Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11184||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Sander August Standal Jensen og Endre Rørvik Holmeset

Informasjon fra deltaker

Tittel *:	ESG Uncertainty and Stock Returns in Europe
Navn på veileder *:	Stig Roar Haukoe Lundeby

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	173
Andre medlemmer i gruppen:	

ESG Uncertainty and Stock Returns in Europe

Supervisor:
Stig Roar Haukoe Lundeby

Programme:
Master of Science in Business – Major in Sustainable Finance

Oslo, July 3rd, 2023

Abstract

In this paper we analyze whether ESG uncertainty can predict stock returns by studying the 3000 largest common stocks within Europe. We observe a tendency of predictability between ESG uncertainty in the aggregated ESG score and future stock returns, but no significant premium when controlling for known risk factors. However, the average ESG rating seems to be a more important predictor of stock returns, as the environmental dimension yields a premium from stocks with low ESG ratings also after controlling for known risk factors. We observe that in general, companies with higher book-to-market values and high volatility stocks have greater ESG uncertainty while firms with higher market capitalization and leverage have the opposite effect. Accordingly, we observe that firm characteristics are an important attribute to consider when examining the disagreement among ESG providers.

Acknowledgements

We would like to give a special thanks to our supervisor Stig Roar Haukoe Lundeby for his valuable support and insightful discussions during our work with this Master Thesis. Thanks to BI Norwegian Business School for five years of high-quality education and relevant competence which we appreciate taking with us into our professional carriers.

Table of Contents

1 Introduction

There has been an exponential growth in sustainable investing and ESG focus the last decade (Larcker et al., 2022). Investors and companies' actions are increasingly affected through ESG reporting requirements and new regulations. Even though ESG ratings may contribute with information on companies' nonfinancial impact (Larcker et al., 2022), there is a significant challenge related to different ESG rating providers and their divergence (Berg et al., 2022). A proper illustration of this issue is how the correlations in the ESG rating system are way less standardized and precise compared to credit ratings (Berg et al., 2022). To our knowledge, there are relatively few studies in Europe regarding ESG-score and stock returns, and especially when taking the rating disagreement effect into account.

The first part of our analysis addresses the level of correlation between the four ESG-rating providers. We found that the average pairwise correlation is 0.54, approximately 7 percentage-points greater than the average pairwise correlation found by Gibson et al. (2021) analyzing US firms, and about the same correlation as Berg et al. (2022). Despite different time periods and companies, the total average correlations are relatively similar. However, the average pairwise correlation is highest in the social pillar (0.429) and lowest for the governance pillar (0.322). The result is somewhat surprising as we expected the environmental dimension to yield a higher correlation compared to the social dimension as environmental components are easier to measure and assess. (Berg et al., 2022).

Next, we find that in general firms with higher book-to-market value and higher stock volatility tend to have higher ESG rating disagreement. We will use the terms ESG rating disagreement and ESG uncertainty interchangeably in this paper. On the other hand, firms with higher leverage and market capitalization tend to have lower ESG uncertainty, while factors such as return-on-assets and momentum do not seem to explain much of disagreement in ESG-ratings. Given the economic magnitude of our findings it appears that firm characteristics do have an effect on the ESG uncertainty, depending on the variable of interest.

In the main part of this thesis, we are studying whether ESG uncertainty can predict future stock returns. We find some evidence that there is a positive relationship between rating disagreement and stock returns in the aggregated ESG dimension. However, the effect seems to be more explained by the average ESG score in the environmental dimension after controlling for known risk factors. The fact that we cannot conclude that ESG uncertainty is a proper predictor for future stock returns is somewhat surprising compared to findings from Gibson et al. (2021) and Avramov et al. (2022).

Finally, we performed portfolio-sorting by dividing the stocks into five quintiles from low to high ESG-rating uncertainty. A long-short strategy by going long on high disagreement portfolio and short on the low disagreement portfolio generates an insignificant alpha after controlling for the known risk factors in the Fama-French 5 factor model. These findings contradict the findings from Gibson et al. (2021) (S&P 500), but more in line with Anselmi et al. (2022) findings for Stoxx600 which finds no evidence that ESG-rating disagreement affects stock returns.

2 Literature review and theory

Our thesis relates to several strands of the research literature. First, it is related to a growing strand of literature that shows large ESG rating disagreement. ESG information has become an important component in sustainable investment decisions. Moreover, there is an increasing number of investors that use ESG ratings as a third-party assessment of ESG performance (Berg et al. 2022). Nevertheless, as ESG ratings are influencing investment decisions, a major challenge is related to disagreement between rating providers. Berg et al. (2022) finds evidence that there is a substantial divergence between rating providers, where measurement contributes 58% to the divergence when raters measure the same attribute, but with different measurements. Secondly, scope contributes with 38%, which is when agencies measure different attributes. The remaining 6% is weight, which is when same attributes are given different weights in the aggregated ESG score.

These findings indicate that the disagreement regarding ESG-ratings is not only due to different definitions, but also fundamental disagreement about the underlying data. Consequently, this highlights that corrugating for the ESG divergence and creating a common methodology is not that trivial. This is further highlighted as the correlation between the rating agencies varies between 0.38 and 0.71, compared to creditworthiness ratings of 0.99 correlation (Berg et al., 2022). Our study contributes by looking specifically at European firms including a wide range of countries and industries.

Another relevant aspect is the relationship between disclosure and ESG rating disagreement. The consensus in the previous literature seems to support Christensen et al. (2021) findings that higher disclosure leads to higher disagreement, given that higher disclosure entails a higher probability that raters use different metrics. On the other hand, there is also evidence in the literature supporting that longer reports in terms of more disclosing reduces disagreement among ESG-providers (Kimbrough et al. 2020). The paper from Kimbrough et al. (2020) also finds that ESG reports containing positive "tones" and "sticky words" increases disagreement. Some of the explanation was related to third parties (e.g., consultancy firms), which are considered to reduce disagreement in the disclosure process.

Next, there are some papers that not only investigate the divergence between ESG-providers and why they disagree, but also how rating disagreement affects financial performance. The literature finds some mixed evidence, and to our knowledge it seems to be a predominance of studies from the US. Gibson et al. (2021) finds that risk averse investors who invest in firms with high ESG rating disagreement demand compensation in terms of a risk premium due to the uncertainty related to the firm´s excepted ESG performance in the future. Avramov et al. (2022) finds that under ESG uncertainty, the market premium increases and demand for those stocks declines. Additionally, Christensen et al. (2021) finds that higher uncertainty leads to higher return volatility and a reduced probability of external financing. Accordingly, our thesis contributes by investigating this topic in the European stock market as our main objective through this paper.

On the other hand, Anselmi & Petrella. (2022) comparing Stoxx600 and S&P 500 found no evidence to support that divergence between providers affects stock returns. Additionally, they find a negative relationship between excess returns and ESG-score, but the premium was linked to firm size rather than performance. Regardless, the evidence and theory in the literature are mixed. According to theory by Pastor et al. (2021) argues that high ESG performance should be associated with lower returns given two rational types of investors (ESG and non-ESG), while Pedersen et al. (2021) argues that ESG performance can be associated with both increased and decreased cost of capital.

Furthermore, empirical findings are also mixed as Liou (2018) finds a negative relationship between low ESG ratings and returns studying US firms. On the other hand, Billio et al. (2022) finds no statistical significance between the alphas from the ESG disagreement portfolio and the ESG agreement portfolio suggesting that ESG performance has no impact on financial performance even when rating disagreement is low. Our thesis contributes to this literature by analyzing whether there is a significant premium associated with high ESG uncertainty as well as how the average ESG performance affects stock returns in Europe.

3 Data

3.1 ESG-data

We have collected yearly ESG-ratings from four rating providers: Bloomberg, Refinitiv, Sustainalytics and S&P Global. We have constructed a portfolio consisting of the 3000 largest common stocks within Europe, ranked by market capitalization as of 01.03.2023. As there are some limitations related to missing data, we have included as many years and companies as possible to obtain sufficient power in our model. However, the fact that our sample covers approximately 30 of the European countries (included Turkey) and 18 years of data, it is reasonable to argue that there are possibilities of heterogeneity in our data, which will be addressed later.

The four ESG-rating providers uses the same rating scale, 0-100. Bloomberg, Refinitiv and S&P Global reports on a scale from 0-100 where 0 denotes low ESG-performance and 100 denotes superior ESG performance. Sustainalytics used the same rating methodology until 2018 (Morningstar, 2019) where they started reporting ESG risk score instead of ESG performance score. Hence, we transformed Sustainalytics scores after 2018 by subtracting each ESG risk score for firm i in year t from 100 to be comparable to the other ratings.

