Sovereign ESG and European stock markets returns.

Student: Sjoerd Dronkers - 488176

March 1, 2023

Abstract

This paper investigates the relationship between sovereign ESG scores and market index performance. This relationship is studied for the ESG index as well as for the individual pillars. The ESG scores are calculated through the new World Bank data. To research a potential relation, the dynamic panel model of Arellano-Bond is used. The research concludes that there is no significant relation between sovereign ESG scores and market returns. The insignificant results also occur when the pillars are used individually. With an exception for South-Europe, where significance is found for the social pillar and the lagged environmental and governance pillars. The social and lagged governance pillar have a positive effect, the lagged environment pillar has a negative effect. The mostly insignificant results might be caused by the model choice and the entropy weighting method. This opens the door for future research to investigate different models and different weighting schemes to calculate the ESG scores.

Supervisor: dr. R. de Bliek

Co-Reader: dr. J.J.G. Lemmen

Bachelor Thesis Economics and Business Economics

Erasmus University Rotterdam

Erasmus School of Economics

Contents

1 Introduction

Over the past years the importance of countries and companies becoming more sustainable has increased. With this increased importance, the number of studies on this topic has risen as well. One of the topics studied within this domain is the relationship between Corporate Financial Performance (CFP) and Environmental, Social and Governance (ESG) index. The rise of interest in this topic can be seen in the paper from Friede, Busch, and Bassen (2015). In this study 2200 individual studies were combined to allow for more generalized statements about CFP and ESG. In this paper it becomes clear that from 1970 onwards the interest in researching the relation between ESG indicators and CFP has increased almost exponentially based on the number of estimated empirical studies. The paper states that 90% of studies find a non-negative relationship between the ESG and CFP. The notion on exponential growth of papers on ESG ratings can also be found in the number of companies that measure and report ESG data (Amel-Zadeh & Serafeim, 2018). Entities investing according to ESG principles, can be defined as people, institutions, or countries that consider the sustainability of their investments besides profitability.

Despite the increase of interest in the relation between CFP and ESG, the question arises why the number of studies of the ESG index on a country level or region level is relatively low (Jiang, Feng, & Yang, 2022). The main reason for the lack of studies lies in the lack of data. With the new World Bank data set, it has become possible to investigate whether a relationship between changes in the ESG index and the returns of the market index of a country exists. T.-T. Li, Wang, Sueyoshi, and Wang (2021) find that the ESG index is almost always approached from the view of an internal governance and thus it will be interesting to see what the relationship will be when approaching with data from an external governance.

The main idea of this paper is to investigate whether a positive relation between ESG ratings and an European country's market index can be found when grouped by region. The countries included in this research are based on the geographical definition of Europe in accordance with the United Nations data (United Nations, 1999). The countries will be grouped on regional level for this research. So, in total we will have four different panels for West, North, East and South Europe. In total eleven market indexes are taken into account, each European region the three largest markets are used. With the exception of Eastern Europe where two indexes were used due to a lack of data. The market indexes of France, Germany, United Kingdom, Denmark, Norway, Sweden, Hungary, Poland, Italy, Portugal and Spain are used. Past research shows that the European continent is interesting to look at in terms of ESG data. For example, in the paper of Van Duuren, Plantinga, and Scholtens (2016) it is discussed that European investors see ESG investing more like fundamental investing than investors in the United States. Amel-Zadeh and Serafeim (2018) conclude similar results: European investors are more likely to use ESG information than investors in the United States.

The World Bank data consists of the three pillars: Environmental, Social and Governance. These pillars together make up the ESG index. First, the environmental pillar is defined as the way a country handles its natural resources and its proneness to climate change. Second, the social pillar contains information about inequality, health, and the demography. Finally, the governance pillar encompasses human rights, governmental issues, innovation, and gender differences. Because the three pillars all consist of multiple indicators, the data will be normalized and combined into a score for each of the three pillars individually.

The data included in this study ranges from 1999 until 2020, 2020 included. From the mid 1990s there

are stable positive correlation patterns (Friede et al., 2015). Since the data set is still a work in progress, the data until 2020 is used for this research. The index returns will be obtained through the WRDS database (Wharton Research Data Services, 2022). The index returns will be annualized on a yearly basis in order to be compatible with the yearly data from the World Bank.

A problem that arises when using ESG data is that there is no consistent methodology to derive the ESG scores. Therefore, companies and countries may receive different ratings by different providers according to Eurosif (2021) and F. Li and Polychronopoulos (2020). This is where the new sovereign data performs better than firm level data. In Bouyé and Menville (2020) it is argued that sovereign ESG scores converge, they find that all methodologies have a high explaining power with the value R^2 ranging from 0.78 and 0.98. To research a possible non negative relation between European index returns and ESG indices, dynamic panel data regressions are used (Arellano & Bond, 1991).

The main goal of researching the relation between the ESG index and European market index returns is to see if growth or decline of the ESG scores influences a growth or decline on the stock market. The secondary goal is to observe whether each of the three pillars have a different influence on the market returns. The layout of the paper is as follows. First, there is a theory section where the most important literature is outlined and explained. Second, the data & methodology section where basic statistics of the samples will be shown and the normalization will be performed. Furthermore, this section will explain which variables are taken into account and which specific techniques will be used to look for a relationship between the samples. The fourth chapter will contain the results from the methods used from chapter two. Following from the results, conclusions will be drawn about the research in chapter five and possible short comings of this paper and future research will be discussed.

2 Literature Review

In this section relevant past literature will be discussed. Initially, past research on the ESG index is described. In the second part the relationship between stock market returns and ESG ratings will be discussed Finally, the hypotheses for this research will be defined.

2.1 ESG

The ESG principle was formally proposed in 2004 and has been widely used in developed countries as researched by T.-T. Li et al. (2021). They find that since 2004 the ESG index has matured and is widely used in Europe, though the literature has focused mostly on the governance pillar and less on the environmental and social pillar. Furthermore, in this paper it is found that ESG research in the corporate sector is focused on for example increasing shareholder value. In this paper we focus on the sovereign ESG index and we follow the description given by Zhang, Zhao, and Lau (2022). Sovereign ESG provides a picture of a country's business climate and helps investors to make informed investment decisions incorporating sustainability. In Carroll (1979) we can see that corporate social responsibility is first seen around the 1930s, however this is one of the first papers that proposes a model. In this paper the total social responsibility is divided into four responsibilities: discretionary, ethical, legal and economical. The social and governance pillars are already quite visible here, but the environmental pillar is embedded into the legal and economical pillars and does not have its own pillar yet. The model can be used as a framework for managers to see how socially responsible their firm is. CSR, ESG and sustainability can be seen as synonyms throughout economic literature. In Gratcheva, Emery, and Wang (2021) we see how ESG investing differs from traditional investing and impact investing. Traditional investing only focuses on returns and disregards ESG and sustainability factors completely. Impact investing is an investment strategy that takes environmental and social factors into account when making investment decisions. Furthermore, this strategy does not only focus on market returns but on social returns as well. In ESG investing the aspects of the three pillars are integrated into the investment strategies, but it is only focused on market returns. The difference between impact investing and ESG investing is that impact investing focuses on social returns and ESG investing focuses on market returns. One can rank the three investment forms in the following order: traditional investing, ESG investing and impact investing, where traditional investing is the least sustainable and impact investing the most.

In Friede et al. (2015) it is visible that most of the past papers focus on the economic and social pillars of the ESG index and less on the governance pillar. Even though not all three pillars are used regularly, they find that the large majority of studies find a positive relationship between CFP and the ESG index. Support for this can be found in Alshehhi, Nobanee, and Khare (2018), here they observe a trend that the definition of sustainability for companies is shifting. The trend found is the shift towards more socially focused sustainability and less inclusive from an environmental standpoint. Larcker, Tayan, and Watts (2022) find that governance is not supposed to be in the ESG index, a company can have a good governance score while only focusing on shareholder optimization. It should be noted, that these results hold for the North-American markets.

