MSc in International Finance

ESG INTEGRATION IN MULTI FACTOR MODELS

CHRISTOPHER CUTOLO

Work project carried out under the supervision of:

Professor: Antonino Emanuele Rizzo

19-05-2021

ABSTRACT

Over the last years ESG has become one of the major trends within the financial world. Investors are expressing more interest in environmental, social and governance issues and experts are routinely integrating sustainability considerations when building portfolios. Several studies have been conducted and papers have been published trying to analyze ESG's integration into the asset management world. Results show lack of uniformity and coherence when it comes to understanding the impact of socially responsible investments. We construct factor mimicking portfolios based on E, S and G scores to study the relationship between sustainability and financial performance in the American market. The results show that long short portfolios constructed in this fashion display negative returns over the period that spans from July 2002 until June 2020. This is mainly because of the significant positive returns of stocks with low sustainability scores. Moreover, regression analyses are conducted using the classic Fama MacBeth procedure (1973), implementing the Newey West correction to account for heteroskedasticity and autocorrelation. Incorporating both the Fama-French five factor model and other control variables, ambiguous and mainly insignificant results are obtained.

KEYWORDS Asset Pricing, ESG, factor models, Fama MacBeth procedure

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1 Table of Contents

2 Introduction 4

3 Theoretical framework 5

3.1 CAPM 5

3.2 APT 6

3.2.1 FF3M 6

3.2.2 QMJ 7

3.2.3 RMW 7

3.2.4 CMA 8

3.3 ESG 8

4 Literature review 11

5 Data collection and methodology 16

6 Statistical adjustments 19

6.1 Stationarity check 19

6.2 Winsorization 21

7 Empirical findings 22

7.1 Summary Stats 22

7.2 Control Regressions 24

7.3 Fama-MacBeth procedure 26

8 Conclusion 30

9 References 32

10 Appendix 34

2 Introduction

Investors over the last years have been shifting towards factor investing in search for higher returns while keeping costs at a minimum. This particular investing style involves targeting specific drivers of return across different asset classes. These drivers, commonly known as factors, can be defined as any sort of characteristic that is peculiar to a certain asset class or sector that helps in explaining the risk and return profile of said investment. It is a strategy that aims to achieve superior returns over an extended time horizon by harvesting a long-run risk premium. It can help investors in meeting their objectives and desires such as increasing returns, reducing risk and even improve a portfolio diversification.

Besides factor investing, another trend that has gained a lot of traction in the last 5 to 10 years is sustainable investing. The rise of ESG investing started in 2014 with the emission of the first green bonds and since then it has only increased in popularity. Morgan Stanley's 2020 Sustainable Signals survey found that around 85 per cent of investors are interested in sustainable matters when considering investment decisions, up from the 71 per cent that was registered in 2015. What once was considered as niche within the huge world of financial markets has now become mainstream, with more and more investors realizing that in order to understand a company it is not possible to take into consideration only its financial performance. You need to consider ESG information to have a much better understanding of a company's mission and strategy as a whole. To give more perspective to the relevance of this trend we can look at the flow levels for ESG mutual funds. 2019 set a record into ESG funds, with approximately four times the flow levels of 2018. Furthermore, 2020 registered yet another record with an increase of 100% with respect to the previous year. There were several elements that contributed to this shift in the investment paradigm. First and foremost, governments started to allocate more money towards sustainable projects and investments, secondly, events like the World Economic Forum (WEF) really put ESG under a spotlight. Finally there has a been a general commitment by the CEOs to improve the sustainability profile of their companies (eg. Larry Fink).

Nowadays the two trends are showing signs of convergence, and the result is what we can call sustainable factor investing, a phenomenon whose implications for the investment sphere and beyond are potentially incredible. The objective of this paper is exactly to investigate ESG in the context of a multi factor model, in order to understand if the environmental, social and governance pillars are individually priced in the U.S. market.

3 Theoretical framework

This section of the thesis serves as an introduction of the theories within the financial economics field that have laid the foundations of factor investing. These models are presented to provide a solid background on asset pricing theories and in order to give context to the procedure that will be applied in the next chapter. We will start by quickly covering the first asset pricing model, Capital Asset Pricing Model (CAPM), we will continue explaining the other factors present in the model.