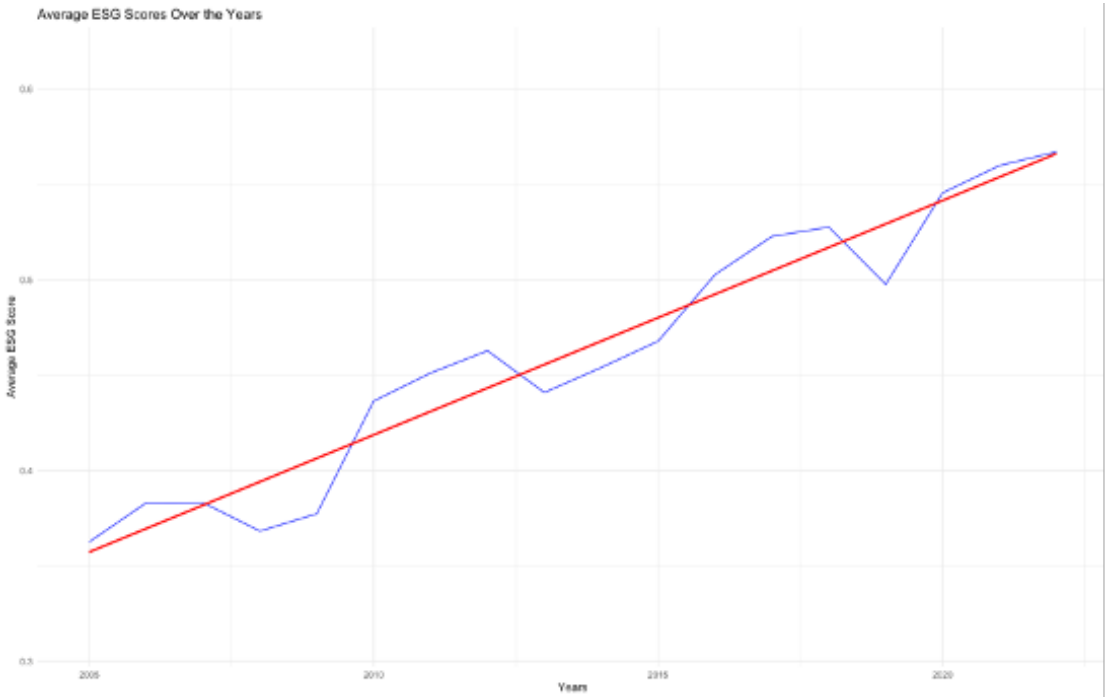
According to summary statistics reported in Table 1, we observe that Bloomberg and Refinitiv reports the highest number of observations. Sustainalytics reports the third most observations and started reporting in 2009. Lastly S&P Global is the provider that reports the least observation and has no ratings prior of 2016. Therefore, our sample has two providers in the period from 2004-2008, three providers between 2009-2015 and all four providers as of 2016. We therefore have more power from 2010, which also will be tested in the robustness check later in this paper.

Furthermore, as we can observe from Graph 1, the average ESG score is upward trending. However, there could be several potential explanations related to this effect. One explanation might be that ESG practices actually are improving, but it is hardly the only reason should we believe findings from Larker et al. (2022) studying aggregate ESG scores for Russel 1000. The paper argues for instance that rating improvements can be driven by changes in index composition in terms of higher rated companies as Microsoft. Another explanation can be changes in weight composition and increased disclosing of. Hence a large fraction of new technology in the information technology sector can increase average score.

We normalize the ratings by calculating z-scores by subtracting the cross-sectional mean from each observation within each year and dividing it by the cross-sectional standard deviation within that specific year. This approach allows for the removal of the “trend” component we discussed above that could influence the results, providing a measure of relative deviation from the average within each year. We apply the following formula for z-score:

$$Z - scores = \frac{(ESG\ rating_{i,t} - Cross\ sectional\ mean_t)}{Cross\ sectional\ standard\ deviation_t} \quad (1)$$

Graph 1



Note: The graph shows the average ESG-score 2005-2022. The blue line is the average ESG-score between all four raters. The red line illustrates the upward trend in ESG scores.

3.3 Stock returns

Stock return data are collected from the Bloomberg Terminal. All data are in Euros. We gathered stock price per share (P_t) and dividend per share (D_t) from Bloomberg Terminal. The short-term risk-free rate (R_{f_t}) is collected from Kenneth French Library data (French, 2023). We calculate the excess returns as follows:

$$R_{e_{i,t}} = R_{i,t} - R_{f_t} \quad (2)$$

Where $R_{i,t}$ denotes the stock return for company i in year t and R_{f_t} is the risk-free rate. We can derive the gross return each year as follows:

$$R_t = \frac{P_t + D_t}{P_{t-1}} \quad (3)$$

P_t is the price per share (i) at the end of time (t) denoted as P_t, and the dividend per share paid at time t is denoted D_t, divided on the stock price in time t-1.

3.4 Additional data

We obtained data on common risk factors from Kenneth French library (French, 2023). We collected the market risk premium (CAPM), Fama-French 3 factor model and Fama-French 5 factor model. Summary statistics for the risk factors are displayed in Table 1 in section 3.5. Control variables are collected from the Bloomberg Terminal and Refinitiv Eikon and normalized to standard z-scores. We have included the following relevant control variables that can explain excess returns: Return-on-assets (ROA) is the net income before financial costs divided by total assets (Refinitiv). Leverage (LEV) is the financial leverage given by market values. (Bloomberg). Book-to-market (BM) is defined as book value per share divided on market value per share (Bloomberg). Market capitalization equity (ME) is the price per share times the number of outstanding shares (Refinitiv). Volatility (VOL) is the last years volatility of stock prices (Bloomberg). Momentum (MOM) is defined as the real time share price momentum which is the fluctuations in share price for a certain stock (Bloomberg).

3.5 Interpretation of economical magnitude

As we normalize our ESG-data and all our control variables (see previous section) into z-scores, it is important to explain how to interpret the economical magnitude throughout this paper. Firstly, to understand how the independent variables affects the dependent variable, we need to calculate the standard deviation of the standard deviation of the ESG z-scores. Moving towards the interpretation of how one standard deviation increase in ESG uncertainty affects the excess return, we derive the following formula:

$$E\left(R_{e_{i,t}}\right) = E(a_{t+1}) + \lambda^{\sigma} E\left[\sigma_{i,t}\right]$$

$$\Delta E\left(R_{e_{i,t}}\right) = \lambda^{\sigma}\left[E\left(\sigma_{i,t}\right) + SD\left(\sigma_{i,t}\right) - E\left(\sigma_{i,t}\right)\right] \quad (4)$$

$$= \lambda^{\sigma} SD(\sigma_{i,t})$$

Where, $SD(\sigma_{i,t})$ denotes the standard deviation of the standard deviation of the whole sample, and λ^{σ} denotes the variable of interest (e.g., ESG uncertainty).

Another case is related to how accounting variables affects ESG uncertainty. In this case we modified the formula above, by apply the following formula:

$$x = \frac{\lambda^{\mu}}{SD(\sigma_{i,t})} \quad (5)$$

Where x denotes the increase in basis points, the λ^{μ} illustrates the independent variable (e.g., stock volatility) and $SD(\sigma_{i,t})$ denotes the standard deviation of the standard deviation of the whole sample. Throughout this thesis we will use the $SD(\sigma_{i,t})$ for the aggregated and environmental dimension. The respective values are 0.370 for the aggregated dimension, and 0.386 for the environmental dimension.

3.6 Descriptive statistics and correlations

Table 1 - Fama-French factors

	Rm-Rf	SMB	HML	RMW	CMA
Mean	0.064	0.023	-0.011	0.048	-0.011
Sharpe ratio	0.300	0.360	-0.090	0.890	-0.125
Std_Dev	0.212	0.075	0.120	0.054	0.087
t-statistic	1.274	1.508	-0.385	3.775***	-0.531

Note: The table displays the Mean, Sharp ratio, standard deviation and t-statistics for the market risk premium (Rm-Rf), Fama-French 3 factor model, and Fama-French 5 factor model.

