Department of Impresa & Management

Master's degree in Corporate Finance

Chair of Asset Pricing

The cross-sectional variation of European stocks' returns and its relationship with ESG

SUPERVISOR
Prof. Paolo Porchia

CO-SUPERVISOR
Prof. Marco Pirra

CANDIDATE
Lorenzo Patacchiola
695391

ACCADEMIC YEAR
2018/2019

Table of Contents

Chapter 1

Introduction

In the last decade, ESG investing got under the spotlight. Ultimately, also in the finance area. In fact, academics, as well as asset managers, have started considering this strategy and investigating whether it could eventually be profitable to take into account people's personal beliefs and values in setting portfolios. Companies themselves have shown their interest in this matter as well, by disclosing their performance in terms environmental footprint, respect for company's stakeholders, and their adherence to corporate governance best-practices, among other things.

Before diving into the details, it is worth to define ESG as well as its bounds with Asset pricing. If, on the one hand, the acronym stands for Environment, Social, and Governance, the concept behind, on the other, refers to the contribution an individual company makes to sustainability. Thus, ESG is referred to as the comprehensive performance a firm has under the different specifications: environment (e.g. Green-house gases emissions, environmental footprint of its products and processes) social (e.g. Labor conditions it offers to its employees, sensitivity versus stakeholders' interests) governance (e.g. Understanding of minority shareholders' interests, best practices). Companies are increasingly more concerned about the topic. One of the reasons could be that people, hence investors, care more about it, thus firms need to align their performances to the demand expectations. Another reason could be found in countries strengthening regulatory frameworks in these regards. So, corporations need to update their production processes to not fall short on legal requirements. Any of these could impact stocks' performance. If investors get more and more aware of sustainability, their decisions will be likely to be influenced by judgements on ESG basis. Hence, integrating this kind of considerations could improve the performance of portfolios. Given the increased attention paid by companies to these matters, it is worth to study if this phenomenon is also reflected into the stocks' market. EY and the Boston College Center for Corporate Citizenship, for example, have conducted a joint study which has shown that firms with high quality sustainability reporting practices have obtained easier access to capital, as well as higher company value.

Much research, accordingly, has focused on answering multiple research questions regarding the issue. First and foremost, whether the strategy of integrating ESG performance in portfolio setting is profitable. The evidence is mixed: part of the studies does not find positive performance of sustainable portfolios relative to sustainability-indifferent ones. Literature in these regards could be divided into two main areas. Some shifts the attention on the more "managerial" side of the issue, showing concerns regarding the overall strategy of chasing other than financial results. In fact, investing company's resources to achieve results which are not directly reflected into the year-end bottom line would be detrimental to the health of the venture. Eventually, the performance of these companies would be negative with respect to their peers, thus making them an undesirable investment.

Others look at the topic from a portfolio management standpoint. The logic behind would relate to a fundamental aspect of portfolio setting: risk. To minimize it differentiation is the key. Since many strategies to integrate sustainability considerations in portfolio setting include a negative screening process, in which companies are excluded given their non-responsible characteristics (e.g. participation in a "*sin industry*", like tobacco), the net effect is that of reducing the set of available firms in which to invest. Moreover, this is likely to produce a portfolio which concentrates on few industries. The ultimate consequence is that of having similar returns to other portfolios, with increased volatility. Similar studies assert that the negative returns associated with sustainable investing could be considered as the cost of doing the right thing. Meaning that investors should be ready to afford this relative negative performance, in order for their portfolios to reflect their set of principles.

On the other hand, some alternative research finds a positive return in ESG investing. In this matter as well, supporting evidence is provided from different angles. Some supporters of this view highlight a positive relationship between sustainable behavior and returns which could be not as easily measurable as financial ones (e.g. efficiency). Some others find that thanks to their ability to anticipate legislative changes, responsible firms tend to outperform lagging competitors. Also, other studies demonstrate that those firms' risk-adjusted returns are higher.

Not much of the research, though, has produced a step further and analyzed the three commonly recognized sustainability elements, Environment Social and Governance, individually. Most likely, the main reason behind this shortage is the lack of data regarding specific ESG criteria. As of today, it is extremely challenging to find specific responsibility components-driven portfolios, or even proxies. This is most probably due to the fact that sustainable finance is an ever-increasing

phenomenon, but still very recent. Consequently, very few data are available to conduct massive research on that.

Moreover, research lacks studies conducted over European samples. Hence, the intent of this paper is that of investigating whether the three singular sustainability components are priced in the European market. To do so, three portfolios, each of them reflecting one determinant of ESG (accordingly, one portfolio for Environment, one for Social, and one for Governance) have been constructed, in a Fama-French High-minus-Low fashion. The sample chosen is the STOXX Europe 600. In order to determine the sustainability performance of sample stocks, data have been collected on Thompson Reuters database, which offers the percentile ranking of stocks, against each item, since 2001. Before testing the research question, some preliminary analysis has been conducted over the three factor-mimicking portfolios, to determine whether they are good candidates to explain risk. To do so, they have been tested against the Fama-French-Carhart four factors specification. Results show that E and S can be explained by the already defined asset pricing model. G shows no statistically relevant relationship with any of the risk factors. Nonetheless, coherent with previous research, all the ESG items display a negative relationship with the value factor mimicking portfolio. This would suggest that sustainable firms tend to perform similarly to growth stocks.

Finally, to test the research question, a standard Fama-Macbeth procedure has been conducted over a 2002-2018 monthly time period. The four factors specification has been added to the model to have some further controls. Results highlight a negative and statistically different from zero explanatory power of Governance in relation with the cross-sectional variation of average stocks' returns in the sample period. Additional analysis is also performed to investigate further the behavior of the factor-mimicking portfolio.

The following section covers a critical literature review and provide the reason behind the choice of this topic. The second chapter describes more in depth the empirical analysis, and the data collection procedure. The third chapter outlines the model used and economic intuitions behind results. Finally, conclusions are drawn.

Literature review and motivation

ESG investing is a very debated issue among experts, not only from a financial viewpoint, but also from a managerial one. Detractors believe that it is not in the best interest of shareholders to pursue objectives other than the bottom line (Friedman, 1970). In fact, according to this view, involving the

firm in activities which do not directly improve net income is detrimental to the net returns of equity holders. Others do no find any correlation between socially responsible investing and financial return (Di Giuli and Kostovetsky, 2014). The main critique moved to this strategy, nevertheless, is that it would lead investors to choose between return and responsibility. The negative returns coming from this choice, could be interpreted as the "*cost of doing the right thing*". From this viewpoint, in fact, responsibility would carry the cost of excluding firms from portfolios given either their inability to report on socially relevant matters, or their presence in so-called "*sin industries*". Clearly, the choice would not be based on assumptions regarding the marginal risk a firm would add to the portfolio. This, in turn, would reduce the differentiation among stocks, thus leading to a portfolio more exposed to diversifiable risk. The inability of sustainable investing to differentiate among firms seems to be most critical aspect of integrating ESG considerations in the construction of portfolios. Renneboog, Ter Horst and Zhang (2008) hold that "*In order to pursue social objectives, SRI funds employ a set of investment screens that restrict their investment opportunities*". It is worth to notice, nevertheless, that there is an indisputable momentum in responsible investing, as highlighted by several evidences. Eccles and Klimenko (2019) find that asset managers of global investing firms have integrated ESG criteria in their strategies, responding to investors' needs. Furthermore, the second annual survey conducted by the Industry Index Association in November 2018, highlighted an impressive 60% increase in ESG indexes over the previous year. Last year's result still shows a 13.85% positive figure. These numbers seem to confirm the growing concern of investors regarding the inclusion of their personal beliefs in their investments. Considering that investors are likely not to only make efficiency considerations in their investing strategies, it would be valuable to update asset pricing models to make sure these concepts are not overlooked. In fact, if responsibility proves not be a risk factor – in the sense that it cannot explain performance – these findings could at least provide evidence of profitable opportunities. If expectations regarding the performance of these firms increase, so will their prices. Since this increase would not be justified by an intrinsic value, rather by unidentified drivers, stocks would result in having higher alphas that could be exploited by investors. Thus, savvy investors need to take into account sustainability. This mispricing characteristic has been discussed by Mănescu (2011). In her study, she investigates seven different ESG items and their ability to explain the cross-sectional variation of stock returns. The sample is made of US traded firms' monthly returns in a time window of 17 years (1992-2008). Her model also implies the Fama-French-Carhart four factors specification, as control variables. Findings highlight that one of the measures, community relations, has positive effects on risk-adjusted

returns. Though, this feature is attributed by the author to mispricing. To further test the sample, the study splits the available data into two separate time frames: from July 1992 to June 2003 and from July 2003 to June 2008. This additional analysis provides more insights regarding other ESG items, in particular: product safety, human rights, and employee relations. The former two are found to have some negative effects on stock returns in the most recent period. Employee relations, instead, is proven to behave differently: in the 1992-2003 term, figures highlight a positive effect on the dependent variable. The author concludes that this is due to mispricing. In the 2003-2008 season, nevertheless, the measure is noticed to have the opposite effect. The reason behind that, is found by the author in a lower compensation for risk.

Also other studies show some negative relationship between Corporate Social Performance (CSP) and Corporate Financial Performance (CFP). Brammer, Brooks, and Pavelin (2006) investigate the relationship between these two in the UK market. They consider CSP as a resulting measure coming from two variables: the social and the environmental behavior. Accordingly, they find that "sin" firms tend to grant a higher excess return. In fact, comparing the returns provided by a portfolio composed according to the CSP performance and its benchmark, average returns of the investigated asset are statistically significantly lower. Their analysis also breaks down this relationship, to investigate the main drivers of this lower returns trend. Apparently, the negative performance of socially desirable firms is to be attributed to their environmental, rather than social, tilt. Coherent with these results is the assumption that firms with a worse CSP attitude are felt riskier, hence investors require higher returns.

Jang (2019) finds negative relationship between ESG measures and stocks returns in the European market. Using an aggregate measure of sustainability performance, the author has analyzed whether investing into firms which score high in this matter yields some return. She finds negative relationship between the two variables. According to her interpretation, ESG helps in gaining market share but heavily hinders companies' profitability. Hence, it yields negative returns.

Part of the research also find that there is no effect in integrating ESG considerations in portfolio setting. It is the case of Xiao, Faff, Gharghori and Lee (2013), who analyze the cost of investing into a portfolio with a strong ESG tilt, on a global scale, in the context of institutional investors. The ESG factor-mimicking portfolio is obtained by ranking sustainable performance of firms in the sample according to the score assigned by the Sustainability Asset Management (SAM). Results from the Fama-Macbeth regression analysis show that there is no significant relationship between

sustainability and returns, hence employing this trading strategy would not be harmful to the profitability of the portfolio.

An interesting point of view is offered by Halbritter & Dorfleitner (2015), who highlight a key aspect of ESG investing: the definition of ESG ratings. The same parameters are applied to the entire sample, thus levying all the intrinsic differences which arise from the industry the firm operates in. Although sustainability could be considered as a tout-court feature, relevancy of the specific item (e.g. liters of water saved, GHG emissions) with respect to the company's business field is actually a major concern. In fact, a stock ranking high on immaterial elements could be preferred over a more concentrated peer, thus reducing the quality of the selection. For this reason, this current paper takes advantage of a database offered by Thompson Reuters, that considers the relevance, with respect to the industry, of the particular measure in computing the percentile score. Other studies have already highlighted this bias and tried to address it by implementing models that took into account the relevancy of the specific items being measured, ESG-scoring wise. Kahn et al. (2016) developed an innovative database by classifying materiality according to the industry. They have defined materiality combining knowledge from the Sustainability Accounting Standards Board (SASB) and the firm-level ratings provided by KLD. Accordingly, they find that firms positioning well on material items consistently outperform peers with a lower score. Moreover, they highlight that firms doing good on immaterial items are as profitable as firms with worse rankings in the same areas. Henriksson et al. (2019) build on that, exploiting deeper the concept of "materiality". In their study, they identify whether a specific item is industry-relevant by analyzing the SASB's materiality map. This is constituted by a set of sustainable topics which are defined to be significant for a particular sector or not. Sectors are identified by the SASB according to the Sustainability Industry Classification System (SICS).