3.1 CAPM

Arguably the most important question within the field of modern financial literature has always been and probably will always be: “What are the factors that drive stock returns?”. The oldest and the most famous asset pricing model that tries to tackle this issue is the Capital Asset Pricing Model (CAPM), developed in the early 1960s (Lintner, 1965; Mossin 1966; Sharpe, 1964 and Treynor, 1961). Within the CAPM world there are only two drivers/factors that have an influence on stocks’ returns, the first one being the stock’s exposure to the market (described and captured by the Beta). This component which is also known as “systematic risk” cannot be diversified away, and investors should be appropriately compensated for bearing this risk. The other component is “idiosyncratic risk”, a type of risk which is unique to the individual security / asset class / sector. It has been shown that this type of risk can be eliminated using appropriate diversification. The basic intuition of the CAPM framework is synthesized within the formula:

$$r_a = r_f + \beta_a * (r_m - r_f)$$

where r_a stands for the expected return of a certain asset a, r_f is the risk-free rate[1], r_m is the expected market return and β_a represents the systemic risk factor / market exposure. Thus according to CAPM, the Beta is the only relevant measure of a stock’s risk and is the only factor that investors should care about. To be more specific, it shows how much the price of a specific security changes with respect to changes in the price of the market portfolio. Higher Beta implies higher risk and investors should be rewarded accordingly. This framework creates a linear stock/security market line, which exhibits the securities’ expected returns contingent to the level of risk that is related to each asset (Figure 1 in the appendix). As previously anticipated a value of one for the Beta factor implies expected returns perfectly in line with the market. Should there be any deviation from the SML arbitrage opportunities would be present due to overvaluation or undervaluation of the securities. This framework has been widely criticized mostly due to its extremely unrealistic assumptions, in particular the hypothesis of no transaction costs / frictions in the market, investors’

[1] The risk-free rate is usually derived from the U.S. 3-month T-Bill

behavior which is assumed to be always perfectly rational and finally the hypothesis that investors have access to credit at risk-free rates. In reality, it will be impossible for investors to borrow at risk-free rates, implying that the CAPM is probably overestimating the returns of assets. This feature has been proven to be true by empirical evidence, Baenz (1981), who discovered that the CAPM model overestimates by a significant margin the returns of large caps, while underestimating returns of small companies.

3.2 APT

In 1976, Steven Ross came up with a different theory addressing the issue of what drives stocks returns. In particular, according to his "Arbitrage pricing theory", expected returns of financial securities can be thought and modeled as a function of several factors. The peculiar aspect of Ross's theory is that he did not specify the meaning of these factors. He stated that these components can be thought as of any characteristic that is peculiar to a certain group of securities, sector or asset class that is important in explaining their return and risk. This implies that the number of factors is likely going to depend on the market and the timeframe considered. According to APT, the expected return of a portfolio can be computed using the expression:

$$r_a = r_f + \beta_1 f_1 + \beta_2 f_2 + \cdots + \beta_n f_n$$

with the expected return (r_a) depends on the risk-free rate (r_f), the sensitivity of every asset's return to a factor (βi) and the risk premium related to that factor (f_i). This framework, due to the inter-changeability of the factors, represented the first instance in which multi-factor "thinking" was adopted by financial literature. It is also true that this flexibility represents the main limit of this model since factors cannot be directly observed in the market. As a matter of fact, there have been several debates about how to proceed with regards to their definition and estimation. As of today, three main categories of factors are identified:

1) Macroeconomic factors: changes in inflation, movements in the yield curve etc.
2) Fundamental factors: They capture common stock characteristics such as valuation ratios, sector and industry membership, technical indicators etc.
3) Statistical factors: factors are identified with using advanced statistical methods (PCA).

3.2.1 FF3M

The two most renowned pioneers in the field of factor investing are Eugene Fama and Kenneth French. In 1992 Fama, French came out with their famous three factor model which still as of today constitutes one of the pillars of this landscape. In this paper, they explained the returns of the US equity market via the implementation of three factors, namely market, value and size. Their proposed regression goes as follows:

$$r_i - r_f = r_f + \beta_m MKT + \beta_s SMB + \beta_v HML + \alpha$$

with r_i being the portfolio's return and r_f the risk-free rate return.

