates of ESG score volatility-sorted portfolio excess returns on 7 risk factors

	α	Mkt. - Rf	SMB	HML	RMW	CMA	Mom.	Low Vol.	R^2
				Panel A: Regression Coefficients					
1	0.460	0.679	-0.035	0.102	0.670	0.150	-0.092	-0.284	0.790
2	0.351	0.743	-0.004	0.015	0.539	0.199	-0.127	-0.185	0.808
3	0.419	0.719	-0.101	0.119	0.620	0.047	-0.136	-0.255	0.805
4	0.431	0.734	-0.102	0.029	0.455	0.137	-0.122	-0.285	0.813
5	0.199	0.741	-0.040	0.171	0.658	-0.011	-0.077	-0.270	0.806
1 - 5	0.262	-0.063	0.005	-0.070	0.012	0.160	-0.015	-0.014	0.109
				Panel B: t-statistics					
1	1.897	7.897	-0.156	0.363	1.964	0.694	-1.268	-2.667	
2	1.628	8.866	-0.019	0.055	1.703	0.949	-1.943	-1.646	
3	1.696	7.583	-0.429	0.428	1.694	0.194	-1.847	-2.126	
4	1.854	7.501	-0.442	0.108	1.229	0.547	-1.681	-2.359	
5	0.790	7.524	-0.156	0.572	1.768	-0.050	-1.112	-2.507	
1 - 5	2.523	-1.808	0.071	-0.550	0.103	1.111	-0.235	-0.309	

Table 21: Robustness of abnormal returns to sorting by volatility using rank neutralised ESG scores

	18-Month		24-Month		30-Month		36-Month	
	α	t-stat	α	t-stat	α	t-stat	α	t-stat
1	0.003	1.394	0.005	1.897	0.004	1.773	0.005	1.873
2	0.004	1.773	0.004	1.628	0.004	1.498	0.003	1.188
3	0.003	1.235	0.004	1.696	0.003	1.205	0.005	2.069
4	0.003	1.157	0.004	1.854	0.005	1.857	0.003	1.074
5	0.003	1.014	0.002	0.790	0.002	0.678	0.003	0.906
1 - 5	0.001	0.558	0.003	2.523	0.003	2.602	0.002	2.686
GRS test statistic	0.797 (0.554)		1.518 (0.192)		1.736 (0.135)		1.544 (0.186)	

The table shows a robustness analysis for the significance of the pricing errors with respect to the choice of the ESG volatility estimation period. ESG volatility is computed using the rank neutralised ESG scores. It reports the pricing errors for the portfolios sorted by 18-month, 24-month, 30-month, and 36-month ESG volatility respectively. Additionally, the table reports the Gibbons et al. (1989) test for joint significance of the pricing errors for every ESG volatility computation period. The corresponding p-values are reported below the test statistic in parentheses. For the estimation of the average abnormal returns, α, the excess portfolio returns are regressed on a constant and seven risk factors. The incorporated risk factors 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).