Additional research is inconclusive regarding the pricing of ESG risk. Fiskerstrand et al. (2019) apply a Fama-Macbeth procedure to test whether the hypothesis holds in the Norwegian market. Due to the lack of data regarding ESG performance, the authors have derived the performance of each stock in this aspect. In particular, after having collected data for Norwegian stocks between 2009 and 2018, they measured the sensitivity of each stock towards the Dow Jones Nordic Sustainability index (DJNSI) through a time series regression of each stock versus the DJNSI, and then they have sorted the resulting betas into two portfolios, applying a Fama-French High minus Low approach. Excess returns of this portfolio have been computed. Thus, they obtained a fifth risk factor to add to the Carhart specification. Their results show no explanatory power of the ESG factor with respect

to the cross-sectional variation of excess stock returns in the Norwegian case. A similar result is provided by Limkriangkrai, Koh, and Durand (2017), that investigate the effects of ESG ratings on both stock returns and financing decision in the Australian market. In particular, they investigate the performance of a portfolio sorted on the basis of ESG considerations. ESG data are collected by Regnan, which is an Australian company investigating over ESG performance of the top companies, by market capitalization, listed in the S&P ASX200. The firm offers a rating for each company, in each of the relevant ESG constituents. The authors, accordingly, create for each component, two separate portfolios, made of the companies that score high/low. These are used to derive three factor-mimicking portfolios, which are then tested. This paper highlights an additional key aspect of ESG research: the need to segregate the effects of each sustainable field, when studying the subject. They assert, in fact, that a company may highly invest in one of the three areas, leaving behind the other two. For this reason, a "*more fine-grained analysis of the (ESG) ratings may be advantageous to better understand the impact of ESG activities on a firm's financial performance"*. This relatively straightforward intuition plays a major role in the research field, as it widens the perspective from which investigation should be took. The results, though, seem to confirm that investing into sustainability delivers no abnormal return.

Part of the literature, on the contrary, finds positive relationships between sustainable investing and good performance on key portfolio performance variables (i.e. volatility and returns). First of all, Albuquerque, Durnev, and Koskinen (2014) find that investing in companies with good CSR performance reduces systematic risk. Among the intuitions the authors give, customer loyalty plays a major role. One of the hypotheses of the paper, confirmed by results, is that CSR is associated with customer loyalty. The authors state that higher customer loyalty means less elastic price demand, which in turn has positive effects on profit margin. Moreover, this demand curve property results in more stable cashflows, also during economic shocks.

One another interpretation of Albuquerque, Durnev, and Koskinen (2014)'s findings is the ability of these firms to anticipate regulatory framework changes and so not fall short on new directives, but instead developing a competitive advantage over peers. This intuition is supported by the results obtained in Eccles et al. (2014), who analyze the performance of U.S. firms that were first-movers in implementing CSR policies – particularly focusing on social and environmental issues, even respecting to regulatory framework, and find that they outperformed their laggard competitors. Their model applies a singular approach: they define the independent variable as the early adoption, and its extent, of sustainable practices, whereas the dependent variable is the tracked financial

performance. Hence, the study is able to highlight the long-term perspective of ESG investing: the underlying hypothesis, in fact, is that sustainability integration is able to provide long-term profitability. Results, eventually, show that sustainability laggards tend to underperform first movers. In their interpretation, higher expectations which investors have on such companies justify the outperformance relative to their peers. Interesting results are provided by Annér and Jakobsson van Stam (2018), who give evidence of sustainability effects on stocks' cross-sectional returns, by investigating whether ESG measures are priced in the Swedish market. Their study answers to two research questions: firstly, whether stock returns are affected by ESG measures. Secondly, if the effect is given by mispricing or a compensation for risk. After having created a combined ESG score and having used individual measures, they have applied a Fama-Macbeth procedure to test the first hypothesis. Results show that the aggregate sustainable figure has no explanatory power on stocks returns, in line with previous expectations. Interesting insights, however, are offered on some sustainable measures. According to their results, in fact, community and product responsibility are found to be able to explain some variation in returns. Galema, Plantinga, and Scholtens (2008), also, find that SRI is able to explain stock returns. They investigate the relation between social investing and US stocks' returns, finding non statistically different from zero alphas. Their results confirm that sustainable investing is able to command a risk premium. An interesting assumption they make, confirmed by results, is that firms with a social consciousness have a higher demand. Thus, they tend to be characterized by greater market values. That is, lower book-to-market ratios. To add on this point, Serafeim (2018) provides additional results. In his research, he combines the ESG performances of US stocks with the public sentiment around them, in order to find the relationship between what investors feel about the sustainable performance of a company and its returns. Findings strengthen the point. In fact, he highlights that firms with a positive public sentiment result in a stronger association between ESG performance and market valuation. Since this behavior would be explained by a greater interest of investors into sustainability, a similar pattern should be expected in the European market. Hence, a similar assumption is made also for the studied sample. In fact, this effect can be observed in more than one research paper, thus a closer look at it is given also in this study.

Khan (2019), constructing on his previous results regarding the materiality of ESG items, finds that ESG metrics are able to provide significant positive alphas, worldwide. The paper shows that the combined sustainable score has informative power regarding the future performance of stocks.

The current literature leaves out two aspects from the scope of the analysis. No study seems to pay attention to the European market as a whole. That is, limited literature shifts the attention on the behavior of European stocks, when it comes to sustainability. Indeed, much more effort is put either on individual European cases or on American stocks. Nevertheless, what holds for US stocks may not be true for continental companies. In first instance, ESG has much to do with regulatory frameworks. It is true that globally there is a general tendency towards homogeneity in these matters: the numerous agreements signed with the purpose of improving countries performance in diminishing pollution, as well as laws and regulations strengthening the legal requirements in the field of corporate governance – following the Cadbury code, the SOX is the main reference – give strong evidence of that. The act, in fact, has been enacted into different regulatory systems, like the Japanese version, J-SOX, and the many European directives. Still, no full congruence has been achieved and sustainability, as a general term, is integrated into business cycles at different extents. Moreover, the European case is even more interesting, given its heterogeneity. In fact, unlike the other relevant markets, the old continent financial field is composed of several countries, that physiologically behave differently when it comes to ESG. Hence, it is worth to investigate further its dynamics in this composite market.

The second area not yet investigated in the current literature is the individual behavior of the three components of E S and G. What different studies have done, as outlined above, is to investigate the effects of ESG comprehensively. Results, indeed, may be only partial. In fact, should no effect be found, it would be hard to determine whether any of the items is incapable of producing consequences. Thus, the Environment, the Social, and the Governance elements should be analyzed independently.

For these reasons, this paper draws its attention towards investigating whether Environment, Social, and Governance are priced, individually, in the European market.

Chapter 2

In this chapter the ESG data collection methodology will be explained. Then, I will show how the factor-mimicking portfolios have been constructed, as well as how their excess returns have been computed. Finally, the Fama-Macbeth procedure used to test the research question of whether Environment, Social, and Governance are priced in Europe will be displayed.

One of the main reasons why the current literature regarding sustainable investing provides such a mixed evidence is surely the challenge of correctly testing the actual sustainability of a firm. In fact, this feature is not easily derivable as other characteristics (e.g. size or book to market ratio). Reporting around these matters is still developing. Moreover, it is even difficult to test the truthfulness of the statements, as a proper auditing procedure has not been drafted. One of the consequences, for example, is the existence of green washers: firms which pretend to be eco-friendly, to mislead consumers about their environmental performance for marketing and advertising purposes.

As of today, the most reliable measure of sustainability is given by third party entities which are directly involved in estimating the performance of companies, under the sustainable point of view. These are agencies that have been involved in this activity since the very beginning of the century, hence have developed appropriate expertise to correctly and properly evaluate companies' performances on sustainability matters. As stated by Limkriangkrai, M., Koh, S., & Durand, R. B. (2017), in fact, when it comes to these issues it is important to obtain a "*fine-grained*" analysis of the performance of the company under each aspect (i.e. Environment, Social, and Governance). Moreover, as highlighted by Halbritter & Dorfleitner (2015), a major role is also played by the specific aspects under which companies are analyzed. Set differently, among the several activities which a company may undertake in order to improve its social performance, it is important to consider the industry which the company operates in. That is, two companies, in two different industries, may perform similarly in the same ESG aspect, but if this is only relevant to one of the two industries, then the value added for the two is likely to be very different. This concept is defined as "materiality" by Kahn et al. (2016), first, and then exploited by Henriksson et al. (2019) later on. For the purpose of this study, accordingly, the Thompson Reuters' ranking has been selected to collect data regarding ESG performance of companies in the sample. In particular, scores were obtained by: ENVSSCORE (Environment), SOCSCORE (Social), and CGVSCORE (Governance). Data regarding sustainable practices and performance of companies were initially collected by ASSET4, a pioneer in this aspect. In 2009, though, Thompson Reuters has acquired the Swiss EGS data provider, enlarging its information availability. This instrument has different features which make it an appropriate tool to estimate sustainable performance of companies. In first instance, unlike many others currently available, this database has been collecting data since the end of the 90s, and thus it offers a rare availability of scores (data before 2003, year in which ASSET4 was founded, have been retrospectively acquired). Then, it is able to provide data which take into account the relevancy

of the specific item with respect to the industry the company belongs to. That is, more than 400 ESG measures are identified and grouped into 10 categories, under which firms are evaluated. These measures are chosen on the basis of the comparability and the relevancy with regards to the industry. Hence, the score given to each company takes into account the materiality of each individual item. The result is a score ranging from 0 to 100, where 0 is the minimum and 100 is the maximum, which has the very nice property of already incorporating the relevancy of the specific measures with respect to the industry the firm belongs to. Hence, using this score allows to already account for industry effect. For example, Schindler AG, one of the market leaders in elevators industry, in 2003 ranked among the most virtuous in terms of Environment scoring (95.1), whereas in terms of Governance it was in the bottom quintile (10.74). In the same year, Nokia ranked among the top performers under every specification: Environment (96.68), Social (90.18), and Governance (84.04).

Since the purpose of the study is that of investigating whether sustainability is priced in the European market, factor-mimicking portfolios for Environment, Social, and Governance were needed. As of today, no such an asset is available in the market. Probably, this is due to the already outlined challenges in properly evaluating sustainable performance. Thus, three portfolios have been drawn using the STOXX Europe 600 as stock sample. Accordingly, through the ISIN of each stock, their respective scores (under Environment, Social, and Governance) have been downloaded by Thompson Reuters from the end of 2001 to the beginning of 2018. In order to form the three factor-mimicking portfolios, then, per each year, the stocks have been ordered from the top to the bottom scorer. In case no score was available for a company in a specific year, this has been deleted. The correction was done in order to have an as much unbiased as possible sample. In fact, companies for which a score is not available are not necessarily laggards on sustainability, rather they may have not properly disclosed their performance. This could depend on the inability of the firm itself to communicate its performance, but also on a country factor. In fact, reporting over sustainability matters differs from country to country. Since the scope of the paper is to test the research question throughout Europe, in aggregate, regulatory differences may arise. Hąbek & Wolniak (2013) assert that if, on the one hand, it is true that the several directives are in place to regulate both mandatory and voluntary disclosures, on the other hand some countries are much more concerned than others in these matters. To make their point, they highlight that France, for example, requires listed companies to disclose on CSR matters since 2001. Denmark has applied the same regulatory requirement, since 2010. Hence, it is plausible to assume that some companies

have not met the bar, not necessarily because they reject to, but also for a lack of external (regulatory) pressure. By consequence, it has been deemed appropriate to rely on the scores given by Thompson Reuters, and just exclude the unrated companies, instead of considering them as ESG poor performs. Once the companies have been ordered, under every specification, according to their score, the sample has been divided into quintiles. For the purpose of obtaining the three factor-mimicking portfolios, only the top and the bottom quintiles were considered. The procedure has been performed for each of the three portfolios.