The current state of ESG indexes is researched in Pagano, Sinclair, and Yang (2018). In the period from 1995 until 2016, ESG investments had an annual growth rate of 13.25 percent. Furthermore, they find that there is a bias towards geographical regions. In certain countries or regions there is more readily available

research and the amount of covered companies is greater.

An important problem to keep in mind, is the difference the provider of the ESG data makes. As found in F. Li and Polychronopoulos (2020), there is a lack of robust options for data to determine ESG ratings. Furthermore, the rating of a stock can differ quite significantly when looked at by different rating agencies. In Berg, Koelbel, and Rigobon (2022) the divergence between different ESG ratings becomes clear when six different rating agencies have a correlation between 0.38 to 0.71, this implies that the companies compute the ESG rating differently. This lack of correlation was also found by Dimson, Marsh, and Staunton (2020). Even though the large difference in use of different indicators between the six agencies researched, there are some indicators that all of the agencies used. The common indicators are: Biodiversity, Employee Development, Energy, Green Products, Health and Safety, Labor Practices, Product Safety, Remuneration, Supply Chain, and Water. Therefore, a big issue when calculating the ESG score for a country, will be the lack of a general rating scheme (e.g. the weighting scheme for indices). In contrary to firm level data, we can see that sovereign ESG ratings do converge between different rating agencies (Bouyé & Menville, 2020). A solution for the weighting scheme is the use of entropy weighting or principal component analysis (pca). In this paper entropy weighting is used as it outperforms pca within this context (Jiang et al., 2022). Entropy weighting is an objective assignment method, this way dependency on the rating schemes of different agencies is avoided.

2.2 Relation between market performance and ESG

In this subsection literature about the market returns in Europe is discussed first, followed by the relation between ESG indexes and market returns.

2.2.1 Index Returns

In Bessler and Yang (2003) we can see that national stock markets are co integrated within Europe (and from outside Europe). This means that stock markets within Europe influence each other in the case of innovations and price movements. This becomes clearer in the paper of Égert and Kočenda (2007) where they look at interdependencies between Western and Eastern European stock markets. Using vector autoregression they find that there are spillover effects of returns and volatility between Eastern European countries, between Western markets and from Western markets towards Eastern markets. However, there are no significant spillover effects from Eastern markets towards Western markets. The spillover effects increased during the years of economic crisis from 2008 to 2012 (Harkmann, 2014). Within the Western market the effect was the strongest between the three largest stock markets (France, England and Germany). Also, the effect becomes stronger when looking at a larger time frame of returns (e.g. daily versus yearly) (Dajcman, 2013).

In general, there were positive annual gains on European indexes between 1997 to 2004 (Chukwuogor-Ndu, 2006). Even though there is growth in general, between 1998 to 2002 most of the market indexes experienced a year of negative returns.

2.2.2 Relation

In the introduction some arguments to focus on Europe were already discussed, an additional argument is found in Morgenstern, Coqueret, and Kelly (2022). Their paper states that there is a slightly higher signif-

icance of ESG factors than in other developed or emerging markets, consistent with the attention the ESG receives in Europe. Furthermore, it is argued that there is less dispersion for sovereign ESG data and the lack of standardization is less problematic than when using firm level data.

In section 2.1 we already saw that the importance of each pillar differs. The paper of Drei et al. (2019) find similar results when looking at financial performance and the three pillars. They looked into the effect of ESG investing on asset pricing in Europe and North-America between 2010 and 2019, using portfolios that track the complete stock market. The main conclusion was that between 2010 and 2013 the use of ESG in an investing strategy was penalized, but between 2013 and 2019 it was a source for excess returns. In Europe they found evidence for a cycle. In the beginning the environment pillar dominated, then the governance pillar took over and at the end of their sample period the social pillar had the largest effect.

In De Spiegeleer, Höcht, Jakubowski, Reyners, and Schoutens (2021) equity portfolios are used with either a low or high ESG score to check if there is enhanced performance. This is done using stocks from the STOXX Europe 600. They find that there is no clear evidence to support enhanced performance. The paper of Morea, Mango, Cardi, Paccione, and Bittucci (2022) finds similar results in the STOXX Europe 50. The ESG impacts the stock performance but not significantly. As a risk factor for the STOXX Europe 600, the ESG factors are significant at 5%. Furthermore, including these factors in an asset pricing model outperforms the Fama and French (1993) three factor model (Maiti, 2021). The risk factor of the ESG index has an insurance-like character, a high ESG score reduces a firms' risk (Bannier, Bofinger, & Rock, 2019). Lower score ESG firms have stronger positive returns but they need to offer a corresponding risk premium, this result is stronger in more volatile markets. In Sassen, Hinze, and Hardeck (2016) they look at the relation between Corporate Social Performance (CSP) and risk. Using firm level panel data they find that total and idiosyncratic risk decreases as CSP increases. Furthermore, they find that environment performance in general decreases idiosyncratic risk but systematic and total risk only decreased in environmental sensitive industries. Social performance decreases all three of the risk measures and they find that governance performance is not significant for firm risk. The paper of Shanaev and Ghimire (2022) investigates the relation between ESG rating changes and stock returns of US firms between 2016-2021. They find that increases in ESG index ratings lead to small but significant excess stock returns. A decrease of the ESG index does have a negative influence on stock returns and lead to negative stock returns. However, they state that the effect of the ESG rating changes is modest.

The paper of Marti, Rovira-Val, and Drescher (2015) looked at CSR strategies of companies in the STOXX Europe 600. They conclude that during a period of crisis companies that implemented CSR outperformed companies that did not implement CSR. Furthermore, they found that companies located in more developed countries performed better than companies in less developed countries. In United Nations Development Program (2022) we can see that Western and Northern European countries are slightly better developed than Southern and Eastern European countries. This raises the question whether this effect of development on the ESG index can also be seen on a country or regional level versus on a company level.

2.3 Hypotheses

As we can conclude from the literature there has been a lot of analysis on corporate level whether responsible investing has a positive effect on a firm's value. This is also shown in Jiang et al. (2022) where they conclude that the importance of the ESG on micro-firm level has been shown but macroeconomic entities are significantly under researched. They conclude that current mainstream ESG ratings are still based on corporate data and that sovereign ESG research remains at an early stage.

Since there are no prior papers researching this relation, for the first hypothesis we look at the conclusion drawn on corporate level data and project it onto sovereign data.

Hypothesis 1 *There is a positive relation between a country's ESG score and it's leading stock market index.*

In Shanaev and Ghimire (2022) they conclude that future research could consist of separation of the different pillars and look at the individual effects. The paper of Giese, Nagy, and Lee (2021) deconstructs the ESG index into the different pillars and looks at the effect of the individual pillars. They find that governance has the largest influence on the short term, the social and environmental pillars develop slowly but have effects on the long term. In Drei et al. (2019) they find that the individual effect of each pillar is negative between 2010 and 2013, but positive between 2014 and 2019. The notion that the deconstruction of the ESG index leads to different results for the effect of each pillar, leads to the second set of hypotheses.

Hypothesis 2a *The environmental pillar has a negative relation with the market returns.*

Hypothesis 2b *The social pillar has a positive relation with the market returns.*

Hypothesis 2c *The governance pillar has a positive relation with the market returns.*

In the second hypothesis we will look at the different effects each individual pillar has on the stock market returns.

3 Data & Methodology

In this section an overview of the data is provided as well as the normalization technique used to prepare the data for analysis. To calculate the weights of the indicators, the entropy weight method is used. In the second part descriptive statistics will be shown to get a better understanding of the data. In the third part the methodology behind the regressions and statistical tests are explained.