MKT represents the market factor, already implemented by Sharpe in the CAPM model. The SMB factor stands for "Small minus big" and is related to the size of the companies. This factor is included because of the outperformance of small caps relative to large caps, as a consequence the SMB factor is built as a long-short portfolio, going long on companies with small capitalizations and shorting companies with large caps. Instead HML stands for "High minus low" and is related to the fact that value stocks tend to outperform growth stocks. The same procedure is applied in this instance with the HML portfolio that is built by going long on companies with low price-to-fundamental ratios (proxy for value stocks) and shorting companies with higher ratios (growth stocks).

3.2.2 QMJ

The quality minus junk factor was firstly introduced by Frazzini, Asness in 2013. The basic rationale behind the factor is that, all else being equal, investors should be willing to pay more for "quality" stocks that exhibit certain specific traits. The factor mimicking portfolio is constructed in the classic Fama-French style, long-short with a long position in the top 30% of high-quality stocks and a short position in the bottom 30%. In particular there are three key variables that are used to define quality:

1. Safety: investors should pay a higher price for stocks that require lower rate of returns.
2. Growth: investors should pay a higher price for stocks that exhibit growing profits, with growth being measured as the prior five-year growth in the profitability measures
3. Profitability: definition is profit per units of value and investors should pay more for profitable companies. Profits are measured looking at margins, earnings, cash flows etc.

3.2.3 RMW

The Robust minus Weak factor (RMW) was introduced by Fama-French during later stages of their research. This factor was constructed to account for the difference in returns between companies with strong and weak operating profitability. This last measure is defined as "the revenues minus cost of good solds, minus selling, general and administrative expenses, minus interest expense, all divided by book equity[2]". This RMW component is often linked to the classic value factor, even though those address slightly different dimension. The robust minus weak factor can be intended as a measure of how productive firms' assets are, and investors should pay a higher price for companies that exhibit this characteristic.

3.2.4 CMA

As RMW, the Conservative minus Aggressive factor (CMA) was introduced by Fama-French. This factor investigates the difference between firms with low (conservative) and high (aggressive) investment policies. This factor is measured by focusing on the annual changes in property, plant and equipment (PPE) plus the annual change in inventories all divided by the book value of total assets. Previous studies document a negative relationship between several forms of investments and stock returns." Firms experiencing rapid growth by raising external financing and making capital investments subsequently have low stock returns, whereas firms experiencing contraction via divestiture, share repurchase, and debt retirement enjoy high future returns." (Watanabe et M. Desban al, 2013 [50], p. 529). The economic rationale and for these findings are several but one of the most popular includes "over-investment". Managers may over-value companies with large investments by over-valuing its forecasted cash-flows and the final low return comes as a direct consequence of a market adjustment.

3.3 ESG

ESG stands for Environmental, Social and Governance, and refers to three central components that are used to measure an investment's sustainability. Its origin lies in the "Triple Bottom Line" concept, first developed during the late 1980s. The focused was all on the three P, People, Planet and Profits. It stated that firms should invest time and resources on each of these components instead of focusing only on profits. A decade later this general idea was given a more rigorous framework and structure to it under the name of ESG, which as of today represents the pillar of sustainable investing. The basic idea that revolves around ESG investing is the belief that companies are more likely to be successful from a financial point of view, delivering substantial positive returns, if they focus on value creation for all the stakeholders, instead of acting only to satisfy shareholders. When analyzing an investment opportunity it has become extremely important to understand what is the firm's impact on the environment, does the company serve society (if yes how and why?), has the company a solid governance structure that is going to prioritize stakeholder's objectives? A crucial step when performing this type of analysis is also consideration

[2] This definition is in line with the one of "Hou, Xue and Zhang (2015)"

of future trends, visualizing possible changes and disruptions that could have drastic implications on a company's ESG considerations and its profitability.