Table 2 - Normalized ESG ratings

Panel A: Aggregated ESG				
	N	Refinitiv	Sustainalytics	Bloomberg
Refinitiv	14245			
Sustainalytics	10127	0.487		
Bloomberg	1489	0.660	0.415	
S&P Global	5211	0.641	0.437	0.574
Average correlation			**0.536**	

Panel B: Environmental				
	N	Refinitiv	Sustainalytics	Bloomberg
Refinitiv	13439			
Sustainalytics	7079	0.341		
Bloomberg	13651	0.603	0.167	
S&P Global	6205	0.554	0.361	0.409
Average correlation			**0.406**	

Panel C: Social				
	N	Refinitiv	Sustainalytics	Bloomberg
Refinitiv	14140			
Sustainalytics	7204	0.388		
Bloomberg	14447	0.536	0.348	
S&P Global	5211	0.582	0.265	0.456
Average correlation			**0.429**	

Panel D: Goverance				
	N	Refinitiv	Sustainalytics	Bloomberg
Refinitiv	14312			
Sustainalytics	7238	0.388		
Bloomberg	14702	0.536	0.348	
S&P Global	5175	0.582	0.265	0.456
Average correlation			**0.322**	

Note: The table displays the number of observations (N) as well as the average pairwise correlations between the four providers.

Table 2 Panel A shows summary statistics and Pearson correlation between the aggregated ESG ratings from the four different providers. Comparing the providers, it is Blomberg and Refinitiv that report the highest average pairwise correlation of 0.66.

On the contrary, Sustainalytics is the provider that disagree the most with the other providers. Accordingly, it is Sustainalytics and Bloomberg that agrees the least with a pairwise corelation of 0.42.

A potential explanation for the deviation between Sustainalytics and the other providers is that Sustainalytics changed their rating methodology after 2018 (Morningstar, 2019). Secondly, when examining Sustainalytics risk rating score, Sustainalytics report that a risk score above 40 is categorized as “severe”. Although they report on a scale from 0-100, Sustainalytics seems to be more restrictive in the upper parts of the rating system (resulting in few low ESG scores when reversing the risk score). Hence, we performed a robustness check by removing ESG-ratings from Sustainalytics after 2018. Interestingly, we saw a significant increase in correlation, with an average score of 0.64 (total sample). Nevertheless, the inference of our analysis is the same after controlling for this change in rating-methodology (see appendix, exhibit 2).

From the separate dimensions it is the social dimension that somewhat surprisingly obtains the highest correlation with 0.43. Although environmental issues like carbon accounting by no means is a straightforward calculation and exercise, it is still somewhat more tangible and systematic regulation driven with for instance standards and guidance from Greenhouse Gas Protocol (Greenhouse Gas Protocol, n.d.) compared to the two other dimensions. Accordingly, a correlation of 0.41 and marginally lower than social dimension with 2 percentage points is somewhat surprising. Especially since the social dimension with the example from Berg at al. (2022) in measuring employee satisfaction appears more subjective and rater dependent. However, this subjectivity obstacle in the social dimension does not seem to be that present in the European stock market given our findings.

Lastly, the governance dimension displays the lowest correlation of 0.32. Given the previous discussion of more subjectivity in the social and governance dimension considering the findings from Gibson et al (2021), the low governance correlation was expected. However, as the study is based on US firms, it can appear to be higher consensus among the providers on European firms since our correlation reported is

way higher than what Gibson et al. (2021) found at 0.16. A potential explanation for the higher governance consensus among the ESG providers could be due to the fundamental difference in board structure between the continents. While two tier board separate board and management it also increases the monitoring and transparency, which could explain the increased consensus in Europe compared to US with one-tier boards (Goergen, 2018).

The average correlation between the providers in our sample is 0.54, which is 7 percentage points higher than the study of US firms from Gibson at al. (2021). Therefore, given our sample period it tends to be higher agreement and correlation in ESG scores between European firms compared to US firms. However, compared to the credit rating agencies Moody and S&P, Berg at al. (2022) found that these providers exceeded a corelation of 0.99 which is well above what we reported. This also supports the importance of a common framework and more standardization in the ESG-reporting, which is lacking compared to the more established creditworthiness reported from the credit rating agencies.

4 Analysis and empirical results

4.1 Determinants of ESG uncertainty

4.1.1 Financial and accounting variables

Prior to investigating into whether ESG uncertainty can predict future stock returns, it is interesting to see how typical firm characteristics measured by common accounting variables, firm-level attributes and valuation affects the rating uncertainty variable.

In order to test different firm characteristics on ESG rating disagreement, we perform pooled panel data regressions with ESG uncertainty as the dependent variable and different firm characteristics and valuation metrics as explanatory variables. The standard errors are robust to heterogeneity (or non-normality), as addressed in section 3. In Table 2 we report the regression results for ESG rating disagreement for the

aggregated ESG rating, as well as the separate ESG-dimensions. We apply the following regression:

$$\sigma_{i,t} = \alpha_{i,t} + \gamma X_{i,t} + \lambda(D1_{i,t} * D2_{i,t})\, \varepsilon_{i,t} \qquad (6)$$

Where ESG uncertainty is denoted as $\sigma_{i,t}$, and the included firm specific effects are denoted as $X_{i,t}$ (described below table 3). We also included an interaction term for industry*year specific effects $D1_{i,t} * D2_{i,t}$

Table 3 - Determinants of ESG uncertainty

Dependent variable:	**ESG Rating Disagreement**			
Pillars	ESG	E	S	G
Regression	1	2	3	4
ROA	-0.005	-0.004	0.002	-0.011
	(-1.328)	(-1.035)	(0.556)	(-3.225)***
LEV	-0.005	-0.005	-0.002	-0.006
	(-1.311)	(-1.903)*	(-0.733)	(-2.217)**
BM	0.050	0.048	-0.077	0.058
	(2.180)**	(3.966)***	(-2.765)***	(1.837)*
ME	-0.007	-0.009	0.008	-0.015
	(-2.305)**	(-2.857)***	(2.459)**	(-5.112)***
VOL	0.012	-0.005	0.018	0.008
	(2.518)**	(-1.085)	(3.399)***	(1.448)
MOM	-0.004	-0.006	-0.000	0.003
	(-1.172)	(-2.355)***	(-0.118)	(0.719)
Industry*Year FE	YES	YES	YES	YES
Adjusted R^2	0.071	0.132	0.027	0.036

Note: Table 2 shows the results of the of the pooled panel regression for the aggregated- and separate ESG dimensions where the dependent ESG disagreement (StDev) is regressed on different firm characteristics: ROA (return on assets), Leverage, BM (book-to-market), ME (market Capitalization of equity), VOL (volatility) and MOM (momentum). Additionally, we include Industry*Year fixed effects (FE). Newey-West t-statistics with standard errors robust for potential heteroscedasticity or non-normality are reported in parathesis.

Findings

According to Table 3, we find that higher book-to-market (BM) leads to higher ESG rating disagreement. Secondly, higher market capitalization (ME) tends to decrease disagreement, except for the social dimension where we observe the opposite effect. Volatility (VOL) displays the same pattern, where higher volatility leads to higher ESG rating disagreement, except in the environmental- and governance dimensions which yields insignificant results. Further, higher leverage (LEV) tends to reduce disagreement, but it is only significant in the environmental and governance-dimensions. Lastly, return on assets (ROA) and momentum (MOM) are mostly insignificant with no clear relationship with the dependent variable ESG rating disagreement.

The different firm characteristics are also reasonable given the economic magnitude displayed in Table 3. For instance, in aggregated ESG, moving from a firm with average volatility to a firm with above average volatility will explain approximately 3.4 basis points of the ESG disagreement. Secondly, moving from a firm with average book-to-market to a firm above average book-to-market in the environmental dimension will explain 33.0 basis points of the ESG disagreement. Hence, given the firm characteristics and effect on the ESG disagreement, the findings appear intuitive and reasonable as argued in the explanation part below.

Potential explanations:

We find that some of the most important firm characteristics and accounting variables seem to affect the uncertainty of ESG ratings. From Table 3, firms with higher book-to-market value tend to have higher ESG rating disagreement in most dimensions. These findings are consistent with what Gibson at al. (2021) finds studying US firms. Since higher book-to-market firms have lower valuations, it is reasonable to argue that those firms accordingly have a higher expected return in the future and higher ESG uncertainty. Nevertheless, we will analyze this relationship more closely in the next section.

However, the book-to-market coefficient in the social dimension is negative significant, indicating a negative relationship between book-to-market and ESG uncertainty. Although our finding in the social dimension contradicts the findings from Gibson et al. (2021), firms with higher book-to-market are considered to be in a more mature phase with limited growth opportunities. Hence, it is reasonable to argue that companies in a more stable phase have more resources available to improve other aspects than pure financial performance, like for instance the social aspect as well as improve transparency which will reduce rating disagreement and potentially explain the observed result in the social dimension.

Next, higher market capitalization reduces ESG uncertainty, except for the social dimensions, which is somewhat surprising compared to findings in Gibson et al. (2021). In one way it is intuitive that larger and more complex firms are analyzed more thoroughly by ESG data providers and therefore higher disagreement, as discussed in Gibson et al. (2021). On the other hand, larger companies are often easier to monitor as transparency is better compared to smaller companies, which could explain our findings in Table 3. According to theory European firms are typically characterized with concentrated control and strong ownership, which gives monitoring incentives (Goergen, 2018).