Environment	Social	Governance
Emissions & Pollution	Education & skills	Human Rights
Natural capital	Employment	Government Effectiveness
Energy use & security	Demography	Stability & Rule of Law
Climate risk	Poverty & Inequality	Economic Environment
Food Security	Health & nutrition	Gender
	Access to Services	Innovation

Table 1: Key themes within the three pillars in the World Bank database.

In table 1 we can see the key themes within each pillar. Each of the indicators will be normalized (see chapter 3.1), this way the indicators can be combined into a single score for every pillar, these scores are calculated on a yearly basis.

The market data from every index is retrieved using WRDS. In total there are eleven countries with 252 months of returns, the returns are annualized using equation 1, where y_i is the yearly return of country i and r_j is the return of the market index in month j.

$$y_i = \prod_{j=1}^{12}(1 + r_j) - 1 \tag{1}$$

After the data is annualized, we end up with 21 years of returns.

3.1 ESG Normalization & Weighting

This subsection follows the normalization and weighting methods from Jiang et al. (2022), this includes determining which indicators are used and which indicators are positive or negative. The method starts to differ after equation 5 since the ESG score is split into different scores for each individual pillar. For positive indicators (p_{pos}) equation 2 is used and for negative indicators (p_{neg}) equation 3, we end up with the normalized indicator p. For the ESG calculations the following indices are defined: k is the indicator, i is the year and j is the country. The total number of indicators is K, the total number of years equals I and the number of countries equals J.

$$p_{ij,pos} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{2}$$

$$p_{ij,neg} = \frac{x_{min} - x}{x_{min} - x_{max}} \tag{3}$$

In equation 2 and 3, x_{min} is the minimum value of an indicator, x_{max} is the maximum value of an indicator and x is the indicator value. x_{min} and x_{max} are the respective minimum and maximum value found by looking at every value of an indicator across every country. This means x_{max} could be from another country than x itself. For the rest of the paper $p_{ij,pos}$ and $p_{ij,neg}$ will simply be denoted as p_{ij}. The full overview of

whether an indicator has a positive or negative weighting can be found in appendix A.

To determine the weights of each indicator, the entropy weight method is used. The entropy weight method measures value spread, if the spread is greater, more information can be derived from an indicator (Zhu, Tian, & Yan, 2020).

$$entropy_k = -\frac{1}{ln(q)} \sum_i \sum_j s_{ijk} ln(s_{ijk}) \tag{4}$$

In equation 4 we define: $q = \#countries \cdot \#years$ and $s_{ijk} = \frac{p_{ijk}}{\sum_i \sum_j p_{ijk}}$. Here, s_{ijk} scales the value of one normalized indicator in comparison to all normalized values of every country in every year. This way we can see the relative performance of a country in a specific year compared to other years and other countries.

$$weight_k = \frac{1 - entropy_k}{\sum_k (1 - entropy_k)} \tag{5}$$

Using $1 - entropy_k$, the entropy calculations become positively correlated with the weights. More specifically, smaller entropy values contain more information and thus we want to give these a higher weight. The weights do not change throughout the years or between countries. The weight calculation is done through the Python library Crispyn (Baczkiewicz & Watróbski, 2022). The K indicators and their respective weights are grouped under the pillars E, S or G, the indicators change for each year i. For country j pillar E is made up of l indicators denoted by: $e_{1ij}, .., e_{Lij}$, S is made up of m indicators: $s_{1ij}, ..., s_{Mij}$ and finally G consists of n indicators: $g_{1ij}, ..., g_{Nij}$. The respective weights are denoted by w_{e_l}, w_{s_m} and w_{g_n}. By definition $l+m+n = K$.

$$E_{ij} = \sum_{l=1}^{L} w_{e_l} e_{lij} \tag{6}$$

$$S_{ij} = \sum_{m=1}^{M} w_{s_m} s_{mij} \tag{7}$$

$$G_{ij} = \sum_{n=1}^{N} w_{g_n} g_{nij} \tag{8}$$

$$ESG_{ij} = E_{ij} + S_{ij} + G_{ij} \tag{9}$$

The ESG score for a country j in year i becomes equal to equation 9 and we have: $0 \leq ESG_{ij} \leq 1$. In equations 6, 7 and 8 the score per pillar is calculated for each country in each year. This way we can see the different effects of the pillars.

3.2 Descriptive Statistics

In this subsection the returns and the ESG score of every country will be shown. These returns will be cumulative and indexed (starting from 100), this way some trends can already be seen. After this, the ESG ratings of every country throughout the years will be shown. Both the index returns and ESG ratings will be shown per country and grouped by their respective region.

3.2.1 Index Returns

The index returns are retrieved from the WRDS database, the data consists of portfolios that follow the market indexes listed in table 2 (Wharton Research Data Services, 2022). The correlation of the portfolios with their respective market index can also be seen here. The correlations of the portfolios are between 0.962 (Germany) and 0.996 (United Kingdom).

Country	Index Name	Correlation
West		
France	MSCI - France	0.990
Germany	FT - Germany Index	0.962
United Kingdom	FTSE 350 Index	0.996
North		
Denmark	FT - Denmark Index	0.972
Norway	FT - Norway Index	0.968
Sweden	FT - Sweden Index	0.977
East		
Hungary	Budapest Stock Exchange Index	0.971
Poland	Warsaw W.I.G Index	0.975
South		
Italy	MSCI - Italy Index	0.985
Portugal	PSI 20 Index	0.979
Spain	FT - Spain Index	0.989

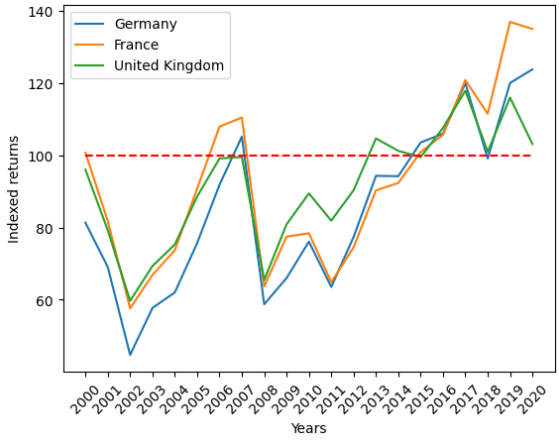
Table 2: The market indexes that are used for this research. The first column contains the country, second is the index name and the third states the correlation between the WRDS portfolio and the index.

In figure 1 we can see that eight countries experienced growth and three countries experienced a decline. The red dashed line is equal to 100, this way it is easier to see whether countries experienced overall growth or decline. If a country is indexed above 100, there is growth since 1999. In all countries the effect of the 2008 crisis is clearly visible.

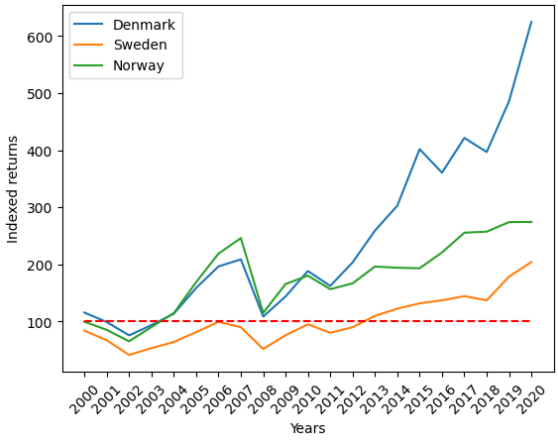
In figure 2 the average returns per European region are visible. Until 2011 all the regions seem to follow the same pattern. What stands out is the breakout growth of Northern Europe after 2011. In contrast to West, North and East, the Southern region does not display growth over the course of the ten years after 2011. Besides this, South-Europe is the only region where there is a overall decline between 2000 and 2020. Except for North-Europe, every region experiences a decline in market performance in 2019 and 2020.

