ESG and the cross-section of expected return

Auteur : Cloostermans, Bryan
Promoteur(s) : Lambert, Marie
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences économiques, orientation générale, à finalité spécialisée en macroeconomics and finance
Année académique : 2021-2022
URI/URL : http://hdl.handle.net/2268.2/15987

ESG AND THE CROSS-SECTION OF EXPECTED RETURNS

Jury :
Supervisor :
Marie LAMBERT
Readers :
Pierrick CLERC
Henry-Jean Gathon

Master thesis by
Bryan CLOOSTERMANS
For a Master in Economics – Macroeconomics & Finance
Academic year 2021/2022

Contents

Acknowledgements

I would like to express my great gratitude to my promoter, Ms. Marie Lambert, for her guidance and support in both the academic and professional context. She continues to provide me with valuable advice and always pushed me to go beyond what I thought I could do.

I also would like to thank Mr. Pierrick Clerc and Mr. Henry-Jean Gathon for the time they will devote to the reading and the assessment of my work, as well as for their upcoming feedback.

Lastly, I would like to thank Corentin Werenne for his strong support during my studies at HEC and for the redaction of my master thesis. The frequent exchanges that we had allowed me to exchange on the issues encountered during my analysis. I believe the quality of this thesis was significantly improved thanks to his contribution and feedback.

1 Introduction

Environmental, Social, and Governance (ESG) criteria become increasingly used in the investment industry by both the retail investor and the institutional investor. While Environmental criteria assess the performance of a company in terms of environmental friendliness, Social criteria look at how a company manages its relationships with employees, suppliers, customers but also the community as a whole. Lastly, Governance criteria assess the rights of shareholders, a company's internal controls but also how the leadership team manages the company. The increased awareness of these non-financial data and sustainable investments comes from multiple sources. First, legal authorities begin to implement regulations around ESG. For instance, in Europe, investment firms received the recommendation from the European Bank Authority to implement ESG in their investment process. With the implementation of the EU Taxonomy, investment firms will have the obligation to disclose their ESG exposure and assess the ESG appetite of their clients. Secondly, investors and analysts put pressure on public firms to report this additional information, either because they believe sustainable investments can bring additional returns or simply for ethical reasons. Being currently employed as an equity analyst in the sell-side industry while writing this thesis, I can confirm that investors have a particular attraction for ESG issues and that a significant part of client requests are related to ESG and particularly the environmental impact of companies.

One of the main concerns that investors have with ESG implementation in investment portfolios is its impact on investment performance. Does investing in sustainable companies will lower investors' returns? Does investing in such companies will provide out-performance compared to a benchmark? These are typical questions of interest investment firms or retail investors have. These questions are legitimate, as one could expect that poor ESG companies could face new sort of risks not already included in current Asset Pricing Models. Some risks such as environmental risks or strike risks could increase the required rate of return of poor ESG companies and thus could command higher returns. On the other hand, one could expect that a sustainable company benefits from higher returns by increasing customers' satisfaction, and avoiding some types of risks, ...

To track the effects that such non-financial data have on returns, one has to incorporate ESG metrics in asset pricing and analyze the impact on the cross-section of returns. Numerous data providers, such as Refinitiv, Bloomberg, and MSCI, report ESG scores and pillar scores for publicly traded companies. As these scores are easily accessible with financial-data providers, they are widely used in the literature and will be the basis for the analysis performed in this thesis.

This thesis will be ordered as follow: in section 2, I will quickly review the current literature on Asset Pricing and ESG in Asset Pricing Models. I will clearly mention the possible discrepancy between professional research and academic research on the subject. Section 3 will present the data used in this thesis. In section 4, I will present the methodology used in this master thesis to assess the effect of ESG in the cross-section of expected returns. At the same time, I will describe the explanatory variables used as control variables and their implication. Next, section 5 will first present some additional descriptive statistics about the variables used in the model before proceeding into the analysis of the model and its results. Section 6 will be dedicated to some robustness checks to assess the reliability of the results. I will conclude in sections 7 and 8 with a large discussion on the results obtained, the limitations encountered with the literature on ESG and asset pricing models, and will conclude this master thesis with a big summary of the results obtained.

2 Literature review

The literature around asset pricing models is quite large. It first started with modern portfolio theory and the Capital Asset Pricing Model of Sharpe (1964) linking a stock's required rate of return to its systematic risk, i.e., its non-diversifiable risk or the market risk. Later, Fama and Fench (1992) showed that aside from systematic risk, size and book-to-market value also capture part of stock returns and could be used as a good predictor of stocks' performance. Nowadays, dozens of studies try, each year, to show the relevance of new factors in predicting stocks' returns. Harvey et al. (2015) found that, in 2012, 316 factors were identified as "significant" pricing factors. All these factors reflect risks associated with companies. Recently, many authors have tried to include Environmental, Social, and Governance risks to the list of priced risks.

Adding an ESG factor to asset pricing models is not counter-intuitive. To illustrate that, Friede et al. (2015) regroups the conclusions of more than 2000 empirical studies on the effect of ESG on Corporate Financial Performance (CFP) and found that a large majority of these papers find a positive or neutral effect of ESG on CFP. It means that companies investing in "ESG" either do not materially impact their financials (profitability, solvability, ...) or even improve their financials. On the other side, few studies find evidence of a negative impact of ESG investment on a company's financials. Even so, the authors end their discussion by informing the reader that investors are not able to harvest ESG alpha due to the implementation costs of ESG portfolios.

Something important to note is that we see some disagreement around this new potential risk factor. Particularly, we can see a discrepancy between the significance of such factors found in professional research written by investment professionals and academic research. Indeed, while the latter often conclude with neutral or negative results of ESG premium for holding good ESG stocks, investment professionals more often find neutral or positive results.

2.1 Professional Research

Being a subject of particular interest to asset managers and investment professionals, we can find many studies from the professional world around an ESG factor. These studies often argue for the out-performance of good ESG stocks, i.e., stocks outperforming the market in their ESG scores. Khan (2019) suggests potential investment value in ESG companies. This study is of particular interest as it addresses the materiality issue faced with ESG rating and the effect of corporate governance. Indeed, some environmental risks may be more important in specific industries than in others. Let's take for instance the social score around "Health Safety". This metric will be much more important for Construction companies than for Insurance companies or Engineering companies (which are often classified in the same category as Construction companies as "Construction Engineering"). The author also computes a modified governance score considering country-specific features about shareholders' rights, and thus creates a new composite ESG Score. While this method allows for new ESG scores to reflect better the reality of companies, ones could cast doubt around this method. Such scores being not available publicly, we could hardly assume that the market is pricing risks according to these new synthetic scores, a discussion we follow in section 7.

Melas et al. (2017) use ESG integration in the context of passive investors[1], active investors[2] and factor investors[3]. The authors show that ESG integration generally improves the risk-adjusted performance of stocks. In addition, the study explains that empirically: 1) larger firms tend to have better ESG scores; 2) European companies tend to have better ESG than Japan or US companies; and 3) ESG scores seem to be auto-correlated up to 36 months. Another interesting result of this study is the tradeoff shown by the authors for an increase in ESG exposition in one's portfolio. For example, they find that when increasing the ESG rating of a portfolio by 20% (compared to the benchmark) one could expect an increase in the volatility of the portfolio of around 20 basis points (and 50 bps for 30% improvement in ESG rating). This increased volatility can be explained by decreased diversification when investors increase ESG exposure to their portfolios. Nevertheless, this study focuses on realized return instead of expected return, as we will see later.

Similarly, Bennani et al. (2018) analyzed the impact of ESG investing in the context of active/passive management and factor investing. More interestingly, they found that the results depend on the time period. During the period between 2010 and 2013, they did not find any evidence of reward for investing in sustainable companies, while they found significant excess returns of good ESG stocks from 2014 to 2017. In the next section, some academics show that this is the time period bias that explains the improved performance found by professionals. Also, they show that the effect is more present in North America and Eurozone than in other, less developed, markets. Eventually, the authors note that even though they find a positive relationship between ESG scores and performance, infinitely increasing the ESG exposure in optimized portfolios can have detrimental effects on performance. The latter result can also be explained by a lack of diversification.

Madhavan et al. (2021) found strong relationships between fund alphas and factor ESG scores. In addition, they found that a high E (environmental score) tends to have high quality and momentum factor loading (i.e., these stocks are exposed to

[1]Investors want to capture broad equity risk premium cheaply and efficiently.
[2]Investors want to keep the flexibility to find alpha opportunities with active stock selection.
[3]Investors want to maintain high exposure to targeted factors.

quality and momentum risk factors). We can expect that the ESG out-performance can be linked to other factors and is thus not a ”standalone” effect.

Without going into the details of each paper, other investment firms address the ESG questions and have positive views, such as Stanley (2015). Fulton et al. (2012) studied numerous research done on the subject and concluded that, at the security level, ESG can provide higher returns. Lastly, the latest sustainable reality report of Morgan Stanley (2021) again supports the conclusion of previously stated authors.

Nevertheless, we still find some professional research arguing against the out-performance of good ESG companies. Wang and Sargis (2020) used a global factor model and a portfolio method similar to Fama and Fench (1992) and found that there was no significant out-performance of good ESG firms. However, the authors explain that an investor would not significantly underperform by choosing to hold high ESG score companies[4]. In addition, the authors found a moderate correlation between the ESG factor and size factor, meaning that good ESG companies tend to be larger firms, something already supported by their peers. The study shows no other important correlation among the following factors: Value, Profitability, and Investment, meaning that we could reject the hypothesis that ESG could be a proxy for quality (it contrasts with Madhavan et al. (2021)).

2.2 Academic Research

Academic research appears much less optimistic about the presence of a significant ESG factor. Naffa and Fain (2021) analyzed the effect of ESG integration following the Fama-Macbeth method and even addressed the endogeneity issue (which is rarely addressed by other authors) using a GMM-IV estimator (a generalized method of moments regression using instrument variable to correct for endogenous variables). The authors did not find any evidence of added value when adding an ESG factor along with the classic Fama-French 5 factors.

One important consideration is that numerous research papers (especially professional ones) focus on realized returns instead of expected returns. The distinction is quite important as an asset can, at the same time, outperform while having a lower expected return. Pastor et al. (2021) show the difference using the German bond market as an example. While we cannot observe expected returns for stocks, expected returns for bonds can be approximated by their Yield To Maturity (YTM). They specifically show that ”green” German bonds[5] showing the same characteristics as normal German bonds outperformed during their analysis time frame while having lower YTM. In other words, one investor buying the 2 bonds at the same

[4]Nevertheless, they found that investors investing only in US/Canada would slightly underperform.

[5]Bonds for which the proceeds will be used in selected green projects.

time would have higher returns on the green bonds (due to demand from investors pushing green bonds price up) while the YTM (yields received from an investor if he holds the bond until maturity and reinvests the coupons at the YTM) would be lower for the green bond, thus meaning underperformance of green bond if hold until maturity.

