
Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: www.tandfonline.com/journals/rquf20

# The non-linear ESG premium

Runfeng Yang, Juan-Angel Jimenez-Martin & Massimiliano Caporin

**To cite this article:** Runfeng Yang, Juan-Angel Jimenez-Martin & Massimiliano Caporin (2025) The non-linear ESG premium, Quantitative Finance, 25:5, 817-840, DOI: 10.1080/14697688.2025.2492794

**To link to this article:** https://doi.org/10.1080/14697688.2025.2492794

*Quantitative Finance*, 2025
Vol. 25, No. 5, 817–840, https://doi.org/10.1080/14697688.2025.2492794



# The non-linear ESG premium

RUNFENG YANG †, JUAN-ANGEL JIMENEZ-MARTIN *‡ and MASSIMILIANO CAPORIN §

†Business School, Central South University, Changsha, People's Republic of China
‡Instituto Complutense de Análisis Económico (ICAE), Universidad Complutense de Madrid, Madrid, Spain
§Department of Statistical Sciences, University of Padova, Padua, Italy

(*Received 11 July 2024; accepted 26 March 2025*)

We examine if the ESG performance is being priced in the cross-section of stock returns in a non-linear manner. We consider square and cubic forms of ESG scores when estimating the ESG premium. We find that the premiums from ESG, E, and S scores deviate from linearity, the extent of which depends on types of ESG scores and sample periods. We associate the non-linearity with the cross-sectional distribution of ESG scores. We find that investors ESG sentiment interacts with the cross-sectional distribution of ESG scores in driving linear ESG factors. A change in the ESG data provider will also change the characteristic of non-linear ESG factors. However, the non-linearity still exists using the common sample from different data providers. The above finding also applies in the corporate bond market, in the European market and when extending the sample back to 2004.

*Keywords*: ESG; ESG factor; Factor model; Fama/MacBeth cross-sectional regression; Non-linearity

*JEL Classifications*: G12, C51

## 1. Introduction

The growth of Environmental, Social, and Governance (ESG) investments has seen a remarkable increase in recent years, particularly during the COVID-19 crisis.† Despite the high demand for ESG investments, the relationship between ESG performance and stock return, especially the potential mechanisms through which ESG influences a company's performance, still encounters contradicting evidence (Clark *et al.* 2015, Gillan *et al.* 2021). In addition, the literature discusses multiple channels explaining how the ESG performance is realized in the asset pricing.‡

If the observed relationship between ESG performance and stock performance is a mixture of different channels, and in the meantime we still have mixed empirical evidence, a sensible conjecture is that there exists a non-linear relationship between ESG performance and stock performance. In fact, recent empirical research found that the way ESG activities affect company performance can be both indirect and non-linear (Jahmane and Gaies 2020, Franco *et al.* 2020). However, the mainstream literature on ESG–stock price relationship builds on linear assumptions (e.g. Pástor *et al.* 2021, Lioui and Tarelli 2022).

In this paper, we examine, from an empirical perspective, if the ESG performance is being priced in the cross-section of stocks in a non-linear manner, and study possible drivers behind the non-linearity. In particular, we consider ESG score distributions, in addition to the ESG score level, to capture the non-linear ESG factor. Both linear, square and cubic forms of scores are included in the Fama-MacBeth cross-sectional regression, so that the non-linear part captures how the ESG–stock relationship can deviate from the linear setting.

There are two types of interpretations for the coefficients of the square and cubic forms of ESG score. The first explanation is that the coefficient is a linear combination of the left-hand-side asset returns (a factor-mimicking portfolio), with the weight being the level, square or cubic forms of ESG scores. The second explanation is that, when the ESG score is standardized cross-sectional to have zero mean and variance of one, coefficients of $ESG^2/ESG^3$ measure how cross-sectional standard deviation/skewness of ESG scores are being priced.

Our empirical analysis is based on the US market. We first construct non-linear ESG risk factors using ESG scores and pillar scores (E, S and G scores). The pattern of non-linear risk factors is quite similar among ESG, E and S scores, while for the G score we find a different pattern, probably because the G score is less related to the sustainability. We find that the non-linearity exists for ESG, E and S scores—as we increase

*Corresponding author. Email: juanangel@ccee.ucm.es
† https://www.moodys.com/research/Moodys-ESG-investing-a-boon-for-asset-managers-as-product–PBC_1 265808.
‡ See, for example, discussions in Becchetti *et al.* (2015), Giese *et al.* (2019) and Pástor *et al.* (2022).

from linear to square to cubic form of ESG score, the overall return predictability increases, and this is especially the case for the E score. However, we find limited evidence for G score to exhibit non-linearity.

We then capture and test the non-linearity through the spanning regression proposed by Barillas and Shanken (2016). Specifically, we run the regression between the non-linear ESG risk factors (dependent) and linear risk factors (independent), the intercept of which represents the additional information brought by the non-linear factor. We find that the choice of applying the cubic form or the squared form of ESG scores depends on the types of the ESG scores, and that the non-linearity does not necessarily come from the highest order. For example, for the E score, the intercept is significant with the squared form but not the cubic form. Further, we perform the test under a dynamic setting. We find that, while the non-linearity are significant most of the periods for ESG, E and S scores, there are still periods when the with insignificant results, which implies that the choice of non-linear setting also depends on how ESG factor performs.

Regarding possible determinants of the detected non-linearity, our first conjecture is that investors ESG preference, one of the main channel behind the pricing of ESG performance in the stock market (Pástor *et al.* 2021), does not increase linearly with ESG scores. To test the conjecture, we study if the non-linearity is related to the cross-sectional distribution of ESG scores. We find that changes in the cross-section standard deviation and skewness of ESG scores are associated with changes in the non-linearity, which implies that the cross-sectional distribution of ESG scores is being priced. In addition, investors ESG sentiment interacts with the cross-sectional distribution of ESG scores in driving linear ESG factors, which implies that preference is being 'distorted' by the cross-sectional distribution of ESG scores. That is, investors choices, when making decisions based on ESG scores, might also be altered by how ESG scores are distributed.

The non-linearity may also be related to the uncertainty of the ESG score. All types of ESG scores are an approximation of the unknown true ESG performance. Even if the relationship between true ESG performance and stock performance is linear, measurement errors in the errors in ESG scores will introduce non-linearity into the relationship. We test this conjecture using two different ESG data providers and compare the results. We find that changing the ESG score data provider will change the characteristics of non-linear ESG factors: the patterns of non-linear factors are different between the two providers, with different levels and signs. Therefore, sample selection is critical to the results on non-linearity.

Finally, we repeat the analysis using common companies with similar ESG scores in different databases. The idea is to see if the non-linearity still exists when the uncertainty in the ESG score is minimized. The results with the common data sets, where we use companies with similar ESG scores in both databases, are different from the previous results, implying that part of the non-linearity is due to ESG uncertainty. However, we still observe non-linearity with the common sample, which means that the presence of ESG score uncertainty changes the estimation of the non-linear ESG factors, but does not prevent us from using the non-linear setting.

Taken together, our results have several important implications. First, the existence of non-linearity provides novel insights into the ESG pricing mechanism. Specifically, while investors exhibit ESG preferences that influence the cross-section of stock prices, the way investors value ESG-related stocks may not increase linearly with their ESG preferences. This non-linearity can arise from the distribution of ESG scores. Second, in capital markets, investors aiming to capture ESG-related risk factors must consider the potential non-linearity in the relationship to better capture ESG pricing and ESG-related risk. Furthermore, our analysis highlights that differences in ESG databases represent another source of non-linearity, stemming from ESG score uncertainty.

We examine the robustness of the results under several alternative settings. First, we check if non-linearity exists during earlier periods (from 2004 to 2012), when ESG-related investment was less popular and developed. Although we find weaker linear and non-linear risk factors than those in the current period, non-linearity still persists. Next, we examine whether non-linearity is a phenomenon specific to the US market. We repeat the analysis in the European market and find that non-linearity still exists. Finally, we extend our analysis to the bond market and study whether ESG performance is priced in the cross-section of corporate bond returns. We find that ESG factors are weaker in the bond market than in the stock market. Yet, there is still evidence of non-linearities.

**Related literature**: Our results add new evidence to the literature of the relationship between ESG and stock performance. We are the first to formally discuss the non-linear relationship between ESG and stock performance. In terms of how ESG performance affects the stock price, recent studies find that investors preference over high-ESG (green) companies is the major force (Pástor *et al.* 2021, 2022). However, all of them assume a linear relationship between investor's reference and stock performance. Differently, we find that there might be a non-linear relationship, the ESG preference channel might be altered by the distribution of ESG scores.

Second, we propose a new non-linear setting to construct the ESG factor, which adds to the vast research on constructing the ESG factor based on linear settings. In general, three methods are found in studies that capture the ESG factor. The first is the Fama/French (FF) two-sort method (Becchetti *et al.* 2018, Maiti 2020, Jarjir *et al.* 2022). Under the FF method, the ESG factor is constructed as long-short portfolios: long on stocks with low-ESG companies and short on stocks with high-ESG companies. The second approach is to use the Fama/MacBeth (FM) cross-sectional regression (Lioui and Tarelli 2022, Naffa and Fain 2022). Under the FM method, in each time $t$, a cross-sectional regression is carried out between company return and firm variables like size or ESG score. Another approach is the mimicking-portfolio approach (MP) discussed in Lamont (2001). Under the MP method, a time-series regression is conducted between the ESG score (dependent) and a selection of portfolio returns (independent). The selection of portfolios aims to encompass the actual ESG factor in the return space. Then, the beta in the time-series regression serves as the weight to construct a tracking portfolio. Engle *et al.* (2020) applied the MP method to construct a tracking portfolio that replicates a climate news index. One thing common to all three methods is that they assume a linear relationship (or semi-linear for the FF method) between ESG and stock performance. Our new method allows investors to flexibly include the cubic or squared forms of ESG score (or even higher monomials) in capturing the ESG factor.

Finally, we relate our work to the vast literature on ESG activities and firm performance. We contribute to this line of research by providing novel evidence on the existence of a non-linear relationship. Multiple mechanisms are found in the literature on how ESG activities might affect firm performance. First, the resource-based view (Surroca *et al.* 2010) says that intangible assets, such as innovation, human resource and reputation, serve as a mediator between ESG and corporate finance performance, and those intangible assets created by ESG activities will finally increase shareholder value. Second, compared to high-ESG companies, low-ESG companies have higher stakeholder risk, the risk brought by having increased conflicts with stakeholders (Becchetti *et al.* 2015). Similarly, the stakeholder theory says that companies with low ESG performance should have more conflicts with stakeholders (governments, NGOs, consumers, and so forth) and thus are more prone to receive punishments from stakeholders such as boycotts, fines or increased taxation by authorities. Third, investors may also react to ESG-related news (Krueger 2015, Pástor *et al.* 2022). Our results further evidenced the existence of multiple channels discussed above.

The paper develops as follows: section 2 contains a detailed discussion of how we capture and test the non-linear ESG factor. Section 3 describes the data we are going to use. Section 4 presents the result of ESG factors and the test results. Section 5 further studies the potential drivers of non-linear ESG factors. Section 6 discusses the results under alternative settings. We conclude the paper in section 7.

## 2. Methodology

### *2.1. ESG performance, non-linearity and the cross-section of stock returns: a conceptual framework*

Consider the following utility function for investors with ESG preferences (Pástor *et al.* 2021)

$$V\left(\tilde{W}_{1i}, X_i\right) = -e^{-A_i \tilde{W}_{1i} - b_i' X_i}, \tag{1}$$

where $A_i$ is the risk aversion of agent $i$, $\tilde{W}_{1i}$ is the wealth at time 1, $X_i$ is the portfolio weight vector. Further, $b_i'$, the ESG benefit vector, reflects the utility derived from holding ESG-related assets: an investor with stronger ESG preferences will experience higher utility as the ESG performance of the portfolio ($b_i' X_i$) improves. The ESG benefit vector is independent of wealth and risk aversion. Moreover, the ESG benefit vector incorporates both agent-specific and firm-specific components, defined as:

$$b_i = d_i g, \tag{2}$$

where $g$ is an $N \times 1$ vector of company ESG performance and $d_i$ is a scalar is a scalar measuring the intensity of ESG preferences for agent $i$. In this framework, ESG performance is priced in the capital market because investors derive utility from sustainability preferences. The relationship between a company's ESG performance ($g$) and the investor's ESG benefit ($b_i$) is linear.

We conjecture that the relationship between ESG performance and ESG benefit is, in fact, non-linear, and one potential source of this non-linearity is the distribution of ESG scores. Incorporating this into the model, the benefit vector becomes:

$$b_i^* = d_i f\left(g, \text{var}(g), \text{skew}(g)\right), \tag{3}$$

where var($g$) represents the cross-sectional variance and skew($g$) represents the cross-sectional skewness of ESG scores. The distribution of ESG scores is important for the following reasons. First, the level of investor preference for companies does not necessarily increase linearly with the level of ESG performance. For example, investors may not differentiate between companies with ESG scores of 20 and 35 (out of 100), but they may have a strong preference for companies whose ESG scores exceed 80. The screening process for constructing an ESG portfolio operates similarly: fund managers set an ESG performance threshold and retain companies with ESG scores above this threshold. The ESG benefit may diminish when all companies exhibit similar ESG performance, as investors cannot differentiate between firms with high and low ESG scores. These factors are directly related to the distribution of ESG scores.

Second, uncertainties in ESG scores arise due to discrepancies across ESG data providers, which is reflected as varying ESG scores. All measures of ESG scores approximate the unknown true ESG performance. Berg *et al.* (2022) show significant divergence in ESG scores across providers. Avramov *et al.* (2022) demonstrate that uncertainties in ESG scores reduce ESG preferences, leading to declining demand for high-ESG stocks. Measurement errors in ESG scores add non-linearity to the relationship between ESG performance and stock performance, as linear changes in true ESG performance may manifest as non-linear changes in ESG scores.

More generally, recent empirical research has found that ESG activities' impact on company performance can be both indirect and non-linear (Jahmane and Gaies 2020). ESG performance may influence company performance by affecting daily operations (Freeman 1984, Krueger 2015, Lins *et al.* 2017, Albuquerque *et al.* 2019), for example, through ESG-related events such as government fines or new ESG policies. The complexity arises because it is uncertain how and when these ESG-related events will impact the company. For instance, a company's likelihood or severity of government fines does not necessarily decrease linearly with improved ESG performance.

There is a growing literature on applying non-linear methods in asset pricing. Freyberger *et al.* (2020) use nonparametric regression to study the nonlinear predictive relationship in the US stock market. Kirby (2019) considers the non-linear firm characteristic in addition to linear in the cross-sectional regression, and find that well-known asset pricing models fail to explain the non-linear effects. Gu *et al.* (2020) use machine learning approach to capture latent factors in the cross-section of stock returns. Dittmar (2002) explored nonlinear pricing kernels (square and cubic form) with preference for kurtosis in explaining the cross-sectional stock returns. All these studies find a larger return predictability using a non-linear model than linear models.

The relationship between ESG score distribution and stock returns may not be determined *apriori*, because the relationship between ESG performance and stock returns evolves over time due to shifting ESG sentiment in financial markets. Pástor *et al.* (2022) and Ardia *et al.* (2023) show that unexpected increases in investors' green concerns drive up the returns of green companies, whereas during normal periods, green companies exhibit lower expected returns. Kahn and Kotchen (2011) also identify associations between climate concerns and business cycles. This pattern is similar to that observed for market risk variance and skewness. The literature presents mixed evidence on the pricing of market skewness (e.g. Kapadia 2006, Adrian and Rosenberg 2008). Chang *et al.* (2013) find that the pricing of variance is not constant and depends on the sample period and test methods. Therefore, whether and how the ESG score distribution affects stock returns remains an empirical question.

In the following analysis, we address two questions: (1) Is the ESG performance being priced in the cross-section of stock returns in a non-linear manner? (2) if so, what drives the non-linearity? Our research design consists of two parts. We first describe how we capture the non-linear ESG factor and how we capture and test the non-linear effect. We then provide a regression analysis of the possible drivers behind the non-linearity.

### 2.2. *Capturing the non-linear ESG factor*

To incorporate the ESG score distribution in capturing ESG pricing, we follow in Kirby (2019) to consider the **square** and **cubic** form of ESG scores. This form of ESG scores might not be the best choice to fully capture the non-linearity. In fact, we do not have a prior knowledge about which nonlinear form should be the most appropriate. However, our goal is not finding a non-linear format that maximizes return predictability (otherwise we could use machine learning models). Instead, we focus on a non-linear form that is intuitive and that might be associated with economic channels that we are able to identify (ESG score distribution).

Choosing the highest order to be cubic is also arbitrary. In fact, one can choose as many orders as desired under our setting and may also provide significant results. These higher-order factors may even survive the out-of-sample test. The issue here is finding an economic motivation as to why the factor should exist, which further guarantees that the observed pattern of pricing abnormality persist in the future, as is suggested by Fama and French (2018).