Closing monthly prices for the stocks have been downloaded from Bloomberg. Given the nature of the research, it is likely that bank holidays were considered when downloading days. To avoid biases coming from that, a specific Bloomberg function was used, which gives back the closing price at the specified date, but if data are not available for that stock in that day, it gives the same piece of information at the closest previous date available. Any correction regarding missing data, notwithstanding this formula, is explained later on in the section.

The monthly frequency has been chosen since it is perceived as a convenient trade-off between avoiding a non-trading bias and collecting enough data to obtain valuable and significant results. A recent research conducted by PWC (PricewaterhouseCoopers, L. L. P. (2015). Global financial markets liquidity study. *report prepared for the Global Financial Markets Association and the Institute of International Finance*) has confirmed how the European stock markets, in general, are deemed to be less liquid than bigger and less fragmented markets, like the US one. This feature makes European assets more likely to fall short on the non-trading bias, which arises when stocks trading volumes are low. Using daily, or even weekly prices, would result in zero, or close to zero, returns for many of the stocks in the sample, thus hindering the validity of the results.

Monthly returns were calculated in first instance, as:

$$\frac{P_{t+1}}{P_t} - 1$$

Where *t* refers to a specific month.

Then, excess returns were computed subtracting the risk-free rate. The procedure has been applied both to the highest and lowest quintiles, resulting in monthly excess returns for both sustainability

leaders and laggards. Finally, portfolios' monthly excess returns were computed subtracting returns of the bottom quintile from those of the top.

With respect to the Fama-French-Carhart four factors, monthly excess returns have been downloaded from Kenneth French's database (available at https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

Factor-mimicking portfolios have been rebalanced each year (this is consistent with the frequency with which firms disclose ESG performance). In order to reflect the effect of ESG performance over companies returns, the monthly excess returns were computed in year *t* (not to be confused with the *t* referring to the monthly returns), using quintiles formed with *t-1* EGS scores (i.e. ESG score have been taken at the end of 2001, the quintiles for year 2002 were formed according to these). Both the top and bottom quintiles have been computed as equally weighted portfolios. Eventually, the monthly excess returns have been obtained by subtracting the bottom quintile excess returns to the top quintile ones, following Fama, French (1993). The procedure has been applied to each portfolio, throughout the entire sample. The excel spreadsheet containing the seven (Fama-French-Carhart 4 factors plus E S G portfolios) independent variables has been uploaded into MATLAB. Thus, a total of 1,428 firm-month prices made up the independent variables' matrix. To obtain the matrix of the dependent variables, a very similar procedure has been followed. In order to avoid any survivorship bias all the firms that have been listed at least once in the index have been considered. Monthly excess returns have been then ordered in an *N x T* matrix, where:

N = Number of firms on the sample (1245)

T = Number of months in the sample (204)

Unfortunately, prices were not available for all firms throughout the entire time period, so firms for which at least on monthly data was missing have been deleted. After the corrections, the total number of firms in the sample amounted to 865. That makes a total of 176,460 firm-month data for the dependent variables.

Before testing the research question, descriptive statistics of the analyzed portfolios has been investigated. This is intended to obtain more details regarding the performance and the significance of the three portfolios. Hence, the monthly excess returns of E, S, and G have been examined.

To investigate whether E, S, and G are priced in the European market, a standard Fama-Macbeth procedure has been applied. This model allows to test the variables under investigation while simultaneously controlling for other effects (Bali, Engle, and Murray (2016)). The Fama-French-Carhart specification has been added to the model, as control variables. Thus, the total number of independent variables is 7:

K = Risk factors (7)

Hence, the matrixes were uploaded into MATLAB. The first time-series regression, which gives back the sensitivity of each stock versus the factor-mimicking portfolios, is:

$$y_1 = \alpha_1 + \beta_1(Mkt - Rf_1) + \beta_1 SMB_1 + \beta_1 HML_1 + \beta_1 MOM_1 + \beta_1 E_1 + \beta_1 S_1 + \beta_1 G_1 + e_1$$
$$y_2 = \alpha_2 + \beta_2(Mkt - Rf_2) + \beta_2 SMB_2 + \beta_2 HML_2 + \beta_2 MOM_2 + \beta_2 E_2 + \beta_2 S_2 + \beta_2 G_2 + e_2$$

.

.

.

$$y_T = \alpha_T + \beta_T(Mkt - Rf_T) + \beta_T SMB_T + \beta_T HML_T + \beta_T MOM_T + \beta_T E_T + \beta_T S_T + \beta_T G_T + e_T$$

To run the *N* time series regressions, a standard OLS formula returning the β estimates has been used:

$$\hat{\beta} = (X'X)^{-1}X'y$$

Where:

X= Matrix of factor-mimicking portfolios time-series excess returns (*T x K*)

y= Column vector of the n^{th} stock's monthly excess returns (*T x 1*)

The OLS formula is iterated *N* times.

The resulting betas reflect each company's sensitivity versus the portfolios. As such, the effect of sustainability is intrinsically measured by means of how the 3 portfolios are constructed: they are rebalanced each year, on the basis of stocks' scores in each specification at previous year end.

The column vectors are transposed and stored in an *N x K* matrix, called *Beta*, having in each column the time series coefficient estimates per each firm, for any factor-mimicking portfolio:

$$Beta = \begin{bmatrix} \hat{\beta}_{1,Mkt-Rf} & \cdots & \hat{\beta}_{1,G} \\ \vdots & \ddots & \vdots \\ \hat{\beta}_{N,Mkt-Rf} & \cdots & \hat{\beta}_{N,G} \end{bmatrix}$$

The second step of the procedure involves a cross-sectional regression of each of these estimates on the *N* stocks monthly excess returns. To save computations, a vector of time-series average stocks monthly excess returns is derived, which is then used as dependent variable in the upcoming cross-sectional regression:

$$Y = \begin{bmatrix} E[y_1] \\ E[y_2] \\ \vdots \\ E[y_N] \end{bmatrix}$$

Where,

$$E[y_i] = \frac{1}{T}\sum_{t=1}^{T} y_t$$

This vector is regressed against the *N x K* Beta matrix using the MATLAB function "*Fitlm*". The function runs a regression between the two input variables (i.e. the matrix containing the time-series beta estimates and the stocks' monthly excess returns in the average month), returning intercept and coefficient estimates, plus the t-stat and the p-value of them; details useful in understanding the validity and significance of the results. The aim of this step, as suggested by Bali, Engle, and Murray, (2016), is to understand whether there is any significantly different from zero relation between the factor-mimicking portfolios and the excess returns in the average month. Moreover, if the model is able to explain the cross-sectional variation of stock returns, the resulting alpha should be not statistically different from zero. In fact, firms are not supposed to earn abnormal returns not due to any of the risk factors in the model. In case this happens, it can be concluded that part of the risk is not explained by the model itself.

Chapter 3

Overview of the factor-mimicking portfolios characteristics

	E	S	G
mean	−0.0032793	−0.0039636	−0.001661
median	−0.0031337	−0.0022438	−0.002089
max	0.043844	0.049257	0.082672
min	−0.093121	−0.11514	−0.069749
std	0.018947	0.020342	0.019513
mode	−0.093121	−0.11514	−0.069749
kurtosis	6.022	8.2936	5.3142
skewness	−0.5132	−1.2369	0.079493
range	0.13697	0.1644	0.15242
sharpe	−0.17308	−0.19485	−0.085121
t-test	−2.4721	−2.783	−1.2158

Table 1 - Descriptive statistics of E-S-G portfolios

Table 1 contains the descriptive statistics of the three investigated portfolios, E S and G, drawn by the average monthly returns over the analysed period. The distribution is approximately normal, even though all of the modes are quite far from the mean and the median. Also respective kurtosis measures confirm the occurrence of outliers: in fact, all the three distributions are leptokurtic, with values above 3. Their skewnesses provide further details regarding the shape of the distributions: the first two tend to have more “negative events”, being left tailed, meaning that among the returns which fall farther from the mean, the majority is negative. G, on the other hand, is right tailed, given its positive skewness, with more returns that fall in the positive extreme of the distribution. The value is not considerably different from zero. In fact, it can be asserted that among the three, G is the one resembling a normal distribution the most.

With a sufficiently large sample size, 204 per each portfolio, (i.e. monthly time series data from January 2002 until December 2018), anyways, it is possible to take advantage of the central limit theorem to assume the normalilty attributes of these distributions.

A first glimpse of the performance of these ESG investing is given by the average monthly return across the defined time period, which is slighly negative, even though not particularly different from zero (above all, the G factor). The table also provides a t-test for the mean of each portfolio. E and S have a statistically different from zero average monthly return, as confirmed by the relative statistics. The same cannot be said for G, for which the t-stat is approximately -1.215.

According to these results, investors would have earned negative (in excess of the risk free rate) returns should they have invested into Environment and Social aware firms. In fact, the comulative returns for the Environment, Social , and Governance factor-mimicking portfolios are respectively: -50.96%, -57.40%, and -31.47%. This could be indicative of a general tendendecy by European investors of not valuing ESG investments, which lead them not to invest into these firms. Maintaining the assumption that prices are determined by expectations, negative returns could be interpreted as investors not believing that firms investing into actvities that produce positive externalities for their stakeholders, other than shareholders, will produce financial value. Or at least, that it is more profitable to invest into firms which do not pay much attention to their social performance. The above mentioned results seem to confirm the position that, on average, for European firms, corporate financial performance is not related to corporate social performance. Another interpretation that could be given to these numbers, though, relates to perceived riskiness of the most sustainable firms. In fact, sustainable firms, given their characteristics, could be perceived as less risky assets, thus worthing a lower return. On the contrary, *irresponsible* firms might be felt as riskier, overall, hence investors require a higher return for investing into them. Since the portfolio performance is derived by taking a long position in the former and a short in the latter, the net effect would be a negative performance. To further test this hypothesis, the average return of the High scored and the Low scored portions of each portfolio have been individually analysed during the financial recession. The underlying hypothesys is that riskier firms would perform worse relative to less volitile peers, in a period like that, hence confirming that sustainable firms are perceived as safer. Results do not support this hypothesys. In fact, sin firms consistently outperform those in the high portion of the portfolio, under each specification. Moreover, monthly average returns for *unethical* firms are all statistically significantly different from zero at the 10% significance level, whereas the same cannot be said for counterparties. With respect to G, it is not possible to

statistically infer that an investor would have earned an average return different from the risk free rate if he invested into this portfolio, as the mean monthly excess return is not statistically different from zero. Since the portfolios are constructed in a HML fashion, these numbers do not offer appropriate granularity to determine whether the negative performance of the portfolios is due to sustainable firms performing poorly or to firms ranking low in these aspects, but performing better. By decomposing the portfolios, more details can be drawn. Hence, additional analysis is performed on the sample. With respect to E, the negative performance of the firms ranking high in terms of environment seems to be more relevant, in absolute terms, relative to less aware firms (with average monthly excess returns of -0.12% and 0.10% respectively). S and G results, on the contrary, seem to be driven by the higher returns of low-sustainablity portfolios (0.17% and 0.65% respectively). These results, anyways, do not allow to infer any detail regarding the population, since t-stats for all of the figures are not sufficient to determine a statistically significant difference from zero for all the portfolios.