Pástor et al. (2021) show theoretically that Green firms should have lower expected returns than brown firms, while possibly outperforming brown firms in a certain time frame. Their model argues that to have a difference in returns for brown and green stocks, climate concern (i.e., the importance that an investor gives to climate issues) is essential and should be heterogeneous between investors. Without heterogeneous climate concerns, green firms would not be different from brown firms. Green firms can outperform thanks to increasing demand for these types of stocks, but this results in a lower cost of capital and expected return afterward. Cornell (2021) has the same reasoning and argues that premiums can only arise from 1) risks; 2)behavioral biases; or 3) market friction. Preference towards green firms can be thought of as a behavioral bias leading to out-performance in the short term. This preference leads to increased prices for responsible companies and thus higher realized returns compared to other stocks. He continues explaining that, due to investors preferences, the discount rate or cost of capital (which represents the required rate of return for holding the shares of a firm) will be lowered. Eventually, the return of these stocks will underperform in the long run, as the expected return will become the lowered required rate of return once at equilibrium. This observation is similar to conclusions obtained by Chava (2014) and El Ghoul et al. (2011).

Following the above discussion, Pastor et al. (2021) tested empirically their model and found that in both time-series and cross-section analysis, the expected return of green bonds is negative once controlled for "Shocks" affecting the returns of green bonds. They controlled for Climate concerns, Earnings shocks, and ESG flows. They found that climate concerns significantly explained the out-performance of green stocks. Once controlled for climate concerns, green stocks do not exhibit significant positive expected returns. Increasing climate concern drives demand for green stocks and results in short-term out-performance. Ardia et al. (2020) constructed the Climate Concern Index used by Pastor et al. (2021) and found similar conclusions. The out-performance shown in some studies is simply the result of the sample period. The same analyses, once the shocks are no longer present in the market, should show lower expected and realized returns for green stocks.

By back-testing the results of studies showing higher performance of good ESG stocks, Bruno et al. (2021) demonstrate a non-significant conclusion for an ESG premium. Instead, ESG out-performance seems to be linked with other factors, such as quality factors and sector biases. They observe the following results: First, out-performance holds when ignoring estimation risk only. Second, ESG strategies do not protect investors from downside risk. Third, similarly to the authors abovementioned and ESG out-performance is the result of sample-period and investors' preferences in the short run. In that way, the authors confronted the professional studies and pointed out the unreliability of these researches.

Recent empirical works even show negative ESG premium. Bolton and Kacperczyk (2020) and Bolton and Kacperczyk (2021) conclude that investors demand higher returns when investing in carbon-intensive stocks, thus resulting in a negative premium for good ESG firms. Hsu et al. (2020) found similar results. In the same way, an older study, Hong and Kacperczyk (2009) found that unethical stocks (sin stocks) such as alcohol, tobacco, or gaming stocks are less held by norm-constrained institutions (due to investment mandates prohibiting investment firms to hold sin stocks). Eventually, these stocks benefit from higher performance. This conclusion is quite intuitive. Good ESG companies represent a lower risk for an investor. The returns being positively related to risk, we can expect lower returns for responsible companies and higher returns for the bad ones. The complete opposite situation to what investment professionals (especially the ones selling ESG funds) try to prove. Nevertheless, contrary to the results found by Pastor, these analyses focus on ex-post returns instead of ex-ante.

While we observe more negative conclusions in the academic world than in the professional one, we still find some positive conclusions. Maiti (2021) finds out-performance of portfolios constructed on ESG and pillar factors while providing a better Sharpe ratio (on pillars). In addition, he concludes that a Fama-French method with 3 factors: Market, Size, and ESG performs better than the classical 3 factors: Market, Size, and Value. Kempf and Osthoff (2007) find a similar conclusion for long-short portfolios on sustainability scores. Pollard et al. (2018) also demonstrate that ESG not only increases the social responsibility of investors but is also an alpha generator. They present the ESG premia as a factor to use along with other mainstream factors. Clark et al. (2015) aggregated results from 41 studies and found that 80% showed a significant positive relationship between sustainability and returns. Eventually, Statman and Glushkov (2009) observe out-performance of tilted portfolios toward Best-in-Class social responsibility. Still, these studies focus on realized returns (ex-post) instead of expected returns (ex-ante).

To conclude, current literature seems to disagree on the subject. Specifically, numerous studies focus too much on ex-post results to provide a conclusion on the impact of ESG on investment. This thesis fits directly into the literature as it analyzes the impact of ESG on returns on an ex-ante basis. I will follow the methodology of Pastor et al. (2021) described above but will differ by using a European dataset.

An important remark concerns the data itself. I will address this point in more detail in section 7 when discussing the results obtained, however, it is important to note that the empirical results are very dependent on the data collected (ESG Scores) which differ from one data provider to another.

3 DATA

In order to assess the impact of ESG on stocks' returns, some measure of ESG has to be chosen. This thesis will use the ESG Scores and Pillars scores from Refinitiv. According to Refinitiv, their database covers *"over 70% of the global market cap, ..., with historical data back to 2002"* LSEG (2021). Refinitiv scoring covers more than 10,000 companies and more than 2,000 belongs to European markets. Through annual reports, company websites, CSR (Corporate Social Responsibility) reports, ... analysts score each company according to 450 ESG measures and all Refinitiv ESG scores are updated on a weekly basis. Nevertheless, we observe that ESG scores remain constant for 12 months straight, at least for the European dataset used in the thesis. Refinitiv also provides ESG controversy scores that include recent controversies faced by companies such as worker strikes, bribery, ... Note that this thesis will not use these metrics updated more frequently due to the lack of availability of companies with frequent controversy scores in Europe.

ESG Pillars scores are constructed on 11 subcategories: Resource Use, Emissions, Innovation, Workforce, Human rights, Community, Product responsibility, Management, Shareholders, and CSR strategy. A detailed figure (1) is available in section 8. ESG Scores from Refinitiv directly take into account materiality issues and assign more weight to relevant categories depending on the company's business. These weights are publicly available and will be used later as described in section 4.

The dataset used includes all companies (active and inactive[6]) quoted in European markets with at least 1 ESG score available during the last 5 years. The time frame of this study will consist of 11 years from January 2010 up to December 2020. As shown by figure 2, before 2018, less than 1,000 companies had an environmental "E" score (it will be addressed later in section 7) and few companies had scores updated after the 31 December 2020 at the time of collecting the data. Eventually, this dataset includes 2412 companies over 132 monthly periods. Table 1 presents the summary statistics of some variables used in the thesis.

To control for Climate concern, I will use the Media Climate Change Concern (MCCC) from Ardia et al. (2020). The data are publicly available on the Sentometrics Re-

[6]inactive companies include previous public companies now private, merger, bankrupt companies, ... to avoid survivorship bias

search's website[7]. This index measures the daily climate change concern from January 2003 through June 2018 using eight U.S. newspapers. Unfortunately, there exists no such index using European newspapers. Still, we can make the hypothesis that, with free access to information, the climate concern index would be similar using European Newspapers. For each article related to climate change, the authors compute a measure that depends on the total fraction of words related to risk and the difference between negative and positive words. Eventually, they aggregate the measure, correct it for heterogeneity across newspapers and take the square root of the measure to reflect decreasing marginal concern for an additional piece of news on the same day.

Following the methodology of Pastor et al. (2021) (PST henceforth), I will take $MCCC_t$ as the time average of the MCCC index in month t. We can fairly assume that news has an impact on climate concern not only contemporaneously, but also with a lag. Still following PST, I will compute a monthly climate concern variable C_t as:

$$C_t = \sum_{\tau=0}^{T} \rho^{\tau} MCCC_{t-\tau} \tag{1}$$

with τ measuring how long climate news persists in investors' memories. PST assumes that the half-life of a news story is equal to 1 year. Thus, it implies that $\tau = 0.94$[8]. In addition, they set $T = 36$ months as after 36 months the effect of more lags have small effects on C_t $(0.94^{36} \approx 0.1)$. I will follow the same assumptions as PST. Figures 3 represents the evolution of climate concern from 2010 to 2018. As expected the figure follows the same trend as PST, but interestingly we see a downward trend before 2013 that was not visible in PST's paper as they used a smaller timeframe. Summary statistics are available in table 1.

As part of the methodology, in the panel regression, we will analyze how green factor returns respond to unanticipated changes in climate concerns. The change in climate concerns is defined as $\Delta C_t = C_t - C_{t-1}$ or:

$$\Delta C_t = MCCC_t - \sum_{\tau=1}^{T} MCCC_{t-\tau} \rho^{\tau-1}(1-\tau) - \rho^T MCCC_{t-1-T} \tag{2}$$

This methodology is different from the one used by Ardia et al. (2020) that used the prediction errors from an AR(1) model. Pastor et al. (2021) justifies its equation by

[7] https://sentometrics-research.com

[8] $0.5 = \tau^{11}$

stating that they found similar results and found a 94% correlation with the AR(1) error series. I will follow PST's methodology throughout this thesis for its simplicity of use.

Table 1: Summary Statistics for the ESG score, pillars scores, monthly returns and climate concerns.
This table presents summary statistics for the dataset. Each month, the mean (Mean), standard deviation (SD), minimum (Min), 25th percentile (25%), median (Median), 75th percentile (75%) and maximum (Max) values of the cross-sectional distribution of each variable are calculated. The table presents the time-series means for each cross-sectional value. The column labeled n indicates the average number of stocks for which the given variable is available. ESG is the ESG score from Refinitiv. E, S and G are pillars scores, r_{t+1} is the one-month-ahead stock return and C_t represents the climate concern.

	Mean	SD	Min	25%	Median	75%	Max	n
ESG	51.14	20.79	1.12	35.64	51.84	67.57	93.86	1182
E	51.60	25.96	0.42	30.32	52.36	74.03	98.58	1114
S	53.33	23.92	0.58	35.09	54.07	72.96	97.98	1181
G	50.08	22.99	0.95	31.43	50.59	68.86	96.98	1182
r_{t+1}	0.010	0.121	-0.604	-0.040	0.004	0.052	2.392	2030
C_t	7.79	1.79	5.23	6.15	7.71	9.06	9.06	

Interestingly, we observe that in the dataset used in this thesis, Environmental "E" and Social "S" scores are highly positively correlated, as shown in figure 4. This significant correlation would mean that greener companies also tend to give more importance to their social impact. An important conclusion is that any analysis using synthetic scores that aims at quantifying the impact of both environmental and social scores on returns could be misleading, as the impact could be impacted by the multicollinearity of the two measures. While using both measures in the regression analysis would lead to unbiased estimators, the inference statistics would be degraded due to high multicollinearity. As one could expect, correlations between the governance pillar "G" and the two other pillars are significantly positive, but at a lower level. A high score in Governance does not automatically imply that a company outperforms in terms of social impact or environmental impact.

A part of the methodology that will be explained later in section 4 requires working with Market Adjusted Excess Return $\tilde{r}_t^e$ and is equal to:

$$\tilde{r}_t^e \equiv \tilde{r}_t - \beta_{m,t-1}\tilde{r}_{mt} \tag{3}$$

where $\tilde{r}_t$ and $\tilde{r}_{mt}$ are respectively a vector of monthly stock return in excess of the risk-free rate and the market return in excess of the risk-free rate. The 3-months

German Government Bond acts as a proxy for the risk-free rate, while the Stoxx600 acts as a proxy for the market return. $\beta_{m,t-1}$ is the vector of sensitivity of stock returns with respect to the market return. Following PST, I construct the beta from rolling monthly regressions using no more than 60 months but no less than 36 months of data ending in month t. Stocks for which less than 36 months are available at time t will thus be excluded from the computation of the Green factor in section 4. Summary statistics for computed betas and market-adjusted excess return are available in table 2.