About the adoption of ESG considerations and matters Europe is in the clear lead with respect to the United States. Europe has assured a clear dominance within these markets, mainly due to the promotion of green investing, also via regulation, whilst on the American side Washington's politics and efforts in promoting sustainable investing have been extremely limited. Moreover, a solid infrastructure has facilitated the inflow of money into these markets. Europe has established as a clear leader especially in the debt markets. Considering green bond emission in 2020, European issuers were able to raise approximately $160 billion with respect to the $60 billion raised by the U.S. and the $34 of Asia. One of the major challenges in the upcoming future for the United States will involve changing the regulatory mindset. As of today, current regulation assumes that sustainable investing has a negative impactive on the risk-return profile. An example emerges from the Department of Labor's new 'Financial Factors in Selecting Plan Investments' rule which prohibits funds with an explicit sustainability mandate from being a default option when it comes to pension plans, specifically for defined benefit retirement plans. This prohibition includes also all the funds that decide to exclude specific sectors because of sustainability considerations or that include non-financial factors in their decision-making process. Furthermore, the issue of disparity and lack of coherence when it comes to reporting on ESG persists all over the world but especially in the US. Investors are hoping that the Securities and Exchange commission is going to put consistent effort to standardize US's firms ESG disclosures requirements and criteria.

To better understand how the three E, S and G factors are built this section will offer a brief analysis of the major elements that influence each component's score. Few decades ago environmental considerations were seen just as a side consideration, but now because of issues such as carbon emissions, extreme pollution that are causing harm also from an economic point of view investors have started to pay more attention to these issues. The E component of ESG addresses these matters, analyzing how a company decides to utilize their resources, the effect of its actions on the environment etc. Specific criteria include:

- *Climate Change* -> Climate change adds another layer to environmental risk. Greenhouse gases are one of the causes of climate change and are the most monitored indicator to assess a company's behavior with respect to their emission levels and what tools are in place to tackle set levels.
- *Water* -> This indicator is more pertinent to industries for which water is more crucial than others. Some are more exposed to risks linked to the exploitation of this resource.
- *Waste and pollution levels*

Social criteria address an incredibly wide range of potential issues and matters but at the end all of them can be brought back to social relationships. Definitely the most relevant relationship for a company is the company's relationship with its employees. Some of the elements that are considered when trying to understand how a company behaves towards their employees are:

- *Human rights* – > In 1948, the universal declaration of human rights was published, laying the foundation for the affirmation equal rights for human beings. Moreover, in 2008 the U.N. affirmed three principles intended to address the topic of businesses and human rights. Governments must protect citizens against abuses committed by companies, they have to ensure that companies are respecting all human rights and they need to guarantee an effective compensation for victims in case of abuse. Thus, has the firm a published human rights policy, is it applied equally and efficiently across all regions in which it operates?
- *Human capital retention and management* -> how much time/resources is the company investing in its employees? Training and education activities together with performance evaluation that trigger positive feedback loops can enhance the overall businesses' quality and improve workers' productivity. Moreover, considering COVID-19, what has the firm done to retain people during the pandemic, how has it handled unavoidable pay cuts?
- *Salary / retribution* -> Is employee retribution fair, given the amount of work and taking into consideration comparable jobs at other companies / sectors?
- *Diversity, equity and inclusion* -> what are the firm's goals and metrics when it comes to diversity and inclusion and how will these goals be achieved?

Corporate governance in the context of ESG relates to how a company's top management behaves. Do the executives prioritize the interests of the different stakeholders, from the employees to the shareholders and customers? Are there any potential conflicts of interest that could cause the top management to shift their interest away from those of the stakeholders? Aspects considered when analyzing corporate governance profile of a company include:

- *Remuneration*: Transparency on remuneration policy, on the performance objectives and on the criteria of reference, transparency with regards to top-managers remuneration and approval by shareholders, are among the determining aspects when assessing a firm's structure.
- *Effect of investors in the company* -> Investor behavior can limit or enhance the process of effecting corporate changes.
- *Corruption*: The risk of corruption may vary from one industrial sector to another. It goes without saying that an effective anti-corruption program can strengthen reputation and increase credibility towards stakeholders.