A probable scenario is that the sample results are affected by the 2006-2008 financial crisis, which by the way seems to have beaten sustainable portfolios much harder: in fact, Mkt-Rf SMB HML and MOM portfolios still maintain a positive avearge monthly return throughout the 2002-2018 period (relative descriptive statistics can be found in Table 2). It is worth to notice that, with the only exception of MOM portfolio, none of the analysed average excess returns are statistically different from zero as confirmed by the respective t-stats. This suggests some doubts on the ability of the Fama French original specifcation to explain risk, further evidence is given in the cross-sectional analysis.

Individually studying standard deviations, E seems to be the least volatile, with a figure below those of the others. This is surely a nice property for risk averse investors, especially if looking at its pre-crisis performance (though, few years are available to have a proper analysis): average monthly excess return was 0.29% from 2002 to the end of 2005 with a standard deviation of 0.019, leading to a Sharpe ratio of approximately 15.6%, which in annualized figures amounts to 54.04%. This result is below, but still in line with the trend of Mkt-Rf and MOM (respectively, 79.67% and 90%). To give a sense of the magnitude of these results, S and G, on the same annualized item, score -4.15% and -5.2% respectively (results are not reported).

S standard deviation is the highest among the three, thus indicating a higher volatility attached to it. Cross-checking the results of Table 1 with those in Table 3 and Table 4, containing the summary statistics of the individual time-series regressions of factor-mimicking portfolios E and S vs the Fama-

French-Carhart specification, it can be concluded that much of the risk associated to them can actually be explained by the already defined risk factors. In fact, even if R^2 is not particularly high in both cases, results show a non irrelevant loading especially on SMB, with coefficients of 0.503 and 0.515, respectively, which are statistically significant at 95% confidence interval – as highlighted by the p-values of both.

G looks volitile as well, even though to a lesser degree with respect to S, in the universe of these constructed portfolios.

A visual representation of the volatility and the overall performance of the ESG portfolios is given in Figure 1, which plots the cumulative returns for all the three assets between 2002 until 2018. The graph clearly shows that the portfolios have generated positive returns in the early years of the 2000s. Since the beginning of 2008, though, the portfolios have consistently negatively performed. The plot also show the similar behaviour E and S have had throughout the whole period. In fact, with the only exception of the very beginning of the sample, the two portfolios have shown an almost identical trend (also magnitudes look very similar). This consideration strenghtens the point that the two sustainability matters, in the sample, look particularly related. Hence, it is possible to assume that firms that score high (low) in one item have a strong (weak) performance also in the other. The chart, in addition, displays how volatile the three portfolios are. One example of such a property is the 2008 drop of the three assets. All of the three lost more that 10 percentage points, but S (with a monthly stantard deviation of 0.02, the highest of the three) has lost almost 200 basis points.

In terms of risk adjusted returns, eventually, from Table 1, it can be stated by the Sharpe ratio (computed as average monthly return over standard deviation) that portfolios had negative performances, driven by below-zero average returns. In annualized terms, they have realized -247.21%, -278.30%, and -121.58%.

A closer look at the dynamics of the portfolios yields valuable insights regarding their nature. Since E and S seem fairly similar in their behaviour, an analysis of the their correlation is performed. As expected, the two portfolios show a meaningful correlation (0.72). The assumption behind this result is that there is no clear dinstiction between an environmentally strong firm and a socially resposnible one. This blurred line may be caused by a bias in the rating process (i.e. scores given to each firm). Else, it could be explained by a mutual characteristic which makes the two items behave similarly. That is, firms with good (bad) environmental performance, tend to have good (bad) social performance as well. By consequence, their ratings under the two matters are very similar. Thus,

portfolios are very similar in their composition. Moreover, the two factor-mimicking portfolios show some sort of correlation also with SMB, versus which E (S) scores 0.49 (0.43). This is a first evidence of both a relation between these two portfolios and also with the Fama-French-Carhart specification. Governance does not show any particularly relevant correlation neither with the other two ESG items, nor with the four factors specification. Hence, returns movements do not seem to be comparable to any of the other factors. The intuition is that the Governance attribute identifies a more defined characteristic of firms.

In conclusion, E and S show some similarity and their mutual heavy and significant (negative) sensitivity towards the SMB factor-mimicking portfolios does not make Environment and Social scorings eligible as risk factors, since their performance seems already to be partially explained by the factor itself. G, which has a slightly different performance, seems not to be explained by any of the Fama-French-Carhart factors, for alpha=5%. Hence, Governance may be a candidate as individual risk factor, in investigating whether sustainability is priced in the European market.

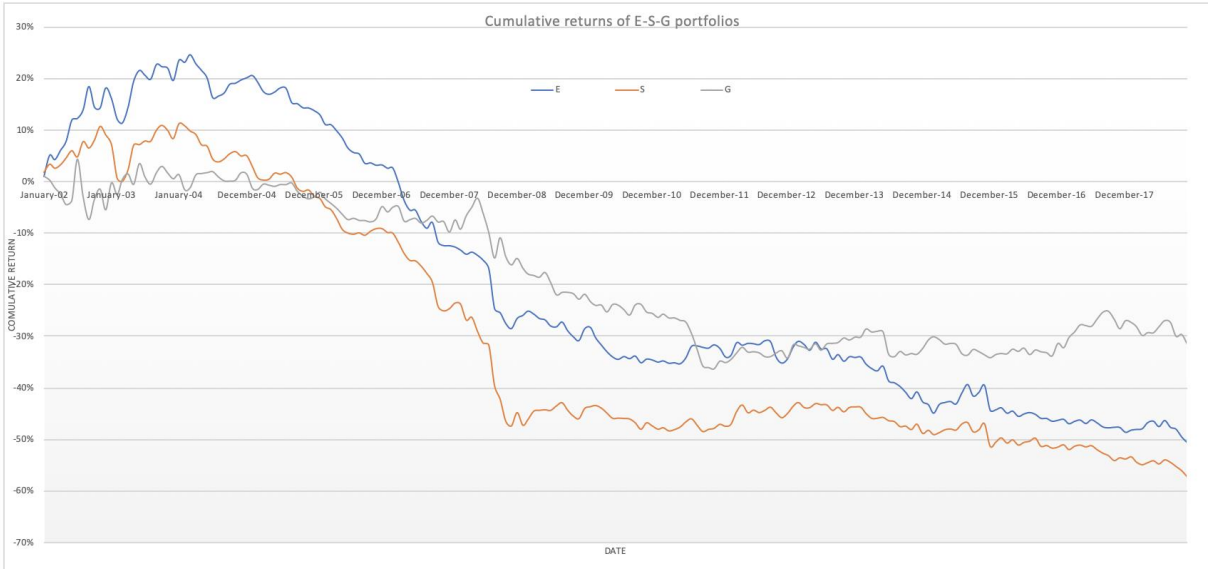

Figure 1 - Cumulative returns of E, S, G factor-mimicking portfolios 2002-2018

	MktRF	SMB	HML	WML
mean	0.005576	0.0021539	0.0021926	0.0087353
median	0.00645	0.002	0.00265	0.01115
max	0.1374	0.0486	0.0829	0.1369
min	−0.2206	−0.0682	−0.0472	−0.2627
std	0.052573	0.018714	0.021537	0.039783
mode	−0.1	−0.0077	0.0049	0.0064
kurtosis	4.6916	3.6618	3.786	14.449
skewness	−0.59824	−0.31357	0.35011	−1.7577
range	0.358	0.1168	0.1301	0.3996
sharpe	0.10606	0.1151	0.10181	0.21957
t-test	1.5149	1.6439	1.4541	3.1361

Table 2 - Descriptive statistics Fama-French-Carhart four factors specification

Analysis of the sustainability factor-mimicking portfolios in relation to the Fama-French four factors model

The three tables below (namely Table 3, Table 4, and Table 5) show the results of the time-series regressions of each portfolio versus the Fama-French-Carhart Factors during the mentioned time period. The coefficient estimates represent each portfolio sensitivity versus each of the four portfolios in the specification.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	−0.0046658	0.0012213	−3.8202	0.00017814
Mkt-Rf	0.05231	0.02642	1.9799	0.049088
SMB	0.50335	0.062507	8.0526	7.278e−14
HML	−0.11107	0.060444	−1.8376	0.067607
MOM	0.029094	0.033476	0.86912	0.38583

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0164
R-squared: 0.261, Adjusted R-Squared 0.246
F-statistic vs. constant model: 17.6, p-value = 2.25e−12

Table 3 - Summary statistics time series regression of E against the 4 Fama-French-Carhart factors

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	-0.0049308	0.0012806	-3.8504	0.00015893
Mkt-Rf	0.12152	0.027702	4.3867	1.8646e-05
SMB	0.51547	0.06554	7.865	2.2983e-13
HML	-0.14068	0.063377	-2.2198	0.027564
MOM	-0.058648	0.0351	-1.6709	0.096319

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0172
R-squared: 0.295, Adjusted R-Squared 0.281
F-statistic vs. constant model: 20.8, p-value = 2.28e-14

Table 4 - Summary statistics time series regression of S against the 4 Fama-French-Carhart factors

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	-0.0021648	0.0014327	-1.5111	0.13236
Mkt-Rf	0.056061	0.030992	1.8089	0.071973
SMB	0.14382	0.073323	1.9615	0.051216
HML	-0.13557	0.070903	-1.9121	0.057303
MOM	0.020462	0.039268	0.52107	0.6029

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0193
R-squared: 0.0415, Adjusted R-Squared 0.0222
F-statistic vs. constant model: 2.15, p-value = 0.0756

Table 5 - Summary statistics time series regression of G against the Fama-French-Carhart specification

A few words regarding the choice of R_f: this parameter has been selected since data regarding Fama-French-Carhart portfolios have been downloaded directly from Kenneth French's website. In the details, it is clearly stated that the risk-free rate is to be intended as the one-month T-bill rate. Hence, for the sake of consistency across data, the same rate has been applied to the investigated assets - and to the sample stocks – excess returns. Moreover, according to OECD data, US and Europe (in the meaning of European Union) inflation rates have been similar throughout the

considered time window, making the real interest rates equivalent. Hence, it would not make any significant difference using one instead of the other.

Results seem to confirm the initial intuition derived by the diagnostic tests run on the E S G portfolios: E and S exhibit a significant loading on the already defined asset pricing model, whereas G seems not to be explained by it. In favor of this view, the first two regressions both highlight a higher adjusted R^2 with respect to G (0.246 and 0.281, respectively, against 0.022). Moreover, the whole model seems to be significant for the first two portfolios, with p-values considerably below the rejection threshold, whereas the G figure lies above 0.05. Going deeper in the analysis, E shows significant correlation with the first two portfolios in the FFC model (Mkt-Rf and SMB): 0.05 and 0.50 estimates are significant at the 95 % level of confidence. This would lead to the conclusion that adding E as a risk factor would not help to explain better the cross-sectional variation of stock returns. This result indicates that environment aware firms tend to perform similarly to small stocks. The same holds for S: in fact, a relevant portion of its behavior can be attributed again to SMB, with an estimate of almost 0.515 (where the p-value is less than 0.05). Moreover, the excess return of the market portfolio seems to play a role in explaining the behavior of this asset with a statistically significant estimate of 0.12. Interestingly, it seems to have a significant negative loading on HML, equal to -0.14. This is a common characteristic of the investigated portfolios – at differing confidence levels -, which could lead to the conclusion that on average growth stocks tend to outperform value stocks in terms of sustainability.

To conclude, for the purpose of this study to investigate whether the sustainability is priced, results referring to E and S would be of a limited relevancy, as these preliminary tests demonstrate that with a very limited probability of committing a type I Error, we should reject the null hypothesis that there is no relationship between the FFC factor-mimicking portfolios and the two responsible portfolios constructed.