3.2.2 ESG Ratings

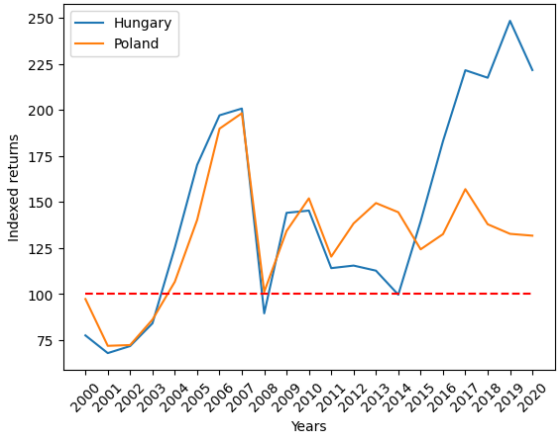
In table 3 we can see that in total 54 indicators are used. We can see that environment makes up the largest part of the index followed by social and governance. These results are comparable to Jiang et al. (2022) except for the number of indicators. This difference can be explained by the number of countries in the sample, for three indicators every country had the same value throughout the years. This allows for zero values in the entropy weighting method and are thus removed from the sample.



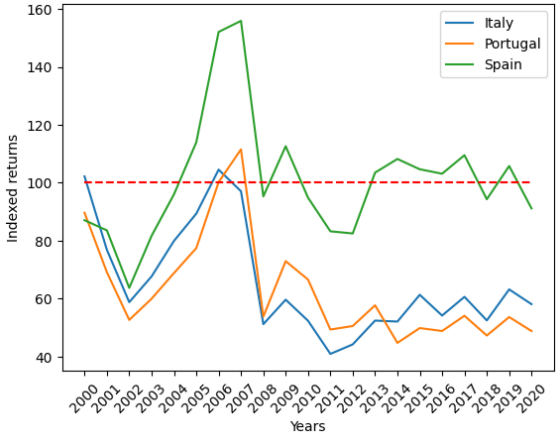
(a) West-Europe.



(b) North-Europe.



(c) East-Europe.



(d) South-Europe.

Figure 1: Returns of every country within their European region. The starting point is 01-01-2000 and they are indexed at 100 on 01-01-1999. Each sub graph contains the countries belonging to that European region. The red dashed line is equal to 100 and indicates whether a country has seen growth or decline since the start.

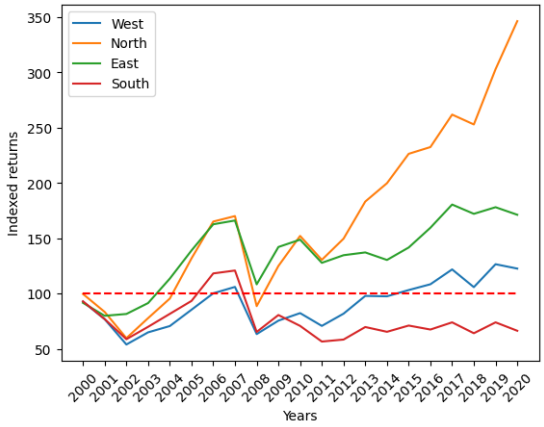


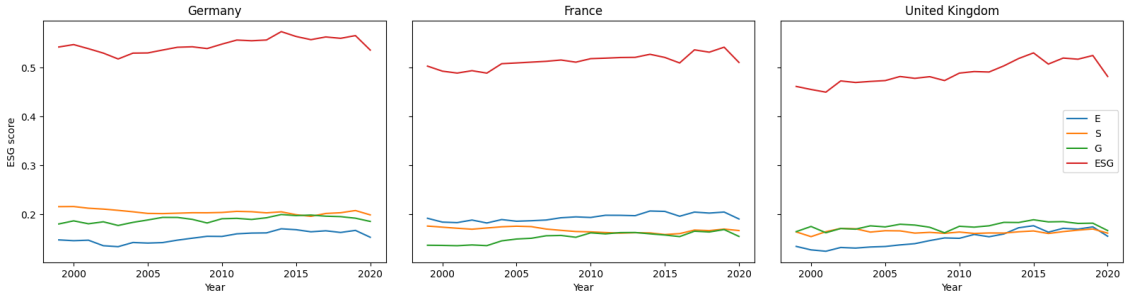
Figure 2: Average index returns per European region. Each region contains the average of the returns per country as seen in figure 1. The starting point is 01-01-2000 and they are indexed at 100 on 01-01-1999. The red dashed line is equal to 100 and indicates whether a region has seen growth or decline since the start.

	Environment	Social	Governance	ESG
Weight	36.2	32.8	31.0	100
Indicators	22	17	16	55

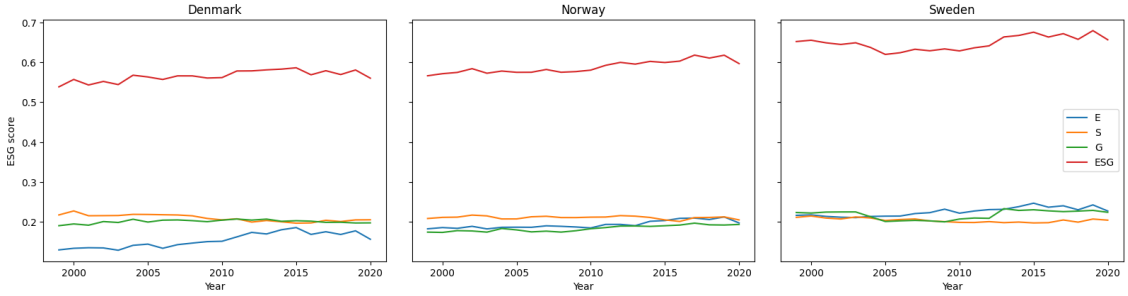
Table 3: The weights and number of indicators for the ESG index. In the first three columns for each of the pillars and in the fourth for the total ESG index.

In figure 3a we can see the ESG scores and the three individual pillar scores for the Western European countries. In all three of the countries we can see that the growth made in the beginning is lost in the last two years. In Germany the scores remained roughly the same throughout the years besides the small decline in the social score. In France we can see a decline in the environmental and social score but an increase in the governance score. In the United Kingdom the social and governance scores initially start higher than the environmental score, but they converge throughout the years. Figure 3b shows the ESG scores for North-Europe, we can see that they achieve relatively high scores. In Denmark and Norway the social score is the highest. In Sweden the environmental score is the highest scoring pillar. We can see that Denmark has a low score for the environmental pillar compared to Sweden and Norway. The ESG scores for East-Europe can be seen in figure 3c. On average this is the worst performing region of Europe. In Poland we can see growth between 2004 and 2016 but in the last four years, there is a strong decline. The last set of graphs are those from South-Europe which can be seen in figure 3d. In every country we can see that the highest scores are on the environmental pillar. The lowest scoring pillar is governance. Italy is the worst performing country in this region.

Overall, we can see that Northern Europe has the best ESG scores, followed by Western Europe, then Southern Europe and Eastern Europe. The largest change in ESG scores was seen in Eastern Europe. When comparing the ESG graphs to the market return graphs, we see that the ESG ratings are more stable than the market returns.

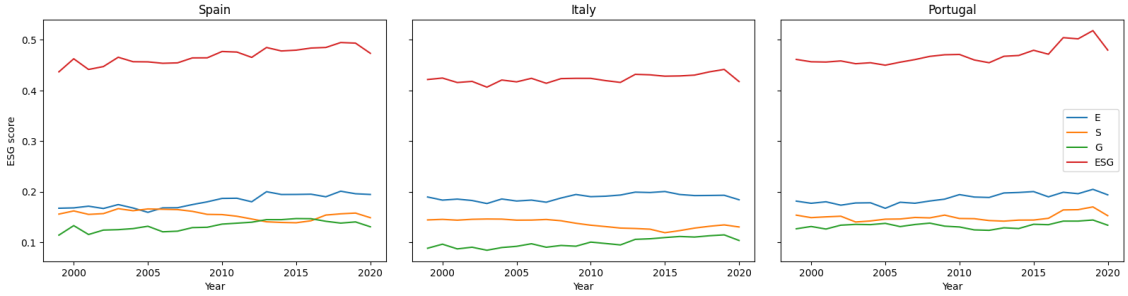


(a) West-Europe.