Table 2: Summary Statistics for rolling monthly beta and market adjusted excess return.

	Mean	SD	Min	25%	Median	75%	Max	n
Beta	0.89	0.57	-5.19	0.53	0.84	1.21	4.33	1838
$\tilde{r}_t^e$	0.006	0.119	-0.599	-0.043	0.001	0.047	2.253	1838

4 Methodology

The analysis will be separated into 3 sub-analyses: 2 time-series and 1 cross-section (panel-data). First, I will analyze the return characteristics based on portfolio methods, similar to Fama French. The second time series analysis will be based on the theoretical Green Factor constructed by PST and will be tested empirically. Lastly, I will confirm the 2 previous analyses through panel data regressions. In each regression, we will observe the impact of ESG integration on the returns and analysis the drivers of performance. This section addresses the particularities of the model: the construction of the greenness score and the ESG factor.

4.1 Greenness score

The analyses rely on a stock's greenness measure from PST. First, I construct an unadjusted greenness score for each firm i at the beginning of the month t as:

$$G_{i,t-1} = -(100 - E_{i,t-1}) * Eweight_{i,t-1} \tag{4}$$

where E represents the environmental pillar score from Refinitiv. $Eweight$ represents the weight of the E pillar in the ESG score and is industry-dependent. Refinitiv lists each stock in one of the 53 industries, and each industry has a specific weight for each pillar. When E is not available at time t but an E score was available in the 12 last months, the latest value is used as a proxy (this was done to increase the amount of data used in the analysis and was also done by PST). The unadjusted greenness score is limited between 0 and -100 and the higher the score the higher the greenness of a firm.

The weight plays an important role in a firm's greenness as it captures the environmental impact of different industries. Let's take 2 companies (A and B) with respectively 50 and 45 as E score and 50% and 15% as $Eweight$. Their unadjusted greenness score will be -25 and -8.25 thus reflecting that company A has a larger environmental impact than company B. Numerous studies don't take into account the importance of environmental issues for specific companies, thus leading to biased conclusions.

The final greenness measure that I will use is

$$g_{i,t} = G_{i,t} - \bar{G}_t \tag{5}$$

where $G_{i,t}$ is the unadjusted greenness score from equation 4 and $\bar{G}_t$ is the time value-weighted average of the unadjusted greenness score at time t. Companies having an unadjusted greenness score larger than the time average will thus have a positive greenness score, while the ones with a lower unadjusted score than the average will have negative greenness. Summary statistics are available in table 3.

Table 3: Summary Statistics for the Greenness score.

	Mean	SD	Min	25%	Median	75%	Max	n
$g_{i,t}$	-0.45	8.29	-30.34	-5.61	1.09	5.94	12.12	1130

An important hypothesis is implied by this greenness measure: $\bar{G}_t$ represents the greenness of the market portfolio in the theoretical model of PST. Hence, $g_{i,t}$ measures the company's greenness score relative to the market portfolio[9]. According to PST, the difference in performance related to greenness can only happen if investors have heterogeneous preferences for companies' green impact. In a world where investors have homogeneous preferences, stock returns for green and brown stocks should not differ from the market portfolio. By construction, the product of weights by the greenness scores should be equal to 0 as imposed by PST:

$$w_t^{'} * g_t = 0 \tag{6}$$

[9]As data providers do not compute E score for every company, this assumption is not true empirically. Nevertheless, we make the hypothesis that unscored companies would not deflect the average score $\bar{G}_t$

where $w_t^{'}$ is the transposed vector of market capitalizations and g_t the vector of greenness scores at time t. Lastly, table 12 shows the average greenness score by industry on the 31st December 2020. Similarly to PST, less than 50% of the industries exhibit positive average greenness (14 industries). Among the greenest companies, we find financial firms and service companies while the brownest companies include precious metals and energy companies, also similar to PST findings.

4.2 Green Factor

PST construct theoretically their Green Factor as:

$$\hat{f}_g = \frac{g'\tilde{r}^e}{g'g} \tag{7}$$

and it is simply the slope from a cross-sectional regression of market-adjusted excess returns $\tilde{r}^e \equiv \tilde{r} - \beta_m \tilde{r}_m$ on the stocks' greenness without an intercept. To interpret time-serie of the green factor, one has to repeat this cross-section regression for each period. The Green factor becomes:

$$\hat{f}_{gt} = \frac{g_{t-1}^{'}\tilde{r}_t^e}{g_{t-1}^{'}g_{t-1}} \tag{8}$$

where $\tilde{r}_t^e \equiv \tilde{r}_t - \beta_{m,t-1}\tilde{r}_{mt}$. The interpretation is that the Green Factor is simply the greenness weighted average of market-adjusted stock returns instead of market-weighted, as usually done in Fama and French models (FFM). Positive g values indicate "green" stocks, as their greenness is superior to the mean greenness and belongs to the long leg. On the contrary, stocks with a negative value of g are considered "brown" stocks and belong to the short leg of the portfolio. The factor is different from what is usually done in the finance literature. While FFM are built empirically, PST construct their factor on a theoretical model where stocks are priced in equilibrium by the market portfolio as well as the green factor and is thus less arbitrary. Also, factors constructed on FFM are the difference between the returns of the long and short legs. Interestingly, while the Green Factor is not the difference in returns, it can be expressed as:

$$\hat{f}_{gt} = \hat{f}_{green,t} + \hat{f}_{brown,t} \tag{9}$$

Where $\hat{f}_{green,t} = \frac{g_{t-1}^{'+}\tilde{r}_t^{e+}}{g_{t-1}^{'}g_{t-1}}$ and $\hat{f}_{brown,t} = \frac{g_{t-1}^{'-}\tilde{r}_t^{e-}}{g_{t-1}^{'}g_{t-1}}$.

5 Analysis

As stated earlier, the analysis will be separated into 3 stages. Section 5.1 will cover the first time-series analysis following a methodology similar to Fama-French. Section 5.2 will address the second time-series methodology with PST's Green Factor. Lastly, section 5.3 will analyze the effect of ESG on a cross-section basis (panel data).

5.1 Green Minus Brown

I first start this section with the Green Minus Brown (GMB) factor (similar to FFM). The factor is the value-weighted difference of returns between the first and third tertiles. We observe that for our European dataset, the brown stocks seem to outperform the green stocks. Figure 5 shows the cumulative performance of the GMB. We see that ”brown” stocks outperformed the ”green” stocks between 2010 and 2020, especially during the first half of the decade. The average monthly return of the GMB portfolio was -0.26% as shown in table 4. Nevertheless, with a p-value of 0.4312, the average return of the portfolio is not significantly negative. It contrasts with PST results where, on their US dataset using MSCI, they found significant positive returns for ”green” stocks.

Table 4: GMB returns Summary Statistics

	Min	25%	Mean	Median	75%	Max	n
GMB	-0.0898	-0.0222	-0.0026	0.0007	0.01622	0.1262	131

Before entering into the decomposition of the green and brown returns for the GMB factor, it may be interesting to check for a possible change-in-mean during the 10 years time period. Indeed, one could expect a change-in-mean happening around 2015/2016 when looking at figure 5. A change-in-mean of the GMB factor could mean that the previous conclusion is biased i.e., unsignificant outperformance of brown stock may be statistically biased. To assess the possible change-in-mean, I will use the CUSUM test proposed by Wenger et al. (2018). One of the parameter d correspond to the estimated long-memory parameter, and we can expect this parameter to be quite low as shown by figure 6. To estimate the long-memory parameter, I used the local Whittle estimator by Robinson (1995) with a bandwidth of $T^{0.65}$ (T=131 here), which is usually used in the literature, and obtained a d parameter of 0.083. The CUSUM test eventually gives a p-value of 0.897 and tells us that the previous mean test was not impacted by a possible change-in-mean. To give an indication of the portfolio performance, we computed the Sharpe ratio of the GMB portfolio. With a Sharpe ratio of -0.071, the GMB portfolio underperforms the benchmark, which had an average monthly Sharpe ratio of 0.098.

Figure 7 represents the cumulative return of the top tertile (green stock) and bottom tertile (brown stock). We indeed observe that, except for 2012 and 2016, ”green” and ”brown” companies do not exhibit large differences in returns. Table 5 shows the returns characteristics of the 2 tertile portfolios. With a p-value of 0.0631 and 0.0155 for the green and brown stocks respectively, it seems that green stocks do not exhibit returns significantly different from 0 while brown stocks show statistically significant positive returns. Again, none of the portfolios shows significant evidence for a change-in-mean. In terms of the Sharpe ratio, the Green stocks had an average monthly Sharpe ratio of 0.21 while the brown stocks had an average Sharpe ratio of 0.28. The two portfolios independently outperformed the benchmark portfolio in terms of the Sharpe ratio.

Table 5: Green vs. Brown returns Summary Statistics

	Min	25%	Median	Mean	75%	Max	n
Green	-0.1450	-0.0196	0.0059	0.0079	0.0312	0.1697	131
Brown	-0.1575	-0.0189	0.0093	0.0105	0.0405	0.1500	131

PST identifies 3 different ”shocks” that could influence the distribution of returns of green and brown stocks: Climate concerns, Earnings, and capital flow. As stated previously, climate concerns from investors may impact investors’ preferences toward green stocks and thus lead to short-term outperformance. The second one is related to the change in earnings around the announcement date due to the higher preference of customers toward ESG products, thus driving earnings. Lastly, ESG flow represents the capital flow to ESG funds. On these 3 shocks, Climate Concerns seem the most important and almost always appear statistically significant. Earnings shocks appear significant at the stock level (cross-sectional regression) while capital flow does not appear significant in their dataset.

I first reduce the dataset from 2010 to June 2018 as the Climate Concerns index is no longer available after June 2018. After controlling for Climate Concerns, we observe in table 6 that, contrarily to PST, the shock does not appear significant at a 5% level or lower. The table regresses the monthly return of the GMB (1), the green portfolio (2), and the brown portfolio (3). Data for the 2 remaining shocks of PST not being readily available, they were not controlled for. We conclude that climate change does not have any impact on the portfolio returns in Europe. In section 6.1 assesses the robustness of the obtained in this section.

Lastly, PST regressed the GMB time-series on the Market return (Mkt-Rf), the SMB (Size) and HML (Value) factor of Fama and French (1993), the HMD (momentum) factor of Carhart (1997) and the RMW (profitability) and CMA (Investment) factors of Fama and French (2015). By doing that, they identified that the outperformance of their GMB portfolio was still significant after controlling for different investment

Table 6: GMB Factor | Climate Concerns Impact

The dependent variables are: the GMB factor (1), the market-hedged return of the top third green stocks (2), and the market-hedged of the bottom third green stocks. The standard errors are in parentheses.

	(1)	(2)	(3)
Constant	-0.0039	0.0047	0.0086
	(0.0038)	*(0.0045)*	*(0.0046)*
Climate Concerns (Same Month)	-0.0274	0.0289	0.0564
	(0.0273)	*(0.0324)*	*(0.0335)*
Climate Concerns (Lag)	0.0276	0.0334	0.0058
	(0.0276)	*(0.0324)*	*(0.0336)*
R^2	*0.0143*	*0.0330*	*0.0383*
Observations	*100*	*100*	*100*

factors. As shown in table 7 no coefficient appears significant and the GMB alpha never appears significant event after controlling for different factors.