Nevertheless, G's explanatory power should be further investigated, as time-series analysis versus the already mentioned model provides different results. In first instance, it demonstrates no important loading versus any of the Fama-French-Carhart portfolios, differently from the other two. Moreover, all of the results have higher p-values, leading to the conclusion that it is not possible to reject H_0, that is: there seems to be no relationship between the Governance factor-mimicking portfolios and the 4 regressors.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.001592	0.00066782	2.3839	0.017348
E	−0.0024225	0.00062154	−3.8976	0.00010475
S	−0.00065199	0.00060491	−1.0778	0.28141
G	−0.0019721	0.00058473	−3.3726	0.00077793
Mkt-Rf	0.0075634	0.0010194	7.4196	2.8285e−13
SMB	0.0038247	0.00041274	9.2665	1.5165e−19
HML	−0.0024748	0.00044369	−5.5777	3.2671e−08
MOM	0.01245	0.00082903	15.017	1.9997e−45

Number of observations: 865, Error degrees of freedom: 857
Root Mean Squared Error: 0.00719
R-squared: 0.44, Adjusted R-Squared 0.435
F-statistic vs. constant model: 96, p-value = 2.74e-103

Table 6 - Cross sectional regression of average monthly excess returns vs time-series betas of 7 factor-mimicking portfolios

Table 6 shows the summary output of the cross-sectional regression of the time series mean of the excess returns of the stocks in the sample, and the time series betas of the factor-mimicking portfolios as regressors.

Results seem to be in line with expectations. E commands a statistically different from zero premium, even though of a very limited degree. The same holds for the 4 FFC portfolios, among which MOM ranks first in terms of weight. Surprisingly, S is not statistically significant, as opposed to previous results which have shown dependence between the social risk factor and the 4 factors specification. To further investigate this phenomenon, an additional regression is run taking as regressors only these 5 independent variables -of which results are not displayed. Accordingly, the coefficient estimate of S becomes statistically significant, even if of limited amount (it leads a 0.14% premium, with a p-value of 0.00012). It could be asserted that, given these values, there is some linear combination which is explained by one of the other two studied sustainability components. Hence, further analysis is performed. By adding as control variable Environment, S loses its statistical relevance. This confirms the initial hypothesis of some sort of relationship between E and S – results of these last study are summarized in Table 7. They strengthen the point of a relevant correlation between the two firm characteristics.

In general, in any case, the model does not seem to explain much of the cross-sectional variation of the average excess returns of the stocks in the sample. In fact, the alpha is positive and statistically different from zero. If the model was able to explain the cross-sectional variation of stock returns, we would expect abnormal excess returns to be indistinguishable from zero. Further analysis is done to understand whether also the original specification still has explanatory power regarding cross sectional risk or not. Results are shown Table 8, which contains the summary output of the cross-sectional regression of the average excess returns against the sensitivities of the stocks in the sample versus the Fama-French-Carhart specification across the time window 2002-2018 (i.e. time series betas). Both the models do show a statistically significant different from zero alpha (Intercept), which denotes the inability of the two specifications to explain risk. In fact, firms have earned abnormal excess returns, not described by the model. All the results in Table 8, statistically significant at the 95% confidence interval, show low coefficients. The main intuition behind this is that since these risk factors were first established in 2002, as of today, they could be not as relevant as they used to be. In fact, as the results of the original study performed by Eugene Fama and Kenneth French were made public, many may have undertaken similar investing strategies. This spread behavior is most likely to have reduced returns from investing into them. Accordingly, risk attached to the factors has diminished as well. This could lead to the conclusion that the FFC specification is not representative of risk anymore.

Going deeper in the analysis of Table 6, some more research is worth for G. In fact, as already stated, the portfolio has shown a different behavior if compared to the other two, as it seems not to be already incorporated into the already established model. Under the null that there is no relationship between the portfolio and the cross-sectional variation of excess returns, we can reject H_0: Governance commands a negative premium of 0.0019, with a t-stat of -3.37. Accordingly, we can statistically infer that G has power to describe the cross-sectional variation of stock returns, given its statistically significant coefficient estimate. In particular, it commands a negative, small, premium. Results in Table 9, which describe the regression of the factor-mimicking portfolio against the average cross-sectional monthly returns confirm this. In fact, when the only independent variable is G, its coefficient estimate is still negative (-0.0039) and statistically significant (t-stat equal to -5.53).To collect further information, an additional analysis has been performed: a cross sectional regression of sample average excess returns against G, adding the Fama-French-Carhart specification as control variables. Results are displayed in Table 10 and show that G does not lose its statistical significance (t-stat of-3.39), and the coefficient estimate is still negative (-0.0019). Thus,

under different specifications, G consistently has statistically relevant negative premiums. Thus, for higher performances of firms under governance matters, we can expect lower returns, also when other risk factors are taken into account.

The main intuition we can derive based on these results is that this negative relation can be meant as the "*cost of doing the right thing*". The result is in line with other research papers on the topic and it could be interpreted as an investors' behavior. That is, they would be willing to give up some profits in order to hold portfolios which reflect their personal set of values. This conclusion adds evidence to behavioral theories which assert that investors are biased in their choices, hence they fail to maximize the efficiency of their portfolios. Specifically, investors would be better off by investing into firms with lower performance on Governance.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.0017605	0.00066996	2.6277	0.0087493
E	-0.0019718	0.00061066	-3.2289	0.0012899
S	-1.6507e-05	0.00057828	-0.028545	0.97723
Mkt-Rf	0.0074057	0.0010244	7.229	1.0764e-12
SMB	0.0040619	0.00040915	9.9276	4.6269e-22
HML	-0.0027805	0.00043695	-6.3633	3.2111e-10
MOM	0.01268	0.00083119	15.256	1.1318e-46

Number of observations: 865, Error degrees of freedom: 858
Root Mean Squared Error: 0.00723
R-squared: 0.432, Adjusted R-Squared 0.428
F-statistic vs. constant model: 109, p-value = 6.85e-102

Table 7 - Cross sectional regression of average excess returns vs time-series betas of E S and FFC specification

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.002068	0.00066016	3.1327	0.001791
Mkt-Rf	0.0079906	0.001029	7.7652	2.314e-14
SMB	0.0042226	0.00040909	10.322	1.2548e-23
HML	-0.0032569	0.00043191	-7.5406	1.188e-13
MOM	0.011623	0.00079814	14.563	4.1819e-43

Number of observations: 865, Error degrees of freedom: 860
Root Mean Squared Error: 0.00733
R-squared: 0.415, Adjusted R-Squared 0.413
F-statistic vs. constant model: 153, p-value = 1.13e-98

Table 8 - Cross sectional regression of average excess returns vs time-series betas of FFC specification

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.005154	0.00032652	15.785	1.6804e-49
G	-0.0039048	0.00070487	-5.5398	4.0227e-08

Number of observations: 865, Error degrees of freedom: 863
Root Mean Squared Error: 0.0094
R-squared: 0.0343, Adjusted R-Squared 0.0332
F-statistic vs. constant model: 30.7, p-value = 4.02e-08

Table 9 - Cross sectional regression of average excess returns against G time-series betas

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.0017413	0.00066318	2.6257	0.0088002
G	−0.001909	0.00056244	−3.3941	0.0007201
Mkt-Rf	0.0082013	0.0010247	8.0039	3.8998e−15
SMB	0.0040505	0.00040976	9.8852	6.7525e−22
HML	−0.0029668	0.00043772	−6.7778	2.2664e−11
MOM	0.011608	0.00079332	14.632	1.8712e−43

Number of observations: 865, Error degrees of freedom: 859
Root Mean Squared Error: 0.00728
R-squared: 0.423, Adjusted R-Squared 0.42
F-statistic vs. constant model: 126, p-value = 5.02e−100

Table 10 - Cross sectional regression of average excess returns against G and FFC factor-mimicking portfolios time-series betas

Chapter 4

Conclusion

In this research paper I have investigated if the Environmental, Social, and Governance factors are priced in the European stock market. To do so, I have run a Fama-Macbeth procedure to examine the cross-sectional variation of stock returns and its relationship with ESG factor mimicking portfolios, identified by means of third-party scores. Moreover, I have also studied the three portfolios in relation to the Fama-French-Carhart four factors specification, to understand the dynamics of the investigated assets.

In general, ESG investing has not been profitable throughout the chosen time period (i.e. 2002-2018). All of the three factor-mimicking portfolios have performed negatively in the average month. Their cumulative returns confirm a strong negative performance. This is true especially after the financial depression happened between 2006 and 2008, after which none of them has been able to provide a positive return, in aggregate terms. That is, if an investor had invested into these portfolios, it would have constantly earned less than its initial investment. Further analysis of this

phenomenon has shown that the negative performance of the Environment factor-mimicking is explained by the returns of virtuous firms which are worse, in absolute terms, with respect to returns yielded by *sin* firms. The other two portfolios are characterized by a better performance of the bottom quintiles.

Preliminary tests over the assets have highlighted a statistically relevant relationship between E, S and the Carhart specification. Hence these two should not be considered as capable of explaining excess returns. Their explanatory power, in fact, is already enclosed in the four factors specification. In particular, concerning E, SMB has a strong and statistically significant coefficient estimate. With respect to S, both the Market factor and SMB have a robust and statistically significant coefficient estimate. Moreover, E and S are proven to have a strong correlation. This is assumed to be explained either by a bias in the scoring process, or by a common element which drives both the items.

The same cannot be told for G, which shows no significant dependence on any of the factor-mimicking portfolios for alpha = 5%. This result is also reinforced by the low correlation of Governance with any other portfolio. The time series regressions of the factor-mimicking portfolios over the Carhart specification have highlighted, even though at differing confidence levels, an overall negative relationship between sustainability and the HML factor. This result is in line with previous literature. The main intuition behind it is that investors' demand for responsible firms is relevant. Hence, book-to-market ratios are lower for highly sustainable companies. By definition, their performance resembles that of growth firms.

The model applied seems not to be able to entirely explain the cross-sectional variation of stocks' returns. In fact, the resulting intercept of the cross-sectional regression between stocks returns in the average month and the betas representing the sensitivities of the sample stocks versus the factor-mimicking portfolios is positive and statistically different from zero, thus firms earn abnormal returns not explained by the control variables. Moreover, S has a non-statistically significant coefficient estimate. Further investigation demonstrates that there exists some sort of linear relationship between E and S (Bali, Engle, and Murray, (2016)). This adds to previous outcomes obtained concerning their correlation. Results for G, which seems to be the best candidate as risk factor given its independence on the Carhart specification, show a statistically significant negative cross-sectional relationship between the factor-mimicking portfolio and European average stocks excess returns. Further tests, to check the robustness of the result, strengthen this negative relationship: the portfolio is found to have a negative and significant coefficient estimate under different specifications.

After having investigated the research question of whether the Environmental, the Social, and the Governance factors are priced in the European stock market, this study concludes that G commands a slightly negative monthly premium, which is interpreted – in line with previous research on the topic – as the cost of incorporating governance-related considerations in portfolio setting. On the other hand, since E and S seem to be already explained by previously defined model, they do not command any risk premium which is not already enclosed in the Fama-French-Carhart specification.

According to the results this paper adds to the literature, future research may want to investigate further the dynamics between Environment and Social factors to understand the roots of their mutual relationship. Both the assumptions could be analyzed: on the one hand if by applying a different scoring process, this relationship vanishes. On the other, questioning whether the factors Environment and Social truly identify something distinctive or not. Moreover, it could also dig deeper in the meaning of the negative premium which the Governance factor-mimicking portfolio demonstrates: is it really the case that investors have to give up some profits to include virtuous companies in their portfolios?