Moreover, we also find that increased volatility leads to higher rating disagreement, especially in the aggregated ESG and social dimension. In general, it is reasonable to argue that firms exposed to higher volatility in prices and unstable financial performance may not have as many resources to disclose and focus on ESG related activities that can reduce transparency and increase rating disagreement. Kim & Koo. (2023) finds a positive correlation between idiosyncratic volatility in a firm's stock returns and ESG disagreement, suggesting uncertainty increases information asymmetry. Hence, this finding supports our argument.

Further, higher leverage tends to result in lower ESG rating disagreement, with significant results in the environmental and governance-dimension. A rational explanation for these findings is that firms with higher leverage have a higher share of creditors on the ownership side. Accordingly, as the capital provided is at risk, a

subsequent higher degree of monitoring and engagement from the creditors is expected, as debt can be seen as a disciplining mechanism according to theory (Goergen, 2018). This could increase the transparency among the stakeholders and outwardly, hence reduce the ESG rating disagreement.

Lastly, we find a negative significant relationship between ROA and ESG uncertainty in the governance dimension. ROA yields mostly negative coefficient and is only significant in the governance dimension. It is reasonable to argue that firms with higher ROA have more cash available to spend on development of for instance new routines and disclosure as discussed in Gibson et al. (2021). Moreover, Michelon & Parbonetti. (2012) finds that good corporate governance is associated with better disclosure of sustainability. Hence, it's reasonable to argue that good governance routines can result in more transparency that affects rating disagreement.

4.1.2 Country and industry specific ESG uncertainty

In this subsection we will briefly look at how ESG disagreement varies across countries and industries. We apply the following regression with dummies for country or years:

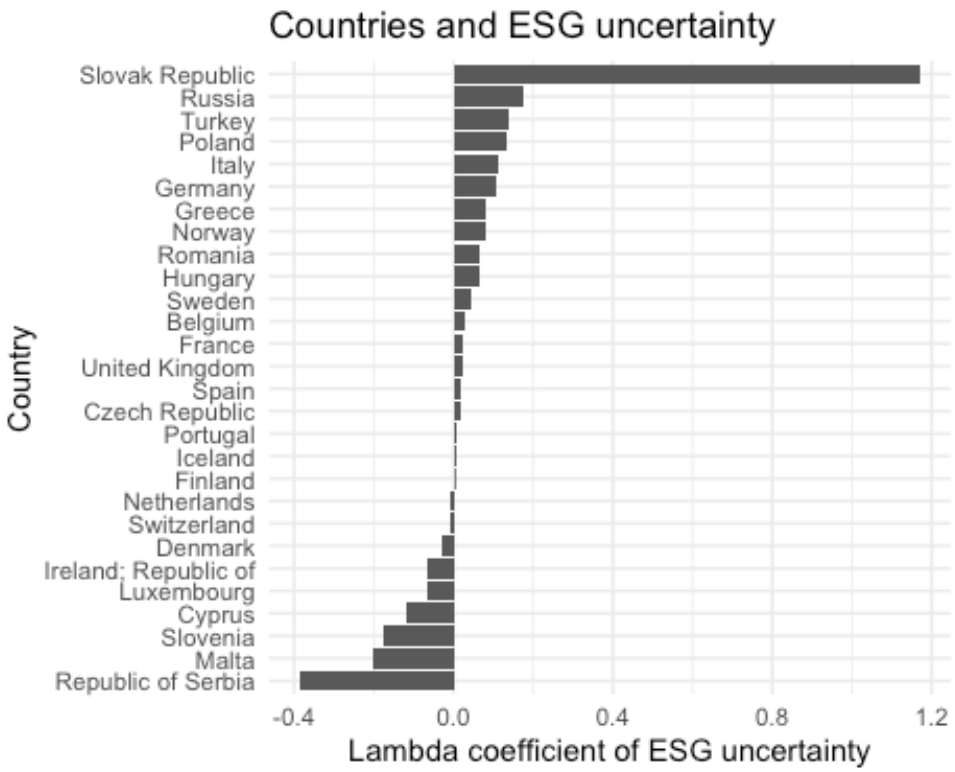
$$\sigma_{i,t} = \alpha_{i,t} + \lambda(D_{i,t})\, \mathcal{E}_{i,t} \tag{7}$$

Where $\sigma_{i,t}$ is the ESG uncertainty and $D_{i,t}$ represents the dummies for countries or industries.

According to Graph 1, we observe that eastern countries as Russia and Poland have greater ESG disagreement compared to countries as Finland, Netherlands, Denmark and Spain. In order to interpret the economical magnitude, we observe that for instance Denmark with a λ coefficient of approximately -0.02, yields a below average individual effect of ESG uncertainty by approximately 5.263 basis points (See section 3.5 for methodology). The result is expected as we believe it is reasonable to argue that more developing economies (for instance Russia) has a less developed ESG practice compared to a typical industrial country as Sweden. Accordingly, this is also supported by Singhania & Saini. (2021), who classifies countries as Norway,

Sweeden and Denmark as countries with well-developed ESG frameworks compared to countries with early-stage framework such as Russia.

Graph 1

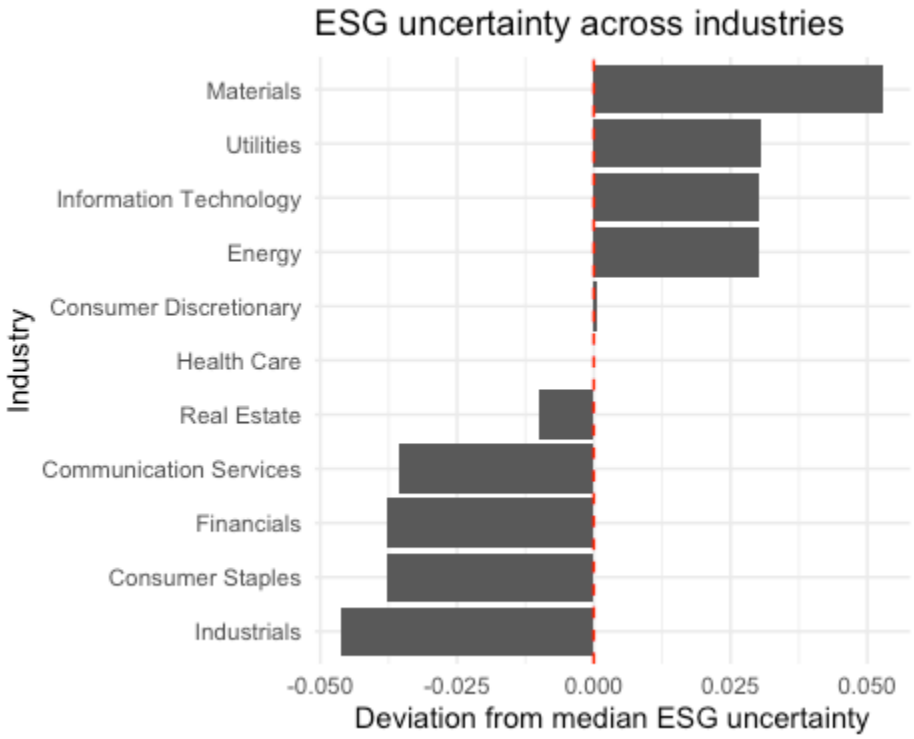


Note: The graph displays the country specific effects on ESG uncertainty in the aggregated ESG dimension. Notice that Slovak Republic and Serbia has few companies represented in our dataset and might causes the beta coefficient to be skewed, hence we find it most intuitive to focus on the other countries more towards the center of the graph when interpreting the results.

In Graph 2 we observe that the industries like materials, information technology and utilities, have greater ESG rating disagreement compared to industrials and consumer staples. It is reasonable to argue that companies in materials for instance have higher ESG uncertainty as companies within high emitting industries are differentiating in disclosure practice. Berg et al. (2022) also finds a low correlation within water and hazardous Materials. Despite the high disagreement it could be large fluctuations due to the industry categories wideness. often face greater ESG disagreement as a consequence of market volatility and fast changing environment. It is also intuitive

that the energy sector is in rapid change due to the green shift, and there will be huge differences in use of metrices and disclosure methodology. This argument can be linked to our evidence of volatility and ESG uncertainty in section 4.1.1. In contrast, one explanation why industries like financials and consumer staples have less disagreement among ESG providers may be due to relatively stable business model and lower stock volatility, also discussed in section 4.1.1.