Specifically, based on the cross-sectional regression setting discussed in Lioui and Tarelli (2022), we propose the following baseline non-linear setting to capture the ESG premium:

$$r_{i,t} = f_{0,t} + \sum_{l=1}^{L} f_{l,ESG,t} ESG^{l}_{i,t-1} + \varepsilon_{i,t}, \tag{4}$$

where $L$ represents the monomial with the highest degree and we set $L = 1, 2, \text{and} 3$. $ESG_{i,t-1}$ refers to the ESG score of company $i$ at the beginning of period $t$. $r_{i,t}$ is the excess return in the current period. At each period $t$, we conduct the cross-sectional regression of equation (4) and get the coefficient $f_{l,ESG,t}$. We then stack $f_{l,ESG,t}$ over time to get the premium time series (**the linear/nonlinear ESG factor**).†

The analytical expression for the estimator of the coefficient in the cross-sectional regression has several important traits that determine how we interpret the factor. As is shown in equation (A7) in the Appendix, under the least square estimator, $f_{l,ESG,t}$ is always a zero-cost long-short portfolio consisting of all assets in the cross section. If $l = 1$, it is the linear setting discussed in the literature, with the solution being (in matrix form):

$$\begin{cases} f_{0,t} & = \dfrac{1}{N^{\text{Assets}}_{t-1}} R_t \\ f_{1,ESG,t} & = \dfrac{(\boldsymbol{ESG}_{t-1})'}{(\boldsymbol{ESG}_{t-1})'\,\boldsymbol{ESG}_{t-1}} R_t \end{cases}, \tag{5}$$

where $N^{\text{Assets}}_{t-1}$ is the number companies at $(t-1)$, $R_t$ is the vector of company returns and $\boldsymbol{ESG}_{t-1}$ is the vector of ESG scores. In equation (5), since the ESG score is being standardized cross-sectionally, $(\boldsymbol{ESG}_{t-1})'\boldsymbol{ESG}_{t-1} = N^{\text{Assets}}$ (because the variance of $\boldsymbol{ESG}_{t-1}$ is one), and thus $f_{1,ESG}$ is a zero-cost long-short portfolio that longs companies with high ESG score and shorts companies with low ESG score.

When $L > 1$, it means that, in addition to the linear part, the ESG score of a company is affecting the cross-section return in a non-linear manner. As is shown in equation (A10), when $L > 1$, both linear and non-linear factors are a linear combination of $\sum_{i=1}^{n} ESG^{l}_{i,t-1} r_{i,t}$, which means that instead of using the linear form of ESG score, we apply the squared and cubic forms of the ESG score to the weight of the portfolio. Therefore, the non-linear setting measures how the ESG impact on asset pricing can deviate from the linear setting.

Since the ESG score is standardized, $E(ESG^{l}_{i,t-1})$ can be regarded as moments of cross-sectional distribution of the ESG score.‡ In particular, $E(ESG^{2}_{i,t-1})$ is the cross-section variance at time $t-1$ and $E(ESG^{3}_{i,t-1})$ is the cross-section skewness. A higher variance means that companies are more spread out in terms of ESG performance.§ A positive skewness, given that the mean is zero, could imply that, in a given cross-section, more companies have high ESG score or there are companies who have extremely high-ESG scores. Therefore, $f_{2,ESG}$ and $f_{3,ESG}$ can be interpreted as how the cross-sectional distribution of ESG scores is affecting the

† While *ESG* and $ESG^3$ could have high correlation, the multi-collinearity is not a problem under our setting, because we do not use the standard error in the cross-sectional regression for the inference of the non-linear ESG factor.

‡ Note that for each cross-sectional regression, we have

$$E\left(r_{i,t}\right) = \hat{f}_{0,t} + \sum_{l=1}^{L} \hat{f}_{l,ESG,t} E\left(ESG^{l}_{i,t-1}\right).$$

§ If the ESG score is standardized cross-sectional such that $E(ESG^{2}_{i,t-1}) = 1$, the actual cross-section variance is then reflected in the estimated beta ($\hat{f}_{2,ESG,t}$).

company return, thus captures the relationship described in the ESG benefit function (3).

In addition to the baseline setting, we apply the following regression:

$$r_{i,t} = f_{0,t} + \sum_{l=1}^{L} f_{l,ESG,t} ESG_{i,t-1}^{l} + \sum_{j=1}^{J} f_{j,t} x_{ij,t-1} + u_{i,t}, \quad (6)$$

where $L = 1, 2, \text{or} 3$, and $x_{ij,t-1}$ are a set of other firm characteristics. The setting is different from equation (4) in that we have added more controls in the cross-sectional regression. Previous analyses show that the ESG score interacts with other corporate variables (see, for example, the discussion of Drempetic *et al.* 2019, Gillan *et al.* 2021). Lioui and Tarelli (2022) also suggest adding firm variables as controls in the cross-sectional regression when capturing the linear ESG factor.

To measure the level of deviation from linearity, we refer to Kirby (2019) to calculate the Local Marginal Effects (LME) as (take $L = 3$ for example):

$$\widehat{LME_{i,t}} = \frac{\partial r_{i,t}}{\partial ESG_{i,t-1}} = \hat{f}_{1,ESG,t} + 2\hat{f}_{2,ESG,t} ESG_{i,t-1} + 3\hat{f}_{3,ESG,t} ESG_{i,t-1}^{2}. \quad (7)$$

Equation (7) means that the local marginal effect is a combination of both linear and non-linear factors. If $L = 1$, LME will be the value of $\hat{f}_{ESG,t}$. If non-linearity is a significant trait, then the LME should vary across different ESG scores for $L = 2$ and $L = 3$. The LME for each company $i$ in the whole period is estimated as:

$$\widehat{LME_{i,t}} = E\left(\widehat{LME_{i,t}}\right) \approx E\left(\hat{f}_{1,ESG,t}\right) + 2E\left(\hat{f}_{2,ESG,t}\right) E\left(ESG_{i,t-1}\right) + 3E\left(\hat{f}_{3,ESG,t}\right) E\left(ESG_{i,t-1}^{2}\right), \quad (8)$$

where $E(\hat{f}_{1,ESG,t})$ and others are the average return of ESG factor over time; $E(ESG_{i,t})$ is the average score of company $i$ over the sample period. The approximation comes from the finding that firm characteristics (in our case, the ESG score) poorly predict the short-term time-series variation in the stock returns (in our case, the long-short portfolio $\hat{f}_{2,ESG,t}$ and $\hat{f}_{3,ESG,t}$), as discussed in Welch and Goyal (2007) and Kirby (2019).

### *2.3. Spanning regression and non-linear effects*

In total, we have three sets of factors to choose from if we consider different $L$s in equation (6):

$$\begin{cases} \left\{f_{1,ESG,t}\right\}, & \text{for } L = 1 \\ \left\{f_{1,ESG,t}, f_{2,ESG,t}\right\}, & \text{for } L = 2\,. \\ \left\{f_{1,ESG,t}, f_{2,ESG,t}, f_{3,ESG,t}\right\}, & \text{for } L = 3 \end{cases} \quad (9)$$

In order to choose between different sets of factors, we have, at least, two methods (Fama and French 2018). The first approach is to choose a set of assets, and then run a system regression using different sets of factors to explain the asset return. The best set of factors should produce the smallest intercept in the regression. The idea is that the best set of factors should be able to span the return universe created by the set of assets and reach the largest Sharpe ratio.

Another approach is called the spanning regression, proposed by Barillas and Shanken (2016). They suggest running a regression between the factors to be tested (as the dependent variable) and a set of known factors (such as Fama/French 6 factors or the market risk factor). The best factor should produce non-zero intercept in the regression. The spanning regression aims at testing whether factors on the left side of the regression contribute to the explanation of return universe provided by the possible combination of factors on the right side, that is, whether the factor on the left side adds to the maximum Sharpe ratio of factors on the right side of the regression.

We base our discussion on the spanning regression. Our research design consists of two steps. In the first step, we compare between different sets of ESG factors under different $L$s (as demonstrated in equation (9)). Specifically, assume that for some type of ESG score, we have a set of ESG factors, estimated under some $L$, to be tested. For each factor within the set, we run the following regression:

$$f_{l,n,t} = \alpha_{l,n} + \sum_{j=1}^{J} \beta_{l,n,j} f_{j,t} + \varepsilon_{l,n,t}, \quad (10)$$

where $l = 1, 2 \ldots, L$ (in total, we run $L$ regressions) and $f_{jt}$ $(j = 1, 2, \ldots, J)$ are common risk factors. $n$ is the type of ESG risk factors ($n$ = ESG, E,S or G). We consider two sets of common risk factors. The first is the market risk factor (the basic CAPM model, serving as the benchmark for testing the pricing error), and the second is Fama/French 5 factors (namely, market, size, value, profitability and investment) plus momentum factor, all from Kenneth R. French data base.†

The $\alpha_{l,n}$ in the spanning regression is the additional information (the pricing error) brought by the ESG factor to common risk factors. Then, we test if the intercept of the $L$ regressions is jointly zero:

$$\mathbb{H}_0 : \alpha_{l,n} = 0, \quad \forall\, l.$$

We use the Gibbons, Ross, and Shanken (1989) test (GRS, Gibbons *et al.* 1989) for the joint test. Rejecting the null hypothesis means that the set of ESG factors carries unique information in addition to the common risk factor structure. Accordingly, the best set of ESG factors should produce the largest intercept and thus the smallest $p$-value in the GRS test. In this step, we choose between the ESG factors under different $L$s.

Assume that set $\{\alpha_{1,n}, \alpha_{2,n}, \alpha_{3,n}\}$ in the first step is jointly non-zero. In the second step, we further check if the significant pricing error is driven by non-linear factors. We carry out the spanning regression between factors under the same $L$. For example, the following spanning regression will be carried out

† https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

for $L = 3$:

$$\begin{cases} f_{1,n,t} = \delta_{1,n} + \sum_{j=1}^{J} \beta_{1,n,j} f_{j,t} + \varepsilon_{1,n,t} \\ f_{2,n,t} = \delta_{2,n} + \beta_{1,n,j} f_{1,n,t} + \sum_{j=1}^{J} \beta_{2,n,j} f_{j,t} + \varepsilon_{2,n,t} \\ f_{3,n,t} = \delta_{3,n} + \beta_{1,n,j} f_{1,n,t} + \beta_{2,n,j} f_{2,n,t} + \sum_{j=1}^{J} \beta_{3,n,j} f_{j,t} + \varepsilon_{3,n,t} \end{cases}, \tag{11}$$

The purpose of this analysis is to determine, given a certain $L(L > 1)$, whether the non-linear parts of the ESG factors add additional information beyond the linear components.

The key difference between $\alpha_{l,n}$ from the first step and $\delta_{l,n}$ from the second step lies in the interpretation. Specifically, $\alpha_{l,n}$ represents the overall pricing error contributed by the set of factors $\{f_{1,n,t}, f_{2,n,t}, f_{3,n,t}\}$ relative to the common risk factors, whereas $\delta_{l,n}$ captures the incremental pricing error from non-linear risk factor within the same $L$. For instance, $\delta_{2,n}$ measures the additional pricing error from $f_{2,n,t}$ over $f_{1,n,t}$, showing whether the square ESG factor adds explanatory power beyond the linear one. If $\delta_{2,n} \neq 0$ while $\delta_{3,n} = 0$, it means that the additional pr comes specifically from the variance (square form of ESG scores) rather than higher-order interactions.

In sum, with the above process, we aim to answer the following questions: 1. whether we should choose a higher $L$ over a lower $L$; 2. if so, is the additional explanation power brought by the non-linear part? For simplicity of expression and to avoid confusion, we use $f_{l,n,t}$ ($l = 1, 2$ or 3 and $n =$ ESG, E, S, G) to express the group of '$f_{l,ESG,t}$, $f_{l,E,t}$, $f_{l,S,t}$ and $f_{l,G,t}$' and we use $f_{l,ESG,t}$ to refer to the factor calculated using the ESG score and so on.

## 3. Data

### 3.1. ESG scores

We use companies in the US market for the analysis. Our main result is based on ESG score from Eikon (Refinitiv). Refinitiv is a leading global provider of ESG data, encompassing over 10 000 companies across 76 countries. We use the ESG combined score, which evaluates a company's overall sustainability performance, ranging from 0 to 100, with higher values indicating superior environmental, social, and governance practices.

Apart from the ESG combined score, we also use the environmental, social, and governance pillar scores. The environmental score (the E score) measures a company's performance in environmental-related issues. The score is consisted of three sub-scores: emission, energy use and innovation. Each sub-score has more than 15 areas that are evaluated. The social score (the S score) measures a company's performance in contributing to society. It contains four sub-scores: workforce, human rights, community and product responsibility. Each sub-score is also calculated by evaluating dozens of issues. The governance score (the G score) measures the internal management quality of a company. We have obtained monthly ESG score data from 31st December 2012 to 30th November 2021 (108 months). Earlier data was omitted due to limited company disclosure of ESG data. The score is updated biannually or annually, depending on the fiscal year of the company. Finally, we follow Kirby (2019) and Lioui and Tarelli (2022) to first convert ESG scores into percentile ranks and then standardize the rank.

Additionally, we use the ESG score from Bloomberg for a comparison of results. The Bloomberg ESG score is calibrated using various data sources offered on the Bloomberg Terminal service, mainly company-reported sustainability information and financial fundamentals data.† Bloomberg offers E, S and G scores separately and we use the average of the three as the ESG score for one company. The score ranges from 0 to 10 and combines a matrix of indicators of sustainability similar to that of Eikon. The ESG score used in our analysis dates back to 2015 and we only have annual data. We assume that the ESG information is updated at the beginning of the year (so that we can convert the annual data to monthly data from 31st January 2015 to 31st December 2021).

### 3.2. Company data

We download the month-end total price index from 2012/12/31 to 2021/12/31 from Refinitiv. The monthly return is calculated as $R_{it} = \frac{RI_{i,t} - RI_{i,t-1}}{RI_{i,t-1}}$ for each company $i$ and month $t$. In total, we have 109 monthly price observations and 2205 companies in our sample. The number of companies used in the cross-sectional regression is time-varying, because only part of the company has ESG score data. We exclude penny companies in constructing factors, defined as companies with a price smaller than 5 USD at month $t$ (Amihud 2002, Fama and French 2008). There are more companies in Eikon than in Bloomberg (see figure A.1 in the Online Appendix B). In addition, Eikon has a longer history of ESG data. For this reason, our main analysis is based on ESG scores from Eikon.

As discussed in the previous section, we also control for several firm characteristics in the cross-sectional regression: market value ($MV_{i,t}$), calculated as the log of market capitalization; book-to-market ratio ($BM_{i,t}$); operating profit ($OP_{i,t}$), calculated as the percentage of operating profit to the market value of a company; momentum factor ($MOM_{i,t}$), calculated as the ratio between price of month $t$ to the price of month $(t - 12)$; and the investment growth ($INV_{i,t}$), defined as the growth in the total assets for the last year. All variables will be updated monthly (if there is a change).

### 3.3. Summary statistics

Table 1 shows the summary statistics of variables used in the cross-sectional regression. The 25% quantile of $E$ score is smaller than other scores, which implies that a large part of companies have a low $E$ score. Apart from that, the statistics of other ESG scores are quite similar. The monthly mean excess return is 1.62%, with the standard deviation being 13.52%.

Table 2 shows the correlation matrix between all company variables. As can be seen, ESG scores have a relatively high correlation with market value (firm size) compared

† https://data.bloomberglp.com/professional/sites/10/ESG_Environmental-Social-Scores.pdf.

Table 1. Summary statistics of ESG score and firm characteristics.

| | N | Mean | Std. | 25% | 50% | 75% |
|---|---|---|---|---|---|---|
| ESG | 166 259 | 0.50 | 0.29 | 0.25 | 0.50 | 0.75 |
| E | 166 259 | 0.50 | 0.29 | 0.22 | 0.50 | 0.75 |
| S | 166 039 | 0.50 | 0.29 | 0.25 | 0.50 | 0.75 |
| G | 166 220 | 0.50 | 0.29 | 0.25 | 0.50 | 0.75 |
| MV | 219 638 | 7.61 | 1.87 | 6.26 | 7.55 | 8.87 |
| BM | 218 762 | 0.61 | 1.76 | 0.24 | 0.46 | 0.76 |
| OP | 218 669 | 0.08 | 1.71 | 0.03 | 0.07 | 0.11 |
| INV | 218 363 | 0.10 | 0.30 | − 0.01 | 0.06 | 0.14 |
| MOM | 219 217 | 0.10 | 0.41 | − 0.09 | 0.10 | 0.29 |
| Excess Return (%) | 219 359 | 1.62 | 13.52 | − 4.14 | 1.12 | 6.44 |

Notes: The table presents the summary statistics for the ESG score, company characteristics, and excess returns (monthly). All data come from Refinitiv database. The ESG scores are expressed as percentile ranks normalized to have a mean of 0.5, which is calculated as $(\frac{X_t - 0.5\times 1_{N^{\text{asset}}}}{N^{\text{asset}}})$($X$ is the cross-section rank from 1 to $N^{\text{asset}}$ at every month $t$).