References

- Friedman, M. (1970). The Social Responsibility of Business is to increase its Profits. *The New York Times*.
- Di Giuli, A. & Kostovetsky, L. (2014). Are red or blue companies more likely to go green? Politics and corporate social responsibility. *Journal of financial economics*, vol. 111, issue 1, 158-180
- Albuquerque, R. Durnev, A. Koskinen, Y. (2014) Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence.
- Eccles, R., & G., Serafeim. 2013. The Performance Frontier: Innovating for a Sustainable Strategy. *Harvard Business Review* 91, no. 5: 50–60.
- Eccles, R. & Klimenko, S. (2019). Shareholders are getting serious about sustainability. *Harvard Business Review*. May–June 2019 (pp.106–116)
- Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. Journal of Corporate Finance, 14(3), 302-322. doi:10.1016/j.jcorpfin.2008.03.009
- https://www.businesswire.com/news/home/20191015006005/en/Index-Industry-Association's-Annual-Survey-Finds-2.96 (Accessed on 10/11/2019)
- https://www.businesswire.com/news/home/20181114005124/en/Index-Industry-Association-Survey-Reveals-3.7-Million (Accessed on 10/11/2019)
- Fiskerstrand, S. Fjeldavli, S., Leirvik, T., Antoniuk, Y., & Nenadić, O. (2019): Sustainable investments in the Norwegian stock market, Journal of Sustainable Finance & Investment, DOI: 10.1080/20430795.2019.1677441
- Halbritter, G., & Dorfleitner, G. (2015). The wages of social responsibility — where are they? A critical review of ESG investing. Review of Financial Economics, 26(1), 25-35. doi:10.1016/j.rfe.2015.03.004
- http://zeerovery.nl/blogfiles/esg-scores-methodology.pdf (accessed on 05/01/2020)
- Khan, Mozaffar and Serafeim, George and Yoon, Aaron. (2016) Corporate Sustainability: First Evidence on Materiality. The Accounting Review, 91 (6), 1697-1724.
- Henriksson, R., Livnat, J., Pfeifer, P., & Stumpp, M. (2019). Integrating ESG in Portfolio Construction. The Journal of Portfolio Management, 45(4), 67-81.

- Limkriangkrai, M., Koh, S., & Durand, R. B. (2017). Environmental, social, and governance (ESG) profiles, stock returns, and financial policy: Australian evidence. International Review of Finance, 17(3), 461-471.
- Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. Financial management, 35(3), 97-116.
- Mănescu, C. (2011). Stock returns in relation to environmental, social and governance performance: Mispricing or compensation for risk?. Sustainable development, 19(2), 95-118.
- Annér, L., & Jakobsson van Stam, N. (2018). The effects of Environmental, Social and Governance measures on the cross section of stock return, a compensation for risk or mispricing? Evidence from the Swedish stock market.
- Thomson Reuters (2018) Thomson Reuters esg scores. url: https://financial.thomsonreuters.com/content/dam/openweb/documents/pdf/financial/esg-scores-methodology.pdf.
- Fama, E. F. and French, K. R. (1993) Common risk factors in the returns on stocks and bonds, Journal of financial Economics 33, 3–56.
- https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
- Bali, T. G., Engle, R. F., and Murray, S. (2016) Empirical asset pricing – the cross section of stock returns. New Jersey: John Wiley Sons, Inc.
- Hąbek, P., & Wolniak, R. (2013). European Union regulatory requirements relating to Sustainability Reporting. The case of Sweden. *Zeszyty Naukowe/Akademia Morska w Szczecinie*.
- Galema, R., Plantinga, A., & Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance*, *32*(12), 2646-2654.
- Jang, Y. E. (2019). Do ESG scores matter in the market? Environmental, Social and Governance performance in relation to stock returns and profitability in European Market.
- Serafeim, G. (2018). Public Sentiment and the Price of Corporate Sustainability. *Harvard Business School Accounting & Management Unit Working Paper*, (19-044).
- Xiao, Y., Faff, R., Gharghori, P., & Lee, D. (2013). An empirical study of the world price of sustainability. *Journal of business ethics*, *114*(2), 297-310.
- Khan, M. (2019). Corporate Governance, ESG, and Stock Returns around the World. *Financial Analysts Journal*, *75*(4), 103-123.
- PricewaterhouseCoopers, L. L. P. (2015). Global financial markets liquidity study. *report prepared for the Global Financial Markets Association and the Institute of International Finance*.

- EY and Boston College Center for Corporate Citizenship, B. c. C. s. o. m. (2013) Value of sustainable reporting.

In the last decade, ESG investing got under the spotlight. Ultimately, also in the finance area. In fact, academics, as well as asset managers, have started considering this strategy and investigating whether it could eventually be profitable to take into account people's personal beliefs and values in setting portfolios. Companies themselves have shown their interest in this matter as well, by disclosing their performance in terms environmental footprint, respect for company's stakeholders, and their adherence to corporate governance best-practices, among other things.

Before diving into the details, it is worth to define ESG as well as its bounds with Asset pricing. If, on the one hand, the acronym stands for Environment, Social, and Governance, the concept behind, on the other, refers to the contribution an individual company makes to sustainability. Thus, ESG is referred to as the comprehensive performance a firm has under the different specifications: environment (e.g. Green-house gases emissions, environmental footprint of its products and processes) social (e.g. Labor conditions it offers to its employees, sensitivity versus stakeholders' interests) governance (e.g. Understanding of minority shareholders' interests, best practices). Companies are increasingly more concerned about the topic. One of the reasons could be that people, hence investors, care more about it, thus firms need to align their performances to the demand expectations. Another reason could be found in countries strengthening regulatory frameworks in these regards. So, corporations need to update their production processes to not fall short on legal requirements. Any of these could impact stocks' performance. If investors get more and more aware of sustainability, their decisions will be likely to be influenced by judgements on ESG basis. Hence, integrating this kind of considerations could improve the performance of portfolios. Given the increased attention paid by companies to these matters, it is worth to study if this phenomenon is also reflected into the stocks' market. EY and the Boston College Center for Corporate Citizenship, for example, have conducted a joint study which has shown that firms with high quality sustainability reporting practices have obtained easier access to capital, as well as higher company value.

Much research, accordingly, has focused on answering multiple research questions regarding the issue. First and foremost, whether the strategy of integrating ESG performance in portfolio setting is profitable. The evidence is mixed: part of the studies does not find positive performance of sustainable portfolios relative to sustainability-indifferent ones. Literature in these regards could be divided into two main areas. Some shifts the attention on the more "managerial" side of the issue, showing concerns regarding the overall strategy of chasing other than financial results. In fact, investing company's resources to achieve results which are not directly reflected into the year-end

bottom line would be detrimental to the health of the venture. Eventually, the performance of these companies would be negative with respect to their peers, thus making them an undesirable investment.

Others look at the topic from a portfolio management standpoint. The logic behind would relate to a fundamental aspect of portfolio setting: risk. To minimize it differentiation is the key. Since many strategies to integrate sustainability considerations in portfolio setting include a negative screening process, in which companies are excluded given their non-responsible characteristics (e.g. participation in a "*sin industry*", like tobacco), the net effect is that of reducing the set of available firms in which to invest. Moreover, this is likely to produce a portfolio which concentrates on few industries. The ultimate consequence is that of having similar returns to other portfolios, with increased volatility. Similar studies assert that the negative returns associated with sustainable investing could be considered as the cost of doing the right thing. Meaning that investors should be ready to afford this relative negative performance, in order for their portfolios to reflect their set of principles.

On the other hand, some alternative research finds a positive return in ESG investing. In this matter as well, supporting evidence is provided from different angles. Some supporters of this view highlight a positive relationship between sustainable behavior and returns which could be not as easily measurable as financial ones (e.g. efficiency). Some others find that thanks to their ability to anticipate legislative changes, responsible firms tend to outperform lagging competitors. Also, other studies demonstrate that those firms' risk-adjusted returns are higher.

Not much of the research, though, has produced a step further and analyzed the three commonly recognized sustainability elements, Environment Social and Governance, individually. Most likely, the main reason behind this shortage is the lack of data regarding specific ESG criteria. As of today, it is extremely challenging to find specific responsibility components-driven portfolios, or even proxies. This is most probably due to the fact that sustainable finance is an ever-increasing phenomenon, but still very recent. Consequently, very few data are available to conduct massive research on that.

Moreover, research lacks studies conducted over European samples. Hence, the intent of this paper is that of investigating whether the three singular sustainability components are priced in the European market. To do so, three portfolios, each of them reflecting one determinant of ESG (accordingly, one portfolio for Environment, one for Social, and one for Governance) have been constructed, in a Fama-French High-minus-Low fashion. The sample chosen is the STOXX Europe

600. In order to determine the sustainability performance of sample stocks, data have been collected on Thompson Reuters database, which offers the percentile ranking of stocks, against each item, since 2001. Before testing the research question, some preliminary analysis has been conducted over the three factor-mimicking portfolios, to determine whether they are good candidates to explain risk. To do so, they have been tested against the Fama-French-Carhart four factors specification. Results show that E and S can be explained by the already defined asset pricing model. G shows no statistically relevant relationship with any of the risk factors. Nonetheless, coherent with previous research, all the ESG items display a negative relationship with the value factor mimicking portfolio. This would suggest that sustainable firms tend to perform similarly to growth stocks.

Finally, to test the research question, a standard Fama-Macbeth procedure has been conducted over a 2002-2018 monthly time period. The four factors specification has been added to the model to have some further controls. Results highlight a negative and statistically different from zero explanatory power of Governance in relation with the cross-sectional variation of average stocks' returns in the sample period. Additional analysis is also performed to investigate further the behavior of the factor-mimicking portfolio.

Once data over ESG performance have been collected, portfolios were set according to previous year's sustainability score, per each element. Prices for every company that has been listed on the STOXX 600 – to avoid any survivorship bias - have been collected throughout the period, and monthly returns were computed. Sustainability factor-mimicking portfolios have been constructed in an HML fashion (i.e. top scores minus bottom scorers). Then, the Fama-Macbeth procedure was applied.

The first time-series regression, which gives back the sensitivity of each stock versus the portfolios, is:

$$y_1 = \alpha_1 + \beta_1 Mkt - Rf_1 + \beta_1 SMB_1 + \beta_1 HML_1 + \beta_1 MOM_1 + \beta_1 E_1 + \beta_1 S_1 + \beta_1 G_1 + e_1$$
$$y_2 = \alpha_2 + \beta_2 Mkt - Rf_2 + \beta_2 SMB_2 + \beta_2 HML_2 + \beta_2 MOM_2 + \beta_2 E_2 + \beta_2 S_2 + \beta_2 G_2 + e_2$$

.

.

.

$$y_T = \alpha_T + \beta_T Mkt - Rf_T + \beta_T SMB_T + \beta_T HML_T + \beta_T MOM_T + \beta_T E_T + \beta_T S_T + \beta_T G_T + e_T$$

The resulting betas reflect each company's sensitivity versus the factor-mimicking portfolios. As such, the effect of sustainability is intrinsically measured by means of how the 3 portfolios are constructed: they are rebalanced each year, on the basis of stocks' scores in each specification at previous year end.

The column vectors are transposed and stored in an *N x K* matrix, called *Beta*, having in each column the time series coefficient estimates per each firm, for any factor-mimicking portfolio:

$$Beta = \begin{bmatrix} \hat{\beta}_{1,Mkt-Rf} & \cdots & \hat{\beta}_{1,G} \\ \vdots & \ddots & \vdots \\ \hat{\beta}_{N,Mkt-Rf} & \cdots & \hat{\beta}_{N,G} \end{bmatrix}$$

The second step of the procedure involves a cross-sectional regression of each of these estimates on the *N* stocks monthly excess returns. To save computations, a vector of time-series average stocks monthly excess returns is derived, which is then used as dependent variable in the upcoming cross-sectional regression:

$$Y = \begin{bmatrix} E[y_1] \\ E[y_2] \\ \vdots \\ E[y_N] \end{bmatrix}$$

Where,

$$E[y_i] = \frac{1}{T}\sum_{t=1}^{T} y_t$$

This vector is regressed against the *N x K* Beta matrix. The aim of this step, as suggested by Bali, Engle, and Murray, (2016), is to understand whether there is any significantly different from zero relation between the portfolios and the excess returns in the average month. Moreover, if the model is able to explain the cross-sectional variation of stock returns, the resulting alpha should be not statistically different from zero. In fact, firms are not supposed to earn abnormal returns not due to any of the variables in the model. In case this happens, it can be concluded that part of the risk is not explained by the model itself.