Graph 2



Notice: We have chosen the median ESG uncertainty industry "Health care" as the baseline. The graph displays the deviation of ESG uncertainty of the other industries.

4.2 ESG rating disagreement and stock returns

In this subsection we will focus on our main research question and examine whether ESG uncertainty can predict future stock returns. We perform a pooled panel regressions with robust standard errors. We use yearly excess stock returns at time t as the dependent variable in the regressions. As our main variable of interest, we use the disagreement at time t-1 between the providers each year which are measured through standard deviation, denoted as StDev.

Secondly, since the disagreement is a product of the ESG ratings reported by the providers, we find it interesting to control for average ESG score, denoted as Avg_ESG. Additionally, the ESG agenda has been quite prominent in recent years, especially in the EU after the Paris-agreement in 2015 with active regulators and substantial yearly increase of green bond financing (European Parlament, 2022).

We also control for industry*year and country*year fixed effects. As our sample is based on the European continent, it includes around 30 different countries. Accordingly, we observe a substantial level of heterogeneity (Graph 3 in the appendix), across countries due to for instance different regulations or economic conditions. On the same note, the different companies are also operating in different industries where for instance the legal framework and firm size and characteristics varies, which we illustrated in previous section. We control for heterogeneity by not only considering the cross-sectional differences in ESG uncertainty, but also including interaction terms for industry and year, as well as country and year, as fixed effects. This approach allows us to account for differences and trends across industries and countries over time.

Finally, we perform the following pooled regression model:

$$R_{i,t} = \alpha_{i,t} + \lambda_1 StDev_{i,t-1} + \lambda_2 Avg\_ESG_{i,t-1} + \lambda_3 (D1_{i,t} * D2_{i,t}) + \mathcal{E}_{i,t} \qquad (8)$$

Where Excess returns is denoted as $R_{i,t}$, the lagged ESG uncertainty is denoted as $StDev_{i,t-1}$, and the lagged average ESG score is denoted as $Avg\_ESG_{i,t-1}$. In addition, we include the interaction term with dummy variables for industry/country times year. Notice, that in order to avoid look-ahead bias we have lagged both the ESG uncertainty variable (StDev) and the average ESG variable (Avg_ESG).

Table 4

Dependent variable	**Excess returns**					
Regressions	1	2	3	4	5	6
Panel A: Aggregated ESG						
StDev	0.020	0.014	0.028	0.023	0.025	0.020
	(1.720)*	(1.222)	(2.731)***	(2.214)**	(2.252)**	(2.055)**
Avg_ESG		-0.029		-0.024		-0.028
		(-6.408)***		(-6.107)***		(-7.018)***
Adjusted R^2	0.000	0.003	0.298	0.300	0.312	0.315
Panel B: Environmental						
StDev	-0.01	-0.027	0.002	0.011	-0.006	0.005
	(-0.984)	(-0.246)	(0.251)	(1.168)	(-0.651)	(0.589)
Avg_ESG		-0.017		-0.021		-0.027
		(-3.509)***		(-4.997)***		(-6.427)***
Adjusted R^2	0.000	0.000	0.300	0.302	0.314	0.316
Panel C: Social						
Standard deviation	-0.003	0.003	-0.007	-0.001	0.001	0.008
	(-0.239)	(0.263)	(-0.757)	(-0.117)	(0.142)	(0.869)
Avg_ESG		-0.019		-0.020		-0.024
		(-4.098)***		(-4.971)***		(-5.815)***
Adjusted R^2	0.000	0.001	0.303	0.304	0.317	0.319
Panel D: Governance						
StDev	0.005	-0.007	0.000	-0.008	0.010	0.000
	(0.575)	(-0.726)	(0.056)	(-0.997)	(1.200)	(0.113)
Avg_ESG		-0.029		-0.022		-0.026
		(-5.613)***		(-4.793)***		(-5.694)***
Adjusted R^2	0.000	0.002	0.310	0.302	0.317	0.318
Industry*Year FE	NO	NO	YES	YES	NO	NO
Country*Year FE	NO	NO	NO	NO	YES	YES

Note: Panel A to D displays the results of the pooled panel regression of yearly excess return at time t regressed on ESG uncertainty t-1 (StDev) for aggregated ESG as well as the separate dimensions (E, S and G) at time t-1. We also include interaction terms for industry-year fixed as well as a country-year fixed effects. Newey-West t-statistics are reported in parenthesis and standard errors are robust for heteroscedasticity.

4.2.1 Findings

From Table 3 Panel A, the ESG disagreement (denoted as StDev) in regression one has a coefficient of 2.0 basis points (with a t-statistic of 1.720). Further, when including average ESG (denoted as Avg_ESG) the ESG disagreement coefficient

becomes insignificant and drops to 1.4 basis points (with t-statistic of 1.222). However, when controlling for country- and industry-fixed effects, the ESG disagreement coefficients deliver significant results ranging between 2.0 and 2.8 basis points (with t-statistics varying between 2.055 and 2.731), indicating a positive relationship. For the average ESG, the coefficient is negative significant with alphas ranging between -2.9 and -2.4 basis points with (t-statistic ranging from -7.018 and -6.107), indicating a strong negative relationship.

The separate dimensions (Panel B to D) report ambiguous and insignificant findings regarding the ESG disagreement. The coefficients in the three separate dimensions varies between -1.0 to 1.1 basis points (with t-statistics ranging between -0.984 and 1.168). Hence, the separate dimensions do not display a clear relationship between ESG disagreement and excess return. On the contrary, the average ESG coefficient is also negative significant in the separate dimensions varying between -2.9 to -1.7 basis points (with t-statistics ranging between -5.694 and -4.098).

4.2.2 Potential explanations

From Table 5, the pattern for the different explanatory variables ESG disagreement (StDev) and average ESG (Avg_ESG) varies. As stated in Panel A, ESG rating disagreement is mostly significant in aggregated ESG, with no significance in the separate dimensions. Comparing the different E, S and G dimensions, it is the environmental dimension that drives the aggregated ESG coefficient the most, although it is insignificant. The average ESG coefficient has negative significant coefficients in Table 5, both for the aggregated ESG and the separate dimensions, which reiterate the strong negative relationship between average ESG and excess returns.

Further, when including explanatory variables and control variables, the variable ESG uncertainty in aggregated ESG in Panel A remains significant. From Panel A it is when only average ESG is included that the coefficient yields insignificance. Hence, this indicates that the average ESG score explains most of the positive relationship that the disagreement has on excess returns. However, the ESG rating disagreement variable tend to have a positive relationship with our dependent variable excess

returns. Therefore, our findings align with our preliminary expectations and previous research by Avramov et al. (2022) and Gibson et al. (2021) that finds a positive link between ESG disagreement and stock returns.

A potential explanation for the positive relationship in the aggregated ESG in Panel A between ESG disagreement and excess return could be risk theory. As the disagreement between ESG providers may be perceived as a source of uncertainty in the spirit of Knightian risk as discussed by Gibson et al. (2021), it could trigger a risk premium. As risk averse investors would avoid or demand a risk premium for holding firms that are exposed to the uncertainty of ESG disagreement, it could potentially explain the positive relationship we find in Panel A.

Nonetheless, although there is a positive significant relationship in aggregated ESG, the result from the separate dimensions is more ambiguous. Even though the findings are insignificant, we find it surprising that the ESG disagreement variable display a negative relationship with excess return in certain regressions in the separate dimensions. Accordingly, our findings in the separate dimensions seem to be more in line with evidence from Billio et al. 2021 who finds that ESG disagreement has no impact on financial performance.

A potential explanation for the absent positive relationship between ESG disagreement in the separate dimensions and excess returns could be twofold. Firstly, European investors might lack trust in ESG providers' ability to assess firms on a separate dimension basis, due to the low correlations reported by Gibson at al. (2021) and Berg at al. (2022) as well as our findings. Secondly, it is time-consuming for the investors to gather and analyze the ESG data reported, especially in the separate dimensions. Hence, when investors can utilize the metric aggregated ESG score with the highest consensus from ESG providers at the same time as it is supposed to reflect the overall ESG performance of a firm, it is reasonable to argue that investors are mainly leaning on the aggregated ESG dimension, explaining the significance.

Lastly, the average ESG which is consistently negative and significant in Table 5 indicates that increased ESG performance reduces returns. A theoretical explanation

for the highly negative relationship between average ESG and excess returns is discussed in the paper from Pastor et al. (2021). The paper argues that higher ESG performance should be associated with lower returns given two rational types of non-ESG and ESG investors. More precisely, the effect of a significant portion of ESG investors holding green stocks lowers the cost of capital and the premium of holding brown stocks will be subsequently higher. Pedersen et al. (2021) argues that the cost of capital related to green stocks would be higher if ESG investors holds a large wealth fraction, and accordingly higher expected returns for brown stocks.