5.2 Green Factor

We start this new section with the construction of the Green Factor following section 4.2. Figure 9 shows the relation between the Green Factor and the Climate Concerns index. Interestingly, over the period 2010 - 2020, the green factor had a robust significant average return of -0.0002 (p-value 0.0444) and no apparent change in mean during the period. This result slightly contrasts with the GMB factor but supports results from the robustness check concluding that, in Europe, green stocks tend to perform lower than brown stocks.

As explained earlier, we can easily decompose the green factor into two legs: the long green leg and the brown short leg (eq. 9). Interestingly, we find similar conclusions as in the GMB framework. The green leg shows a very unsignificant (both economically and statistically) return, while the brown leg appears significant at the 10% level. Still, we clearly see in figure 10 that the performance of the ESG factor is mainly driven by the brown leg. Similar to the GMB factor, the results are highly dependent on the breakpoint used. In this context, the entire sample was used, and the breakpoint was a greenness score of 0 following the PST methodology. Nevertheless, by digging into the distribution of the sample, it appears that

Table 7: GMB Performance

The dependent variable is the GMB factor. Independent variables are: the market excess return, the SMB factor, the HML factor, the UMD factor, the RMW factor, and the CMA factor. Data were recovered from the Kenneth R. French website and correspond to the European value of factors. Standard Errors are in parentheses

	(1)	(2)	(3)	(4)	(5)
Constant	-0.0026	-0.0022	-0.0031	-0.0036	-0.0027
	(0.0033)	*(0.0033)*	*(0.0034)*	*(0.0035)*	*(0.0035)*
Mkt-Rf		-0.0597	-0.0513	-0.0442	-0.0408
		(0.0642)	*(0.0752)*	*(0.0765)*	*(0.0821)*
SMB			0.2892	0.2903	0.2985
			(0.1953)	*(0.1959)*	*(0.2002)*
HML			-0.0416	0.0002	-0.1702
			(0.1440)	*(0.1641)*	*(0.2636)*
UMD				0.0670	
				(0.1252)	
RMW					-0.1804
					(0.3444)
CMA					0.1107
					(0.3671)
R^2	*0.0000*	*0.0067*	*0.0373*	0.0262	0.0268
Observations	*131*	*131*	*131*	*131*	*131*

some stocks frequently change from green to brown (and inversely) as their greenness score is close to 0. These stocks can influence the results obtained in that section. Indeed, while their weight is lower than ”true” green and brown stocks, 18.5% of the sample has a greenness score between -2 and 2 and can potentially interfere with the results. We will analyze this issue in section 6. The monthly Sharpe ratio of the green factor was -0.28 while the green leg scored -0.12 and the brown leg scored 0.23.

PST wanted to see if the recent underperformance of value stocks can be explained by the performance of green stocks. The reasoning for this question comes from

Table 8: Green Factor

	Green Factor	Green Leg	Brown Leg
Mean	-0.0002	0.0000	-0.0002
t-stat	*-2.0304*	*0.0571*	*-1.6614*
p-value	*0.0443*	*0.9546*	*0.0990*
Observations	*130*	*130*	*130*

the equilibrium setting of PST. In their framework, expected returns are modeled with a two-factor model that includes the market excess return and an ESG factor (the green factor). Nevertheless, the situation is quite different in Europe. While we can see an underperformance of value stocks between 2010 and 2020 in figure 11, the conclusion is not clear during the period analyzed by PST (2012-2020). In Europe, the average return appears unsignificant (robust t-stat of -0.9922) in that time period. Still, we will practice the same exercise as PST to see how the green factor influenced both value stock's returns and momentum stock's returns. First, concerning the momentum factor, we observe that the green factor does not have any predicting power. Indeed, the coefficient appears unsignificant, and the regression does not lead to increased R^2 or decreased alpha value. On the other side, the results for the value factor appear much more satisfying. But first, let's address the reflection we could have on the coefficient (9.82). While it seems, at first sight, to be a mistake, especially with a PST value of -0.803, the result is correct and simply comes from the environmental scores[10]. We observe that the market-adjusted monthly alpha of HML is 0.46 bps and is statistically significant (but maybe not economically significant). Taking the equilibrium setting of PST, we see that the HML's alpha becomes unsignificant while the Green factor appears to significantly explain the HML's returns. This is contrary to PST's conclusion but in line with what we found earlier.

We continue now with the effect of climate concerns on the Green Factor. PST found that an increase in climate concerns increased its green factor performance, even when controlling for other effects. It is intuitively correct to assume that climate concerns can have an impact on green stocks' returns. The higher the climate concerns, the higher the demand for products and services from green firms and the higher the demand to hold green stocks; both have positive effects

[10]PST uses MSCI score which is scaled on a 0 to 10 basis. We used Refinitiv scores, which are scaled on a 0 to 100 basis. This explains the 10-fold difference between PST and this paper. Returns (for Panel and Green factor) also have to be compared knowing this 10x difference. We can see that the average greenness g (table 12 from this paper ranges from 3.3 to -23.3 while PST ranges from 0.87 to -3.78

Table 9: GMB Performance

I estimate monthly time-series regressions of either Value stocks (HML from Fama French) or Momentum (UMD from Carhart four-factor model) on the excess market return and the green factor. Robust standard errors are in parentheses.

	Value		Momentum	
Constant	-0.0046	-0.0036	0.0098	0.0098
	(0.0022)	*(0.0029)*	*(0.0022)*	*(0.0025)*
Mkt-Rf	0.1559	0.1997	-0.1960	-0.1980
	(0.0573)	*(0.0715)*	*(0.0558)*	*(0.0669)*
Green Factor		9.8243		-0.1771
		(3.5738)		*(3.0289)*
Observations	*98*	*98*	*98*	*98*
R^2	0.0626	0.1827	0.0736	0.0743

on green stocks' performance. Table 10 shows the time-series monthly regression of the green factor and its legs on the change in climate concerns. None of the coefficients appears significant, which is again in opposition with PST findings.

Next, we analyze the impact of climate concerns on the green factor's negative alpha. To capture the green factor's alpha, we extract the returns net of its exposure to the Fama and French (1993) 3 factors model. The time-series regression of the green factor on Market return, HML, and SMB is available in the appendix (table 15). Interestingly, we observe that the green factor and its legs are negatively impacted by the market excess return. It is explained by the short position in brown stocks[11] that drives the negative relationship of the green factor. The green factor is negatively related to the SMB factor, meaning that the green factors seem more exposed to big capitalization. For the HML factor, nothing significant comes out of the regression. We can now extract the alphas net from Fama and French (1993) 3 factors. The results are available in the appendix in table 16 but again, the conclusion is similar to what we previously said. Nothing appears significant enough to drive conclusions.

[11]the brown leg represents a short position, meaning that a long position in the brown leg is positively related to the excess market return

Table 10: Decomposition of Green factor returns

I estimate monthly time-series regressions of the green factor (1), the green leg (2), and the brown leg (3) on the change in climate concerns from Ardia et al. (2020). Robust standard errors are in parentheses. None of the coefficients appears significant.

	(1)	(2)	(3)
Constant	-0.0039	-0.0000	-0.0002
	(0.0037)	*(-0.0000)*	*(0.0001)*
Climate Concerns (Same Month)	-0.0275	0.0005	0.0004
	(0.0273)	*(0.0006)*	*(0.0007)*
Climate Concerns (Lag)	0.0275	-0.0003	0.0003
	(0.0274)	*(0.0006)*	*(0.0007)*
R^2	0.0373	0.0008	0.0107
Observations	100	100	100

5.3 Panel Data

So far, all the empirical analyses were based on the time series of green-versus-brown portfolio returns. This section addresses the individual stock returns by using panel regression to show that the conclusions obtained in time series analysis remain acceptable.

Before entering the results of the panel regression analysis, it is important to mention the treatment performed on extreme values/outliers. There exist two techniques commonly used in empirical asset pricing to treat the outliers of a dataset: 1) Winsorization and 2) Truncation. Winsorization consists of setting the value of a variable that is above/below a certain threshold to the value of that threshold. The threshold is usually chosen very low, such as the percentile 1% or 0.5%. Truncation on the contrary considers the value of variables above/below the threshold to be missing. For instance, let's imagine that the 0.5% (and 0.995%) threshold of returns is -20% (and 35%). While winsorization would change all values below -20% (or above 35%) as -20% (or 35%), truncation would change all values as missing values (NA values) and would thus not be considered in panel regression. Following Bali et al. (2016), I decided to winsorize outliers returns from the dataset at the 0.5% (and 0.995%) level. The explanatory variables used in the panel regressions were constructed synthetically with no significant outliers, I thus did not threat the features of the models.

Table 11 presents the results from the panel regression. The dependent variable is the winsorized individual stock returns, while the explanatory variables consist of the greenness of stocks as well as interaction terms between greenness and the change in climate concern (ΔC_t). Before entering into the results, let's answer the question of the variability of observations. Indeed, we see observations going from over 260,000 to around 90,000 while PST's results have a large 150,000 observations when including ΔC_t. It can be explained by 2 things: first, the large 260,000 is explained by the larger time frame used when only greenness is considered (from 2010 to 2020 while PST's dataset begins in November 2012). Second, the observations present when including ΔC_t include much fewer stocks than PST's dataset. As shown in figure 2, the number of stocks with an E score available during the period covered by the climate concern index was around 800 and jumped to around 1,500 after 2018. Between 2010 and 2018, Refinitiv's dataset covered as much as 1/3 of stocks covered by MSCI's dataset.

Regression (1) of table 11 shows the panel regression of returns on the greenness score. The coefficient appears moderately significant, with a p-value of 0.064 using robust standard errors (p-value of 1.75e-05 using non-robust s-e). While it remains difficult to infer any conclusion on such a p-value level, we observe that it confirms previous results that stocks' greenness has an unsignificant (or negative) impact on stock returns. Once again, this conclusion is at odd with PST, who found a significant positive relationship between stocks' returns and greenness. In regression (2), I included an interaction term between the greenness score and the contemporary change in climate concerns. Interestingly, both coefficients appear significant with robust p-values for the greenness and interaction term of 0.037 and 0.02 respectively. With a greenness coefficient negative, it seems that the greener the stock, the lower the returns, leading thus to an ESG discount for green stocks and a premium for brown stocks, thus remunerating the higher risk. At the same time, the interaction term appears positive, thus leading to the conclusion that positive change in climate concerns indeed contributes to the performance of green stocks. This second conclusion is shared by PST. Regression (3) is similar to regression (2) but includes an interaction term with lagged ΔC_t instead of a contemporary ΔC_t. Results are quite similar to regression (2) but no longer significant at the 5% level (p-values of 0.06 and 0.08). Lastly, regression (4) includes both terms from regression (2) and (3). In this situation, we observe that the greenness score appears significantly negative (p-value 0.045) while the interaction term from (2) is significantly positive at the 10% level (p-value 0.058). The interaction term is unsignificant with a p-value of 0.5. This last regression tries to confirm the results from (2) of a green discount but with a positive impact of unexpected change in climate concern on green stocks.