A preliminary analysis of the three newly derived ESG portfolios has been done.

	E	S	G
mean	−0.0032793	−0.0039636	−0.001661
median	−0.0031337	−0.0022438	−0.002089
max	0.043844	0.049257	0.082672
min	−0.093121	−0.11514	−0.069749
std	0.018947	0.020342	0.019513
mode	−0.093121	−0.11514	−0.069749
kurtosis	6.022	8.2936	5.3142
skewness	−0.5132	−1.2369	0.079493
range	0.13697	0.1644	0.15242
sharpe	−0.17308	−0.19485	−0.085121
t-test	−2.4721	−2.783	−1.2158

Table 11 - Descriptive statistics of E-S-G portfolios

Table 1 contains the descriptive statistics of the three investigated factor-mimicking portfolios, E S and G, drawn by the average monthly returns over the analysed period. The distribution is approximately normal, even though all the modes are quite far from the mean and the median. All the three distributions are leptokurtic, with values above 3. Their skewnesses provide further details regarding the shape of the distributions: the first two tend to have more “negative events”, being left tailed, meaning that among the returns which fall farther from the mean, the majority is negative. G, on the other hand, is right tailed, given its positive skewness, with more returns that fall in the positive extreme of the distribution. The value is not considerably different from zero. In fact, it can be asserted that among the three, G is the one resembling a normal distribution the most.

With a sufficiently large sample size, 204 per each factor, (i.e. monthly time series data from January 2002 until December 2018), anyways, it is possible to take advantage of the central limit theorem to assume the normalilty attributes of these distributions.

The average monthly return across the defined time period is negative, even though not particularly different from zero (above all, the G factor). The table also provides a t-test for the mean of each portfolio. E and S have a statistically different from zero average monthly return, as confirmed by the relative statistics. The same cannot be said for G, for which the t-stat is approximately -1.215.

According to these results, investors would have earned negative (in excess of the risk free rate) returns should they have invested into Environment and Social aware firms. In fact, the comulative returns for the factors Environment, Social , and Governance are respectively: -50.96%, -57.40%, and -31.47%. This could be indicative of a general tendendecy by European investors of not valuing ESG investments, which lead them not to invest into these firms. Maintaining the assumption that prices are determined by expectations, negative returns could be interpreted as investors not believing that firms investing into actvities that produce positive externalities for their stakeholders, other than shareholders, will produce financial value. The above mentioned results seem to confirm the position that, on average, for European firms, corporate financial performance is not related to corporate social performance. With respect to G, it is not possible to statistically infer that an investor would have earned an average return different from the risk free rate if he invested into this portfolio, as the mean monthly excess return is not statistically different from zero.

A probable scenario is that the sample results are affected by the 2006-2008 financial crisis, which by the way seems to have beaten sustainable portfolios much harder: in fact, Mkt-Rf SMB HML and MOM portfolios still maintain a positive avearge monthly return throughout the 2002-2018 period (descriptive statistics of them can be found in Table 2). It is worth to notice that, with the only exception of MOM portfolio, none of the analysed average excess returns are statistically different from zero as confirmed by the respective t-stats. This suggests some doubts on the ability of the Fama French original specifcation to explain risk, further evidence is given in the cross-sectional analysis.

Individually studying standard deviations, E seems to be the least volatile, with a figure below those of the others. S standard deviation is the highest among the three, thus indicating a higher volatility attached to it. Cross-checking the results of Table 1 with those in Table 3 and Table 4, containing the summary statistics of the individual time-series regressions of E and S vs the FFC specification, it can be concluded that much of the risk associated to them can actually be explained by the already defined model. In fact, even if R^2 is not particularly high in both cases, results show a non irrelevant loading especially on SMB, with coefficients of 0.503 and 0.515, respectively, which are statistically significant at 95% confidence interval – as highlighted by the p-values of both.

G looks volitile as well, even though to a lesser degree with respect to S, in the universe of these portfolios.

A visual representation of the volatility and the overall performance of the portfolios is given in Figure 1, which plots the cumulative returns for all them between 2002 until 2018. The graph clearly

shows that the portfolios have generated positive returns in the early years of the 2000s. Since the beginning of 2008, though, the portfolios have consistently negatively performed. The plot also show the similar behaviour E and S have had throughout the whole period. In fact, with the only exception of the very beginning of the sample, the two portfolios have shown an almost identical trend (also magnitudes look very similar). This consideration strenghtens the point that the two sustainability matters, in the sample, look particularly related. Hence, it is possible to assume that firms that score high (low) in one item have a strong (weak) performance also in the other. The chart, in addition, displays how volatile the three portfolios are. One example of such a property is the 2008 drop of the three assets. All of the three lost more that 10 percentage points, but S (with a monthly stantard deviation of 0.02, the highest of the three) has lost almost 200 basis points.

In terms of risk adjusted returns, eventually, from Table 1, it can be stated by the Sharpe ratio (computed as average monthly return over standard deviation) that portfolios had negative performances, driven by below-zero average returns. In annualized terms, the three factors have realized -247.21%, -278.30%, and -121.58%.

A closer look at the dynamics of the portfolios yields valuable insights regarding their nature. Since E and S seem fairly similar in their behaviour, an analysis of the their correlation is performed. As expected, the two of them show a meaningful correlation (0.72). The assumption behind this result is that there is no clear dinstiction between an environmentally strong firm and a socially resposnible one. This blurred line may be caused by a bias in the rating process (i.e. scores given to each firm). Else, it could be explained by a mutual characteristic which makes the two items behave similarly. By consequence, their ratings under the two matters are very similar. Thus, portfolios are very similar in their composition. Moreover, the two factors show some sort of correlation also with SMB, versus which E (S) scores 0.49 (0.43). This is a first evidence of both a relation between these two items and also with the Fama-French-Carhart specification. G does not show any particularly relevant correlation neither with the other two ESG items, nor with the four factors specification. Hence, returns movements do not seem to be comparable to any of the other portfolios. The intuition is that the Governance attribute identifies a more defined characteristic of firms.

In conclusion, E and S show some similarity and their mutual heavy and significant (negative) sensitivity towards the SMB factor does not make them eligible as risk factors, since their performance seems already to be partially explained by the Carhart specification. G, which seems not to be explained by any of the Fama-French-Carhart portfolios, for alpha=5%, has a slightly

different performance and seems to be the best candidate as individual risk factor, in investigating whether sustainability is priced in the European market.

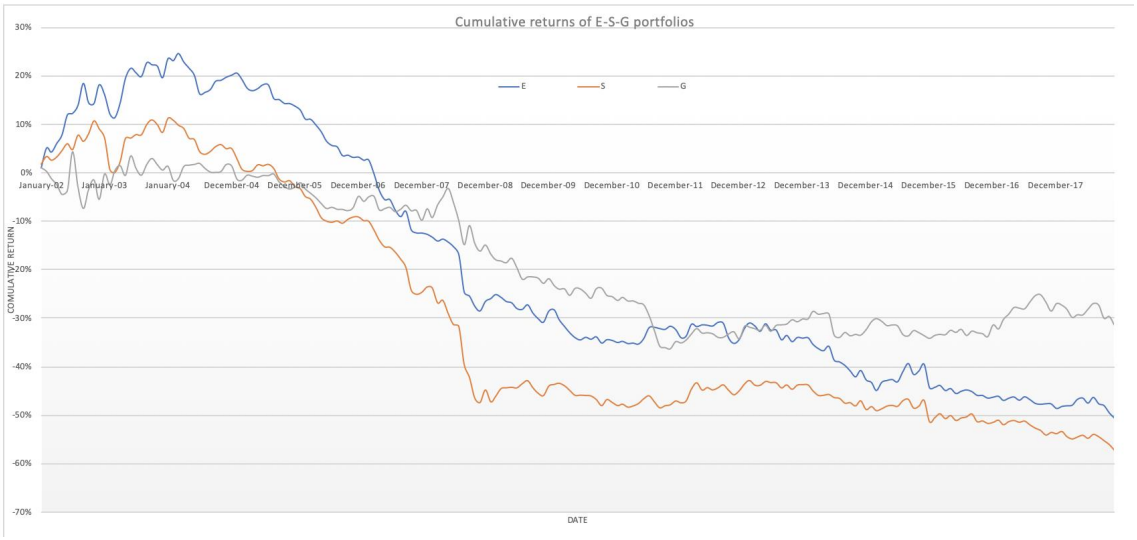

Figure 2 - Cumulative returns of E, S, G factor-mimicking portfolios 2002-2018

	MktRF	SMB	HML	WML
mean	0.005576	0.0021539	0.0021926	0.0087353
median	0.00645	0.002	0.00265	0.01115
max	0.1374	0.0486	0.0829	0.1369
min	−0.2206	−0.0682	−0.0472	−0.2627
std	0.052573	0.018714	0.021537	0.039783
mode	−0.1	−0.0077	0.0049	0.0064
kurtosis	4.6916	3.6618	3.786	14.449
skewness	−0.59824	−0.31357	0.35011	−1.7577
range	0.358	0.1168	0.1301	0.3996
sharpe	0.10606	0.1151	0.10181	0.21957
t-test	1.5149	1.6439	1.4541	3.1361

Table 12 - Descriptive statistics Fama-French-Carhart four factors specification

The three tables below (namely Table 3, Table 4, and Table 5) show the results of the time-series regressions of each portfolio versus the Fama-French-Carhart specification during the mentioned time period. The coefficient estimates represent each portfolio sensitivity versus the factor-mimicking portfolios.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	-0.0046658	0.0012213	-3.8202	0.00017814
Mkt-Rf	0.05231	0.02642	1.9799	0.049088
SMB	0.50335	0.062507	8.0526	7.278e-14
HML	-0.11107	0.060444	-1.8376	0.067607
MOM	0.029094	0.033476	0.86912	0.38583

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0164
R-squared: 0.261, Adjusted R-Squared 0.246
F-statistic vs. constant model: 17.6, p-value = 2.25e-12

Table 13 - Summary statistics time series regression of E against the 4 Fama-French-Carhart factors

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	-0.0049308	0.0012806	-3.8504	0.00015893
Mkt-Rf	0.12152	0.027702	4.3867	1.8646e-05
SMB	0.51547	0.06554	7.865	2.2983e-13
HML	-0.14068	0.063377	-2.2198	0.027564
MOM	-0.058648	0.0351	-1.6709	0.096319

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0172
R-squared: 0.295, Adjusted R-Squared 0.281
F-statistic vs. constant model: 20.8, p-value = 2.28e-14

Table 14 - Summary statistics time series regression of S against the 4 Fama-French-Carhart factors

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	-0.0021648	0.0014327	-1.5111	0.13236
Mkt-Rf	0.056061	0.030992	1.8089	0.071973
SMB	0.14382	0.073323	1.9615	0.051216
HML	-0.13557	0.070903	-1.9121	0.057303
MOM	0.020462	0.039268	0.52107	0.6029

Number of observations: 204, Error degrees of freedom: 199
Root Mean Squared Error: 0.0193
R-squared: 0.0415, Adjusted R-Squared 0.0222
F-statistic vs. constant model: 2.15, p-value = 0.0756

Table 15 - Summary statistics time series regression of G against the 4 Fama-French-Carhart factors

Results seem to confirm the initial intuition derived by the diagnostic tests run on the E S G portfolios: E and S exhibit a significant loading on the already defined asset pricing model, whereas

G seems not to be explained by it. In favor of this view, the first two regressions both highlight a higher adjusted R^2 with respect to G (0.246 and 0.281, respectively, against 0.022). Moreover, the whole model seems to be significant for the first two portfolios, with p-values considerably below the rejection threshold, whereas the G figure lies above 0.05. Going deeper in the analysis, E shows significant correlation with the first two portfolios in the FFC model (Mkt-Rf and SMB): 0.05 and 0.50 estimates are significant at the 95 % level of confidence. This would lead to the conclusion that adding E would not help to explain better the cross-sectional variation of stock returns. The same holds for S: in fact, a relevant portion of its behavior can be attributed again to SMB, with an estimate of almost 0.515 (where the p-value is less than 0.05). Moreover, the excess return of the market portfolio seems to play a role in explaining the behavior of this asset with a statistically significant estimate of 0.12. Interestingly, it seems to have a significant negative loading on HML, equal to -0.14. This is a common characteristic of the sustainability portfolios, at differing confidence levels, which could lead to the conclusion that on average growth stocks tend to outperform value stocks in terms of sustainability.