Table 2. Correlation between firm characteristics.

| | ESG | E | S | G | MV | BM | OP | INV | MOM |
|---|---|---|---|---|---|---|---|---|---|
| ESG | 1 | | | | | | | | |
| E | 0.76* | 1 | | | | | | | |
| S | 0.83* | 0.70* | 1 | | | | | | |
| G | 0.67* | 0.40* | 0.36* | 1 | | | | | |
| MV | 0.54* | 0.59* | 0.59* | 0.31* | 1 | | | | |
| BM | − 0.06* | − 0.04* | − 0.06* | − 0.03* | − 0.1* | 1 | | | |
| OP | 0.00 | 0.00 | 0.00 | 0.00 | 0.01* | 0.50* | 1 | | |
| INV | − 0.09* | − 0.11* | − 0.06* | − 0.08* | 0.00 | 0.03* | 0.03* | 1 | |
| MOM | − 0.01* | − 0.01* | − 0.01* | − 0.02* | 0.14* | − 0.07* | 0.02* | 0.06* | 1 |

Notes: This table presents the statistics of variables used in our analysis. *denotes significance of pair-wise correlations at the 5% level.

to other characteristics, which is in-line with the literature (e.g. Drempetic *et al.* 2019), where studies find that larger companies usually have higher ESG scores, because they have more resources devoted to improving their sustainability performance.

To see if the non-linear part adds to the predictability in the cross-sectional regression, we further report in figure 1 the adjusted $R^2$ when $L = 1, 2,$ and3 under the baseline model (4). In general, the level of $R^2$ is low, with most of the time being below 0.05. The difference between $L = 1$ and $L > 1$ is small most of the times, which implies that adding the non-linearity does not add significant predictability in the cross-section of stock returns. However, we observe an increase in the $R^2$ in some periods when $L > 1$. For example, for the $E$ score during the period of 2014 and 2016, we observe a higher predictability when counting for non-linearity, which means that the linear setting might not be the best choice during some periods.

## 4. Linear vs. non-linear factors

### *4.1. Factor statistics*

Table 3 shows factor return statistics from 2013 to 2021. Most non-linear factors have non-significant average returns over the sample period. In line with Lioui and Tarelli (2022), the linear factors' returns are non-zero for the ESG and S score. Figure 2 shows the cumulative return of factors under $L = 3$. There are positive shocks of ESG, E, and S factors, especially after the COVID-19 pandemic. A possible explanation is that during crisis times investors turn to high quality investment (the 'flight-to-quality' effect, Dong *et al.* 2019). The patterns of ESG, E, and S linear factors are quite similar. In comparison, the G risk factor has a different pattern than other risk factors, which can be explained by the fact that the G score measures mostly the internal management quality of a company instead of the sustainability. The pattern of non-linear risk factors is also quite similar between ESG, E and S scores.

We then draw the scatter plot of the local marginal effects (LME) estimated as equation (8) in figure 3, with the $x$-axis being the average ESG score of each company over the whole period. Mathematically speaking, we should see non-linearity effect being present from companies with extreme ESG scores (because the term $3E(\hat{f}_{3,ESG,t})E(ESG^2_{i,t-1})$ in equation (8) is large for companies with extreme scores), which is why the deviation is larger when the ESG score becomes more extreme. However, the shape and the curvature are not known a-priori because it also depends on the sign and size $E(\hat{f}_{3,ESG,t})$.

The non-linear effect is large for ESG and E score. When $L = 3$, the shape of LME is concave, indicating an off-setting relationship between linear and non-linear effect. In addition, the LME changes from positive to negative when the scores becomes more extreme – for the ESG score, the value can change from 0.2% to $-0.4\%$, implying a large offsetting effect. While we do observe large non-linearity in ESG, E and S scores, the non-linear effect is smaller for the G score, which implies that the level of non-linear effect is not the same in all

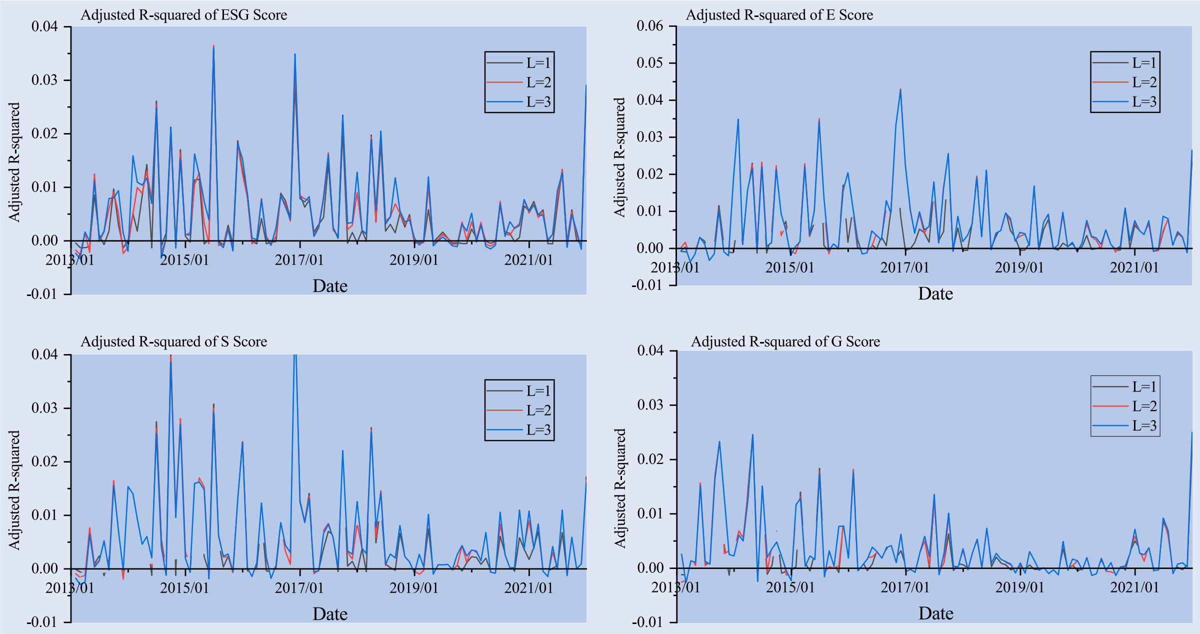


Figure 1. Predictability of non-linear settings.Note: The figure shows the adjusted $R^2$ of when $L = 1$, 2 and 3 under the baseline model (4).

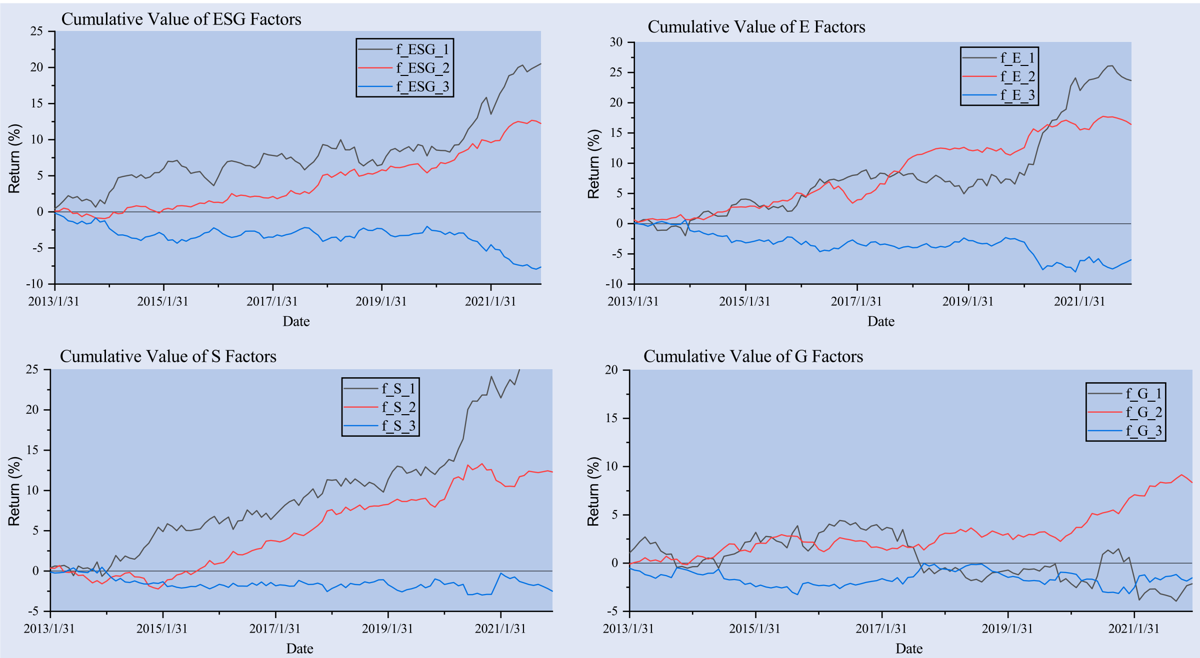


Figure 2. Cumulative value of ESG factors.
Notes: The graph shows the cumulative monthly factor returns: $\sum_{t=1}^{T} f_{l,n,t}$. The factors are calculated using model (6) with $L = 3$.

types of ESG scores. In sum, the above analysis implies that the ESG–stock return relationship does deviate from linearity.

### 4.2. *Choosing the non-linearity?*

**4.2.1. Static test.** Table 4 shows the $p$-value of the GRS test. The test results are significant at the 5% level for all $L$s for the ESG factor. One the one hand, the test results further confirms the findings in the literature about the existence of the ESG factor. On the other hand, the factor becomes stronger under the non-linear setting—we observe a smaller $p$-value as we apply a higher $L$. This is especially the case for the E score, where the test results is significant only when $L > 1$, indicating that applying the non-linear setting has a better return predictability than the linear setting. In comparison, the results are non-significant for the G score, inline with our observation in the previous section.

To see if the increased test performance is due to the non-linear risk factor, we present in table 5 the spanning test results for $L = 3$. For the ESG score, $f_{3,ESG,t}$ has a non-zero intercept, which partly explains the increased GRS test performance from $L = 1$ to $L = 3$. In other words, the non-linear risk factor does capture the cross-section pricing abnormality in addition to the linear one. As for the E score, the intercept is significant for $f_{2,E,t}$ but not $f_{3,E,t}$, which means that the non-linearity

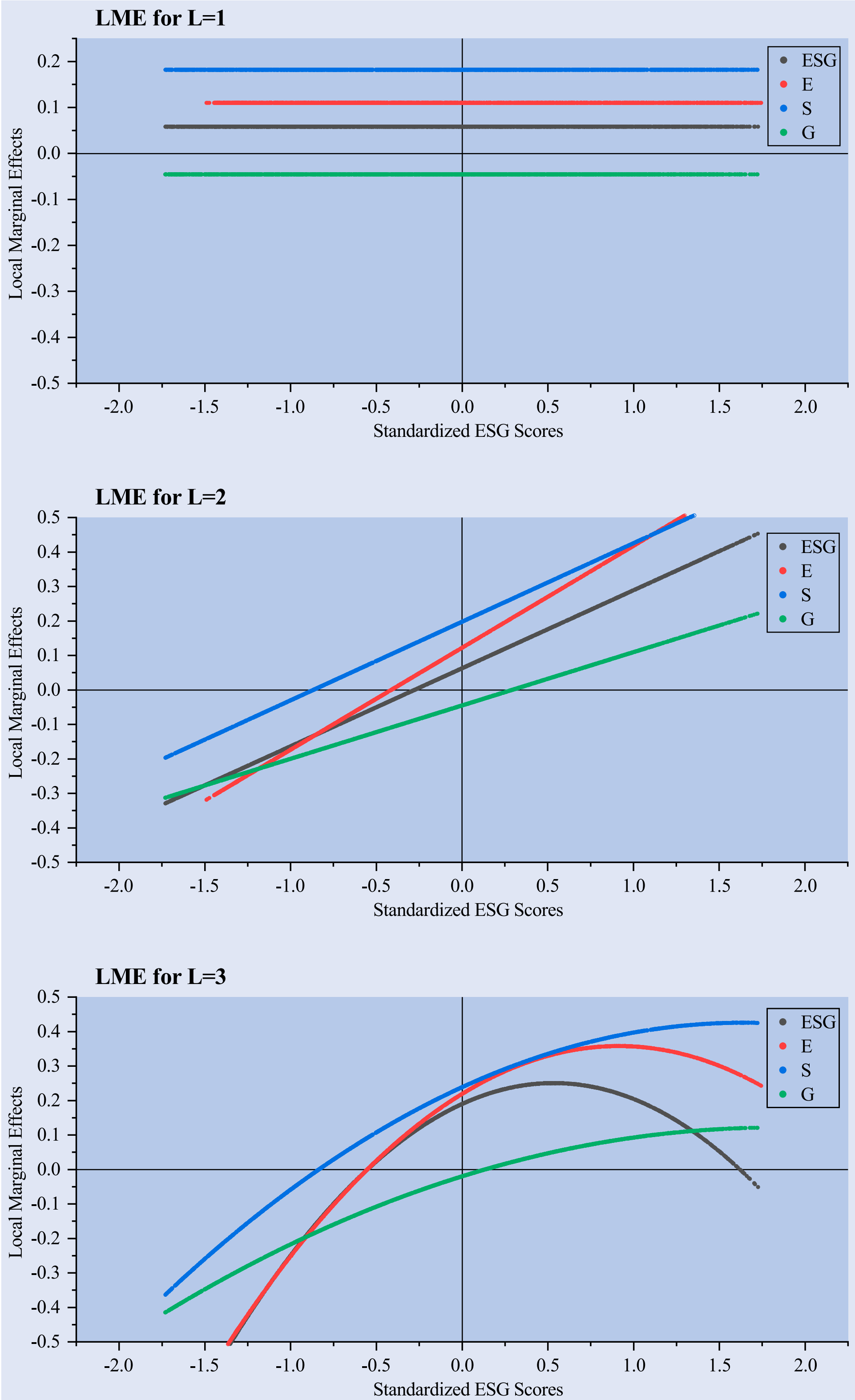


Figure 3. Local marginal effects.Note: The graph presents the local marginal effect for $L = 1$, 2 and 3, calculated as inequation (8).

does not necessarily come from the highest order. In terms of the S score, both $f_{2,S,t}$ and $f_{3,S,t}$ have non-zero intercepts while $f_{1,S,t}$ has an insignificant intercept, which means that increased GRS test performance when $L > 1$ is due to the non-linear part but not the linear part. In addition, $f_{3,ESG,t}$ has a negative relationship with $f_{2,ESG,t}$ and $f_{1,ESG,t}$, which is inline with the results in figure 2.

**4.2.2. Dynamic test.** The previous analysis shows that for the whole sample period, the non-linearity is presented, but not in all situations. This is to some degree explainable: as we have observed in figure 1, adding the non-linearity does not add to the predictability at all times. However, there are some periods where we see a large contribution of predictability. For this reason, we repeat the test in the previous section under

Table 3. Summary statistics of monthly factor returns.

| Mean | $L = 1$ | $L = 2$ | $L = 3$ |
|---|---|---|---|
| $f_{1,ESG,t}$ | 0.0579 | 0.0627 | 0.1900 |
| $f_{2,ESG,t}$ | – | 0.1131 | 0.1134 |
| $f_{3,ESG,t}$ | – | – | – 0.0708 |
| $f_{1,E,t}$ | 0.1099 | 0.1222 | 0.2192 |
| $f_{2,E,t}$ | – | 0.1478 | 0.1521 |
| $f_{3,E,t}$ | – | – | – 0.0555 |
| $f_{1,S,t}$ | 0.1821 | 0.1977 | 0.2387 |
| $f_{2,S,t}$ | – | 0.1139 | 0.1139 |
| $f_{3,S,t}$ | – | – | – 0.0231 |
| $f_{1,G,t}$ | – 0.0456 | – 0.0451 | – 0.0199 |
| $f_{2,G,t}$ | – | 0.0773 | 0.0774 |
| $f_{3,G,t}$ | – | – | – 0.0141 |
| $t$-statistic | $L = 1$ | $L = 2$ | $L = 3$ |
| $f_{1,ESG,t}$ | 1.43 | 1.56 | 2.56 |
| $f_{2,ESG,t}$ | – | 2.94 | 2.94 |
| $f_{3,ESG,t}$ | – | – | – 1.84 |
| $f_{1,E,t}$ | 1.51 | 1.71 | 2.47 |
| $f_{2,E,t}$ | – | 3.14 | 3.04 |
| $f_{3,E,t}$ | – | – | – 1.12 |
| $f_{1,S,t}$ | 3.47 | 3.84 | 2.95 |
| $f_{2,S,t}$ | – | 2.44 | 2.45 |
| $f_{3,S,t}$ | – | – | – 0.63 |
| $f_{1,G,t}$ | – 1.13 | – 1.12 | – 0.27 |
| $f_{2,G,t}$ | – | 2.18 | 2.18 |
| $f_{3,G,t}$ | – | – | – 0.38 |

Notes: The table presents the summary statistics for the monthly factor returns (%) from 2013 to 2021. $f_{ESG,t}$ means the factor estimated using the ESG score. The factors are estimated through model (6). The top panel shows the average return over the period and the bottom panel shows the $t$-statistic, calculated as $(\frac{Mean}{Std}\sqrt{T})$.