Another contrast we see compared to PST's result is the larger significance of the contemporary interaction term instead of the lagged one. In section 7.3, PST analyze the role of firm size and found that smaller companies tend to react with a lag

Table 11: Greenness and individual Stocks Returns

This table shows the results from panel regressions in which the dependent variable is the winsorized individual stocks returns ($r_{i,t}$) for firm i in month t. $g_{i,t-1}$ corresponds to the stock's lagged greenness, ΔC_t is the month's t change in aggregate clime concerns from Ardia et al. (2020)'s index following equation 2. The sample begins in January 2010 and ends in June 2018 when the dependent variables include ΔC_t and in December 2020 when the dependent variables do not include ΔC_t . All regressions include month fixed effects, cluster by month, and use robust standard errors (in parentheses).

	(1)	(2)	(3)	(4)
$g_{i,t-1}$	-0.0001	-0.0002	-0.0002	-0.0002
	(6.72e-05)	*(8.08e-05)*	*(8.26e-05)*	*(8.14e-05)*
$g_{i,t-1} * \Delta C_t$		0.0012		0.0011
		(5.21e-04)		*(5.90e-04)*
$g_{i,t-1} * \Delta C_{t-1}$			0.0008	0.0004
			(4.89e-04)	*(5.32e-04)*
R^2	0.0000	0.0006	0.0004	0.0006
Observations	265,970	90,002	89,300	89,280

compared to larger firms. This can be explained by the proportion of large companies included in Refinitiv's sample, where less than 30% of the sample can be considered "small-cap firms" (< €2bn market cap), and around 35% of stocks can be considered big-capitalization (> €10bn market cap). With such a proportion of big and mid-cap, it is not surprising that stocks in the dataset appear to react to climate concerns within the month.

Overall, the panel regression analysis of individual stocks' returns gives similar conclusions as in the time series analysis context: in the European market, the green stocks do not tend to outperform brown stocks. Some evidence even leads to the outperformance of browner stocks, thus remunerating their higher risk. Nevertheless, the cross-section analysis shows that unexpected changes in climate concerns indeed increase green stocks' performance. This conclusion is similar to PST's findings, stating that the performance of green stocks is partly explained by changes in climate concern.

5.3.1 Industry and climate concerns

In this section, I will analyze the effect industry classification in relation to climate concerns has on the stock returns. I will use a different approach than PST and will use panel regression similar to the previous section but including some interaction terms between industries and the change in climate concerns ΔC_t. The results are available in appendix (section 8) in table 17.

Two different panel regressions are reported. The first one regress the winsorized individual returns on the 1) the greenness, 2) the interaction term between the greenness and the contemporary change in climate concerns, and 3) interaction terms between dummy industry variable and the change in climate concerns. The second one is similar to the first one, but without 1) and 2). Both regressions include automatically the dummy variables for the industry, but their coefficients were not reported in the table. Also, note that similar to the previous regressions, all regressions include month fixed effects, cluster by month, and use robust standard errors.

First, we see in regression (1) that adding the industry dummies as well as the interaction terms do not change the conclusion that we previously found: significant negative underperformance for greener stock and positive impact of unexpected change in climate concerns. At first sight, we do not have significant conclusions, only Banking Services, Biotechnology & Medical Research, Leisure Products, and Semiconductors & Semiconductor Equipment appear to be positively and significantly impacted (at the 5% level) by unexpected changes in climate concerns. No industry appears to have significant negative relation, and surprisingly Construction Materials also join the group of significant positive coefficients while this industry is known as a carbon-intensive industry. Nevertheless, I consider these results unreliable for one main reason: the redundancy of the industry characteristic. Indeed, by construction, the greenness factor already includes some industry-specific components in it. As a reminder, the greenness score was constructed with environmental scores weighted based on the industry. This weight included in the construction of the greenness score may thus already create some multicollinearity that may explain why some carbon-intensive industries do not show significant negative coefficients as intuitively expected. To confirm this intuition behind, I thus decided to run the same regression without the two first features that include the greenness score.

The second regression (available in appendix in table 17) shows very interesting results. Indeed, we find that the following industries' returns appear to be significantly negatively impacted (at the 5%) level by unexpected changes in climate concerns: Beverages, Chemicals, Coal, Collective Investments, Machinery, Equipment & Components, Metals & Mining, Oil & Gas, Professional Business Education, Professional & Commercial Services, and Textiles & Apparel. No coefficient is found to be significantly positive. We can exclude Professional & Business Education industry from the results as only 2 companies were included in the dataset between 2010 and 2018 thus leading to inconclusive results for this industry. These results are very interesting as they include industries for which we can intuitively consider brown industries, but are also confirmed with their lower greenness score in table 12. These industries' returns seem to be negatively impacted when climate concerns increase on a global level.

6 Robustness Check

6.1 Green Minus Brown

In this section, I assess the robustness of the results obtained in section 5.1, especially since the results were quite different from PST. Two main assumptions may impact the results obtained and can easily be adapted: the time frame and the quantiles of the portfolio. While PST used a dataset starting from the end of 2012, we used data starting in January 2010. Nevertheless, using it has no impact on the results obtained earlier. Figure 8 shows the cumulative return of the GMB portfolio, and it appears that the mean return of -0.0005 is not statistically significant (p-value: 0.9124). Also, there is no significant change in the mean (t-stat: 0.896). Again, Brown portfolio average return appears significant positive (0.0105 - p-value: 0.0155) while the green portfolio average return remains unsignificant (0.0079 - p-value: 0.0631). The time-series regression provides very similar results and is available in section 8 in table 13.

I constructed the portfolio with the top and bottom third of the dataset, thus following PST methodology. I will now change this hypothesis by taking the top and bottom decile and top and bottom quintile. Very interestingly, the results obtained with different percentile differs from the results of section 5.1. Both GMB portfolios have now statistically significant negative returns, meaning that brown stocks may have outperformed green stocks. The quintile portfolio has an average return of -0.0081 (p-value: 0.0394) while the decile portfolio's average return is -0.0155 (p-value: 0.0014) and no apparent change in the mean is present. By decomposing each portfolio with its long (green) and short leg (brown), we see unsignificant return for the green portfolio (Quintile: 0.0042 - p-value: 0.3088 | Decile: - 0.0011 - p-value: 0.8023) but significant positive return for the brown portfolio (Quintile: 0.0123 - p-value: 0.0120 | Decile: 0.0144 - p-value: 0.0040). This result is very interesting as we see that the greener the stock the lower (or unsignificant) the return, while the browner the stock the higher the return (and the more significant the return). The GMB methodology completely contradicts numerous research trying to promote the outperformance of Green stocks. However, there are still some limitations in the results obtained in this study, and they will be addressed in section 7. Concerning the time-series regressions, they do not exhibit significant results, suggesting that Climate Concerns do not play a role in the GMB Context. Results are available in section 8 in table 14.

6.2 Cross-section regression

While writing this master thesis, Ardia et al. (2020) updated their study (9 May 2022 version) to test the results obtained by PST. In the context of this new version, the authors published a new version of their climate concern index. This new version has now 10 different newspapers (New York Times, Washington Post, Los Angeles Times, Wall Street Journal, Houston Chronicle, Chicago Tribune, Arizona Republic, USA Today, New York Daily News, and New York Post), and 2 newswires (Associated Press Newswires, and Reuters News). Moreover, the topics used in the index also changed compared to the first index. One could wonder whether these modifications may have an impact on the results obtained in the study. For that purpose, I will analyze the impact of this updated index on the cross-section regression from section 5.3.

First, the shape and the trend of the new index are very similar to the first version of the climate concerns. Second, the conclusions appear very similar to the one found previously. In the contemporaneous version of the interaction term (column (2) of 11), the coefficient appears positive with a coefficient of 7.02e-04 (vs. 0.0012) and a p-value of 8.75e-09. The lag version of the interaction term also appear significant at 4.75e-04 (vs. 0.0008 in column 3 of table 11) and a p-value of 1.00e-04. When including both a lag and contemporaneous version of the interaction term, the results are again the same: the lagged version does not appear significant while the contemporaneous version is significant with a p-value of 8.32e-06 (coefficient of 5.96e-04 vs. 0.0011 in column (4)). Overall no change in the findings obtained earlier, I will thus not test the robustness of the 2 time-series analyses as this new index should have little to no impact on the results.

7 Discussion and limitations

In this section, I address the limitations of the results I obtained, especially since the results are quite different from Pastor et al. (2021). I will first address the differences in scores between data providers. Secondly, I will address the retrospective view of the data. Next, I will address the limitations of the companies and synthetic scores. Lastly, I will cover the limitation of the use of the Climate Concern Index.

Differences in scores from different data providers can be an issue. As Berg et al. (2019) show, two data providers do not necessarily agree on scores given to companies when they both cover the same company. Let's remind the reader that I used a Refinitiv dataset, while PST used an MSCI database for their paper. Berg et al. (2019) show that the divergences can come from 1) Measurement, 2) Scope, and 3) weight given to criteria. First, the paper notes that some indicators are data-provider-specific and Refinitiv is classified as the data provider using the most unclassified indicators (42 indicators). In addition, while some indicators should have identical results between the providers, the authors show some particular results with, for instance, a 59% correlation chairperson separation measurement. This sounds counterintuitive, as this indicator is public information and should be identical for all providers. Some indicators also appear negatively correlated between some providers. Table 2 of Berg et al. (2019) shows the correlation of ESG scores, Environmental scores, Social scores, and Governance scores between the 6 providers used in the study. Interestingly, the Refinitiv-MSCI ESG scores correlation is 38%, the lowest value and well below the average (51%). The correlation of the Environmental pillar (the main indicator used for the greenness score construction) is even lower at 23%, again the lowest value of the table and below the average of 53%. Table 8 of the study shows the source of rating divergence and the authors observed that the divergence between Refinitiv and MSCI came mainly from Scope (68%), and Measurement (38%). The weights did not produce any divergence (-7% contribution). The Scope divergence "*refers to the situation where ratings are based on different sets of attributes. One rating agency may include lobbying activities, while another might not, causing the two ratings to diverge*". The Measurement divergence "*refers to a situation where rating agencies measure the same attribute using different indicators. For example, a firm's labor practices could be evaluated on the basis of workforce turnover, or by the number of labor-related court cases taken against the firm*" Berg et al. (2019). Other studies show the divergence of ESG ranking, such as Dimson et al. (2020). These results may explain the differences obtained between this study and PST. Also, it shows the limitations of our findings that can be not only time-specific but data-provider-specific.

A second important limitation of the findings is the retrospective view used by data providers. Indeed, Refinitiv uses past data to create the ESG scores. It means that a company that, today, publishes its sustainability report for the first time with data from the 2 previous years will also have a score for the 2 previous years. It poses a problem of timing where we assume in our analysis that scores 2 years ago may have impacted returns 2 years ago while the data was simply not available at that time. In 2014, the E.U. published the Non-Financial Reporting Directive, or Directive 2014/95/EU[12] that requires listed companies to publish Non-Financial Results from the accounting period starting on 1st January 2017 (Article 4). The impact of the directive is directly visible in Figure 2 with the large peak of companies with an E score. Nevertheless, part of the data is a retrospective view of companies that started to report on non-financial performance. At the same time, the U.S.A. already required some companies to report greenhouse gas emissions from 2009. The state of California has even required GHG emissions reporting starting in 2006. The retrospective view of Environmental Scores is an important limitation of the findings of this study. These ESG scores offered by data providers may be of great utility to assess the social responsibility of companies, but seems to be of lesser utility as inputs for asset pricing models.