4.3 Factor mimicking portfolios for ESG uncertainty

In previous section we found that there might be some predictability between ESG uncertainty and stock returns. We also found a strong negative relationship between returns and average ESG score. Hence, we want to investigate this evidence further by performing portfolio sorts and cross-sectional regressions.

4.3.1 Univariate portfolio sort

In order to gain a deeper understanding of the development of the mean excess returns we divided the five portfolios into quintiles based on ESG uncertainty. As Table 6 displays, we show the respective mean excess returns, sharp ratio and t-statistics for the five portfolios.

Table 6

Panel A: Aggregated						
Portfolios	Low	P2	P3	P4	High	High-Low
Mean excess returns	0.094	0.121	0.097	0.090	0.130	0.036
Sharp ratio	0.456	0.486	0.416	0.415	0.573	0.506
T-statistic	1.935*	2.062*	1.765	1.761	2.431**	2.147**
Panel B: Enviormental						
Portfolios	Low	P2	P3	P4	High	High-Low
Mean excess returns	0.127	0.077	0.091	0.101	0.102	-0.025
Sharp ratio	0.528	0.346	0.436	0.482	0.488	-0.279
T-statistic	2.240**	1.468	1.849*	2.045**	2.070**	-1.184
Panel C: Social						
Portfolios	Low	P2	P3	P4	High	High-Low
Mean excess returns	0.100	0.105	0.106	0.109	0.118	0.018
Sharp ratio	0.447	0.469	0.472	0.501	0.543	0.185
T-statistic	1.896*	1.989*	2.003**	2.126**	2.304**	0.785
Panel A: Governance						
Portfolios	Low	P2	P3	P4	High	High-Low
Mean excess returns	0.095	0.097	0.111	0.123	0.107	0.012
Sharp ratio	0.444	0.435	0.499	0.535	0.472	0.152
T-statistic	1.884*	1.846*	2.117*	2.270*	2.003*	0.645

Note: Table 6 shows the mean excess returns, sharp ratio and t-statistics for ESG uncertainty portfolis. In the aggregated ESG portfolio Low ESG uncertainty is denoted as Low and high uncertainty is denoted as High.

Findings

According to Table 6, we found that the aggregated high minus low portfolio yields a positive and significant value of 3.6 basis points. In addition, we found a sharp ratio of 0.506, which we consider to be quite good given the time-sample. Furthermore, we found that the long-short portfolio in the environmental-dimension yields a negative value of –2.5 basis points with an insignificant t-stat. For the social and governance pillar we found low and insignificant returns for the long-short portfolios.

Potential explanations

Our findings related to the aggregated ESG score are expected given our findings in section 4.2, where we found a significant premium by performing a long-short portfolio. Hence, we expect as discussed in Gibson et al. (2021) that the uncertainty related to disagreement represents a risk compensated by an increased risk premium. Given our findings so far, we see a tendency of predictability between ESG

uncertainty and stock returns in the aggregated dimension. On the other hand, we find the development of the mean excess returns in the aggregated ESG portfolios a bit surprising as we expected returns to be increasing for each level of ESG uncertainty (Panel A and B). Even though we observe an increase from the low portfolio to the portfolio with the next lowest ESG uncertainty, there is a slight decrease in mean excess returns until the highest uncertainty portfolio. A rational explanation can be related to the respective average ESG-scores, where we see a tendency of the portfolios with higher average ESG having lower average returns as well. For instance, the high portfolio yields an average ESG rating of 44 while the low portfolio yields an average ESG of 47. Moreover, the insignificant results for the separate dimensions E, S and G are not surprising given previous findings from pooled OLS.

4.3.2 Cross-sectional regressions

So far, we have obtained significant results related to the mean excess returns in the high-low portfolio in the aggregated ESG dimension. In order to investigate this further we performed the following general cross-sectional regression (Fama-MacBeth) for ESG uncertainty and average ESG score:

$$R_{i,t} = \alpha_{i,t} + \lambda^{\sigma}{}_{i,t}\sigma_{i,t-1} + \lambda^{\mu}{}_{i,t}\mu_{i,t-1} + \gamma_t + \mathcal{E}_{i,t} \quad (9)$$

Where the dependent variable excess return denoted as $R_{i,t}$ for stock i in year t, is explained by the lagged ESG uncertainty variable $\sigma_{i,t-1}$, as well as the lagged average ESG score variable $\mu_{i,t-1}$. We control for industry fixed effects denoted as γ_t. We obtain estimated coefficients for the ESG uncertainty and average ESG variable each year and store the results in separate vectors. Notice, that we do not include year fixed effects in these cross-sectional regressions as the time variation is captured by the intercept each year.

We find it meaningful to perform cross-sectional regressions as we now also have more power in our analysis since the whole data sample is used. The results of mean excess returns, sharp ratio and t-statistics is displayed in Table 7. The variable of interest in Panel A to C is ESG uncertainty, while the variable of interest in Panel D is

average ESG. In order to stay consistent, we first run a univariate regression in Panel A, before we include control variables as average ESG in Panel B and add industry fixed as well in Panel C. In Panel D, we have controlled both for ESG uncertainty and industry fixed effects.

Table 7 - Cross-sectional regressions

Panel A: ESG Uncertainty Univariate portfolio				
Pillars	**ESG**	**E**	**S**	**G**
Mean excess	0.029	-0.012	0.022	0.005
Sharpe ratio	0.528	-0.169	0.231	0.136
T-statistic	2.234**	-0.716	0.980	0.578

Panel B: ESG Uncertainty controlled for average ESG				
Pillars	**ESG**	**E**	**S**	**G**
Mean excess	0.022	0.004	0.023	-0.001
Sharpe ratio	0.456	0.074	0.277	-0.018
T-statistic	1.936**	0.314	1.175	-0.080

Panel C: ESG uncertainty controlled for average ESG and Industry				
Pillars	**ESG**	**E**	**S**	**G**
Mean excess	0.040	-0.000	0.029	-0.015
Sharpe ratio	0.509	-0.002	0.239	-0.361
T-statistic	2.157**	-0.009	1.013	-1.532

Panel D: Average ESG controlled for ESG uncertainty and Industry				
Pillars	**ESG**	**E**	**S**	**G**
Mean excess	-0.010	-0.027	-0.018	-0.012
Sharpe ratio	-0.138	-0.548	-0.376	-0.201
T-statistic	-0.586	-2.325**	-1.594	-0.855

Findings

According to Panel A in Table 7 we find a significant excess return of 2.9 basis points with a significant t-stat of 2.234. Next, we introduced average ESG as a control variable as displayed in Panel B and obtained a weaker mean excess return for the uncertainty coefficient but still significant. In Panel C we introduced industry fixed effects, and we find a stronger significant premium compared to Panel B and slightly lower compared to Panel A. The results are in line with the pooled OLS and

portfolio sorts in previous section. Furthermore, we observe insignificant results for the separate dimensions in Panel A to C.

As we in section 4.2 and 4.3 found a negative relationship between average ESG score and excess returns, we found it interesting to perform the cross-sectional regression controlled for ESG uncertainty and industry fixed effects. We found negative but insignificant coefficients for the aggregate ESG score. In the environmental dimension we found negative and significant results for the environmental dimension with a negative coefficient of –2.7 basis points, and a significant t-statistic of –2.325.

Potential explanations

As expected, the results still show a positive relationship between ESG uncertainty and excess returns in the aggregated dimension. What we find surprising is the fact that we obtained insignificant coefficients for average ESG in the aggregate ESG dimension (Panel D). The reason why we may obtain different results compared to findings from section 4.2, is the fundamental difference between pooled OLS and cross-sectional regressions. More precisely, pooled OLS utilizes the whole data sample, with data points from both the past and the future in order to estimate the coefficient for the independent variable. On the other hand, the cross-sectional regressions only consider previous years independent variable coefficients. It is reasonable to argue that the latter method is a more proper way to test ESG predictability on stock returns, since investors does not have information one year ahead.

Moreover, we found significant and negative λ for average ESG in the environmental dimension, which is expected given our previous evidence. We find it intuitive that a higher average ESG score reduces excess returns as holding “brown” stocks is associated with a risk premium as supported by theory in Pastor et al. (2021). Another rational explanation can be related to evidence from Bolton & Kacperczyk (2021) which finds that companies with Co2 exposure have lower prices and accordingly higher expected returns. Indeed, Co2 can be seen as a systematic risk, which is

incorporated by investors. Our negative relationship can therefore be justified both from a theoretical and empirical point of view.