Table 4. GRS test results.

| LHS | RHS | ESG | E | S | G |
|---|---|---|---|---|---|
| $\{f_{1,n,t}\}$ | Mkt | 0.501 | 0.775 | 0.020 | 0.234 |
| $\{f_{1,n,t}, f_{2,n,t}\}$ | Mkt | 0.025 | 0.000 | 0.001 | 0.138 |
| $\{f_{1,n,t}, f_{2,n,t}, f_{3,n,t}\}$ | Mkt | 0.055 | 0.001 | 0.001 | 0.238 |
| $\{f_{1,n,t}\}$ | FF6 | 0.469 | 0.251 | 0.018 | 0.076 |
| $\{f_{1,n,t}, f_{2,n,t}\}$ | FF6 | 0.020 | 0.000 | 0.000 | 0.066 |
| $\{f_{1,n,t}, f_{2,n,t}, f_{3,n,t}\}$ | FF6 | 0.028 | 0.002 | 0.000 | 0.147 |

Notes: The table presents the GRS test results for the whole period (2013–2022). LHS means left-hand-side factors in equation (10), which is set of ESG factors to be tested. '$n$' here can be ESG, E, S or G, with each column representing one type of factor. RHS is the right-hand-side factors (independent) used in equation (10). The first three rows are the test with only market risk factor and the last three rows controls for FF6 risk factors. Each column means one type of ESG score. 'FF6' means Fama/French 5 factors (market, size, value, profitability and investment) plus momentum factor.

a rolling window scheme, with a window size of 48 months and a step size of one month.

Figure 4 presents the GRS test results over time, with the FF6 as the RHS factors. The test results changes over time. Not all periods have significant test results. In terms of the ESG factor (Panel A, left graph), we find that in some periods the nonlinear factors ($L = 2$ and $L = 3$) have a more significant test results than the linear setting ($L = 1$). This holds in particular during 2018. Regarding the E factor, the linear factors under $L = 1$ do not produce the significant test result in all periods, while under the non-linear setting we have significant test results in most periods. The same observation applies for the S factor (Panel B, left graph), where we see huge improvements in terms of test results from the linear setting to the non-linear setting. The situation is different for the G factor, where the non-linear setting produced a worse test result than the linear setting.

Figure 5 presents the dynamic spanning test results under $L = 3$. As for the ESG factor (Panel A, left graph), combining the GRS test results shown by the green dotted line, we can see that in most windows, the significant GRS result comes from $f_{3,ESG,t}$, while test results for $f_{1,ESG,t}$ and $f_{2,ESG,t}$ is insignificant. In comparison, for the E factor, the significant results come from $f_{2,E,t}$. As for the S score, most of the time the linear factor does not yield significant results. The results for the G score under the rolling window are similar to the whole period, with the test results being insignificant most of the time. In figures A.2 and A.3 of the Online Appendix B, we also provide the test results using only the market factor. Results are quite similar to that of FF6 factors.

Summarizing, the dynamic analysis under the rolling window further confirms the results of the whole period: we see significant results for the non-linear setting for ESG, E and S score most of the times. In addition, it may not be best to always choose the highest $L$, because in some periods choosing a $L = 2$ or even $L = 1$ is enough. The performance of the non-linear factor is also time-varying: there are periods where the tests suggest no significant improvements. Finally, as a robustness check, we provide in Figures A.4 – A.5 of the Online Appendix A the test results with 60-month window size, and the results are quite similar to the 48-month window.

## 5. Chasing the non-linearity

In previous sections, we have identified and shown the need of applying the non-linear setting when examining the pricing of ESG performance in the cross-section of stock returns. A more important question is what drives the non-linear ESG premium to avoid the concern of data mining. We study two possible drivers of the non-linearity: the ESG sentiment channel and the measurement error in the ESG score.

### 5.1. *ESG sentiment and the cross-sectional distribution of ESG scores*

**5.1.1. The model.** To examine whether the non-linear factor is driven by ESG sentiment, we propose the following logistic regression:

$$\Pr\left(y_{l,n,m} = 1 \mid X'_{n,m}\right) = \frac{\exp\left(\gamma_n + X'_{n,m}\Gamma_n\right)}{1 + \exp\left(\gamma_n + X'_{n,m}\Gamma_n\right)}, \qquad (12)$$

Table 5. Spanning test results.

| | Alpha | $f_{1,ESG,t}$ | $f_{2,ESG,t}$ | RHS | $R^2$ |
|---|---|---|---|---|---|
| Panel A: Spanning test results for the ESG factor | | | | | |
| $f_{1,ESG,t}$ | 0.148* | – | – | Mkt | 0.03 |
| | (1.82) | – | – | | |
| $f_{2,ESG,t}$ | 0.072** | 0.235*** | – | Mkt | 0.20 |
| | (2.27) | (4.50) | – | | |
| $f_{3,ESG,t}$ | 0.014*** | – 0.407*** | – 0.167*** | Mkt | 0.74 |
| | (3.04) | ( – 10.52) | ( – 4.08) | | |
| $f_{1,ESG,t}$ | 0.180* | – | – | FF6 | 0.23 |
| | (1.83) | – | – | | |
| $f_{2,ESG,t}$ | 0.063** | 0.253*** | – | FF6 | 0.26 |
| | (2.31) | (4.02) | – | | |
| $f_{3,ESG,t}$ | 0.015*** | – 0.465*** | – 0.128*** | FF6 | 0.80 |
| | (3.24) | ( – 11.07) | ( – 4.31) | | |
| Panel B: Spanning test results for the E factor | | | | | |
| | Alpha | $f_{1,E,t}$ | $f_{2,E,t}$ | RHS | $R^2$ |
| $f_{1,E,t}$ | 0.117 | – | – | Mkt | 0.11 |
| | (1.32) | – | – | | |
| $f_{2,E,t}$ | 0.184*** | 0.091* | – | Mkt | 0.08 |
| | (3.57) | (1.63) | – | | |
| $f_{3,E,t}$ | 0.068* | – 0.373*** | – 0.357*** | Mkt | 0.58 |
| | (1.85) | ( – 9.76) | ( – 5.41) | | |
| $f_{1,E,t}$ | 0.151* | – | – | FF6 | 0.20 |
| | (1.70) | – | – | | |
| $f_{2,E,t}$ | 0.157*** | 0.083 | – | FF6 | 0.26 |
| | (3.21) | (1.53) | – | | |
| $f_{3,E,t}$ | 0.049 | – 0.379*** | – 0.174*** | FF6 | 0.73 |
| | (1.59) | ( – 11.51) | ( – 2.89) | | |
| Panel C: Spanning test results for the S factor | | | | | |
| | Alpha | $f_{1,S,t}$ | $f_{2,S,t}$ | RHS | $R^2$ |
| $f_{1,S,t}$ | 0.147* | – | – | Mkt | 0.11 |
| | (1.82) | – | – | | |
| $f_{2,S,t}$ | 0.104** | 0.242*** | – | Mkt | 0.18 |
| | (2.27) | (4.50) | – | | |
| $f_{3,S,t}$ | 0.073*** | – 0.318*** | – 0.205*** | Mkt | 0.66 |
| | (3.04) | ( – 10.52) | ( – 4.08) | | |
| $f_{1,S,t}$ | 0.152* | – | – | FF6 | 0.16 |
| | (1.83) | – | – | | |
| $f_{2,S,t}$ | 0.109** | 0.224*** | – | FF6 | 0.21 |
| | (2.31) | (4.02) | – | | |
| $f_{3,S,t}$ | 0.077*** | – 0.328*** | – 0.213*** | FF6 | 0.70 |
| | (3.24) | ( – 11.07) | ( – 4.31) | | |
| Panel D: Spanning test results for the G factor | | | | | |
| | Alpha | $f_{1,G,t}$ | $f_{2,G,t}$ | RHS | $R^2$ |
| $f_{1,G,t}$ | 0.027 | – | – | Mkt | 0.03 |
| | (0.35) | – | – | | |
| $f_{2,G,t}$ | 0.073* | 0.033 | – | Mkt | 0.01 |
| | (1.94) | (0.69) | – | | |
| $f_{3,G,t}$ | – 0.013 | – 0.406 | – 0.271 | Mkt | 0.78 |
| | ( – 0.66) | ( – 17.05) | ( – 5.57) | | |
| $f_{1,G,t}$ | – 0.025 | – | – | FF6 | 0.19 |
| | ( – 0.34) | – | – | | |
| $f_{2,G,t}$ | 0.078** | 0.062 | – | FF6 | 0.12 |
| | (2.07) | (1.24) | – | | |
| $f_{3,G,t}$ | – 0.016 | – 0.441*** | – 0.175*** | FF6 | 0.86 |
| | ( – 1.03) | ( – 20.99) | ( – 4.22) | | |

Notes: The table presents the coefficient in the spanning test results. All ESG factors come from $L = 3$. Panel A is the results for the ESG factor, Panel B for the E factor, Panel C for the S factor and Panel D for the G factor. In each panel, the first three rows is the test with only market return and the last three rows control for FF6 risk factors. We report in parenthesis the $t$-statistics. 'FF6' means Fama/French 5 factors (market, size, value, profitability and investment) plus momentum factor.

*Statistical significance at the 10% level; **Statistical significance at the 5% level; ***Statistical significance at the 1% level.

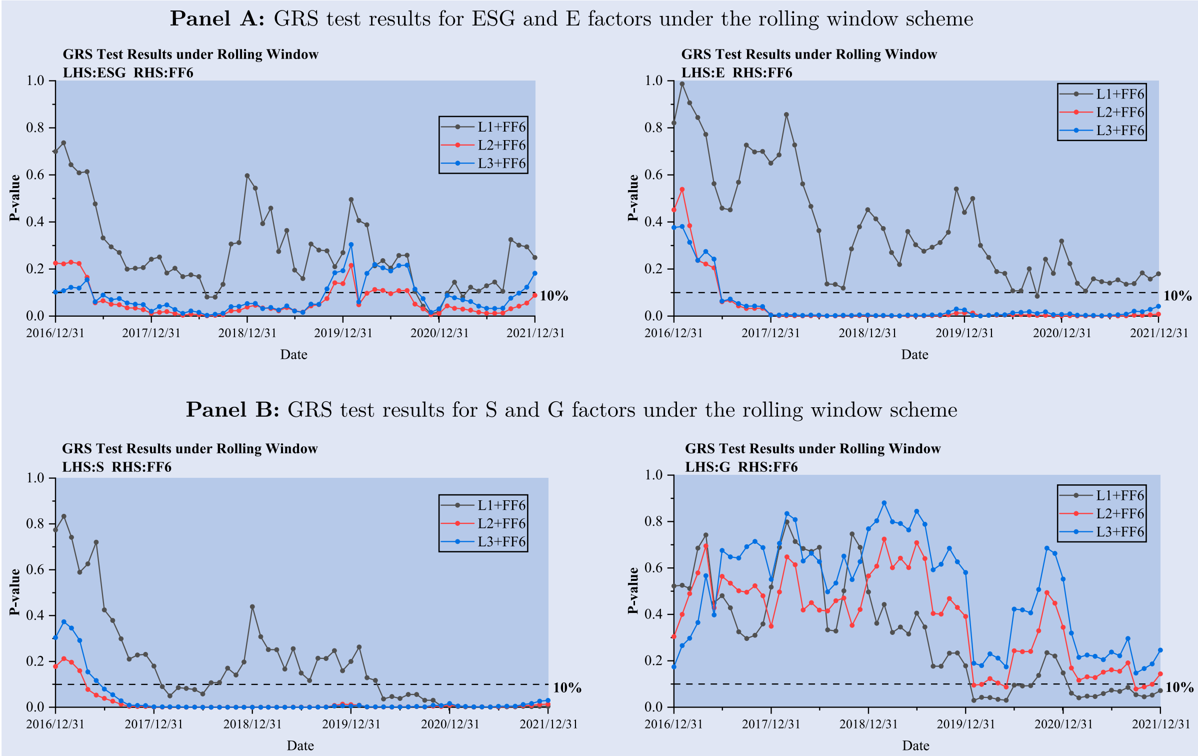


Figure 4. Dynamic GRS test. (a) Panel A: GRS test results for ESG and E factors under the rolling window scheme (b) Panel B: GRS test results for S and G factors under the rolling window scheme.
Notes: The graph presents the GRS test results under the rolling window scheme, with a window size of 48 months and a step size of one month. LHS means left-hand-side factors inequation (10), which is set of ESG factors to be tested. RHS is the right-hand-side factors (independent) used in equation (10). All RHS factors are FF6 factors. 'L1' means we set $L = 1$, with the LHS being $\{f_{1,ESG,t}\}$. 'L2' and 'L3' means $L = 2$ and $L = 3$ respectively, with the corresponding LHS shown in equation (9). Panel A shows the results for ESG and E factors and Panel B shows results for S and G factors.

where $\exp(z)$ denotes the exponential function; $n$ denotes the type of ESG risk factor, and we only consider $n =$ ESG, E and S scores, because as previous results showed, the linear/non-linear G risk factors are mostly insignificant. $y_{l,n,m} = 1$ if, within window $m$, $m = 1, \ldots, M$, the estimated coefficient of $\delta_{l,n,m}$ in the following spanning tests:

$$\begin{cases} f_{1,n,t_m} = \delta_{1,n,m} + \sum_{j=1}^{J} \beta_{1,n,j} f_{j,t_m} + \varepsilon_{1,n,t_m} \\ f_{2,n,t_m} = \delta_{2,n,m} + \beta_{1,n,j} f_{1,n,t_m} + \sum_{j=1}^{J} \beta_{2,n,j} f_{j,t_m} + \varepsilon_{2,n,t_m} \\ f_{3,n,t_m} = \delta_{3,n,m} + \beta_{1,n,j} f_{1,n,t_m} + \beta_{2,n,j} f_{2,n,t_m} \\ + \sum_{j=1}^{J} \beta_{3,n,j} f_{j,t_m} + \varepsilon_{3,n,t_m} \end{cases}, \tag{13}$$

is significant at the 10% level and 0 otherwise ($l = 1$, 2, and 3). $f_{j,t_m}$ $(j = 1, \ldots, J)$ are Fama/French six risk factors and $t_m = 1, \ldots\ldots, 48$. We set the window size to be 48 months and a step size of one month (in total, we have $M = 60$ observations in the time series dimension). $\delta_{1,n}$, $\delta_{2,n}$ and $\delta_{3,n}$ are pricing errors from the spanning regression, which can be interpreted as the unique information brought by the linear/non-linear ESG risk factors. $f_{1,n,t}$, $f_{2,n,t}$ and $f_{3,n,t}$ are all estimated from the non-linear cross-section regression with $L = 3$.

We use the following set of independent variables ($X'_{n,m}\Gamma_n$):

$$X'_{n,m}\Gamma_n = \gamma_{n,1}\triangle RA_m + \gamma_{n,2}\triangle \text{Sent}_m + \gamma_{n,3}\triangle\sigma_{n,m} + \gamma_{n,4}\triangle \text{Skew}_{n,m}$$
$$\gamma_{n,5}\triangle\sigma_{n,m} \times \triangle\text{Sent}_m + \gamma_{n,6}\triangle\text{Skew}_{n,m} \times \triangle\text{Sent}_m + \varepsilon_m. \tag{14}$$

In equation (14), we consider several possible determinants of non-linearity. We follow Lioui and Tarelli (2022) to control for the risk aversion of investors ($RA_m$). The index is constructed by Bekaert *et al.* (2022), which measures the risk aversion level in the market. Since high-ESG companies are considered high-quality companies, investors' taste over high-ESG companies might only be the taste over high-quality investment, which would happen when market uncertainty becomes high (as occurred during the COVID-19 pandemic). If that is the case, shocks in the ESG factor can be explained by shocks in the risk aversion, which then means that risk aversion should contribute to pricing errors. $RA_m$ is the average monthly risk averse index over window $m$.