(b) North-Europe.

(c) East-Europe.



(d) South-Europe.

Figure 3: The ESG scores from 1999 until 2020 for every European region. Each region contains a graph for each country in that region. Each graph consists of the total ESG score for that country and the score for each individual pillar.

3.3 Regression

In this paper we look at multiple countries over multiple years, therefore we use a panel data regression. To be more specific, a dynamic panel model is used. Dynamic panel models differ from static panel models because they can include lags of both the dependent and independent variables. The lags are important because there is evidence of memory in stock market returns (Cheung & Lai, 1995). To verify that a dynamic model is indeed the best model to fit this data sample, a fixed effect regression is conducted to test whether the market return lag is significant. For this data sample we can conclude that the first lag of Index is significant and thus can use a dynamic model. The dynamic model used in this paper is introduced by Arellano and Bond (1991) which uses generalized method of moments (gmm) to estimate coefficients. The general form can be seen in equation 10 and the total error in equation 11. The theory of the Arellano-Bond estimates in this section follows the ideas from Roodman (2009). This general form is also known as the level equation, the equation we try to estimate through gmm. In general, two forms of gmms are used for the Arellano-Bond estimates. The first form of gmm is the difference gmm, this form relies on taking first differences and thereby transforming the data to remove fixed effects (endogeneity). The first differences are used as instruments. The second form is system gmm, this method uses lags as instruments which allows for a larger sample size. The system gmm was first implemented by Blundell and Bond (1998). In this paper we will use the difference gmm as it does not require additional assumptions. Additionaly, the system gmm was tested and the results did not differ significantly from the difference gmm.

$$y_{it} = \alpha y_{it-1} + \beta_i X_{it} + \eta_i + \nu_{it} \tag{10}$$

$$\epsilon_{it} = \eta_i + \nu_{it} \tag{11}$$

In this equation αy_{it-1} is the dynamic component of the regression where α is the $AR(1)$ coefficient. η_i is the time invariant unobserved component and $\nu_i t$ is the idiosyncratic error. To eliminate η_i, the first differences are taken, we end up with equation 12. We eliminate η_i because it is assumed to be related to the variables.

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \beta_i \Delta X_{it} + \Delta \nu_{it} \tag{12}$$

We estimate α_i, β_i and ν_{it} using this differences equation. In the Arellano-Bond equation the use of a constant is not necessary as it does not change the estimates. We add lags for the ESG index as well as for each individual pillar because the effect of the pillars differs over time (Giese et al., 2021). To estimate the level equation, the variables work as instrumental variables for themselves.

$$y_{it} = \alpha_1 y_{it-1} + \alpha_2 y_{it-2} + \beta_{1i} ESG_{it} + \beta_{2i} ESG_{it-1} + \beta_{3i} F_{it} + \beta_{4i} W_{it} + \beta_{5i} CPI_{it} + \beta_{6i} CPI_{it-1} + \epsilon_{it} \tag{13}$$

$$\begin{aligned} y_{it} =& \alpha_1 y_{it-1} + \alpha_2 y_{it-2} + \beta_{1i} E_{it} + \beta_{2i} E_{it-1} + \beta_{3i} S_{it} + \beta_{4i} S_{it-1} + \beta_{5i} G_{it} + \beta_{6i} G_{it-1} + \beta_{7i} F_{it} + \\ & \beta_{8i} W_{it} + \beta_{9i} CPI_{it} + \beta_{10i} CPI_{it-1} + \epsilon_{it} \end{aligned} \tag{14}$$

In equation 13 and 14 we can see the regressions used in this research. Equation 13 uses the ESG index and the lagged ESG index, in equation 14 each pillar is a separate variable. For each pillar the lags are included as well. F, W and CPI are the control variables, more information about these will follow in section 3.3.2. The last term, ϵ_{it} is the error and is defined in equation 11.

3.3.1 Tests for Assumptions

To check the assumptions for the Arellano and Bond estimator two tests are used. The first test is the Arellano-Bond test for autocorrelation in first-differenced errors (Arellano & Bond, 1991). The null hypothesis states that there is no autocorrelation and the alternative hypothesis states that there is autocorrelation. For the Arrelo-Bond estimates to be valid, the null hypothesis should not be rejected for second order autoregression i.e. there is no autocorrelation. The first order is allowed to be autocorrelated.

The second test is the Sargan test of over identifying restrictions (Sargan, 1958). The null hypothesis is that the over identifying instruments are valid, the alternative hypothesis states that they are not. For the Arellano-Bond estimates to be valid, we should not be able to reject the null hypothesis. In case the null hypothesis is rejected, the instruments are not valid and the model is not fitted correctly.

3.3.2 Control Variables

In this paper three control variables are used, all three are from the World Bank database and follow their definitions (The World Bank, 2022). Since we are looking at stock market returns we want to have control variables that have an influence on the market returns. The first control variable is the consumer price index (CPI), the use of CPI as a control variable follows from Anari and Kolari (2001). They find from past evidence and their own research, that there is a strong relation between stock prices and inflation. In the short run the relation is negative and in the long run it is positive. The second control variable is a countries wealth, in Afonso and Sousa (2011) it is argued that wealth has a strong relation with stock returns. The total wealth is calculated as the sum of produced capital, natural capital, human capital and net foreign assets. The third control variable is economic fitness (EF). The fitness is a score for the ability of a country to diversify and produce complex goods. Countries with a high EF tend to have the ability to diversify the products they produce, upgrade to complex goods and have predictable long-term growth.

4 Results

In this section the results will be discussed. We start with the correlation table, then we move on to the results of the regressions and conclude with the tests that were performed and the effect of the control variables.

In table 4 we can see the correlation between all the variables in the regression except for the lags. The first conclusion we can draw is that except for the environment every variable is significantly correlated with the market index returns. We can see that the ESG index, the social pillar, the governance pillar and the CPI have a low positive correlation with the index returns. Something that was to be expected is the highly significant correlation between the three pillars and the ESG index.

Variables	Index	ESG	E	S	G	Fitness	Wealth	CPI
Index	1.000							
ESG	0.300***	1.000						
E	-0.001	0.434***	1.000					
S	0.395***	0.799***	-0.044	1.000				
G	0.255***	0.918***	0.144**	0.740***	1.000			
Fitness	-0.295***	-0.048	-0.128**	-0.109*	0.071	1.000		
Wealth	-0.285***	-0.017	-0.215***	-0.082	0.161**	0.851***	1.000	
CPI	0.297***	0.211***	0.337***	-0.029	0.169***	0.069	0.130**	1.000

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Correlation table, all variables and every country is included.

As for the control variables besides CPI, we can see that there is a significant negative correlation with the index returns. Fitness and Wealth have little to no correlation with the ESG index or the pillars individually. Furthermore, we can see that Fitness and Wealth are very highly positively correlated.

4.1 Regression Results

In table 5 the regressions can be seen; (1) & (2) are for Western Europe, (3) & (4) are for Northern Europe, (5) & (6) are for Eastern Europe and (7) & (8) are for Southern Europe. Two different regressions were fitted for each region: one with the ESG index as a whole (uneven numbers) and one where the three pillars were split up (even numbers). To include every variable in the same order of magnitude the Index and the ESG index are scaled by 100. Furthermore, wealth is divided by 10^{12} and therefore represents the amount in trillions. First, conclusions are drawn for each region individually and then general remarks are made.