Another limitation lies in the number of companies in the dataset. As already mentioned earlier, the PST dataset approximately had around 2,500 companies in its dataset. The dataset used in this study had an average of 1,000 companies with the majority of the period with less than 1,000 companies and a maximum of over 2,000 companies at the end of the period. Nevertheless, the main explanatory variable used in the study: the Climate Concern Index from Ardia et al. (2020) lasted up to June 2018. As shown in Figure 2, it corresponds to the beginning of the increase in the number of companies with an E score (around 1,500). This means that only around 5.9% of the study used more than 1,000 companies when using the Climate Concern Index.

Regarding synthetic scores, these types of scores are often avoided in econometric studies as they introduce some issues. Indeed, while using synthetic scores may appear appealing to capture the effect of multiple indicators with having numerous explanatory variables, it brings its own issue which is the reduction of variability of variables. The larger the variability of variables, the more reliable the results will be. The Refinitiv belongs to the data providers that use the most different indicators in their measure. While it provides a broader view for ESG assessment of companies, these types of explanatory variables are not recommended for OLS regression. Using raw data is often recommended in econometrics (while not often possible due to the availability of economic data or due to the focus of the study) and would partly resolve the first limitation around divergence of ESG ranking. Indeed, Berg et al. (2019) states that ”*the researches should ideally work with raw data that can be independently verified*”.

[12] Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014L0095

The last identifiable limitation lies in the use of the MCCC (Climate Concerns) index from Ardia et al. (2020). The index was constructed based on 8 U.S. newspapers, namely: The Wall Street Journal, The New York Times, The Washington Post, The Los Angeles Times, The Chicago Tribune, USA Today, New York Daily News, and The New York Post. One may wonder if using U.S.-based newspapers reflect the climate concern of European investors. Indeed, it is widely believed that Europe was concerned about climate change way before the U.S. If this reveals to be true, then using Ardia's Climate Concern index may lead to biased results and create an endogeneity issue as the proxy appears to be imperfect.

8 Conclusion

It is time to conclude this thesis on the impact of ESG scores (specifically green scores) on stocks' performance. I analyzed the impact of such metrics by three different methodologies: two time-series and one cross-section. Overall, we observe that the conclusions of the three methods are quite similar while being at odd with PST's results.

We first discovered in the GMB methodology (similar to the Fama-French methodology) that, based on a European dataset with Refinitiv ESG scores, companies defined as green do not significantly outperform brown companies. The brown stock leg can even outperform the green stock leg. It is ad odds with PST and some other research that showed empirically that green stocks tend to outperform. Nevertheless, we noted that this can be the result of the dataset, which can be considered quite restrictive compared to data used by other research. While PST showed that the climate concern index impact the stocks' performance, it did have a significant impact on return on the GMB context.

The second time-series methodology, the green factor, had similar results to the GMB methodology. Let's remind the reader that the main difference between the GMB and the green factor methodology is the weight used in the factor construction. While GMB uses market capitalization as a weight measure, the green factor uses the stocks' greenness as a weight measure. Again, brown stocks tend to outperform (slightly significant results) and the climate concern index does not materially impact the stock performance.

I eventually assessed the previous results through a panel-data regression, thus using individual stocks characteristic to assess ESG score impact on performance. While the results appeared still at odd with PST (brown stocks significantly outperforming green stocks which seems to lead to a risk premium for browner companies), we observed that the climate concern index has a significant impact on green stock performance. The larger the unexpected change in climate concerns, the better the return for the green stocks. This last conclusion is shared with PST.

In a large developed discussion section, I tried to show that results obtained in this study and any other study face a lot of uncertainties. This section is a critique of the current literature and tries to show that the topic of ESG and stocks' performance is not simple to address and that multiple research can find divergent conclusions.

Appendices

Figure 1: Pillar Scores Computation - Refinitiv

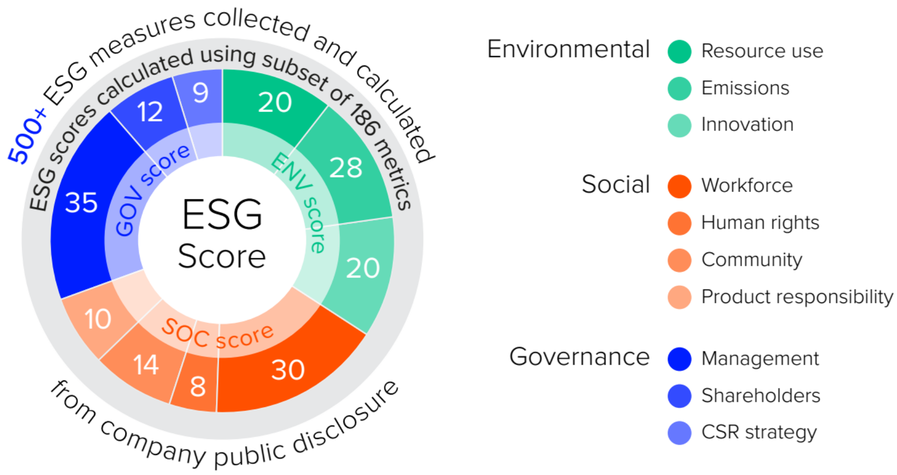

Figure 2: Number of European quoted companies with an environmental score available from 2020 to 2022

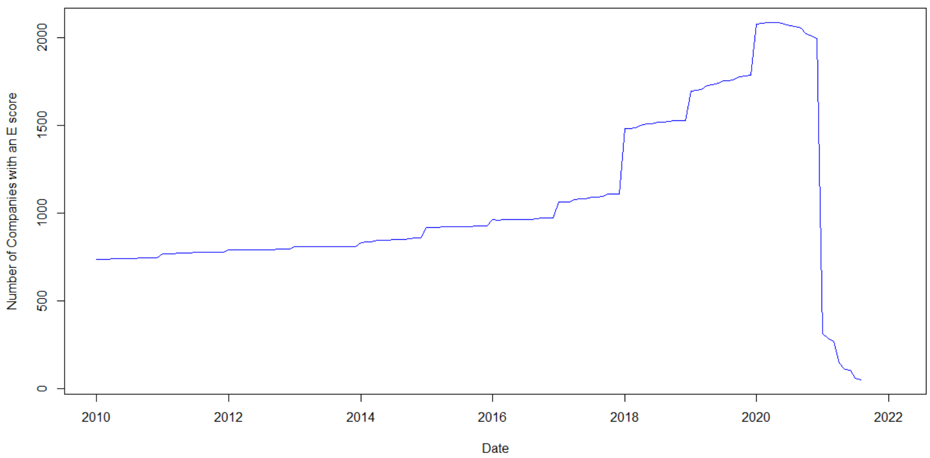

Figure 3: Climate Concern from January 2010 through June 2018

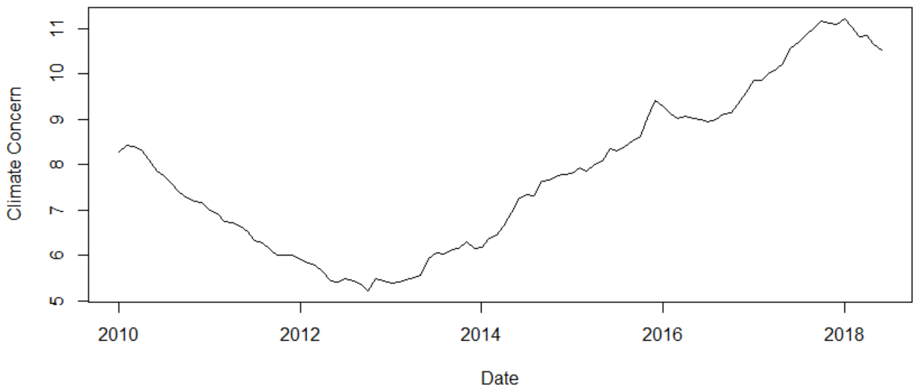

Figure 4: Correlation of ESG Measure

Table 12: Average g represents the average Greenness score from equation 6 by industry at the 31 December 2020.

Rank	Refinitiv Industry	Average g	Rank	Refinitiv Industry	Average g
1	Collective Investments	4.291	28	Homebuilding & Construction Supplies	-4.614
2	Insurance	4.248	29	Biotechnology & Medical Research	-4.817
3	Banking Services	3.789	30	Metals & Mining	-5.045
4	Media & Publishing	3.769	31	Computers, Phones & Household Electronics	-5.422
5	Investment Banking & Investment Services	2.291	32	Passenger Transportation Services	-5.548
6	Textiles & Apparel	2.276	33	Automobiles & Auto Parts	-6.141
7	Leisure Products	1.845	34	Natural Gas Utilities	-6.388
8	Specialty Retailers	1.723	35	Oil & Gas	-6.742
9	Software & IT Services	1.279	36	Semiconductors & Semiconductor Equipment	-6.778
10	Water & Related Utilities	1.030	37	Construction & Engineering	-7.143
11	Healthcare Equipment & Supplies	0.972	38	Real Estate Operations	-7.546
12	Healthcare Providers & Services	0.943	39	Office Equipment	-7.687
13	Telecommunications Services	0.433	40	Containers & Packaging	-8.126
14	Food & Drug Retailing	0.163	41	Freight & Logistics Services	-8.470
15	Pharmaceuticals	-0.296	42	Chemicals	-9.289
16	Communications & Networking	-1.000	43	Electrical Utilities & IPPs	-9.399
17	Beverages	-1.690	44	Household Goods	-9.561
18	Aerospace & Defense	-1.770	45	Construction Materials	-10.144
19	Food & Tobacco	-2.118	46	Machinery, Equipment & Components	-10.401
20	Residential & Commercial REITs	-2.183	47	Consumer Goods Conglomerates	-10.451
21	Personal &Household Products & Services	-2.315	48	Renewable Energy	-12.000
22	Transport Infrastructure	-2.949	49	Electronic Equipment & Parts	-13.090
23	Professional & Commercial Services	-3.327	50	Diversified Industrial Goods Wholesalers	-13.269
24	Hotels & Entertainment Services	-3.508	51	Paper & Forest Products	-13.626
25	Multiline Utilities	-4.105	52	Coal	-20.329
26	Diversified Retail	-4.297	53	Uranium	-23.269
27	Oil & Gas Related Equipment and Services	-4.414			