4 Literature review

The concept of sustainable and responsible investing has been on the forefront of financial literature for almost 30 years now, and its popularity seems only to be increasing**.** Nowadays, the majority of investment professionals and institutional investors understand the existence of a link between sustainability in broad sense and businesses' financial success. As a matter of fact, assets managed with a particular focus on sustainability have grown exponentially and this growth does not look like it is going to stop in a near future. The value of global assets applying environmental, social and governance data to drive investment decisions has almost doubled over four years, and more than tripled over eight years, reaching a total value of $40.5 trillion in 2020. Another interesting statistic that shows willingness from companies to shift towards a more sustainable business model can be found analyzing the Standard and Poor 500: while only 20% of its listed companies published sustainability reports in 2011, this number has reached 85% just six years later, in 2017 (Coppola, 2018). Given the confirmation of this trend, it has become important for the financial industry to understand the effects of sustainable investing on asset prices, portfolio holdings and corporate behavior. Moreover, the main question about ESG investing remains: "Is it possible for corporations to perform well from a financial point of view by also behaving and acting responsibly?"

The literature on Environmental, Social and Governance (ESG) issues is extensive, but far from conclusive and yet there is no general agreement on the matter. As already anticipated one of the fundamental questions to investing based on Environmental, Social and Governance (ESG) considerations is: "How does such investing affect the value of the investor's portfolio?". More specifically, how does this type of information affect the risk-return profile of an investors' portfolio? What impact does it have on the portfolio construction process? Plausible reasons exist for both outperformance and underperformance of ESG investing relative to conventional investing. In general, the main argument for outperformance of ESG based strategies is, in essence, that the stock market underreacts to ESG information. That is, the value effects of a positive ESG event is not sufficiently recognized by the stock market, hence firms with such events tend to be undervalued and a strategy investing in these firms can obtain abnormally high returns. Moreover, a reasonable hypothesis is that the stock market undervalues certain intangibles. The valuation of intangibles is typically more uncertain than tangibles and often intangibles do not appear directly on the balance sheet, hence they are less salient to investors. Evidence of underreaction to intangibles includes R&D costs, patent citations, advertising, and software development costs (Edmans, 2011). Likewise, ESG investments by firms are typically intangibles, and it is possible that the stock market underreacts to the information in ESG-related initiatives. A second reason for why high ESG stocks might outperform the market (and low ESG stocks) is that ESG investing has become increasingly popular over time to investors. That is, a growing demand for a particular set of stocks can push up the prices of those stocks, even in the absence of new fundamental information about the value of those stocks. Demand effects are also a primary reason for why high ESG stocks might exhibit underperformance relative to low ESG stocks. Merton (1987) points out that when a large group of investors ignore certain stocks, say low ESG stocks, they can become undervalued. While this implies initial low returns, subsequently those stocks will have high returns relative to high ESG stocks. Also, firms active in industries often shunned by ESG investors, such as tobacco and weapons industries, have incentives to practice very conservative accounting because their industries fall under considerable scrutiny from regulators (Berman, 2002; Hong and Kaperczyk, 2009).

At the other end of the spectrum we find the line of research which states that it is generally expected that future returns of companies that "enjoy" high ESG scores should be lower. One of the oldest empirical demonstration of underperformance compares the returns of "conventional" and Socially Responsible Investment (SRI) funds. Bauer et al. (2005) find that SRI funds and conventional funds differ in terms of style but produce similar alphas. Renneboog et al. (2008) find that European and Asian SRI funds, mainly internationally oriented, perform worse than classic domestic factor models, but SRI funds do not underperform conventional funds in most countries. In addition to SRI funds, a number of papers, such as Sauer (1997), Statman (2000), Schröder (2004), Statman (2006), Schröder (2007) and Lee and Faff (2009), show us that the performance of "SRI indices" is comparable to conventional indices. However, what is important to notice, is that – despite the lower returns – socially responsible investors enjoy higher utility because of the better ESG scores. If there are sufficient number of investors who prefer good ESG companies and disregard bad ESG firms, the expected returns of the latter should actually be higher. This is coherent with a piece of research provided by **Pastor** et **colleagues** (2020). With the objective of modelling sustainable investing, they show how agents' tastes for green holdings affect asset prices. Agents are willing to pay more for "*greener"* firms, with the effect of lowering the firms' costs of capital. In equilibrium green assets have negative alphas, whereas brown assets have positive alphas. Consequently, agents with stronger ESG preferences, whose portfolios tilt more toward green assets and away from brown assets, earn lower expected returns. Nevertheless, such investors are not unhappy because they derive utility from their holdings. This preference comes from the fact that highly sustainable companies are deemed to be typically more long-term oriented, with a defined and structured process for stake-holders engagement and at the same time are more likely to disclose relevant information. Moreover, green assets outperform when positive shocks hit the ESG factor, which captures shifts in customers' tastes for green products and investors' tastes for green holdings. The ESG factor affects the relative performance of green and brown assets; its positive realizations boost green assets while hurting brown ones. If ESG concerns strengthen unexpectedly but sufficiently, then green assets can outperform brown ones despite having lower expected returns.