For West-Europe (regressions (1) & (2)), when we use the ESG index as a variable we can see that it has an insignificant positive effect on the market returns. If for a moment we assume the estimator was significant, if we increase the scaled ESG index by one, the returns would increase by 1.766. Meaning, an extra return of 1.766%. The ESG score of the year before has an insignificant negative influence on the returns. In regression (2) we can see the regression where the pillars are split and lagged. We can see that both the environment and the lagged environmental pillar are insignificantly negatively related to the index. An increase of one in the score of the environment leads to -3.467% return. We can see that the social and

	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index	(6) Index	(7) Index	(8) Index
L.Index	0.527***	0.412***	0.797***	0.731***	0.754***	0.745***	0.633***	0.540***
	(3.71)	(2.65)	(5.60)	(4.76)	(3.95)	(3.25)	(4.46)	(4.06)
L2.Index	-0.163	-0.0978	0.135	0.127	-0.0645	-0.0732	-0.213	-0.280**
	(-1.20)	(-0.58)	(0.92)	(0.82)	(-0.29)	(-0.29)	(-1.51)	(-2.03)
ESG	1.766		0.142		-3.682		0.630	
	(0.95)		(0.03)		(-0.69)		(0.28)	
L.ESG	-2.056		-2.565		3.527		-1.962	
	(-0.78)		(-0.47)		(0.58)		(-0.89)	
E		-3.467		-4.023		-8.061		-3.040
		(-0.83)		(-0.42)		(-0.72)		(-0.78)
L.E		-4.651		6.207		17.18		-10.74***
		(-0.91)		(0.56)		(1.29)		(-2.83)
S		6.787		18.29		4.622		7.129*
		(0.95)		(1.03)		(0.39)		(1.73)
L.S		-0.176		-22.71		-9.556		-5.821
		(-0.03)		(-1.46)		(-0.86)		(-1.37)
G		5.762		-7.331		-5.649		-5.603
		(1.28)		(-0.67)		(-0.49)		(-1.26)
L.G		-1.705		1.984		9.803		6.452*
		(-0.36)		(0.18)		(0.74)		(1.65)
Fitness	-1.633	-1.530	-6.134	-0.875	18.84	18.72	0.0675	3.575
	(-1.53)	(-1.37)	(-0.57)	(-0.06)	(0.96)	(0.82)	(0.03)	(1.40)
Wealth	2.145**	2.487**	-4.797	-4.069	6.100	-5.521	5.141	-0.675
	(2.20)	(2.42)	(-0.29)	(-0.23)	(0.19)	(-0.16)	(1.22)	(-0.14)
CPI	-4.950**	-4.305*	-17.83***	-17.14***	-2.435	1.966	-5.183**	-5.806***
	(-2.14)	(-1.77)	(-3.33)	(-3.09)	(-0.55)	(0.34)	(-2.28)	(-2.73)
L.CPI	5.489**	5.589**	19.78***	18.96***	3.109	-1.306	4.547**	6.005***
	(2.32)	(2.23)	(3.67)	(3.32)	(0.77)	(-0.23)	(2.15)	(2.96)
Observations	57	57	57	57	38	38	57	57

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: The Arellano-Bond estimates for the regression. Every uneven numbered regression is for the whole ESG index, every even regression uses the individual pillars. Regressions (1) & (2) are West-Europe, (3) & (4) North-Europe, (5) & (6) East-Europe and (7) & (8) South-Europe. The pillars are denoted by their capitalized first letter (E, S & G), the total ESG index is denoted by ESG. The lags are indicated by a L. in front of the variable.

governance pillars have a insignificant positive influence on the returns.

The second region we are going to discuss is North-Europe (regressions (3) and (4)). When looking at the ESG regression we can see that the estimator is insignificant. Not only that but the effect is also very minimal, an increase of one of the ESG index only causes a 0.142% return increase. In the regression with the individual pilalrs, only the social pillar has a positive relation with the index returns.

The East-European countries are the third to be discussed (regressions (5) & (6)). This is the only regression where a negative coefficient is found for the ESG index and the lagged ESG has a positive coefficient. For the other regions the effect is positive. For individual pillars East-Europe has the same results as North-Europe; the environment has a negative coefficient, the social aspect has a positive coefficient and the governance pillar has a negative coefficient as well. However, all estimated coefficients are insignificant.

The last region is South-Europe (regressions (7) & (8)). Here, we can see that not only the first lag of the index returns is significant but the second lag as well. Furthermore, the lags of the environmental pillar and governance pillar are significant. We can see that the environment pillar has a negative relation with next years' index returns, the same goes for the governance pillar. The effect is rather large as well, if a country increases it's environmental score by one, the index returns drop by -10.74%.

In general, we can see that the index returns are influenced by the index returns of the year before. In for example Northern Europe we can see that an increase of 1 in returns, carries over approximately 70% to the next year. Except for the lag of governance, the lag of environment and the social score of Southern Europe, all the ESG scores are insignificant.

In the results we can see that there is no constant and no R^2. The lack of constant is due to the fact that differencing gmm is used, when taking differences the constant is eliminated. Whether there is a constant or not, does not have effect on the estimators values and significance (Roodman, 2009). In contrary to a ordinary least squares regression, the R^2 is not applicable for this regression as a goodness-of-fit. In this paper to test the goodness-of-fit, the Sargan test is used.

An important thing to keep in mind when we are comparing the results of this paper with the current literature, is that the literature uses firm level ESG data. The results found in this paper compare to the literature when looking at the same macro level of the portfolios. The paper of De Spiegeleer et al. (2021) found no clear evidence of enhanced performance when using ESG investing in the STOXX Europe 600. In Morea et al. (2022) they found that the ESG index has an influence on the STOXX Europe 50 but not significantly. This paper concludes the same, it seems that there is some effect of the sovereign ESG index on market indexes but it is not significant. This notion of effect comes from the correlation matrix where we can see that there is a positive correlation between the index returns and the ESG index. The main conclusion of Jiang et al. (2022) is one that follows from these results and the literature review as well. ESG ratings are important on firm-level data but on macroeconomic level it is under researched and that sovereign ESG research is at an early stage. This also leads to a different conclusion than found in Friede et al. (2015), they found that the majority of the literature finds a positive relation between CFP and the ESG index. The individual effects found in Giese et al. (2021) are not found in this research. Since the results are insignificant

we can not draw the conclusion that governance has the most effect on the short term and that environment and social develop more slowly. Even if the results would be significant, this is not the case. The overall conclusion we can draw from this paper in comparison to the literature, is that results from firm level data can not be projected onto macro level data and that further research is needed for the sovereign ESG index.

4.1.1 Tests

All eight regressions have the same results for the Arellano-Bond test for autocorrelation and the Sargan test of over identifying restrictions. The p-values and the corresponding test statistics can be seen in table 6. The Arellano-Bond tests find that there is autocorrelation in the first order differences but there is no autocorrelation in the second order differences (insignificant for $\alpha = 0.1$). In other words, we accept the zero hypothesis for the first order tests but we reject the zero hypothesis for the second order tests. The Sargan tests are all insignificant and thus we can accept the null hypothesis that all the instruments are valid. This is even the case when taking into account the concern of Roodman (2009) that p-values of 0.25 should still be viewed as a potential sign of trouble. From these results we can conclude that the Arellano-Bond model was fitted correctly.