Previous studies show that one of the major drivers of the linear ESG factor is investor's ESG-related concern (Choi *et al.* 2020, Pástor *et al.* 2022). For this reason, we add the climate attention index ($Sent_m$) from Google Trends for the US as a control. The Google Trends index is based on the search volume of a given key word(s) in a certain period and in a certain region. We use the topic 'climate change', 'ESG' and 'CSR' (corporate social responsibility) specified

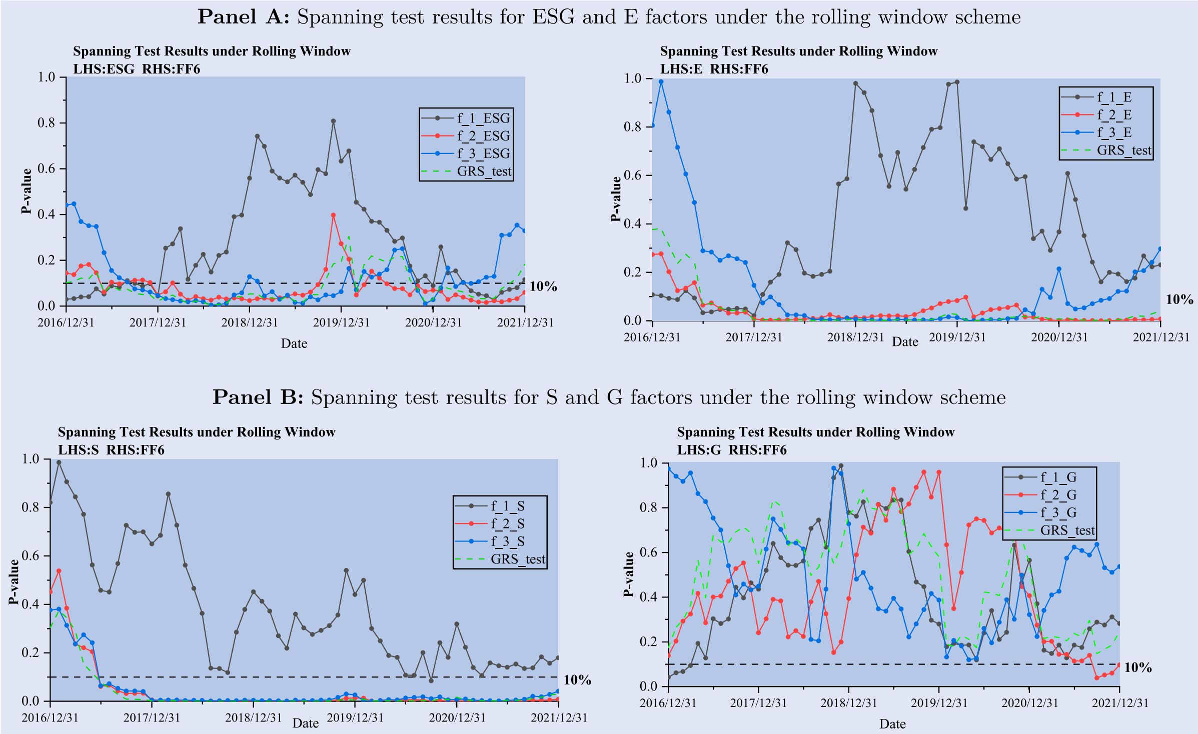


Figure 5. Dynamic spanning test. (a) Panel A: Spanning test results for ESG and E factors under the rolling window scheme (b) Panel B: Spanning test results for S and G factors under the rolling window scheme

Notes: The figure presents the $p$-value of intercept of spanning test under the rolling window scheme, with a window size of 48 months and a step size of one month. LHS means left-hand-side factors inequation (11), which is set of ESG factors to be tested. RHS is the right-hand-side factors (independent) used inequation (11). All RHS factors are FF6 factors. 'f_1_ESG' means $f_{1,ESG,t}$ and so on. All ESG factors comes from $L = 3$. We also provide the $p$-value of GRS test under $L = 3$ in green dotted plots. Panel A shows the results for ESG and E factor and Panel B shows results for S and G factor.

by Google.† A high value of the index is associated with a high search volume related to ESG-related issues, and thus a high ESG sentiment in this region. We download the monthly data. $Sent_m$ is the average monthly sentiment over window $m$.

We consider the ESG score distribution to be a potential driver of the non-linear factor. Specifically, we study the cross-sectional standard deviation ($\sigma_{n,m}$) and skewness ($Skew_{n,m}$). For each month, we calculate cross-section standard deviation and skewness and then estimate the average monthly value of each window $m$. Since the $\sigma_{n,m}$ and $Skew_{n,m}$ are estimated based on the average of rolling windows, $\sigma_{n,m}$ $Skew_{n,m}$will have few changes between adjacent windows, implying potential high auto-correlation and multicollinearity. To ensure a robust estimation, we use the change between adjacent windows (e.g. $\triangle\sigma_{n,m} = \sigma_{n,m} - \sigma_{n,m-1}$) in the regression analysis.

As discussed in section 2, if investors ESG preference drives them to buy/sell high-ESG companies, then such a channel of impact could be altered by how ESG performance is distributed across the market. If all companies in the market have the same ESG score (investors cannot differentiate companies by their ESG performances), then it makes no sense for investors to buy/sell companies based on their ESG score (the channel of climate concern disappears under this scenario). Therefore, it is also possible that the investors ESG sentiment is 'distorted' by the distribution of ESG scores and if that is the case, such distortion would be captured by the non-linear ESG factor. If the above discussion holds, the ESG score distribution is driving the non-linear factor through the channel of climate sentiment. To examine this, we add interaction between the ESG score distribution and climate sentiment.

The sample period is from January 2013 to December 2021. We first present the distribution of ESG scores across time in figure 6. The distribution of ESG scores is different among different sub-scores (the distributions of ESG and S score are quite similar, though). This is especially the case for the E score, where there are some extreme shifts in 2018 and 2019. We then provide in figure 7 the skewness and standard deviation across time. In terms of skewness, the G score has opposite skewness to the other scores. The E score has the highest standard deviation and the standard deviation is quite stable for the ESG score. For every score, there are changes in the shape of distribution over the years.

**5.1.2. Regression results.** Table 6 presents results for $\delta_{1,n}$, the pricing error from the linear risk factor. The coefficient is positive and significant for sentiment, which is in line with previous findings that the linear ESG risk factor is, in fact, driven by sentiment. The pricing error is also affected by the distribution. A higher variance corresponds to lower pricing errors from the linear factor. One explanation is that

† Google Trends allows users to search either based on single keywords or on a topic, which combines multiple keywords related to the topic. For example, the topic 'climate change' includes keywords (search terms) such as 'climate change progress,' 'Paris agreement,' 'climate crisis,' 'climate change debate,' and others.

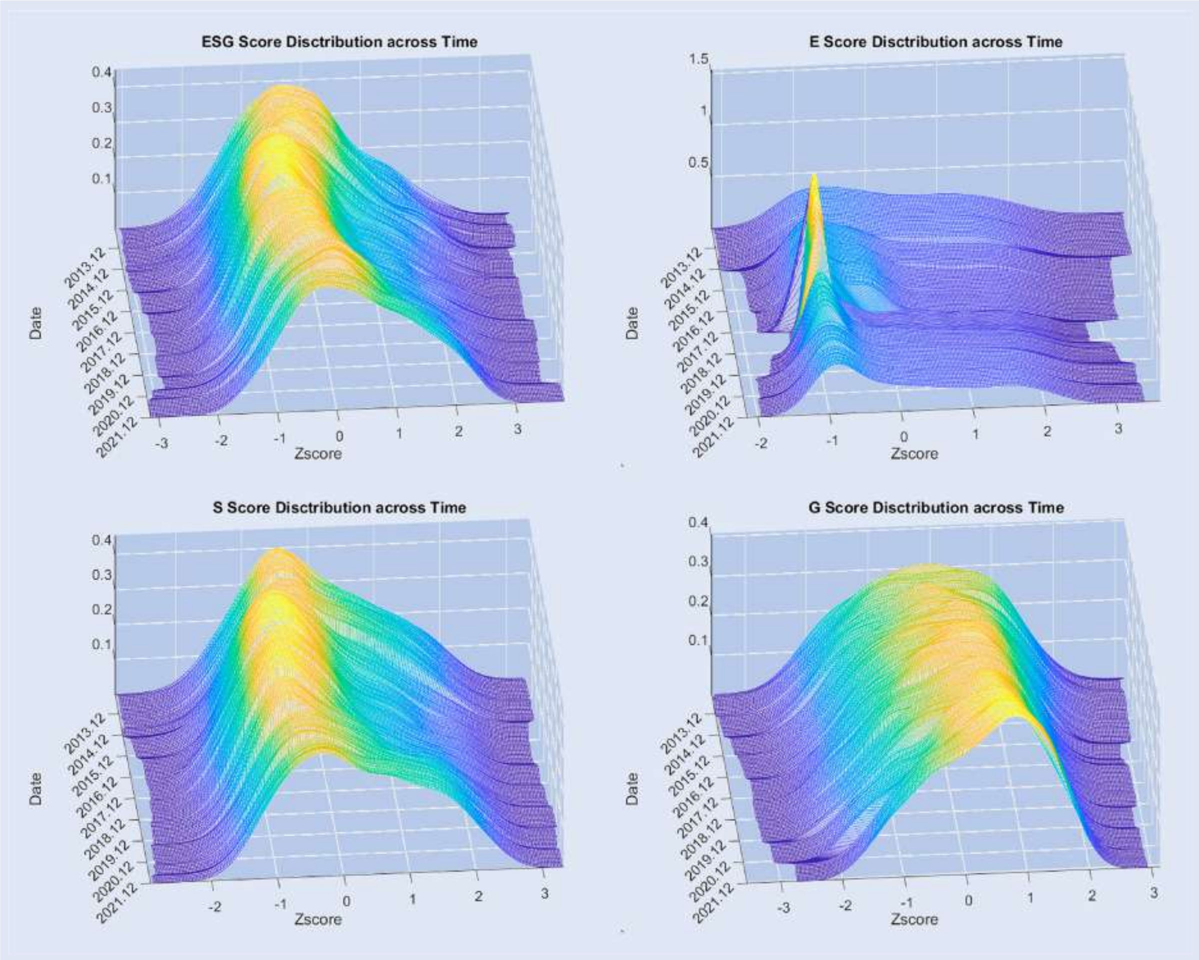


Figure 6. Distribution of ESG scores over time.
Notes: The graph shows the cross section distribution of ESG score in each month. To draw the distribution, we first standardize ESG scores cross-sectionally and draw the empirical distribution based on companies with available ESG scores at month $t$. The ESG scores are from January 31, 2013 to November 31, 2021.

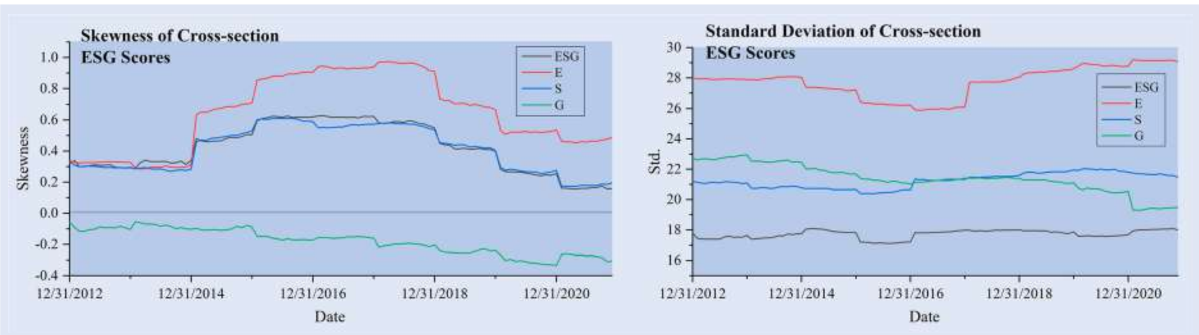


Figure 7. Cross-section skewness and standard deviation over time.
Notes: The graph shows cross-section skewness and standard deviation. The standard deviation is calculated using the original value of the ESG score. The ESG scores are from January 31, 2013 to November 31, 2021.

if the distribution becomes more spread out, the linear risk factor should perform poorly in capturing the ESG pricing. Regarding the interaction effect, we find that the coefficient is negative for the variance, which further indicates that how sentiment influences ESG pricing is 'distorted' by the ESG score distribution. Higher skewness corresponds to higher pricing error.

Mathematically speaking, even under $L = 1$, we should find impact from cross-sectional distribution of ESG scores, because as equation (5) shows, the weight of the long-short portfolio of $f_{l,n,t}$ is based on the ESG score, which means that $f_{1,X,t}$ also changes to the cross-sectional distribution of ESG scores. However, there are several points worth noticing from the empirical results. First, it seems that positive changes in the skewness is associated with higher pricing error, while positive changes in the standard deviation indicates lower pricing errors. Second, we find evidence of cross-sectional distribution interacting with the ESG attention.

Table 6. Regression results for $\delta_{1,n}$.

| Dependent | $\Pr(\delta_{1,n} \neq 0 \mid X'_{n,m}B)$ | | | |
|---|---|---|---|---|
| Column | (1) | (2) | (3) | (4) |
| $\triangle RA$ | − 0.0396 | − 0.0370 | − 0.0396 | − 0.0422 |
| | ( − 0.82) | ( − 0.68) | ( − 0.87) | ( − 0.69) |
| $\triangle Sent$ | 0.1211*** | 0.2159*** | 0.0997*** | 0.2021*** |
| | (4.88) | (5.65) | (3.96) | (4.43) |
| $\triangle\sigma$ | − 0.1623*** | – | − 0.1626*** | – |
| | ( − 5.48) | | ( − 5.51) | |
| $\triangle Skew$ | – | 0.1857*** | – | 0.1776*** |
| | | (4.29) | | (3.73) |
| $\triangle\sigma \times \triangle Sent$ | – | – | − 0.0567* | – |
| | | | ( − 1.68) | |
| $\triangle Skew \times \triangle Sent$ | – | – | – | − 0.0162 |
| | | | | ( − 0.49) |
| $N$ | 180 | 180 | 180 | 180 |
| Adjusted $R^2$ | 0.24 | 0.16 | 0.25 | 0.16 |

Notes: The table presents the regression results for the pricing error from the linear factor, estimated from $L = 3$ in the cross-section regression. We consider three types of risk factors: ESG, E and S factors. For each risk factor, we estimate the spanning regression under a rolling window setting and test if the intercept ($\delta$, the pricing error) is non-zero. If the coefficient is non-zero at 10% level, we set $y_{1,n,m} = 1$. For each factor, we have $m = 1, \ldots, M, M = 60$ observations. Then, we put the $y_{n,m}$ for each factor into a pooled logistic regression and report the marginal effect of each independent variable.
*Statistical significance at the 10% level; **Statistical significance at the 5% level; ***Statistical significance at the 1% level.

Table 7 presents the results for $\delta_{2,n}$ and $\delta_{3,n}$, the pricing errors from non-linear factors. Unlike the linear factor, higher variance is associated with higher pricing errors from the non-linear risk factors. In comparison, skewness is only captured by the cubic risk factor ($\delta_{3,n}$), and a negative change in skewness is associated with higher pricing errors. Again, we find significant interaction terms between sentiment and cross-sectional variance.

In summary, all these elements suggest that the channel through which ESG impacts is altered. ESG sentiment is primarily captured by the linear factor. In addition, the interaction term between ESG sentiment and score distribution is significant for both linear and non-linear factors, and the non-linear factor captures the effect from ESG score distributions.

All these elements indicate that the channel of ESG sentiment impact might be altered. Mathematically speaking, from equation (A10) in the Appendix, we can see that higher moments of ESG score distribution are part of both linear and non-linear factors. Economically speaking, if the market takes into account the ESG score distribution in asset pricing, then the distribution itself may also affect the pricing. Let us consider two distinct situations: when all companies in the market have a score of 50 (out of 100), and when half of the company have 0 score and the other half 100. The non-linear factor would certainly be different under the two situations.

### 5.2. *ESG score uncertainty*

**5.2.1. Non-linearity under different data providers.** If non-linear ESG factors are associated with the cross-sectional distribution of the ESG score, then a sensible conjecture would be that we might get a different non-linear ESG factor if we change the ESG provider. To test the conjecture, we construct the ESG factor using the Bloomberg ESG score. Figure A.6 in the online Appendix B shows the cross-section correlation between the ESG score of two data providers. While we do observe a relatively high correlation in terms of the E score, the correlation of the S score and the G score is low. In addition, the correlation of ESG scores is gradually decreasing over time, which means that the divergence between different ESG providers is increasing. Figure A.7 shows the distribution of ESG scores of Bloomberg across time. As can be seen, the distribution is quite different from that of Eikon in figure 6. For example, compared to the Eikon database, fewer companies in the Bloomberg have high-ESG score.

Table 8 shows the factor statistics. Similar to the results using Eikon, the average return for the linear factor at $L = 1$ is positive and is not significant for non-linear factors. Figure 8 presents the cumulative value of the ESG factors. In terms of the linear factor, there are similarities between Eikon and Bloomberg. For example, we see large positive shocks in recent years in both series. However, the non-linear factors are quite different for different pillar scores.

We then present the LME from $L = 1$ to $L = 3$ in figure 9. The results are opposite to that of Eikon in terms of the ESG and S score—at $L = 3$, we observe a convex shape. Finally, table 9 presents the static GRS test results. We see a sharp increase in the significance for the ESG score from $L = 1$ to $L = 3$, which is different from the previous results using Eikon. Therefore, as in the finding in Lioui and Tarelli (2022), changing the data provider will also change the characteristic of ESG factors.