Another interesting study by Angel and Rivoli (1997) predicts that a socially controversial stock that investors shun has a higher expected return, and that the expected return increases with the proportion of socially responsible investors in the market. Brammer, Brooks and Pavelin (2006) demonstrate that for UK companies, firms with good CSR ratings tend to underperform in relation to their poor CSR counterparts and they attribute this finding to the environmental indicators driving this finding.. Gerhard et al. (2015) found no significant return differences between companies featuring high and low ESG rating levels. This is a particularly comprehensive study because it uses different ESG databases and provides recent performance. Similarly, Manescu (2011) also is unable to show that ESG ratings can affect stock performance. Thus, to summarize the above findings (or non-findings), there is no clear-cut agreement or evidence that good ESG are able to generate higher returns, or also for that matter that good ESG firms earn lower returns.

Apparently, the only area of agreement in the literature is about the positive effects of ESG on the cost of capital, implying that companies with better ESG scores tend to be able to borrow more cheaply, have higher credit rankings and lower cost of capital. Some earlier work focused on the effects of governance on the cost of debt financing. For example, Bhojraj and Sengupta (2003) document that a higher percentage of institutional ownership and outside directors is positively correlated with higher bond ratings and lower bond yields. Moreover, Klock, Mansi and Maxwell (2005) showed that corporations with anti-takeover provisions in place have negative and significant effects on bond yields. On a similar note, Chava, Livdan and Purnaanandam (2009) proved that firms that have fewer anti-takeover devices in place pay on average significantly higher spreads on bank loans. Cremers, Nair and Wei (2007) document that institutional ownership can lower the yields on outstanding corporate bonds. It is typically more challenging to show convincingly the effects of factors on the cost of equity capital than on debt capital. Nonetheless, several studies have documented the positive effects of ESG on the cost of equity capital. Ashbaugh-Skaife, Collins and LaFond (2004) find that well governed firms exhibit a cost of equity financing which is approximately 135 basis points (or 88 BP on a risk-adjusted basis) lower compared to poorly governed counterparts. Derwall and Verwijmeren (2007) find that better corporate governance leads to lower cost of equity capital over the period from 2003 to 2005. Ghoul, Guedhami, Kwok and Mishra (2011) find that firms with better CSR quality exhibit lower cost of equity financing for a large sample of US firms. In addition to the effects of good governance practices on firms' cost of capital Sharfman and Fernando (2008) find that firms with better environmental risk management exhibit significantly lower cost of equity capital. Summing up this section, the literature is clear-cut; better ESG companies have lower costs of debt and equity financing. There seem to be clear benefits for firms to improve their ESG scores given the potential benefits in reducing capital costs. This finding is consistent with the capital asset pricing model (CAPM), in which a lower systematic risk – beta – implies lower cost of equity. Similarly, it was found that the average cost of debt of high-ESG-rated companies was lower than that of low-ESG-rated companies. Again, perfectly aligned with general expectations, as the corporate-governance standard, is known to reduce a firm's default risk, which has a direct impact on its cost of debt.