Tests	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Order 1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(-5.378)	(-5.359)	(-4.178)	(-3.986)	(-4.271)	(-3.792)	(-4.952)	(-4.354)
Order 2	0.316	0.223	0.6822	0.895	0.896	0.734	0.889	0.731
	(1.004)	(1.218)	(-0.409)	(-0.132)	(-0.131)	(0.332)	(0.140)	(0.344)
Sargan	0.618	0.674	0.390	0.443	0.588	0.687	0.419	0.423
	(45.45)	(40.225)	(51.125)	(45.694)	(27.667)	(22.041)	(50.383)	(46.120)

Table 6: P-values of the Arellano-Bond and Sargan test which are used to check the assumptions of the Arellano-Bond estimates. Order 1 & 2 are the results for the auto correlation test with the z statistics in parentheses. Sargan are the results for the Sargan test and the chi squared statistics are between the parentheses.

4.1.2 Control Variables

What is remarkable about the control variables are mostly the CPI values, which are significant except for Eastern Europe. They seem to support the relationship found in the literature where the short term effects (CPI) are negative and the longer term effects are positive (lagged CPI). We especially see a strong relationship in Northern Europe. The Wealth variable is positive and significant in Western Europe but insignificant in the rest of Europe. The EF is insignificant in every regression. Even though the EF is insignificant, the parameter is negative in Western and Northern Europe and positive in Eastern and Southern Europe.

5 Conclusion

In this research we looked at the relationship between stock market returns and sovereign ESG scores in Europe. Panel data regressions were run for each European region to see if there is a significant relationship for those countries. This research was conducted due to the novel data set of the World Bank and the new way to studying ESG data on country level instead of firm level. The research into country level ESG data is still very new and general conclusions about the relation with respect to financial performance are yet to be made. This paper looks to jump into this research gap and strives to make some first conclusions.

To look for a relationship between ESG scores and stock market returns the main regression technique is to calculate the Arellona-Bond estimates. This is a dynamic panel data model which allows for lags of the market returns to be included. Through the Arellano-Bond test for autocorrelation and the Sargan test for validity of instruments we can conclude that the model was fitted correctly in this research. The data for the ESG ratings and the instrumental variables was collected from the World Bank. To remain independent of rating methods of different agencies, entropy weighting was used to calculate the ESG scores. In general, we could conclude that in every region there was some small growth in the ESG index. The market returns were retrieved from Wharton Research Data Services.

From the literature review it can be concluded that we are still at the beginning of doing research into sovereign ESG scores. We can not yet conclude that the positive relation found in firm level data between the ESG and performance carries over to country level data. Furthermore, we can see that the lack of a general rating scheme is also present in ESG scores on country level, even though it is a bit better. Resorting to mathematical objective weight methods might not be the best way to derive economical scores for the ESG, as some indicators carry sentiment that pure mathematics can not analyze. The entropy weighting used in this paper analyzes all indicators in the same way, while humans may view one indicator as more important and in need of a heavier weighting.

From the results of the correlation matrix we can see that the ESG, the social pillar and the governance pillar are positively correlated with the index returns, all three are significant. The environmental pillar is not correlated with the index returns. From the regressions we learned that the index returns are heavily and positively influenced by the first lag of the returns i.e. the returns of this year are dependent on last years' returns. The ESG score as a whole has an insignificant but positive relation with the market returns in Western, Northern and Southern Europe. In Eastern Europe we had an insignificant negative relation. For every European region environment had a insignificant negative relation with the index returns. The lag of the environment had a significant negative relation with Southern European index returns. The social pillar had a positive relation for every region, it was only significant for South Europe.

These results do not allow us to accept either hypothesis 1 or 2. For hypothesis 1 we do find that the ESG seems to be positively related to stock market returns, except in East-Europe. However, the results are not significant. For the second hypothesis we looked at the different impacts each pillar would have on the market returns. The environmental has a negative effect on the index returns but we can see in the correlation table that the effect is negligible. There is no consensus in the results about the sign of the lagged environment variable. The social pillar seems to be positively related to the market index scores as we get positive estimates from the regression as well as from the correlation table. The first lag of the social pillar

has a insignificant negative influence on the stock market returns. For the governance pillar, in West-Europe there is a positive relation but in the rest of Europe there is a negative relation. The correlation table finds a positive correlation between governance and index returns.

From an investor standpoint it is too early to start incorporating the sovereign ESG index when investing. Until now, little research has been conducted into the relation between stock returns and sovereign ESG scores. The advice for now is to stick to micro (firm) level ESG ratings where it is known that there is a positive relation and where the ESG index is a meaningful risk measure.

5.1 Discussion & Future Work

The entropy weighting method is heavily dependent on the input data. For this paper only eleven countries are used to calculate the weights, this can lead to more extreme differences than there might be in real life. For future research it might be better to calculate the ESG weights for at least all of Europe and maybe even for the whole world. Furthermore, the missing data values are filled using the mean method this might lead to the following problem. In every chart of the ESG scores we can see a decline in the end. This can have two reasons: either there is a decline in ESG scores (because of Covid-19 for example) or there is too much data missing near the end and the score is pulled down by the mean values. As discussed in the theory section, sovereign data is slowly updated and can take a couple years to be completed. Both reasons for the decline create opportunities for further research. First of all, the use a more sophisticated method than mean value filling can be important. The second is one that only time can tell, whether there was a real decline after 2017.

The Arellano-Bond estimator works better as N (number of observations) tends towards infinity. Due to the use of the differencing gmm we had $N = 57$. Furthermore, this gives rise to a problem with the Arellano-Bond estimators, it is not known for what N the test becomes reliable only that it should be high.

In the theory it was concluded that the pillars increase or decrease in importance throughout the years, this happens in relative short periods of around three years. In this paper we used 21 years of of observations, which might be too much to capture the changing real effect the ESG has in different years. Therefore, future research could investigate the influence of the length of the time period used on the strength of the relation between ESG scores and stock market returns.

References

Afonso, A., & Sousa, R. M. (2011). Consumption, wealth, stock and government bond returns: international evidence. *The Manchester School*, *79*(6), 1294–1232.

Alshehhi, A., Nobanee, H., & Khare, N. (2018). The impact of sustainability practices on corporate financial performance: Literature trends and future research potential. *Sustainability*, *10*(2), 494.

Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use esg information: Evidence from a global survey. *Financial Analysts Journal*, *74*(3), 87–103.

Anari, A., & Kolari, J. (2001). Stock prices and inflation. *Journal of Financial Research*, *24*(4), 587–602.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The review of economic studies*, *58*(2), 277–297.

Baczkiewicz, A., & Watróbski, J. (2022). Crispyn—a python library for determining criteria significance with objective weighting methods. *SoftwareX*, *19*, 101166.

Bannier, C. E., Bofinger, Y., & Rock, B. (2019). *Doing safe by doing good: Esg investing and corporate social responsibility in the us and europe* (Tech. Rep.). CFS Working Paper Series.

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

Bessler, D. A., & Yang, J. (2003). The structure of interdependence in international stock markets. *Journal of international money and finance*, *22*(2), 261–287.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of econometrics*, *87*(1), 115–143.

Bouyé, E., & Menville, D. (2020). The convergence of sovereign environmental, social and governance ratings. *Social and Governance Ratings (December 21, 2020)*.

Carroll, A. B. (1979). A three-dimensional conceptual model of corporate performance. *Academy of management review*, *4*(4), 497–505.

Cheung, Y.-W., & Lai, K. S. (1995). A search for long memory in international stock market returns. *Journal of International Money and Finance*, *14*(4), 597–615.

Chukwuogor-Ndu, C. (2006). Stock market returns analysis, day-of-the-week effect, volatility of returns: Evidence from european financial markets 1997-2004. *International Research Journal of Finance and Economics*, *1*(1), 112–124.

Dajcman, S. (2013). Interdependence between some major european stock markets-a wavelet lead/lag analysis. *Prague economic papers*, *22*(1), 28–49.

De Spiegeleer, J., Höcht, S., Jakubowski, D., Reyners, S., & Schoutens, W. (2021). Esg: a new dimension in portfolio allocation. *Journal of Sustainable Finance & Investment*, 1–41.