If ESG scores are converted into percentile ranks and then standardized into $Z$-scores, then the above results imply that

Table 7. Regression results for $\delta_{2,n}$ and $\delta_{3,i}$.

| Dependent | $\Pr(\delta_{2,n} \neq 0 \mid X'_{i,m}B)$ | | | | $\Pr(\delta_{3,n} \neq 0 \mid X'_{i,m}B)$ | | | |
|---|---|---|---|---|---|---|---|---|
| Column | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $\triangle RA$ | − 0.0755* | 0.0087 | − 0.0577 | 0.0063 | − 0.0384 | − 0.0508 | − 0.0663 | − 0.0514 |
| | (− 1.81) | (0.18) | (− 1.37) | (0.13) | (− 0.92) | (− 1.13) | (− 1.62) | (− 1.14) |
| $\triangle Sent$ | − 0.2441*** | − 0.1771*** | − 0.2609*** | − 0.2332*** | − 0.1342*** | − 0.2171*** | − 0.1306*** | − 0.2282*** |
| | (− 5.91) | (− 3.12) | (− 4.10) | (− 3.05) | (− 3.23) | (− 4.02) | (- − 4.04) | (− 3.70) |
| $\triangle\sigma$ | 0.3051*** | – | 0.3253*** | – | 0.1774*** | – | 0.2262*** | – |
| | (5.50) | | (5.04) | | (4.55) | | (5.78) | |
| $\triangle Skew$ | – | − 0.0508 | – | − 0.0802 | – | − 0.1923*** | – | − 0.1981*** |
| | | (− 0.81) | | (− 1.17) | | (− 3.25) | | (− 3.19) |
| $\triangle\sigma \times \triangle Sent$ | – | – | 0.2312*** | – | – | – | − 0.1282*** | – |
| | | | (3.40) | | | | (− 4.29) | |
| $\triangle Skew \times \triangle Sent$ | – | – | – | − 0.0672 | – | – | – | − 0.0134 |
| | | | | (− 1.14) | | | | (− 0.31) |
| $N$ | 180 | 180 | 180 | 180 | 180 | 180 | 180 | 180 |
| Adjusted $R^2$ | 0.25 | 0.07 | 0.31 | 0.07 | 0.14 | 0.09 | 0.21 | 0.09 |

Notes: The table presents the regression results for the pricing error from the non-linear factor, all estimated from $L = 3$ in the cross-section regression. We consider three types of risk factors: ESG, E and S factors. For each risk factor, we estimate the spanning regression under a rolling window setting and test if the intercept ($\delta$, the pricing error) is non-zero. If the coefficient is non-zero at 10% level, we set $y_{n,m} = 1$. For each factor, we have $m = 1, \ldots, M, M = 60$ observations. Then, we put the $y_{n,m}$ for each factor into a pooled logistic regression and report the marginal effect of each independent variable. Columns (1)–(4) are results for the square factor while Columns (5)–(8) for the cubic factor.

*Statistical significance at the 10% level; **Statistical significance at the 5% level; ***Statistical significance at the 1% level.

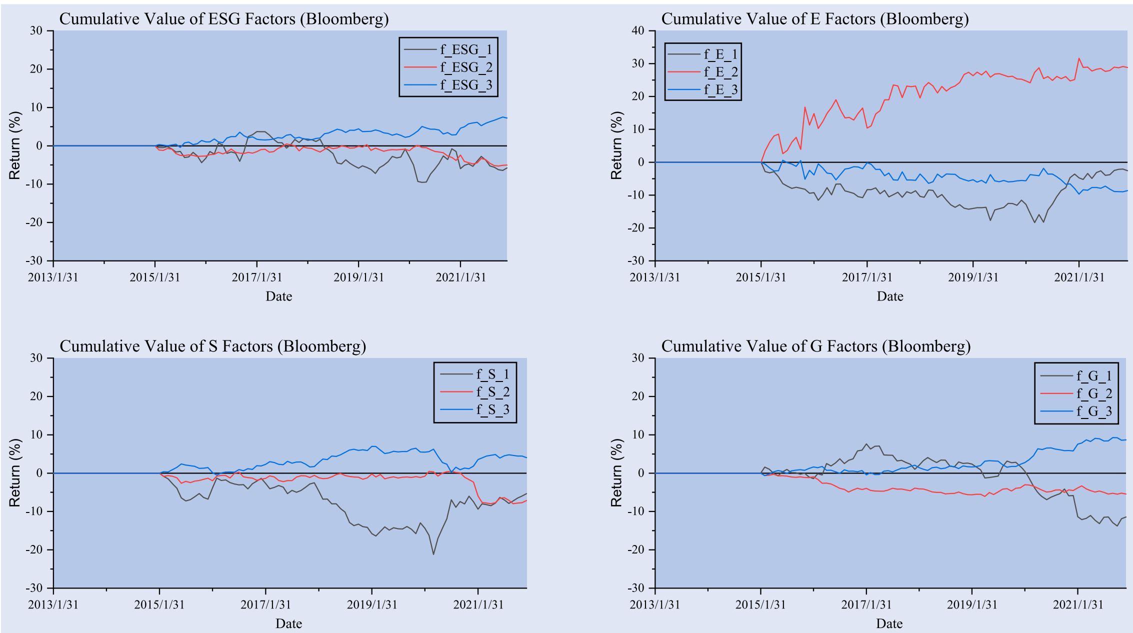


Figure 8. Cumulative Value of ESG Factors under Bloomberg.
Notes: The graph shows the cumulative monthly factor returns: $\sum_{t=1}^{T} f_{l,n,t}$. The factors are calculated using model (6) with $L = 3$ and the ESG score coming from Bloomberg ESG Score.

Eikon and Bloomberg produced different rankings for the same company. In addition, the score distributions might also be different between the two data providers. Such differences come from that fact that, on the one hand, there are measurement errors in the ESG score—for the same indicator, two providers might have different evaluations; on the other hand, scoring methodologies are different between the two (Berg *et al.* 2022). All these factors contribute to numerical differences in the ESG score used in the cross-sectional regression (and then a different result). The variability in ESG rankings across different providers may significantly affect the nonlinear impact of ESG factors on returns. The next section will delve deeper into this issue.

**5.2.2. Taking out ESG uncertainty.** If the results change to the data provider, it means that part of the cross-section variance/skewness is contributed to by the uncertainty of ESG scores, and thus it is hard to make an assertive inference on whether the ESG-stock price relationship truely deviates from the linearity. Avramov *et al.* (2022) show that such uncertainty in the ESG score (defined as the standard deviation

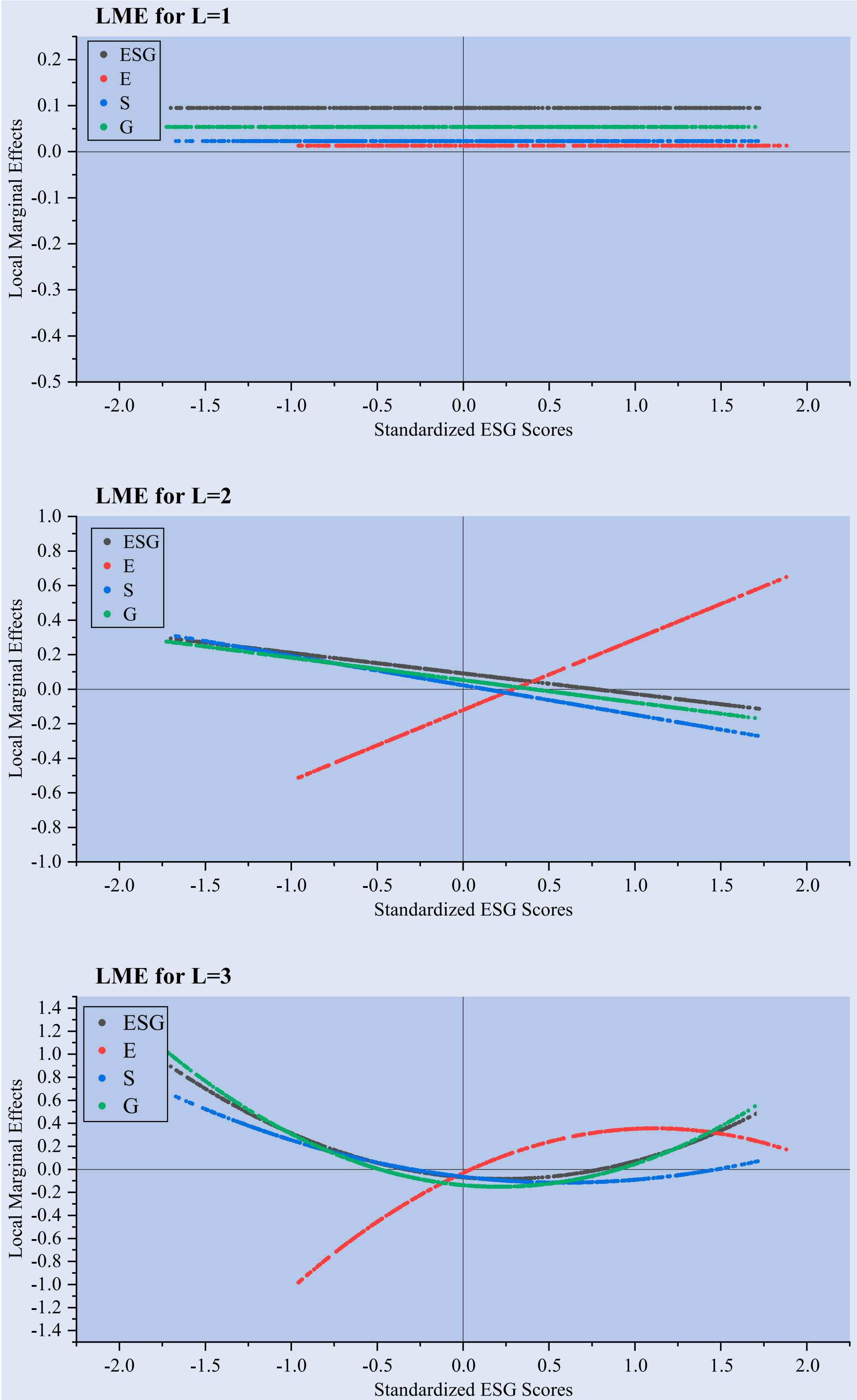


Figure 9. Local marginal effects (Bloomberg).
Notes: The graph presents the local marginal effect for $L = 1$, 2 and 3. The ESG score is from Bloomberg ESG score.

of ESG scores from six major providers) has been priced in the assets. In other words, if the non-linearity simply comes from the measurement error in the ESG score, then it makes no sense considering explicitly non-linearity in the factor construction process. We examine in this sub-section whether the non-linearity still exists after the ESG score certainty has been minimized.

Since we have access to only two ESG score providers, we take out the uncertainty in the ESG score by conducting the cross-sample analysis. Specifically, we repeat the previous

Table 8. Summary statistics of monthly non–linear factor returns.

| Mean | $L = 1$ | $L = 2$ | $L = 3$ |
|---|---|---|---|
| $f_{1,ESG,t}$ | 0.09 | 0.09 | − 0.07 |
| $f_{2,ESG,t}$ | – | − 0.06 | − 0.06 |
| $f_{3,ESG,t}$ | – | – | 0.09 |
| $f_{1,E,t}$ | 0.01 | − 0.12 | − 0.03 |
| $f_{2,E,t}$ | – | 0.20 | 0.35 |
| $f_{3,E,t}$ | – | – | − 0.10 |
| $f_{1,S,t}$ | 0.02 | 0.02 | − 0.06 |
| $f_{2,S,t}$ | – | − 0.09 | − 0.09 |
| $f_{3,S,t}$ | – | – | 0.05 |
| $f_{1,G,t}$ | 0.05 | 0.05 | − 0.14 |
| $f_{2,G,t}$ | – | − 0.06 | − 0.07 |
| $f_{3,G,t}$ | – | – | 0.10 |
| $t$-statistic | $L = 1$ | $L = 2$ | $L = 3$ |
| $f_{1,ESG,t}$ | 0.96 | 0.92 | − 0.45 |
| $f_{2,ESG,t}$ | – | − 0.95 | − 0.97 |
| $f_{3,ESG,t}$ | – | – | 1.45 |
| $f_{1,E,t}$ | 0.12 | − 0.88 | − 0.18 |
| $f_{2,E,t}$ | – | 1.73 | 1.20 |
| $f_{3,E,t}$ | – | – | − 0.68 |
| $f_{1,S,t}$ | 0.24 | 0.23 | − 0.39 |
| $f_{2,S,t}$ | – | − 1.09 | − 1.10 |
| $f_{3,S,t}$ | – | – | 0.68 |
| $f_{1,G,t}$ | 0.74 | 0.73 | − 1.02 |
| $f_{2,G,t}$ | – | − 1.43 | − 1.45 |
| $f_{3,G,t}$ | – | – | 1.76 |

Notes: The table presents the summary statistics for the monthly factor returns (%) from 2013 to 2021. The factors are estimated through model (6). The top panel shows the average return over the period and the bottom panel shows the $t$-statistic, calculated as $(\frac{Mean}{Std}\sqrt{T})$.

analysis using only companies with similar ESG scores in both databases. To do so, we first choose companies that have ESG scores available in both databases. We then rank the companies into percentiles, so that for each company we have one percentile rank in each month $t$ in both databases. Finally, we keep companies whose percentile ranks in the two databases are similar (defined as the difference of percentile rank being smaller than 0.05). In total, we are left with around 100 companies in each database.

We first present the LME at $L = 3$ in figure A.8 in the online Appendix for both databases. We still observe deviations from linearity, with the non-linearity being concave for the S score, which is the same as the previous results. Figure B.7 in the online Appendix B shows the dynamic GRS test for the ESG score and S score for both databases. As can be seen, for the ESG score, the $p$-value is smaller when $L$ is larger than one. The similar happens to the S score in all windows. Therefore, even when taking out the uncertainty in the ESG score, we still observe departures from linearity, which implies that ESG score does present non-linearity in terms of asset pricing. In addition, the pattern of LME using the common sample set is different from previous LMEs of both Eikon and Bloomberg in terms of E and G factors, which means that part of the non-linearity comes from the ESG uncertainty.

In sum, the presence of ESG uncertainty does change the estimation of non-linear ESG factors, but it does not prevent us from using the non-linear setting. First, as the cross-sample analysis shows, non-linearity still exists after using the common sample. Second, as the regression analysis shows, cross-section skewness and standard deviation will interact with climate sentiment, which means that the presence of non-linearity, though it could be small, may also have a huge impact in terms of asset pricing when investors climate sentiment is high. Therefore, a better strategy would be to try to eliminate the noisiness of the ESG score (e.g. Berg *et al.* 2021).

## 6. Robustness and further results

In this section, we explore several alternative settings in addition to the main discussion. The Online Appendix includes the corresponding Tables and Figures.

### 6.1. *Non-linearity during early years*

The concept of 'ESG' is a recent phenomenon (Pollman 2024). However, the idea of sustainable investing dates back to the 1980s, with earlier literature using CSR (corporate social responsibility) as a proxy for ESG-related investing. This trend is evident in figure A.10, where the Google Trends index for the topic 'ESG' shows significant growth

Table 9. GRS test results for the whole period (Bloomberg).

| LHS | RHS | ESG | E | S | G |
|---|---|---|---|---|---|
| $\{f_{1,ESG,t}\}$ | Mkt | 0.982 | 0.376 | 0.368 | 0.385 |
| $\{f_{1,ESG,t}, f_{2,ESG,t}\}$ | Mkt | 0.932 | 0.073 | 0.422 | 0.514 |
| $\{f_{1,ESG,t}, f_{2,ESG,t}, f_{3,ESG,t}\}$ | Mkt | 0.447 | 0.252 | 0.571 | 0.246 |
| $\{f_{1,ESG,t}\}$ | FF6 | 0.550 | 0.434 | 0.713 | 0.297 |
| $\{f_{1,ESG,t}, f_{2,ESG,t}\}$ | FF6 | 0.608 | 0.095 | 0.343 | 0.502 |
| $\{f_{1,ESG,t}, f_{2,ESG,t}, f_{3,ESG,t}\}$ | FF6 | 0.382 | 0.281 | 0.382 | 0.190 |

Notes: The table presents the GRS test results for the whole period. LHS means left-hand-side factors in equation (10), which is set of ESG factors to be tested. RHS is the right-hand-side factors (independent) used in equation (10). The first three rows is the test with only market return and the last three rows controls for FF6 risk factors. Each column means one type of ESG score.

over the past five years. Studies using earlier datasets also provide evidence of ESG factor pricing (e.g. Brammer *et al.* 2006, Krueger 2015, Becchetti *et al.* 2018). Based on this evidence, we predict that non-linearity likely existed during earlier periods, though it may have been weaker compared to more recent periods.