Figure 5: GMB Cumulative return

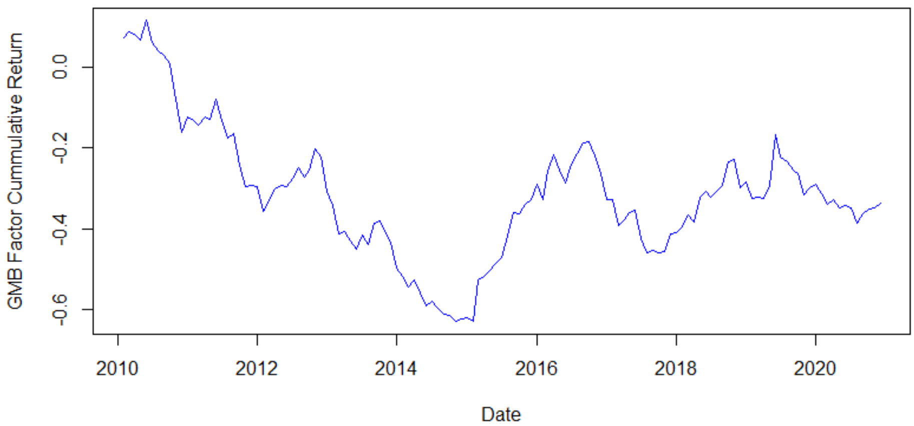

Figure 6: GMB Autocorrelation

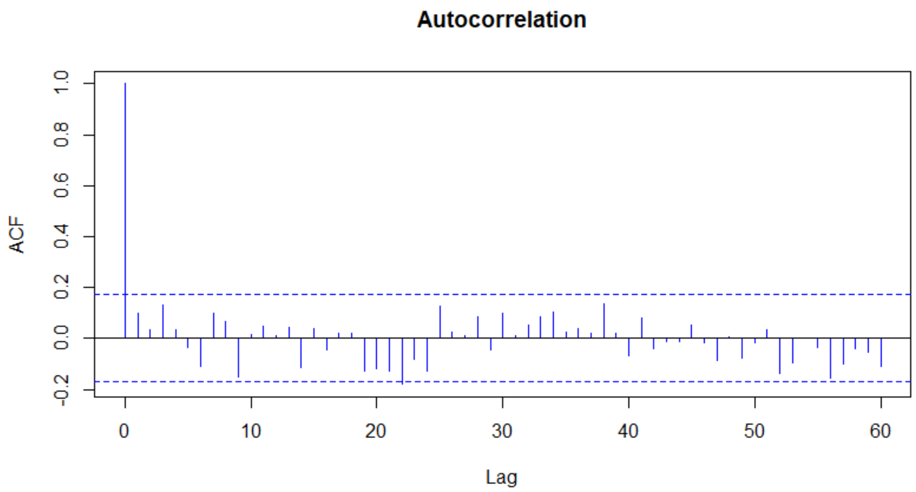

Figure 7: Green vs. Brown stocks

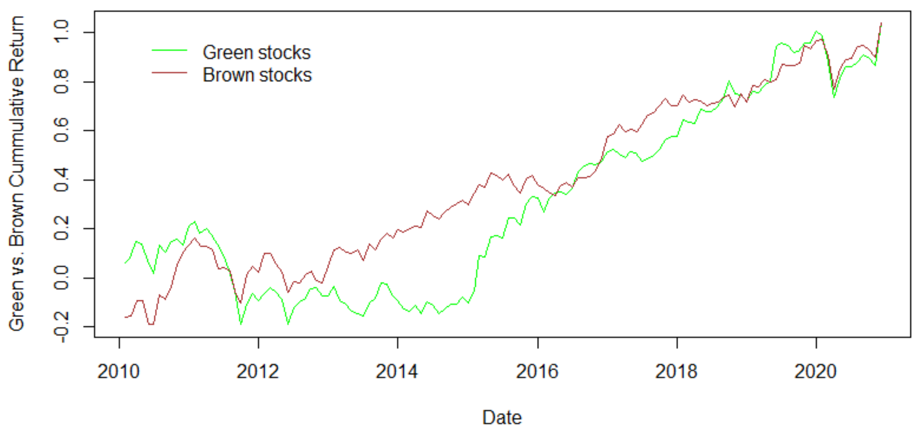

Figure 8: Green vs. Brown stocks - PST Time Frame

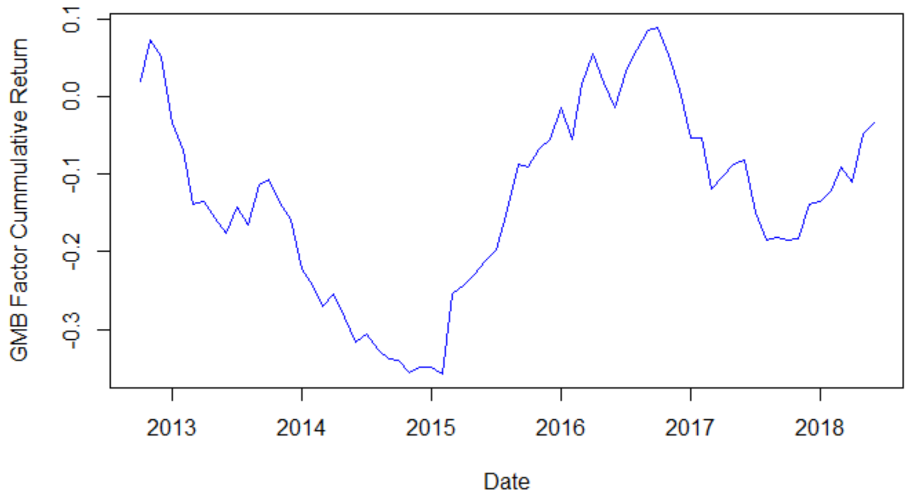

Table 13: GMB Factor - PST Time Frame | Climate Concerns Impact

The dependent variables are: the GMB factor (1), the market-hedged return of the top third green stocks (2) and the market-hedged of the bottom third green stocks. The standard errors are in parentheses.

	(1)	(2)	(3)
Constant	0.0032 *(0.0054)*	0.0102 *(0.0055)*	0.0069 *(0.0052)*
Climate Concerns (Same Month)	-0.0392 *(0.0313)*	0.0051 *(0.0316)*	0.0443 *(0.0303)*
Climate Concerns (Lag)	-0.0123 *(0.0313)*	0.0001 *(0.0316)*	0.0124 *(0.0303)*
R^2	*0.0368*	*0.0004*	*0.0357*
Observations	*68*	*68*	*68*

Table 14: GMB Factor | Climate Concerns Impact

The dependent variables are: the GMB factor (1), the market-hedged return of the top third green stocks (2) and the market-hedged of the bottom third green stocks. The standard errors are in parentheses.

	(1)	(2)	(3)
Panel A: Quintile Portfolio			
Constant	-0.0067	0.0030	0.0097
	(0.0046)	*(0.0045)*	*(0.0052)*
Climate Concerns (Same Month)	-0.0070	0.0370	0.0440
	(0.0338)	*(0.0326)*	*(0.0376)*
Climate Concerns (Lag)	0.0017	0.0024	0.0007
	(0.0339)	*(0.0327)*	*(0.0377)*
R^2	*0.0005*	*0.0172*	*0.0175*
Observations	*100*	*100*	*100*
Panel B: Decile Portfolio			
Constant	-0.0023	0.0061	0.0084
	(0.0043)	*(0.0046)*	*(0.0049)*
Climate Concerns (Same Month)	-0.0131	0.0386	0.0517
	(0.0310)	*(0.0330)*	*(0.0356)*
Climate Concerns (Lag)	0.0027	0.0173	0.0147
	(0.0311)	*(0.0331)*	*(0.0357)*
R^2	*0.0020*	*0.0272*	*0.0346*
Observations	*100*	*100*	*100*

Figure 9: Climate Concerns vs. Green Factor

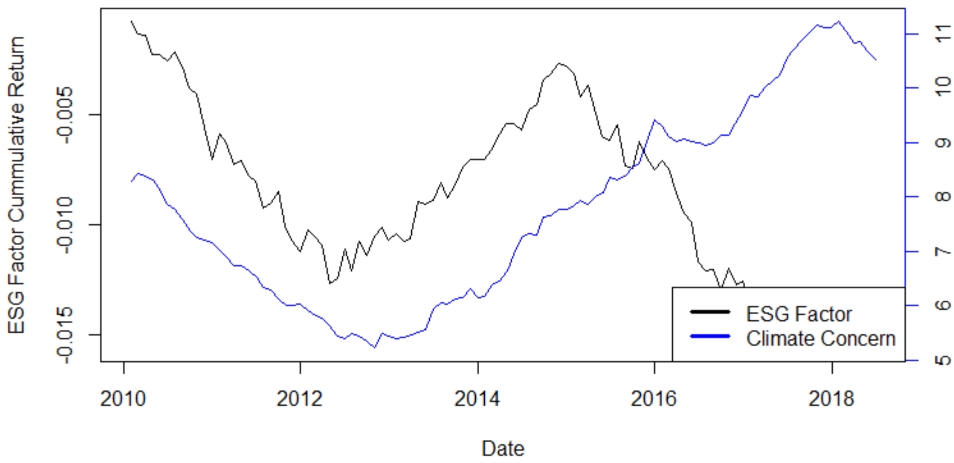

Figure 10: Green Factor Legs

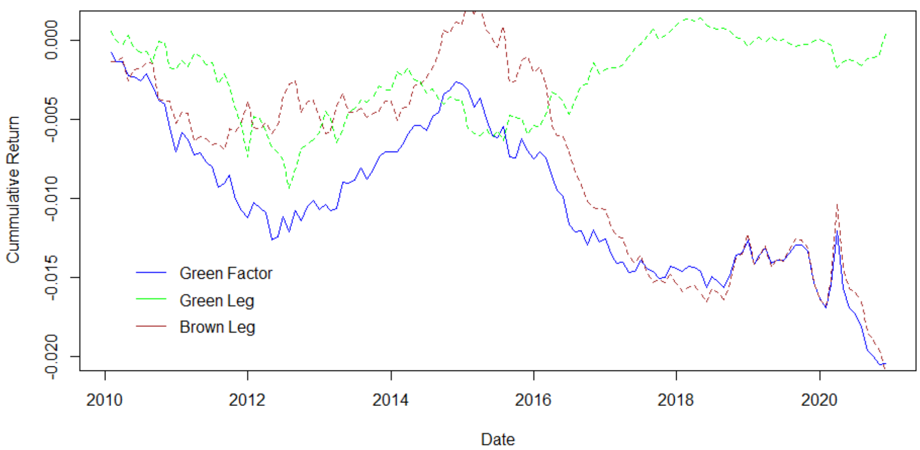

Figure 11: Performance of Value stocks

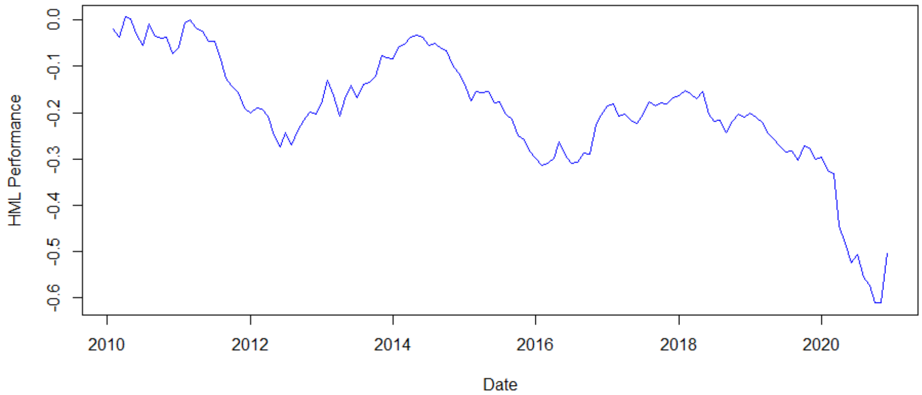

Table 15: Green Factor Fama French decomposition

We estimate monthly time-series regressions of the green factor (1), the green leg (2) and the brown leg (3) on the Fama and French (1993) 3 factors model (Excess market return, SMB factor and HML factors). Robust standard errors are in parentheses.