Dimson, E., Marsh, P., & Staunton, M. (2020). Divergent esg ratings. *The Journal of Portfolio Management*, *47*(1), 75–87.

Drei, A., Le Guenedal, T., Lepetit, F., Mortier, V., Roncalli, T., & Sekine, T. (2019). Esg investing in recent years: New insights from old challenges. *Available at SSRN 3683469*.

Égert, B., & Kočenda, E. (2007). Interdependence between eastern and western european stock markets: Evidence from intraday data. *Economic Systems*, *31*(2), 184–203.

Eurosif. (2021). Eurosif report 2021.

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

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

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

Gratcheva, E., Emery, T., & Wang, D. (2021). Demystifying sovereign esg. *Available at SSRN 3854177*.

Harkmann, K. (2014). Stock market contagion from western europe to central and eastern europe during the crisis years 2008-2012. *Eastern European Economics*, *52*(3), 55–65.

Jiang, P.-C., Feng, G.-F., & Yang, H.-C. (2022). New measurement of sovereign esg index. *Innovation and Green Development*, *1*(2), 100009.

Larcker, D. F., Tayan, B., & Watts, E. M. (2022). Seven myths of esg. *European Financial Management*, *28*(4), 869–882.

Li, F., & Polychronopoulos, A. (2020). What a difference an esg ratings provider makes. *Research affiliates*, *15*.

Li, T.-T., Wang, K., Sueyoshi, T., & Wang, D. D. (2021). Esg: Research progress and future prospects. *Sustainability*, *13*(21), 11663.

Maiti, M. (2021). Is esg the succeeding risk factor? *Journal of Sustainable Finance & Investment*, *11*(3), 199–213.

Marti, C. P., Rovira-Val, M. R., & Drescher, L. G. (2015). Are firms that contribute to sustainable development better financially? *Corporate Social Responsibility and Environmental Management*, *22*(5), 305–319.

Morea, D., Mango, F., Cardi, M., Paccione, C., & Bittucci, L. (2022). Circular economy impact analysis on stock performances: An empirical comparison with the euro stoxx 50® esg index. *Sustainability*, *14*(2), 843.

Morgenstern, C., Coqueret, G., & Kelly, J. (2022). International market exposure to sovereign esg. *Journal of Sustainable Finance & Investment*, 1–20.

Pagano, M. S., Sinclair, G., & Yang, T. (2018). Understanding esg ratings and esg indexes. In *Research handbook of finance and sustainability.* Edward Elgar Publishing.

Roodman, D. (2009). How to do xtabond2: An introduction to difference and system gmm in stata. *The stata journal*, *9*(1), 86–136.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica: Journal of the econometric society*, 393–415.

Sassen, R., Hinze, A.-K., & Hardeck, I. (2016). Impact of esg factors on firm risk in europe. *Journal of business economics*, *86*, 867–904.

Shanaev, S., & Ghimire, B. (2022). When esg meets aaa: The effect of esg rating changes on stock returns. *Finance Research Letters*, *46*, 102302.

The World Bank. (2022). *Databank.* Retrieved 2023-01-10, from `https://databank-worldbank-org.eur.idm.oclc.org/`

United Nations. (1999). *United nations statistics division.*

United Nations Development Program. (2022). Human development report 2021-22: Uncertain times, unsettled lives: Shaping our future in a transforming world.

Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). Esg integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics*, *138*(3), 525–533.

Wharton Research Data Services. (2022). *World indices by wrds.* Retrieved 2023-01-24, from

https://wrds-www-wharton-upenn-edu.eur.idm.oclc.org/pages/grid-items/world-indices-wrds/

Zhang, D., Zhao, Z., & Lau, C. K. M. (2022). Sovereign esg and corporate investment: New insights from the united kingdom. *Technological Forecasting and Social Change*, *183*, 121899.

Zhu, Y., Tian, D., & Yan, F. (2020). Effectiveness of entropy weight method in decision-making. *Mathematical Problems in Engineering*, *2020*.

A ESG indicators and weights

Pillar	Indicator	Sign	Weight
Environment			36.2%
	Adjusted savings: net forest depletion (% of GNI)	-1	0.01%
	Agricultural land (% of land area)	-1	1.43%
	Agriculture, forestry, and fishing, value added (% of GDP)	1	2.25%
	Annual freshwater withdrawals, total (% of internal resources)	1	1.03%
	Annualized average growth rate in per capita real survey mean	-1	1.84%
	CO2 emissions (metric tons per capita)	-1	2.14%
	Electricity production from coal sources (% of total)	-1	1.31%
	Energy imports, net (% of energy use)	-1	0.91%
	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	1	2.07%
	Energy use (kg of oil equivalent per capita)	-1	1.58%
	Food production index (2014-2016 = 100)	1	2.59%
	Forest area (% of land area)	1	2.16%
	Fossil fuel energy consumption (% of total)	-1	2.12%
	GHG net emissions/removals by LUCF (Mt of CO2 equivalent)	-1	2.28%
	Mammal species, threatened	-1	1.44%
	Methane emissions (metric tons of CO2 equivalent per capita)	-1	1.85%
	Nitrous oxide emissions (metric tons of CO2 equivalent per capita)	-1	1.77%
	PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)	-1	1.68%
	Population density (people per sq. km of land area)	-1	1.86%
	Renewable electricity output (% of total electricity output)	1	1.84%
	Renewable energy consumption (% of total final energy consumption)	1	2.00%
	Terrestrial and marine protected areas (% of total territorial area)	1	1.88%
Social			32.8%
	Cause of death, by communicable diseases and maternal, prenatal and nutrition conditions (% of total)	1	2.18%
	Fertility rate, total (births per woman)	-1	1.86%
	Gini index	-1	2.32%
	Government expenditure on education, total (% of government)	1	2.35%
	Hospital beds (per 1,000 people)	1	2.06%
	Income share held by lowest 20%	1	2.30%
	Labor force participation rate, total (% of total population ages 15-64)	1	1.96%
	Life expectancy at birth, total (years)	1	1.75%
	Mortality rate, under-5 (per 1,000 live births)	-1	1.65%

	People using safely managed drinking water services (% of population)	1	1.32%
	People using safely managed sanitation services (% of population)	1	1.23%
	Population ages 65 and above (% of total population)	-1	2.33%
	Poverty headcount ratio at national poverty lines (% of population)	-1	2.15%
	Prevalence of overweight (% of adults)	-1	2.35%
	School enrollment, primary (% gross)	1	2.71%
	Unemployment, total (% of total labor force) (modeled ILO estimate)	-1	1.35%
Governance			31.0%
	Control of Corruption: Estimate	-1	1.80%
	GDP growth (annual %)	1	1.76%
	Government Effectiveness: Estimate	1	2.09%
	Individuals using the Internet (% of population)	1	1.60%
	Net migration	-1	1.99%
	Patent applications, residents	1	1.20%
	Political Stability and Absence of Violence/Terrorism: Estimate	1	2.16%
	Proportion of seats held by women in national parliaments (%)	1	1.85%
	Ratio of female to male labor force participation rate (%) (modeled ILO)	1	2.04%
	Regulatory Quality: Estimate	1	1.98%
	Research and development expenditure (% of GDP)	1	2.22%
	Rule of Law: Estimate	1	1.76%
	School enrollment, primary and secondary (gross), gender parity index (GPI)	1	2.36%
	Scientific and technical journal articles	1	1.68%
	Strength of legal rights index (0=weak to 12=strong)	1	1.59%
	Voice and Accountability: Estimate	1	1.99%

Table 7: The weight of every indicator sorted per pillar. The signs indicate whether a indicator has a positive or negative effect on it's pillar. These weights are consistent from 1999 until 2020. Eleven countries were used as input for the entropy weighting; France, Germany, United Kingdom, Denmark, Norway, Sweden, Hungary, Poland, Italy, Portugal and Spain