We repeat the main analysis using samples from January 2004 to December 2012. Specifically, we construct the ESG risk factors, conduct the dynamic spanning test, and compare the results with those of the most recent period. Although Eikon provides ESG scores dating back to 2002, only a limited number of companies had available data in the cross-section during the first year. To ensure a sufficient number of companies for capturing the ESG risk factor, we start our sample from January 2004, which also includes ESG score data from the fiscal year 2003. Over this period, the cross-section expands from approximately 300 companies at the beginning of the sample to over 700 companies.

Table A.1 presents the summary statistics of the ESG risk factors for the early period (2004–2013, Panel A) and the most recent period (2013–2021, Panel B). The linear risk factor is less significant during the early years, as indicated by smaller $t$-statistics. However, the $t$-statistics for ESG, E, and S factors under $L = 2$ are significant, suggesting the potential presence of non-linearity.

Figure A.11 presents the linear and non-linear ESG risk factors from 2004 to 2013. We observe large positive shocks during the 2008 financial crisis, which is in-line with the observation in Lins *et al.* (2017), who find that during 2008 financial crisis, companies with better ESG performance outperform those with worse performance. Lins *et al.* (2017) further explain that better ESG performance will bring 'social trust' that serves as a protection during crisis times (people will not fire sell those companies). In addition, for time periods before 2008, we observe evidence of pricing only for E risk factors but not others (with more fluctuations than other factors). This is also in-line with earlier surveys that investors mainly focus on the environmental aspects when talking about ESG.

Finally, figure A.12 presents the GRS test and the spanning test over time. We find evidence of non-linearity for the ESG, E, and S factors. In most cases, the linear risk factor is not significant, while the non-linear factors carry unique information. For the G risk factor, results are generally not significant. Thus, while the ESG risk factors are weaker in earlier periods compared to recent years, the non-linear relationship remains evident.

### *6.2. Non-linear ESG factors in the European market*

To determine whether our results represent a phenomenon unique to the U.S. market, we repeat the analysis for the European market, the second-largest market in terms of ESG investment. Since the observed non-linearity arises from the 'distorted' investor preferences influenced by the distribution of ESG scores, we anticipate that the non-linear ESG risk factor will also be present in the European market.

For data collection, we obtain ESG scores, return indices and firm-level characteristics for companies in Europe (European Union and UK) for the period from 2013 to 2022. To mitigate the impact of small-cap companies in the cross-sectional regression, at each time $t$, we exclude companies with a closing price below five euros. Overall, the dataset includes approximately 1,600 companies, with over 700 companies in the early years of the sample and more than 1,300 companies in recent years.

Table A.2 presents the summary statistics of ESG risk factors for Europe (Panel A) and the U.S. (Panel B). Similar to the results for the U.S., we find significant factors for ESG, E, and S components in the European market. Notably, the magnitude of the linear factor is stronger in Europe compared to the U.S., whereas the magnitude of the non-linear factor is weaker. Figure A.13 illustrates the linear and non-linear ESG risk factors for Europe. We observe similar patterns for the linear ESG, E, and S factors, consistent with observations in the U.S. market.

Figure A.14 presents the GRS test and the spanning test over time. Consistent with the results in table A.2, we find limited evidence of non-linearity, as the square and cubic ESG and S factors show no windows with significant test results. However, we observe evidence of non-linearity in the E risk factor, particularly during 2017 and the first half of 2018. Thus, non-linearity persists in other markets, with varying strength and patterns.

### *6.3. Non-linearity in the corporate bond market*

According to model equation (1), ESG is priced in the cross-section of stock returns because investors have a preference for sustainability, a phenomenon examined in various empirical studies (e.g. Serafeim 2020, Dunz *et al.* 2021, Pástor *et al.* 2022). In comparison, the pricing of sustainability in the cross-section of corporate bonds is more nuanced. While the value of stocks depends on the company's overall short- and long-term performance, corporate bond pricing is influenced not only by firm fundamentals but also by factors such as credit risk prior to maturity or the issuer's commitment to sustainability.

In general, there are two types of ESG premiums in the cross-section of corporate bonds. First, there are corporate bonds issued specifically for projects with sustainability goals (green bonds), and there exists a bond premium between green bonds and conventional bonds. Second, firm-level ESG performance may influence general corporate bond pricing. These two cases suggest distinct pricing mechanisms. In the first case, green bonds should command higher prices (lower yields) than conventional bonds, reflecting investors' willingness to pay a premium for the issuer's ESG commitment (e.g. Flammer 2021, Feldhütter *et al.* 2024, Caramichael and Rapp 2024). In the second case, there is no established equilibrium model that comprehensively explains the ESG–corporate bond connection, (credit risk might act as a transmission channel, as suggested by the referee). Nonetheless, we hypothesize that the mechanism is similar to that in the stock market. Huynh and Xia (2021) demonstrate that as investor concern about climate risk increases, companies with superior environmental performance exhibit higher bond prices, driven by the investor taste channel.

In this analysis, we focus on the second case as it aligns with our study, aiming to address the following questions: Does non-linearity persist in the relationship between corporate bond returns and ESG performances? If so, is the magnitude comparable to that observed in the stock market? Our conjecture is that if the ESG–bond return relationship is driven by investor preferences, then the distribution of ESG scores might influence these preferences. However, given the relative illiquidity of the corporate bond market, ESG-related systematic shocks are less likely to be fully incorporated into bond pricing.

The analysis proceeds as follows. We first discusses how we capture both linear and non-linear risk factors in corporate bond returns. We then explain how we test the ESG-related bond risk factor, employing the GRS and spanning tests with a modified factor structure. Next, we describe the data utilized in this study. Finally, presents the empirical results, including comparisons with the stock market.

**6.3.1. Capturing bond ESG risk factor.** Several methods have been proposed in the literature to construct systematic risk factors in the cross-section of corporate bonds. Kelly *et al.* (2023) employ the IPCA (Instrumental Principal Component Analysis) method, using bond-level characteristics as instruments to capture systematic risk factors. Bai *et al.* (2019) and Dickerson *et al.* (2023) use a two-sort method similar to that of Fama and French (1993). Huynh and Xia (2021), on the other hand, utilize panel regression to analyze the relationship between corporate bond returns and climate news risk exposure. To ensure comparability with stock market studies, we use the following cross-sectional regression to construct a non-linear ESG risk factor for the corporate bond market:

$$R^{\text{Bond}}_{i,t} = f_{0,t} + \sum_{l=1}^{L} f^{\text{Bond}}_{k,ESG,t} ESG^{l}_{i,t-1} + \sum_{j=1}^{J} f_{j,t} x_{ij,t-1} + u_{i,t}, \quad (15)$$

where $L$ is the highest-order monomial and we set $L = 1, 2, \text{and} 3$ (skewness at maximum). $R^{\text{Bond}}_{i,t}$ is the monthly corporate bond return at time $t$ for each bond $i$,

$$R^{\text{Bond}}_{i,t} = \frac{P_{i,t} + A_{i,t} + C_{i,t}}{P_{i,t-1} + A_{i,t-1}} - 1. \quad (16)$$

where $A_{i,t}$ is the accrued interest and $C_{i,t}$ is the coupon. $ESG_{i,t}$ are firm-level (issuer) ESG/E/S/G scores for bond $i$. $x_{ij,t-1}$ are a set of bond-level and firm-level characteristics following Bai *et al.* (2021) and Dickerson *et al.* (2023). Bond characteristics include credit rating, illiquidity, time-to-maturity, and the natural logarithm of amount outstanding. Firm-level variables include market value, book-to-market ratio, operating profit, investment growth and one-year momentum. We perform the cross-section regression at each time $t$ and stack the estimated coefficient ($f^{\text{Bond}}_{k,ESG,t}$) over time to get the bond ESG risk factor.

**6.3.2. Testing the non-linearity.** Our procedure to test the existence of linear and non-linear bond ESG risk factors consists of two steps. In the first step, we perform the following GRS test for the whole sample period

$$f^{\text{Bond}}_{i,ESG,t} = \alpha_i + \sum_{j=1}^{J} \beta_{ij} f_{jt} + \varepsilon_{it}. \quad (17)$$

We consider two sets of factor structures in the bond market ($f_{j,t}$): the bond market factor and the four-factor structure discussed in Bai *et al.* (2019), namely, market factor (MKTB), downside Risk factor (DRF), credit Risk factor (CRF), and liquidity risk factor (LRF).† $J$ is the number of common risk factors used in the test. The results, specifically whether $\alpha_i = 0$, will indicate whether the bond ESG risk factor provides additional explanatory power beyond the known factor structure.

In the second step, we perform the spanning regression similar to equation (10) in a dynamic (rolling window setting) setting:

$$\begin{cases} f^{\text{Bond}}_{1,n,t} = \delta_{1,n} + \sum_{j=1}^{J} \beta_{1,n,j} f_{j,t} + \varepsilon_{1,n,t} \\ f^{\text{Bond}}_{2,n,t} = \delta_{2,n} + \beta_{1,n,j} f^{\text{Bond}}_{1,n,t} + \sum_{j=1}^{J} \beta_{2,n,j} f_{j,t} + \varepsilon_{2,n,t} \\ f^{\text{Bond}}_{3,n,t} = \delta_{3,n} + \beta_{1,n,j} f^{\text{Bond}}_{1,n,t} + \beta_{2,n,j} f^{\text{Bond}}_{2,n,t} \\ + \sum_{j=1}^{J} \beta_{3,n,j} f_{j,t} + \varepsilon_{3,n,t} \end{cases}, \quad (18)$$

where $f_{jt}$ are bond risk factors and we test if $\alpha$s are zero; $f^{\text{Bond}}_{1,n,t}$, $f^{\text{Bond}}_{2,n,t}$ and $f^{\text{Bond}}_{3,n,t}$ are estimated using equation (15) under $L = 3$. We set the rolling window size of 48 months and a step size of one month.

**6.3.3. Data.** Our primary source of data is the bond asset pricing database of Dickerson *et al.* (2023), who provide bond returns and issuer information for the US market.‡ The database is compiled from Wharton Research Data Services (WRDS) Bond Database, with standard data cleansing measures in the bond literature.§

We collect data at a monthly frequency, specifically on the last Friday of each month, from January 2013 to December 2021. The CUSIP code is used to match bond identifiers between Eikon and the original database. The cross-section includes more than 15 000 bonds issued by over 600 companies.

**6.3.4. Empirical results.** Table B.3 presents the factor statistics for the entire sample period. Compared to the stock market, the ESG risk factor is considerably weaker in the bond market, both in its linear and non-linear forms, with all factors exhibiting mean values close to zero. Figure B.13 illustrates the cumulative value of the bond ESG risk factor (both linear and non-linear). Consistent with the observations in table B.3, the bond ESG risk factor remains close to zero for most of

† Downloaded from https://openbondassetpricing.com/data/.
‡ https://openbondassetpricing.com/data/, downloaded as of December 2024.
§ For detailed discussion on how to conduct data cleansing for bond returns, interested readers are referred to: https://openbondassetpricing.com/wp-content/uploads/2024/07/DRR-README.pdf.

the sample period. This pattern holds for the all types of ESG factors. We observe shocks around the outbreak of COVID.

Table B.4 present the GRS test for the full sample period. The test results are insignificant for all linear factors. However, we do observe increase in the test statistics when applying the non-linear setting, especially for the S score, where test statistics are significant for $L = 3$. Therefore, the non-linear setting should perform better in capturing the EGS risk factor in the cross-section of corporate bond returns than the linear setting, same as the stock market.

Finally, we perform spanning tests of model (18) under a dynamic setting to see how the non-linearity changes over time. Figure B.14 presents the *p*-value of the spanning test under the rolling window scheme, with a window size of 48 months and a step size of one. We find evidence of non-linearity for E and S factors. For example, for the E risk factor (Panel A, right graph), the significant results of the GRS test (green dotted line) is brought by the cubic E risk factor (f_3_E). The similar applies for the S factor, where we find that square S risk factor carries more information that the others. In comparison, for the G risk factor, before COVID, most of the time the results are driven by the linear G risk factor.

In a nutshell, the above findings are in-line with our predictions. First, the magnitude of ESG risk factor in the cross-section of corporate bond returns are much smaller than that of thioc. Second, despite having weaker evidence of ESG pricing, we still find that existence of non-linearity of ESG pricing.

## 7. Conclusion

We examine empirically if the ESG performance is being priced in the cross-section of stock returns in non-linear manner. We consider the square and cubic forms of ESG score in the cross-sectional regression to capture the ESG premium. Under our setting, the non-linear ESG premium has two interpretations: first, it measures how the ESG impact deviates from linearity and second, it is the price of cross-section standard deviation and skewness of ESG scores.

Our empirical evidence leads to the following remarks: (i) The non-linearity exists for ESG, E and S scores and is the highest for the E score; (ii) however, the choice of square or cubic form of ESG scores depends on characteristic of ESG scores as well as the sample period; (iii) part of the non-linearity comes from the cross-section ESG score distributions, which interacts with the ESG sentiment in affecting the asset pricing; (iv) characteristics of non-linear factors are different among different ESG providers. However, when taking out measurement error, we still see non-linearity.

Future research could extend our analysis in several directions. First, we have explored one form non-linearity, the square and cubic form of ESG score. The form of non-linearity is not limited to that. Further efforts could been carried to find the most suitable form of non-linearity, for instance by more flexible non-linear functions or exploring semi-parametric or non-parametric approaches. However, strong economic motivation should be considered to avoid critics of data mining. Second, while our study is based on empirical findings, further efforts can be made to design equilibrium models like Pástor *et al.* (2021) on potential mechanism of how ESG could be priced in a non-linear way.

## Open Scholarship

This article has earned the Center for Open Science badges for Open Data and Open Materials through Open Practices Disclosure. The data and materials are openly accessible at https://data.mendeley.com/datasets/r3m5x4hgw9/1 and https://data.mendeley.com/datasets/r3m5x4hgw9/1.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This study was funded by the European Union-NextGeneration EU, Mission 4, Component 2, in the framework of the GRINS-Growing Resilient, INclusive and Sustainable project [grant numbers GRINS PE00000018-CUP C93C22005270001]. The authors also acknowledge the financial support from Agencia Estatal de Investigación (AEI), Ministerio de Ciencia e Innovación (MCIN), Spain through [grant numbers TED2021-129891BI00 and PID2023-146832OB-I00] funded by MCIN/AEI/ 10.13039/5011000 11033 and by European Union-FEDER and European Union-NextGenerationEU/PRTR. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

## Supplemental data

Supplemental data for this article can be accessed online at http://dx.doi.org/10.1080/14697688.2025.2492794.

## ORCID

*Runfeng Yang* http://orcid.org/0000-0001-6343-1156
*Juan-Angel Jimenez-Martin* http://orcid.org/0000-0001-5608-3288
*Massimiliano Caporin* http://orcid.org/0000-0001-5014-5951

## References

Adrian, T. and Rosenberg, J., Stock returns and volatility: Pricing the short-run and long-run components of market risk. *J. Finance.*, 2008, **63**, 2997–3030.

Albuquerque, R., Koskinen, Y. and Zhang, C., Corporate social responsibility and firm risk: Theory and empirical evidence. *Manage. Sci.*, 2019, **65**, 4451–4469.

Amihud, Y., Illiquidity and stock returns: Cross-section and time-series effects. *J. Financial Markets*, 2002, **5**, 31–56.

Ardia, D., Bluteau, K., Boudt, K. and Inghelbrecht, K., Climate change concerns and the performance of Green vs. brown stocks. *Manage. Sci.*, 2023, **69**, 7607–7632.

Avramov, D., Cheng, S., Lioui, A. and Tarelli, A., Sustainable investing with ESG rating uncertainty. *J. Financ. Econ.*, 2022, **145**, 642–664.

Bai, J., Bali, T.G. and Wen, Q., Common risk factors in the cross-section of corporate bond returns. *J. Financ. Econ.*, 2019, **131**, 619–642.

Bai, J., Bali, T.G. and Wen, Q., Is there a risk-return tradeoff in the corporate bond market? Time-series and cross-sectional evidence. *J. Financ. Econ.*, 2021, **142**, 1017–1037.

Barillas, F. and Shanken, J., Which alpha? *Rev. Financ. Stud.*, 2016, **30**, 1316–1338.

Becchetti, L., Ciciretti, R. and Dalo, A., Fishing the corporate social responsibility risk factors. *J. Financial Stability*, 2018, **37**, 25–48.

Becchetti, L., Ciciretti, R. and Hasan, I., Corporate social responsibility, stakeholder risk, and idiosyncratic volatility. *J. Corporate Finance*, 2015, **35**, 297–309.

Bekaert, G., Engstrom, E.C. and Xu, N.R., The time variation in risk appetite and uncertainty. *Manage. Sci.*, 2022, **68**, 3975–4004.

Berg, F., Kölbel, J.F. and Rigobon, R., Aggregate confusion: The divergence of ESG ratings. *Rev. Financ.*, 2022, **26**, 1315–1344.

Berg, F., Koelbelz, J.F., Pavlova, A. and Rigobon, R., Esg confusion and stock returns: Tackling the problem of noise. Working Paper, 2021.

Brammer, S., Brooks, C. and Pavelin, S., Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financ. Manage.*, 2006, **35**, 97–116.