4.4 Control for known risk-factors

As displayed in the empirical findings from the factor mimicking portfolios in section 4.3, we found a positive relationship between ESG uncertainty and stock returns. Furthermore, we found a negative relationship between average ESG score in the environmental dimension and returns. We now want to test whether these results can be explained by known risk factors such as the CAPM, Fama-French 3 factor model and Fama-French 5 factor model.

From Table 8, we report alphas from different factors models (CAPM, Fama-French 3 factor model (FF3) and the Fama-French 5 factor model (FF5)) for the univariate and double-sort portfolios for ESG uncertainty, as well as the average ESG from the cross-sectional regression in Panel D. As we found significant results in the aggregated ESG dimension for ESG uncertainty and environmental dimension for average ESG, the focus in the rest of this paper will be related to those findings. Furthermore, we reported the information ratio from the trading strategy high minus low in the aggregated ESG dimension to assess the risk-adjusted performance of the portfolios.

Table 8

Univariate High-low portfolio

Panel A: Aggregated ESG

	CAPM	FF3	FF5
Alpha	0.032 (2.206)**	0.026 (1.894)*	0.027 (1.504)
Rm-Rf	0.068 (1.090)	-0.055 (-0.829)	-0.154 (-1.641)
HML		-0.050 (-0.354)	0.402 (1.585)
SMB		0.554 (2.518)	0.452 (2.080)**
RMW			0.104 (0.354)
CMA			-0.590 (-1.519)
IR	0.456	0.409	0.471

Table 8

Double sort High-low portfolio

Panel B: Aggregated ESG

	CAPM	FF3	FF5
Alpha	0.046	0.036	0.065
	(1.374)	(1.040)	(1.280)
Rm-Rf	0.338	0.175	0.084
	(2.482)**	(0.690)	(0.243)
HML		-0.050	0.258
		(-0.148)	(0.385)
SMB		0.776	0.608
		(0.968)	(0.808)
RMW			-0.433
			(-0.674)
CMA			-0.750
			(-0.841)
IR	0.288	0.240	0.444

Factor mimicing: Dependent variable ESG uncertainty

Panel C: Aggregated ESG

	CAPM	FF3	FF5
Alpha	0.039	0.033	0.035
	(2.212)**	(2.044)**	(1.398)
Risk premium	0.013	-0.078	-0.151
	(0.247)	(-0.651)	(-1.293)
HML		-0.025	0.225
		(-0.203)	(0.722)
SMB			0.355
			(1.490)
RMW			0.039
			(0.135)
CMA			-0.456
			(-1.036)
IR	0.456	0.447	0.504

Factor mimicing: Dependent variable average ESG

Panel D: Enviormental

	CAPM	FF3	FF5
Alpha	-0.020	-0.014	-0.017
	(-2.317)**	(-2.212)**	(-2.077)**
Risk premium	-0.111	-0.004	0.067
	(-2.576)***	(-0.112)	(1.679)*
HML		-0.010	-0.264
		(-0.170)	(-3.313)***
SMB		-0.490	-0.405
		(-4.421)***	(-4.050)***
RMW			-0.028
			(-0.249)
CMA			0.452
			(3.506)***
IR	-0.461	-0.425	-0.651

4.4.1 Findings

From Panel A in Table 8, the long-short for aggregated ESG dimension generates alphas ranging between 2.6 basis to 3.2 basis points (with t-statistics ranging between 1.504 and 2.206). The alpha after controlling for CAPM is significant with the highest statistical and economical magnitude. When regressing on the Fama-French 3 factor model, the alpha is only marginally significant, while Fama-French 5 factor model deliver an insignificant alpha. The information ratio is ranging between 0.456 to 0.471 and is considered to be quite good, but not superior (Informa, 2016). From the cross-sectional regression in Panel C, we obtain similar results as in Panel A. We observe that the generated alphas and information ratio from Panel C is marginally higher compared to Panel A. This is expected given the increased power in the cross-sectional regression (full sample).

By turning our attention towards average ESG, we find a significant alpha after controlling for Fama-French factors in the environmental dimension. The cross-sectional regression yields an alpha ranging from –2.0 to –1.7 basis points, significant on a 5%- significance level. Furthermore, we also observe an increasing information ratio, ranging from –0.651 and –0.425.

4.4.2 Potential explanations

Panel A with aggregated ESG score displays that the long-short portfolio only generates statistical alpha with CAPM and marginally significant alpha in the Fama-French 3 factor model. For the Fama-French 5 factor model, the long-short portfolio generates an insignificant alpha. This indicates that the long-short ESG uncertainty portfolio may not be a sufficient trading strategy, given that the alpha is explained by other risk-factors. Hence, we cannot conclude that there is predictability of stock returns by looking at ESG uncertainty, and therefore chose to turn our focus back to the average ESG score in the environmental dimension.

In line with the evidence from section 4.2 and 4.3, we found a strong relationship between average ESG and excess returns, now with respect to the environmental pillar. The finding is expected given previous research related to “brown stocks” as

discussed in Pastor et al. (2020) and findings from Bolton & Kacperczyk. (2021). It is also expected that the environmental dimension was the biggest driver given the increasing media coverage the last decade and substantial environmental risk (e.g., carbon tax). On the other hand, we did expect a negative relationship between aggregated ESG score as well. Nevertheless, we consider the effect in the environmental dimension to be in line with market expectations and investors incorporation of for instance carbon emissions as a systematic risk (Bolton & Kacperczyck, 2021).

4.5 Robustness checks

As our findings show throughout this thesis, the predictability of excess returns seems to be more explained by the determinant of ESG uncertainty, namely the average ESG score. As discussed earlier, an implication is that we have less data especially before 2010. Hence, we find it intuitive to run the cross-sectional regressions from previous section to test whether we obtain the same results from the subsample 2010-2022 as we believe the results can be biased due to lack of data before 2010. In order to be consistent, we will focus on testing the most important results from the cross-sectional regressions related to the aggregated ESG and environmental dimension in section 4.4.

In Table 9, we display the cross-sectional regressions for the sub-sample 2010-2022, where we tested the average ESG score and ESG uncertainty coefficient estimates on Fama-French 5 factor model. As the Average ESG score turned out to be insignificant in section 4.3, we found it interesting to also test whether this result holds as shown in Panel A.

Table 9 - Robustness checks

Panel A: Aggregated ESG	2010-2022		
Regression	1	2	3
	CAPM	FF3	FF5
Alpha	-0.019 (-2.603)***	-0.015 (-2.520)**	-0.018 (-3.005)***
Risk premium	-0.027 (-0.793)	0.025 (0.775)	0.049 (1.470)
HML		-0.243 (-2.781)***	-0.025 (-2.853)***
SMB		0.056 (1.264)	-0.116 (-1.329)
RMW			0.029 (0.394)
CMA			0.282 (2.580)***
IR	-0.715	-0.736	-1.03

Factor mimicing ESG uncertainty

Panel B: Aggregated ESG	2010-2022		
Regression	1	2	3
	CAPM	FF3	FF5
Alpha	0.026 (2.317)**	0.026 (2.548)**	-0.004 (-0.405)
Risk premium	-0.106 (-1.654)	-0.088 (-1.422)	-0.087 (-1.384)
HML		-0.032 (-0.332)	-0.022 (-0.263)
SMB		-0.082 (-0.376)	0.067 (0.372)
RMW			0.638 (4.463)***
CMA			0.282 (1.555)
IR	0.610	0.613	-0.106

As displayed in Panel A, we found a significant negative alpha after controlling for the Fama-French 5-factor model, which indicates that there is some predictability in the aggregated ESG dimension. Given the high information ratio, it is an indication that the results are strong as we now have at least three providers included from 2010 and all four from 2016, which we believe will strengthen the analysis. However, as we discussed earlier it is intuitive that the average ESG score does affect stock returns.

According to Panel B, we can still conclude that after controlling for the Fama-French 5 factor model, there does not seem to be any premium associated with ESG uncertainty. We observe that the alpha has a significant decrease when introducing the Fama-French 5 factor model. Moreover, the information ratio is also small with a value of -0.106. The significant decrease in the alpha and the low information ratio indicates lack of statical power as the effect is so extensive by including two more degrees of freedom.

4.6 Limitations and future research

Our thesis has some limitations and implications which are important to consider. Firstly, even though we have four rating providers, we have limited data before 2010. In addition, we only have all four providers after 2015. Hence, our data might not have the desired level of power, given that our main objective is to look at ESG disagreement between providers and how it affects stock returns. Indeed, it would have been beneficial with even more providers to obtain a better picture of the uncertainty in the European market. As investigation of this topic to our knowledge is more developed in the US, it would be beneficial for future research to collect more ESG data providers in order to obtain more power also in the years before 2010.