	(1)	(2)	(3)
Constant	0.0000	0.0000	0.0000
	(0.0001)	*(0.0001)*	*(0.0000)*
Mkt - Rf	-0.0040	-0.0018	-0.0051
	(0.0016)	*(0.0015)*	*(0.0018)*
SMB	-0.0252	0.0195	-0.0396
	(0.0041)	*(0.0036)*	*(0.0045)*
HML	0.0015	0.0199	-0.0014
	(0.0031)	*(0.0284)*	*(0.0035)*
R^2	0.3531	0.3770	0.4495
Observations	100	100	100

Table 16: Green Factor's alpha and Climate Concerns impact

We estimate monthly time-series regressions of the green factor (1), the green leg (2) and the brown leg (3) on the change in climate concerns from Ardia et al. (2020). Robust standard errors are in parentheses. None of the coefficient appears significant

	(1)	(2)	(3)
Constant	0.0000	0.0000	-0.0000
	(0.0005)	*(0.0001)*	*(0.0001)*
Climate Concerns (Same Month)	0.0005	0.0004	0.0002
	(0.0005)	*(0.0004)*	*(0.0005)*
Climate Concerns (Lag)	0.0002	-0.0005	0.0007
	(0.0005)	*(0.0004)*	*(0.0005)*
R^2	0.0245	0.0167	0.0350
Observations	100	100	100

Table 17: Industry, Greenness and Climate Concerns

	(1)	*Robust s-e*	(2)	*Robust s-e*
$g_{i,t-1}$	-2.7163	*0.0066*		
$g_{i,t-1} * \Delta C_t$	3.8277	*0.0001*		
Automobiles & Auto Parts	1.8848	*0.0595*	0.0015	*0.0150*
Banking Services	2.0242	*0.0430*	0.0123	*0.0131*
Beverages	-0.7192	*0.4720*	-0.0435	*0.0180*
Biotechnology & Medical Research	3.1999	*0.0014*	0.0209	*0.0155*
Chemicals	-0.1813	*0.8561*	-0.0281	*0.0143*
Coal	-0.3027	*0.7621*	-0.0761	*0.0359*
Collective Investments	-1.1441	*0.2526*	-0.0413	*0.0138*
Communications & Networking	0.2736	*0.7844*	-0.0034	*0.0176*
Computers, Phones & Household Electronics	0.8089	*0.4186*	-0.0212	*0.0191*
Construction & Engineering	0.8223	*0.4109*	-0.0084	*0.0139*
Construction Materials	2.0099	*0.0444*	0.0296	*0.0194*
Consumer Goods Conglomerates	-0.5006	*0.6166*	-0.0178	*0.0221*
Containers & Packaging	-0.1159	*0.9077*	-0.0216	*0.0189*
Diversified Industrial Goods Wholesalers	-0.9063	*0.3648*	-0.0436	*0.0415*
Diversified Retail	0.6375	*0.5238*	-0.0279	*0.0255*
Electrical Utilities & IPPs	0.1546	*0.8771*	-0.0245	*0.0147*
Electronic Equipment & Parts	0.0672	*0.9465*	-0.0297	*0.0181*
Financial Technology (Fintech) & Infrastructure	1.1558	*0.2477*	0.0380	*0.0334*
Food & Drug Retailing	-0.1491	*0.8815*	-0.0294	*0.0172*
Food & Tobacco	-0.2810	*0.7787*	-0.0237	*0.0140*
Freight & Logistics Services	0.4495	*0.6531*	-0.0200	*0.0160*
Healthcare Equipment & Supplies	0.0295	*0.9764*	-0.0214	*0.0145*
Healthcare Providers & Services	-0.6248	*0.5321*	-0.0228	*0.0184*
Holding Companies	0.0761	*0.9393*	-0.0072	*0.0189*
Homebuilding & Construction Supplies	1.0873	*0.2769*	-0.0122	*0.0151*
Hotels & Entertainment Services	0.2033	*0.8389*	-0.0215	*0.0143*
Household Goods	0.7806	*0.4350*	-0.0287	*0.0179*
Insurance	1.5906	*0.1117*	0.0085	*0.0146*
Investment Banking & Investment Services	0.9245	*0.3552*	-0.0143	*0.0136*
Leisure Products	3.3429	*0.0008*	-0.0069	*0.0228*
Machinery, Equipment & Components	-0.5225	*0.6013*	-0.0275	*0.0128*
Media & Publishing	0.5495	*0.5827*	-0.0278	*0.0142*
Metals & Mining	-1.5384	*0.1240*	-0.0578	*0.0138*
Multiline Utilities	1.0798	*0.2802*	-0.0011	*0.0198*
Natural Gas Utilities	1.7019	*0.0888*	0.0349	*0.0338*
Office Equipment	0.5931	*0.5531*	0.0439	*0.0260*
Oil & Gas	-1.6402	*0.1010*	-0.0528	*0.0145*
Oil & Gas Related Equipment and Services	-0.2581	*0.7963*	-0.0245	*0.0152*

Table 17 continued from previous page

Paper & Forest Products	0.8194	*0.4126*	0.0195	*0.0170*
Passenger Transportation Services	1.3195	*0.1870*	0.0185	*0.0176*
Personal &Household Products & Services	-0.2283	*0.8195*	-0.0322	*0.0198*
Pharmaceuticals	-0.2803	*0.7793*	-0.0126	*0.0141*
Professional & Business Education	0.4256	*0.6704*	-0.1812	*0.0439*
Professional & Commercial Services	-0.2765	*0.7822*	-0.0294	*0.0133*
Real Estate Operations	0.3559	*0.7219*	-0.0239	*0.0141*
Renewable Energy	1.2344	*0.2171*	-0.0151	*0.0198*
Residential & Commercial REITs	1.5673	*0.1170*	-0.0068	*0.0145*
Semiconductors & Semiconductor Equipment	2.6702	*0.0076*	0.0060	*0.0160*
Software & IT Services	-0.2147	*0.8300*	-0.0194	*0.0131*
Specialty Retailers	-0.4524	*0.6510*	-0.0237	*0.0146*
Telecommunications Services	0.0228	*0.9818*	-0.0219	*0.0149*
Textiles & Apparel	-0.3964	*0.6918*	-0.0371	*0.0157*
Transport Infrastructure	0.9438	*0.3453*	-0.0273	*0.0189*
Water & Related Utilities	0.2708	*0.7865*	-0.0168	*0.0280*
R^2	0.0060		0.0043	
Observations	90,002		195,127	

References

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. (2020). Climate change concerns and the performance of green versus brown stocks. *National Bank of Belgium, Working Paper Research*, (395).

Bali, T. G., Engle, R. F., and Murray, S. (2016). *Empirical asset pricing: The cross section of stock returns*. John Wiley & Sons.

Bennani, L., Guenedal, T. L., Lepetit, F., Ly, L., Mortier, V., Roncalli, T., and Sekine, T. (2018). How esg investing has impacted the asset pricing in the equity market. *Strategy & Social Policies eJournal*.

Berg, F., Koelbel, J. F., and Rigobon, R. (2019). *Aggregate confusion: The divergence of ESG ratings*. MIT Sloan School of Management Cambridge, MA, USA.

Bolton, P. and Kacperczyk, M. (2020). Do investors care about carbon risk? *Journal of Financial Economics*.

Bolton, P. and Kacperczyk, M. (2021). Global pricing of carbon-transition risk. Technical report, National Bureau of Economic Research.

Bruno, G., Esakia, M., and Goltz, F. (2021). Honey, i shrunk the esg alpha': Risk-adjusting esg portfolio returns. *Scientific Beta, Boston, MA*.

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

Chava, S. (2014). Environmental externalities and cost of capital. *Management science*, 60(9):2223–2247.

Clark, G. L., Feiner, A., and Viehs, M. (2015). From the stockholder to the stakeholder: How sustainability can drive financial outperformance. *Available at SSRN 2508281*.

Cornell, B. (2021). Esg preferences, risk and return. *European Financial Management*, 27(1):12–19.

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

El Ghoul, S., Guedhami, O., Kwok, C. C., and Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, 35(9):2388–2406.

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

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

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

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

Fulton, M., Kahn, B., and Sharples, C. (2012). Sustainable investing: Establishing long-term value and performance. *Available at SSRN 2222740*.

Harvey, C. R., Liu, Y., and Zhu, H. (2015). … and the Cross-Section of Expected Returns. *The Review of Financial Studies*, 29(1):5–68.

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

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2020). The pollution premium. *Available at SSRN 3578215*.

Kempf, A. and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European financial management*, 13(5):908–922.

Khan, M. (2019). Corporate governance, esg, and stock returns around the world. *Financial Analysts Journal*, 75(4):103–123.

LSEG (2021). Environmental, social and governance scores from refinitiv. pages 01–25.

Madhavan, A., Sobczyk, A., and Ang, A. (2021). Toward esg alpha: Analyzing esg exposures through a factor lens. *Financial Analysts Journal*, 77(1):69–88.

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

Melas, D., Nagy, Z., and Kulkarni, P. (2017). Factor investing and esg integration. In Jurczenko, E., editor, *Factor Investing*, pages 389–413. Elsevier.

Naffa, H. and Fain, M. (2021). A factor approach to the performance of esg leaders and laggards. *Finance Research Letters*, page 102073.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Dissecting green returns. Technical report, National Bureau of Economic Research.

Pástor, L., Stambaugh, R. F., and Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.

Pollard, J. L., Sherwood, M. W., and Klobus, R. G. (2018). Establishing esg as risk premia. *Journal of Investment Management*, 16(1):32–43.

Robinson, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *The Annals of statistics*, pages 1630–1661.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, 19(3):425–442.

Stanley, M. (2015). Sustainable reality: Understanding the performance of sustainable investment strategies, morgan stanley institute for sustainable investing.

Stanley, M. (2021). Morgan stanley sustainable reality report 2021.

Statman, M. and Glushkov, D. (2009). The wages of social responsibility. *Financial Analysts Journal*, 65(4):33–46.

Wang, P. and Sargis, M. (2020). Better Minus Worse - Evaluating ESG Effects on Risk and Return. *MorningStar*.

Wenger, K., Leschinski, C., and Sibbertsen, P. (2018). A simple test on structural change in long-memory time series. *Economics Letters*, 163:90–94.

Executive summary

ESG is increasingly becoming a topic of interest for research and investors. While some may be interested to invest in ESG for ethical reasons, some may be more interested in capturing excess returns from this class of asset. In any case, both individuals are interested in the impact of ESG implementation on the performance of stocks. This study analyzes the impact of green scores (Environmental pillar) on stocks' performance and shows that in Europe and using Refinitiv dataset, brown stocks tend to outperform. Following Pastor et al. (2021) methodology and using Ardia et al. (2020) climate concern index, I show that unexpected changes in investor climate concern do not have an impact on the portfolio level but positively impact green stocks' performance at the stock level.

Word count: 14,120