Caramichael, J. and Rapp, A.C., The green corporate bond issuance premium. *J. Bank. Financ.*, 2024, **162**, 107126.

Chang, B.Y., Christoffersen, P. and Jacobs, K., Market skewness risk and the cross section of stock returns. *J. Financ. Econ.*, 2013, **107**, 46–68.

Choi, D., Gao, Z. and Jiang, W., Attention to global warming. *Rev. Financ. Stud.*, 2020, **33**, 1112–1145.

Clark, G.L., Feiner, A. and Viehs, M., Form stockholder to shareholder: How sustainabiliy can drive financial out performance. Resreport, 2015.

Dickerson, A., Mueller, P. and Robotti, C., Priced risk in corporate bonds. *J. Financ. Econ.*, 2023, **150**, 103707.

Dittmar, R.F., Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *J. Finance.*, 2002, **57**, 369–403.

Dong, X., Feng, S., Parida, S. and Wang, Z., Corporate social responsibility exposure and performance of mutual funds. *The Journal of Investing ESG Special Issue*, 2019, **28**, 53–65.

Drempetic, S., Klein, C. and Zwergel, B., The influence of firm size on the ESG score: Corporate sustainability ratings under review. *J. Bus. Ethics.*, 2019, **147**, 333–360.

Dunz, N., Naqvi, A. and Monasterolo, I., Climate sentiments, transition risk, and financial stability in a stock-flow consistent model. *J. Financial Stability*, 2021, **54**, 100872.

Engle, R.F., Giglio, S., Kelly, B. and Lee, H., Hedging climate change news. *Rev. Financ. Stud.*, 2020, **33**, 1184–1216.

Fama, E.F. and French, K.R., Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.*, 1993, **33**, 3–56.

Fama, E.F. and French, K.R., Dissecting anomalies. *J. Finance.*, 2008, **63**, 1653–1678.

Fama, E.F. and French, K.R., Choosing factors. *J. Financ. Econ.*, 2018, **128**, 234–252.

Feldhütter, P., Halskov, K. and Krebbers, A., Pricing of sustainability-linked bonds. *J. Financ. Econ.*, 2024, **162**, 103944.

Flammer, C., Corporate green bonds. *J. Financ. Econ.*, 2021, **142**, 499–516.

Franco, C., Geissler, C., Margotz, V. and Monnier, B., ESG investments: Filtering versus machine learning approaches, 2020.

Freeman, R.E., *Strategic Management: A Stakeholder Approach*, 1984 (Pitman: Boston, MA).

Freyberger, J., Neuhierl, A. and Weber, M., Dissecting characteristics nonparametrically. *Rev. Financ. Stud.*, 2020, **33**, 2326–2377.

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

Giese, G., Lee, L.E., Melas, D., Nagy, Z. and Nishikawa, L., Foundations of ESG investing: How ESG affects equity valuation,risk, and performance. *J. Portfolio Manage.*, 2019, **45**, 1–15.

Gillan, S.L., Koch, A. and Starks, L.T., Firms and social responsibility: A review of ESG and CSR researching corporate finance. *J. Corporate Finance*, 2021, **66**, 101889.

Gu, S., Kelly, B. and Xiu, D., Empirical asset pricing via machine learning. *Rev. Financ. Stud.*, 2020, **33**, 2223–2273.

Huynh, T.D. and Xia, Y., Climate change news risk and corporate bond returns. *J. Financial Quant. Anal.*, 2021, **56**, 1985–2009.

Jahmane, A. and Gaies, B., Corporate social responsibility, financial instability and corporate financial performance: Linear, non-linear and spillover effects-the case of the cac40 companies. *Finance Res. Lett.*, 2020, **34**, 101483.

Jarjir, S.L., Nasreddine, A. and Desbana, M., Corporate social responsibility as a common risk factor. *Global Finance J.*, 2022, **52**, 100577.

Kahn, M.E. and Kotchen, M.J., Business cycle effects on concern about climate change: The chilling effect of recession. *Climate Change Econ.*, 2011, **2**, 257–273.

Kapadia, N., The next Microsoft? Skewness, idiosyncratic volatility, and expected returns, 2006.

Kelly, B., Palhares, G. and Pruitt, S., Modeling corporate bond returns. *J. Finance.*, 2023, **78**, 1967–2008.

Kirby, C., Firm characteristics, cross-sectional regression estimates, and asset pricing tests. *Rev. Asset Pricing Stud.*, 2019, **10**, 290–334.

Krueger, P., Corporate goodness and shareholder wealth. *J. Financ. Econ.*, 2015, **115**, 304–329.

Lamont, O.A., Economic tracking portfolios. *J. Econom.*, 2001, **105**, 161–184.

Lins, K.V., Servaes, H. and Tamayo, A., Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *J. Finance.*, 2017, **72**, 1785–1824.

Lioui, A. and Tarelli, A., Chasing the ESG factor. *J. Banking Finance*, 2022, **139**, 106498.

Maiti, M., Is ESG the succeeding risk factor? *J. Sustainable Finance Investment*, 2020, **11**, 199–213.

Naffa, H. and Fain, M., A factor approach to the performance of ESG leaders and laggards. *Finance Res. Lett.*, 2022, **44**, 102073.

Pástor, L., Stambaugh, R.F. and Taylor, L.A., Sustainable investing in equilibrium. *J. Financ. Econ.*, 2021, **142**, 550–571.

Pástor, L., Stambaugh, R.F. and Taylor, L.A., Dissecting green returns. *J. Financ. Econ.*, 2022, **146**, 403–424.

Pollman, E., The making and meaning of ESG. *Harvard Bus. Law Rev.*, 2024, **14**, 403–454.

Serafeim, G., Public sentiment and the price of corporate sustainability. *Financial Anal. J.*, 2020, **76**, 26–46.

Surroca, J., Tribo, J.A. and Waddock, S., Corporate responsibility and financial performance: The role of intangible resources. *Strategic Manage. J.*, 2010, **31**, 463–490.

Welch, I. and Goyal, A., A comprehensive look at the empirical performance of equity premium prediction. *Rev. Financ. Stud.*, 2007, **21**, 1455–1508.

## Appendix. Discussion on properties of ESG factors under the non-linear setting

Our discussion is based on the baseline model (4) when $L = 3$:

$$\begin{aligned} r_{i,t} &= f_{0,t} + f_{1,ESG,t} ESG_{i,t-1} + f_{2,ESG,t} ESG_{i,t-1}^2 \\ &\quad + f_{3,ESG,t} ESG_{i,t-1}^3 + \varepsilon_{i,t}, \end{aligned} \quad \text{(A1)}$$

The matrix form of the above equation is:

$$R_t = X_{t-1}\Gamma_t + \Phi_t, \quad \text{(A2)}$$

where $R_t$ is the return matrix with a size of $(n \times 1)$, assuming we have $n$ assets cross-sectionally: $R_t = [r_{1t}, r_{2t}, r_{3t}, \cdots, r_{nt}]^T$; $X_{t-1}$ is the variable matrix with a value of time $(t-1)$ with a size of $(n \times 4)$:

$$X_{t-1} = \begin{bmatrix} 1 & ESG_{1,t-1} & ESG_{1,t-1}^2 & ESG_{1,t-1}^3 \\ 1 & ESG_{2,t-1} & ESG_{2,t-1}^2 & ESG_{2,t-1}^3 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & ESG_{n,t-1} & ESG_{n,t-1}^2 & ESG_{n,t-1}^3 \end{bmatrix}_{(n\times 4)}, \quad \text{(A3)}$$

where the $\Gamma_t$ is the matrix of factor returns with a size of $(4 \times 1)$ :$[f_{0,t}, f_{1,ESG,t}, f_{2,ESG,t}, f_{3,ESG,t}]^T$ and the error term $\Phi_t = [\eta_{1t}, \eta_{2t}, \ldots, \eta_{nt}]^T$. And the Ordinary Least Square (OLS) estimation of $\Gamma_t$ is:

$$\Gamma_t = (X_{t-1}^T X_{t-1})^{-1} X_{t-1}^T R_t. \quad \text{(A4)}$$

If we take $\Gamma_t$ as a portfolio (the factor portfolio) and $R_t$ is the component stocks, then $W_t \equiv (X_{t-1}^T X_{t-1})^{-1} X_{t-1}^T$ is the weight for each stock in the $R_t$. The size of $W$ is $(4 \times n)$, i.e. at every time $t$:

$$W_t = \begin{bmatrix} w_{0,1t} & w_{0,2t} & \cdots & w_{0,nt} \\ w_{ESG,1t} & w_{ESG,2t} & \cdots & w_{ESG,nt} \\ w_{ESG2,1t} & w_{ESG2,2t} & \cdots & w_{ESG2,nt} \\ w_{ESG3,1t} & w_{ESG3,2t} & \cdots & w_{ESG3,nt} \end{bmatrix}_{(4\times n)}$$
$$\times \begin{bmatrix} 1 & ESG_{1,t-1} & ESG_{1,t-1}^2 & ESG_{1,t-1}^3 \\ 1 & ESG_{2,t-1} & ESG_{2,t-1}^2 & ESG_{2,t-1}^3 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & ESG_{n,t-1} & ESG_{n,t-1}^2 & ESG_{n,t-1}^3 \end{bmatrix}_{(n\times 4)}. \quad \text{(A5)}$$

Note that $W_t X_{t-1} = (X_{t-1}^T X_{t-1})^{-1} X_{t-1}^T X_{t-1} = I_{(4\times 4)}$. In the meantime, the first column of $X_{t-1}$ is 1, which means that :

$$W_t X_{t-1} = \begin{bmatrix} (w_{0,1t} + \cdots + w_{0,nt}) & (w_{0,1t} ESG_{1,t-1} + .. + w_{0,nt} ESG_{n,t-1}) & \cdots \\ (w_{E,1t} + \cdots + w_{E,nt}) & (w_{ESG,1t} ESG_{1,t-1}^2 + \cdots + w_{ESG,nt} ESG_{n,t-1}^2) & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}_{(4x4)}$$
$$= \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 1 & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}_{(4x4)}, \quad \text{(A6)}$$

which means that

$$\begin{cases} w_{0,1t} + \cdots + w_{0,nt} = \sum_{i=1}^n w_{0,it} = 1 \\ w_{ESG,1t} + \cdots + w_{ESG,nt} = \sum_{i=1}^n w_{E,it} = 0 \\ w_{ESG,1t} ESG_{1,t-1} + \cdots + w_{ESG,nt} ESG_{n,t-1} = 1 \\ w_{ESG,1t} ESG_{1,t-1}^2 + \cdots + w_{ESG,nt} ESG_{n,t-1}^2 = 0 \\ \cdots \end{cases}. \quad \text{(A7)}$$

That is, any combination in the diagonal like in equation (A7) is 1 and others 0. Then, the factor returns in equation (A4) can be expressed as:

$$\Gamma_t = W_t R_t = \begin{bmatrix} w_{0,1t} & w_{0,2t} & \cdots & w_{0,nt} \\ w_{E,1t} & w_{E,2t} & \cdots & w_{E,nt} \\ \cdots & \cdots & \cdots & .. \end{bmatrix}^* \begin{bmatrix} r_{1t} \\ r_{2t} \\ \cdots \\ r_{nt} \end{bmatrix}$$
$$= \begin{bmatrix} (w_{0,1t} r_{1t} + w_{0,2t} r_{2t} + \cdots w_{0,nt} r_{nt}) \\ (w_{ESG,1t} r_{1t} + w_{ESG,2t} r_{2t} + \cdots w_{ESG,nt} r_{nt}) \\ \cdots \end{bmatrix} = \begin{bmatrix} f_{0,t} \\ f_{1,ESG,t} \\ \cdots \end{bmatrix} \quad \text{(A8)}$$

which then means:

$$\begin{cases} w_{0,1t} r_{1t} + w_{0,2t} r_{2t} + \cdots w_{0,nt} r_{nt} = f_{0,t} \\ w_{ESG,1t} r_{1t} + w_{E,2t} r_{2t} + \cdots w_{E,nt} r_{nt} = f_{1,ESG,t} \\ \cdots \end{cases}. \quad \text{(A9)}$$

In sum, $f_{k,ESG,t}$ is always a zero-cost long-short portfolio consisted of all the assets used in the cross section. In addition, the exposure of $f_{k,ESG,t}$ to $ESG^k$ is always 1, according to equation (A7).

To further study what constitute the ESG factor under the non-linear setting, we expand the equation (A4) as:

$$\Gamma_t = (X_{t-1}^T X_{t-1})^{-1} X_{t-1}^T R_t$$
$$= \left( \begin{bmatrix} 1 & \cdots & 1 \\ ESG_{1,t-1} & \cdots & ESG_{n,t-1} \\ ESG_{1,t-1}^2 & \cdots & ESG_{n,t-1}^2 \\ ESG_{1,t-1}^3 & \cdots & ESG_{n,t-1}^3 \end{bmatrix} \times \begin{bmatrix} 1 & ESG_{1,t-1} & ESG_{1,t-1}^2 & ESG_{1,t-1}^3 \\ 1 & ESG_{2,t-1} & ESG_{2,t-1}^2 & ESG_{2,t-1}^3 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & ESG_{n,t-1} & ESG_{n,t-1}^2 & ESG_{n,t-1}^3 \end{bmatrix} \right)^{-1}$$
$$\times \begin{bmatrix} 1 & \cdots & 1 \\ ESG_{1,t-1} & \cdots & ESG_{n,t-1} \\ ESG_{1,t-1}^2 & \cdots & ESG_{n,t-1}^2 \\ ESG_{1,t-1}^3 & \cdots & ESG_{n,t-1}^3 \end{bmatrix} R_t$$
$$= \left( \begin{bmatrix} n & \sum_{i=1}^n ESG_{i,t-1} & \sum_{i=1}^n (ESG_{i,t-1})^2 & \sum_{i=1}^n (ESG_{i,t-1})^3 \\ \sum_{i=1}^n ESG_{i,t-1} & \sum_{i=1}^n (ESG_{i,t-1})^2 & \sum_{i=1}^n (ESG_{i,t-1})^3 & \sum_{i=1}^n (ESG_{i,t-1})^4 \\ \sum_{i=1}^n (ESG_{i,t-1})^2 & \sum_{i=1}^n (ESG_{i,t-1})^3 & \sum_{i=1}^n (ESG_{i,t-1})^4 & \sum_{i=1}^n (ESG_{i,t-1})^5 \\ \sum_{i=1}^n (ESG_{i,t-1})^3 & \sum_{i=1}^n (ESG_{i,t-1})^4 & \sum_{i=1}^n (ESG_{i,t-1})^5 & \sum_{i=1}^n (ESG_{i,t-1})^6 \end{bmatrix} \right)^{-1}$$
$$\times \begin{bmatrix} 1 & \cdots & 1 \\ ESG_{1,t-1} & \cdots & ESG_{n,t-1} \\ ESG_{1,t-1}^2 & \cdots & ESG_{n,t-1}^2 \\ ESG_{1,t-1}^3 & \cdots & ESG_{n,t-1}^3 \end{bmatrix} R_t$$

$$= \left( \begin{bmatrix} n & 0 & n & \sum_{i=1}^{n}(ESG_{i,t-1})^3 \\ 0 & n & \sum_{i=1}^{n}(ESG_{i,t-1})^3 & \sum_{i=1}^{n}(ESG_{i,t-1})^4 \\ n & \sum_{i=1}^{n}(ESG_{i,t-1})^3 & \sum_{i=1}^{n}(ESG_{i,t-1})^4 & \sum_{i=1}^{n}(ESG_{i,t-1})^5 \\ \sum_{i=1}^{n}(ESG_{i,t-1})^3 & \sum_{i=1}^{n}(ESG_{i,t-1})^4 & \sum_{i=1}^{n}(ESG_{i,t-1})^5 & \sum_{i=1}^{n}(ESG_{i,t-1})^6 \end{bmatrix} \right)^{-1} \times \begin{bmatrix} \sum_{i=1}^{n} r_{i,t} \\ \sum_{i=1}^{n} ESG_{i,t-1} r_{i,t} \\ \sum_{i=1}^{n} ESG_{i,t-1}^2 r_{i,t} \\ \sum_{i=1}^{n} ESG_{i,t-1}^3 r_{i,t} \end{bmatrix} = \begin{bmatrix} f_{0,t} \\ f_{1,ESG,t} \\ f_{2,ESG,t} \\ f_{3,ESG,t} \end{bmatrix}. \quad \text{(A10)}$$

$\sum_{i=1}^{n} ESG_{i,t-1}^{k} r_{i,t}$ means a portfolio whose weight is calculated based on $k$th power of ESG scores (if $k = 0$, it is an equal-weighted portfolio). From the above equation, we can see that both linear and non-linear factors are a linear combination of $\sum_{i=1}^{n} ESG_{i,t-1}^{k} r_{i,t}$, whose coefficient depends partly on the cross-sectional distribution of the ESG score (e.g. $\sum_{i=1}^{n}(ESG_{i,t-1})^3$).