To conclude, for the purpose of this study to investigate whether sustainability is priced, results confirm that E and S have limited power to further explain cross-sectional, as these preliminary tests demonstrate that with a very limited probability of committing a type I Error, we should reject the null hypothesis that there is no relationship between the FFC factor-mimicking portfolios and the two responsible portfolios constructed.

Nevertheless, G's explanatory power should be further investigated, as time-series analysis versus the already mentioned model provides different results. In first instance, it demonstrates no important loading versus any factor-mimicking portfolio, differently from the other two. Moreover, all of the results have higher p-values, leading to the conclusion that it is not possible to reject H_0, that is: there seems to be no relationship between the portfolio and the 4 regressors.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.001592	0.00066782	2.3839	0.017348
E	−0.0024225	0.00062154	−3.8976	0.00010475
S	−0.00065199	0.00060491	−1.0778	0.28141
G	−0.0019721	0.00058473	−3.3726	0.00077793
Mkt-Rf	0.0075634	0.0010194	7.4196	2.8285e−13
SMB	0.0038247	0.00041274	9.2665	1.5165e−19
HML	−0.0024748	0.00044369	−5.5777	3.2671e−08
MOM	0.01245	0.00082903	15.017	1.9997e−45

Number of observations: 865, Error degrees of freedom: 857
Root Mean Squared Error: 0.00719
R-squared: 0.44, Adjusted R-Squared 0.435
F-statistic vs. constant model: 96, p-value = 2.74e−103

Table 16 - Cross sectional regression of average monthly excess returns vs time-series betas of 7 factor-mimicking portfolios

Results in Table 6 seem to be in line with expectations. E commands a statistically different from zero premium, even though of a very limited degree. The same holds for the 4 FFC portfolios, among which MOM ranks first in terms of weight. S is not statistically significant, as opposed to previous results which have shown dependence between the social portfolio and the 4 factors specification. To further investigate this phenomenon, additional analysis highlights a linear combination with E. This confirms the initial hypothesis of some sort of relationship between E and S – results of these last study are summarized in Table 7. They strengthen the point of a relevant correlation between the two.

In general, in any case, the model does not seem to explain much of the cross-sectional variation of the average excess returns of the stocks in the sample. In fact, the alpha is positive and statistically different from zero. If the model was able to explain the cross-sectional variation of stock returns, we would expect abnormal excess returns to be indistinguishable from zero. Further analysis is done to understand whether also the original four factor specification still has explanatory power regarding cross sectional risk or not. Results are shown Table 8, which contains the summary output of the cross-sectional regression of the average excess returns against the sensitivities of the stocks in the sample versus the Fama-French-Carhart 4 factors specification across the time window 2002-2018 (i.e. time series betas). Both the models do show a statistically significant different from zero alpha which denotes the inability of the two specifications to explain risk. In fact, firms have earned abnormal excess returns, not described by the model. All the results in Table 8, statistically significant at the 95% confidence interval, show low coefficients. The main intuition behind this is that since these risk factors were first established in 2002, as of today, they could be not as relevant as they used to be. In fact, as the results of the original study performed by Eugene Fama and

Kenneth French were made public, many may have undertaken similar investing strategies. This spread behavior is most likely to have reduced returns from investing into them. Accordingly, risk attached to the factors has diminished as well. This could lead to the conclusion that the FFC specification is not representative of risk anymore.

Going deeper in the analysis of Table 6, some more research is worth for G. Under the null that there is no relationship between the portfolio and the cross-sectional variation of excess returns, we can reject H0: it commands a negative premium of 0.0019, with a t-stat of -3.37. Accordingly, we can statistically infer that G has power to describe the cross-sectional variation of stock returns, given its statistically significant coefficient estimate. In particular, it commands a negative, small, premium. Results in Table 9, which describe the regression of the factor-mimicking portfolio against the average cross-sectional monthly returns confirm this. In fact, when the only independent variable is G, its coefficient estimate is still negative (-0.0039) and statistically significant (t-stat equal to -5.53).To collect further information, an additional analysis has been performed: a cross sectional regression of sample average excess returns against G, adding the 4 FFC specification as control variables. Results are displayed in Table 10 and show that G does not lose its statistical significance (t-stat of-3.39), and the coefficient estimate is still negative (-0.0019). Thus, under different specifications, G consistently has statistically relevant negative premiums. Thus, for higher performances of firms under governance matters, we can expect lower returns, also when other risk factors are taken into account.

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.0017605	0.00066996	2.6277	0.0087493
E	-0.0019718	0.00061066	-3.2289	0.0012899
S	-1.6507e-05	0.00057828	-0.028545	0.97723
Mkt-Rf	0.0074057	0.0010244	7.229	1.0764e-12
SMB	0.0040619	0.00040915	9.9276	4.6269e-22
HML	-0.0027805	0.00043695	-6.3633	3.2111e-10
MOM	0.01268	0.00083119	15.256	1.1318e-46

Number of observations: 865, Error degrees of freedom: 858
Root Mean Squared Error: 0.00723
R-squared: 0.432, Adjusted R-Squared 0.428
F-statistic vs. constant model: 109, p-value = 6.85e-102

Table 17 - Cross sectional regression of average excess returns vs time-series betas of E S and FFC specification

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.002068	0.00066016	3.1327	0.001791
Mkt-Rf	0.0079906	0.001029	7.7652	2.314e-14
SMB	0.0042226	0.00040909	10.322	1.2548e-23
HML	-0.0032569	0.00043191	-7.5406	1.188e-13
MOM	0.011623	0.00079814	14.563	4.1819e-43

Number of observations: 865, Error degrees of freedom: 860
Root Mean Squared Error: 0.00733
R-squared: 0.415, Adjusted R-Squared 0.413
F-statistic vs. constant model: 153, p-value = 1.13e-98

Table 18 - Cross sectional regression of average excess returns vs time-series betas of FFC specification

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.005154	0.00032652	15.785	1.6804e-49
G	-0.0039048	0.00070487	-5.5398	4.0227e-08

Number of observations: 865, Error degrees of freedom: 863
Root Mean Squared Error: 0.0094
R-squared: 0.0343, Adjusted R-Squared 0.0332
F-statistic vs. constant model: 30.7, p-value = 4.02e-08

Table 19 - Cross sectional regression of average excess returns against G time-series betass

Estimated Coefficients:

	Estimate	SE	tStat	pValue
Intercept	0.0017413	0.00066318	2.6257	0.0088002
G	-0.001909	0.00056244	-3.3941	0.0007201
Mkt-Rf	0.0082013	0.0010247	8.0039	3.8998e-15
SMB	0.0040505	0.00040976	9.8852	6.7525e-22
HML	-0.0029668	0.00043772	-6.7778	2.2664e-11
MOM	0.011608	0.00079332	14.632	1.8712e-43

Number of observations: 865, Error degrees of freedom: 859
Root Mean Squared Error: 0.00728
R-squared: 0.423, Adjusted R-Squared 0.42
F-statistic vs. constant model: 126, p-value = 5.02e-100

Table 20 - Cross sectional regression of average excess returns against G and FFC factor-mimicking portfolios time-series betas

In this research paper I have investigated if the Environmental, Social, and Governance factors are priced in the European stock market. To do so, I have run a Fama-Macbeth procedure to examine the cross-sectional variation of stock returns and its relationship with ESG factors, identified by means of third-party scores. Moreover, I have also studied sustainability in relation to the Fama-French-Carhart four factors specification, to understand the dynamics of the investigated portfolios.

In general, ESG investing has not been profitable throughout the chosen time period (i.e. 2002-2018). All of the three factor-mimicking portfolios have performed negatively in the average month. Their cumulative returns confirm a strong negative performance. This is true especially after the financial depression happened between 2006 and 2008, after which none of them has been able to provide a positive return, in aggregate terms. That is, if an investor had invested into these portfolios, it would have constantly earned less than its initial investment. Further analysis of this phenomenon has shown that the negative performance of E is explained by the returns of virtuous firms which are worse, in absolute terms, with respect to returns yielded by *sin* firms. The other two portfolios are characterized by a better performance of the bottom quintiles.

Preliminary tests over the factor-mimicking portfolios have highlighted a statistically relevant relationship between E, S and the Carhart specification. Hence these two should not be considered as capable of explaining excess returns. Their explanatory power, in fact, is already enclosed in the model. In particular, concerning E, SMB has a strong and statistically significant coefficient estimate. With respect to S, both the Market portfolio and SMB have a robust and statistically significant coefficient estimate. Moreover, E and S are proven to have a strong correlation. This is assumed to be explained either by a bias in the scoring process, or by a common element which drives both the items.

The same cannot be told for G, which shows no significant dependence on any of the portfolios for alpha = 5%. This result is also reinforced by the low correlation of Governance with any other factor-mimicking portfolio. The time series regressions of the portfolios over the Carhart specification have highlighted, even though at differing confidence levels, an overall negative relationship between sustainability and the HML portfolio. This result is in line with previous literature. The main intuition behind it is that investors' demand for responsible firms is relevant. Hence, book-to-market ratios are lower for highly sustainable companies. By definition, their performance resembles that of growth firms.

The model applied seems not to be able to entirely explain the cross-sectional variation of stocks' returns as resulting is positive and statistically different from zero, thus firms earn abnormal returns

not explained by the control variables. Moreover, S has a non-statistically significant coefficient estimate. Further investigation demonstrates that there exists some sort of linear relationship between E and S (Bali, Engle, and Murray, (2016)). This adds to previous outcomes obtained concerning their correlation. Results for G show a statistically significant negative cross-sectional relationship between the portfolio and European average stocks excess returns. Further tests, to check the robustness of the result, strengthen this negative relationship: it is found to have a negative and significant coefficient estimate under different specifications.

After having investigated the research question of whether the Environmental, the Social, and the Governance factors are priced in the European stock market, this study concludes that G commands a slightly negative monthly premium, which is interpreted – in line with previous research on the topic – as the cost of incorporating governance-related considerations in portfolio setting. On the other hand, since E and S seem to be already explained by previously defined model, they do not command any risk premium which is not already enclosed in the Fama-French-Carhart specification.

According to the results this paper adds to the literature, future research may want to investigate further the dynamics between Environment and Social portfolios to understand the roots of their mutual relationship. Both the assumptions could be analyzed: on the one hand if by applying a different scoring process, this relationship vanishes. On the other, questioning whether they truly identify something distinctive or not. Moreover, it could also dig deeper in the meaning of the negative premium which Governance demonstrates: is it really the case that investors have to give up some profits to include virtuous companies in their portfolios?