Moreover, it could be interesting in future research to perform analysis of predictability of stock returns by looking at how ESG uncertainty (and ESG score) affects the cost of capital directly. Finally, we recommend testing how the long-short strategy would be affected by the cost of trading and how this affects the premiums related to ESG.

5 Conclusion

The focus on sustainable investing and ESG have increased drastically, especially in the last decade. However, as previous literature from Berg et al. (2021) shows that ESG is a relatively heterogenous environment, as ESG ratings have remarkably low correlations compared to credit ratings. We contribute to the existing literature by studying not only the disagreement between rating providers, but also how company characteristics affect ESG uncertainty as well as its effect on financial performance from a European perspective.

Firstly, we find evidence that correlations between providers in Europe are similar compared to the US (Gibson et al., 2022), but there seems however to be higher agreement in the social and governance dimension which could be explained by higher consensus and more developed ESG practice (Arnell, 2023). Further, we find that different firm characteristics affect the ESG uncertainty differently, depending on the firm characteristic of interest. In the European context, we observe that firms with higher book-to-market and volatility tends to have higher ESG disagreement, while firms with high market capitalization and leverage yields the opposite effect.

For ESG uncertainty as a predictor for future stock returns, the findings are mixed. We find a tendency for predictability in the aggregated ESG score on future stock returns, but no significant premium when controlling for Fama-French 5 factor model. On the same note, we find no predictability for ESG uncertainty on returns in the separate dimensions, which is partly in line with Billio et al. (2021) and contradicts findings from Gibson et al. (2022). Accordingly, ESG uncertainty may not be a sufficient predictor for future returns in Europe.

Nonetheless, we find that average ESG score seems to be a better predictor of future stock returns. However, when controlling for known risk factors the significant premium only remains in the environmental dimension. Regardless, when performing robustness check in the time period 2010-2022, the average ESG turned out to be significant as well, indicating a stronger effect in recent years. Furthermore, we

observe that the environmental dimensions were the main driver. Given the previous findings from Bolton & Kacperczyk. (2021), combined with the potential materializing of environmental risk in the near future, indicates that investors especially regard weak performance in the environmental dimension as a sufficient ESG-risk and subsequent higher premium.

References

Avramov, D., Cheng, S., Lioui, A., Tarelli, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics,* 145(2), 642-664. https://doi.org/10.1016/j.jfineco.2021.09.009

Arnell, S. (2023, April 11). ESG Headwinds: Embraced in Europe, Under Fire In America. *Forbes.* https://www.forbes.com/sites/forbesfinancecouncil/2023/04/11/esg-headwinds-embraced-in-europe-under-fire-in-america/?sh=36d17c7271f2

Berg, F., Kölbel, J., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. MIT Sloan School of Management. http://dx.doi.org/10.2139/ssrn.3438533

Berg, F., Fabisik, K., Sautner, Z. (2021). Is History Repeating Itself? The (Un)predictable Past of ESG Ratings. *ECGI* http://dx.doi.org/10.2139/ssrn.3722087

Bolton, P., Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics,* 142(2), 517-549. https://doi.org/10.1016/j.jfineco.2021.05.008

Christensen, D.M., Serafeim, G., Sikochi. (2021). Anywhere, Why is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings.

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3793804

Fama, E.F & French, K.R. (2018). Choosing factors. *Jorunal of financial economics,* 128(2), 234-252. https://www.sciencedirect.com/science/article/abs/pii/S0304405X18300515

French, K. R. (2023). Data Library. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Gibson, R., Krueger, P., & Schmidt, P. (2021). ESG rating Disagreement and Stock Returns. (Working paper NO. 651/2020). http://dx.doi.org/10.2139/ssrn.3433728

Goergen, M. (2018). Corporate Governance: A Global Perspective. Cengage Learning EMEA.

Greenhouse Gas Protocol. (n.d.) *About us* https://ghgprotocol.org/about-us

Informa. (2016). *Information Ratio.* *https://financialintelligence.informa.com/*

Kim, R., & Koo, B. (2023). The impact of ESG rating disagreement on corporate value. *Journal of Derivatives and Quantitative Studies*. https://doi.org/10.1108/JDQS-01-2023-0001

Larcker, F., Pomorski, L., Tayan, B., Watts, E. (2022). ESG Ratings: A Compass Without Direction. *Rock Center for Corporate governance at Stanford University.* Working Paper Forthcoming. SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4179647

Lioui, A. (2018). Is ESG Risk Priced? SSRN.

https://ssrn.com/abstract=3285091

Michael, D. K., Xu, F., Sijing, W., & Jiarui I. Z. (2022) Does Voluntary ESG Reporting Resolve Disagreement among ESG Rating Agencies?, *European Accounting Review*, DOI: 10.1080/09638180.2022.2088588

https://www.tandfonline.com/doi/full/10.1080/09638180.2022.2088588?casa_token=OHr5RZycTgwAAAAA%3ANbIKyODySDajjXjfaCHhr8OunBYFAHftgj0OE16zioT0r9UpQQ-62B-JJ7jmmHoPgyQ5ljSY752b-A

Michelon, G., & Parbonetti. A. (2012). The effect of corporate governance on sustainability disclosure. *Journal of Management & Governance.* https://link.springer.com/article/10.1007/s10997-010-9160-3

Morningstar. (2020). Fakta om Morningstar Sustainability Rating. https://www.morningstar.no/no/news/202621/fakta-om-morningstar-sustainability-rating.aspx

Pástor, L., Stambaugh, R., Taylor, L. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics,* 142(2), 550-571. https://www.sciencedirect.com/science/article/pii/S0304405X20303512

Reber, B., Gold, A. & Gold, S. (2022). ESG Disclosure and Idiosyncratic Risk in Initial Public Offerings. *J Bus Ethics.* 179, 867–886. https://doi.org/10.1007/s10551-021-04847-8

Sagbakken, S. & Zhang, D. (2022). European sin Stocks. *Journal of asset management,* 23, 1-18. https://link.springer.com/article/10.1057/s41260-021-00247-9

Singhania, M., & Saini, N. (2021). Quantification of ESG Regulations: A Cross-Country Benchmarking Analysis. *The Journal of Business Perspective 26*(2), 163-171.

https://journals.sagepub.com/doi/abs/10.1177/09722629211054173?journalCode=visa

Smith, K. (2022, august 25). Greenwashing and ESG: What you need to know. *Forbes Advisor.* https://www.forbes.com/advisor/investing/greenwashing-esg/

Wright, W., Hamre, E. (2021). The problem with European stock markets. *New Financial (Rethinking capital markets)* https://newfinancial.org/wp-content/uploads/2021/03/2021.03-The-problem-with-European-stock-markets-New-Financial.pdf

Wolf, B. (2020, June 1). ESG gap widens: EU rules become more prescriptive as US proposals wait in the wing. *Thomsonreuters.* https://www.thomsonreuters.com/en-us/posts/investigation-fraud-and-risk/esg-gap-widens/

Zhongfei, C., & Guanxia X. (2022). ESG disclosure and financial performance: Moderating role of ESG investors, *International Review of Financial Analysis*, *83*(), 1057-5219. https://doi.org/10.1016/j.irfa.2022.102291

Appendix

Exhibit 1 - Additional information

$$Sharpe\ ratio = \frac{R_p - R_f}{\sigma_p}$$

$$Mean\ excess\ return = \left(\frac{1}{n}\right) * \sum \left(R_p - R_f\right)$$

Significance level

Significance in this paper is denoted as follows: 10% level, ** , 5% level, and *** 1% level.

Exhibit 2 – Table without Sustainalytics score after 2018

ESG uncertainty

Panel A: Aggregated ESG

Regression	1	2	3
	CAPM	FF3	FF5
Alpha	0.037 (2.150)**	0.035 (2.179)**	0.025 (0.095)
Risk premium	-0.004 (-0.077)	-0.082 (-1.047)	0.163 (-1.330)
HML		0.142 (0.827)	0.438 (1.337)
SMB		0.309 (1.123)	0.286 (1.166)
RMW			0.322 (0.922)
CMA			-0.370 (-0.813)
IR	0.479	0.475	0.349

Average ESG

Panel A: Aggregated ESG

Regression	1	2	3
	CAPM	FF3	FF5
Alpha	-0.020 (-2.448)**	-0.014 (-2.416)**	-0.018. (-2.222)**
Risk premium	-0.104 (-2.485)	0.005 (0.137)	0.072 (1.843)
HML		-0.023 (-0.477)	-0.266 (-3.447)***
SMB		-0.494 (-4.676)***	-0.410 (-4.221)
RMW			-0.010 (-0.085)
CMA			-0.435 (-3.468)***
IR	-0.475	-0.452	-0.710