Pedersen et **colleagues** (2020) provided an interesting insight on how ESG impacts the portfolio construction process. The model considers three types of different investors, *Type-U* ("ESG-unaware") investors are unaware of ESG scores and simply seek to obtain the highest Sharpe Ratio possible. *Type-A* ("ESG-aware") investors also have mean–variance preferences, but they use assets' ESG scores to update their views on risk and expected return[3]. Lastly, *type-M* ("ESG-motivated") investors use ESG information and also have preferences for high ESG scores. Whilst the portfolio construction problem for *type-U* and *type-A* investors is solved through a classical mean-variance analysis approach, for *type-M* investors the solution is slightly different. Since the objective function depends on the ESG scores, the optimal portfolio obviously depends on these scores. The problem is tackled by choosing the portfolio with the highest SR for a given ESG score. What is interesting here is that the ESG-SR frontier can also be computed independent of preferences and then the investor can decide in the end where on the frontier to place herself. This implies that this ESG-SR frontier is incorporating and summarizing only security relevant information. The investor has first to decide where to place himself on the frontier and then to decide how much risk he is willing to take. Moreover, equilibrium security prices and returns are derived. It is shown that expected returns are given by an ESG-adjusted CAPM. When there are many type-U investors and when high ESG predicts high future profits, high-ESG stocks deliver high expected returns. This happens because high-ESG stocks are profitable, yet their prices are not bid up by type-U investors, leading to high future returns. In contrast, if the economy has many type-A investors, then these investors would bid up the prices of high ESG stocks to exactly reflect their expected profits, thus eliminating this market inefficiency. Lastly, if the economy has many type-M investors, then high-ESG stocks would actually deliver low expected returns, because ESG-motivated investors are willing to accept a lower return for a higher ESG portfolio. The optimal

[3] ESG scores in this study affect investors preferences' and more importantly provide information about firms' fundamentals impacting investors' perceptions of risk and expected returns

portfolio turns out to be a combination of the risk-free asset, the tangency portfolio, the minimum-variance portfolio, and the "ESG-tangency portfolio," Another interesting finding – with an apparently counterintuitive result – relates to a common method used nowadays to incorporate ESG into a portfolio: restricting the investment universe by removing the assets with the weakest ESG scores (also known as "negative screening"). It turns out to be that investors who screen out assets with the worst ESG scores and characteristics, may build portfolios that have lower aggregate ESG scores than portfolios of investors who do not impose these negative screens. The most plausible explanation has to do with the fact that unconstrained investors may short poor-ESG assets to hedge out risks. As a matter of fact, not surprisingly, limiting the breadth of the investment universe detracts from financial outcomes.

Annér and Jakobsson van Stam (2018) find evidence of sustainability effects on stocks' cross-sectional returns, by investigating whether ESG measures are priced in the Swedish market. Their paper addresses two specific matters related to ESG, in the first step they try to understand if stock returns are affected by ESG metrics, whilst in the second one the focus on understanding if the effect is given by mispricing or compensation for bearing risk. To test the research questions they apply a classical Fama-Macbeth procedure. Results show that on aggregate the "sustainable metric" does not help in explaining stock returns. However, it turns out that some specific indicators display some degree of explanatory power (eg. Community and product responsibility have a significant effect).

As of this moment in time financial literature has left out some very interesting aspects from the scope of the analysis. As a matter of fact, the standard when analyzing sustainability matters has almost always been to investigate ESG at the aggregate level. Nevertheless, this may result in a partial analysis. If no significance is found at an aggregate level it is going to be very hard to understand whether any of the components taken singularly displays some degree of explanatory power. By analyzing the Environmental, Social and Governance component independently we could determine the specific contribution of each of these factors. Moreover, there is an incredible degree of heterogeneity in the way ESG is defined, measured, reported and scored. This has to do with the presence of several regulatory frameworks that are in place. This variety implies that sustainability is integrated into business at very different extents and levels. Further, the metrics captured often happen to be biased, focusing almost only on reporting some of the positive impact while trying to hide the risks and negative aspects that arise from ESG adoption and integration. For these reasons I have decided to focus my attention towards understanding if Environment, Social, and Governance are priced in the U.S. market.

5 Data collection and methodology

The first part of this chapter is dedicated to the explanation of the data collection process. Another aspect that will be discusses is the methodology used to construct the factor-mimicking portfolios for the Environmental, Social and Governance pillars, as well as how their excess returns have been calculated. The last section will address the Fama-MacBeth procedure that is used to test if the factors are priced in U.S. market.

As already discussed in previous chapters there is mixed evidence when it comes to literature on the integration of ESG in financial decisions and investing. One of the most important reasons of this heterogeneity has to do with ESG data reporting. First of all, to this date, companies are not obliged by any regulator to publish and disclose data for a large range of ESG criteria. As a consequence, data availability when referring to sustainability remains a huge problem. Moreover, the lack of reliability, consistency and comparability constitutes another challenge. For example, even if a company decides to publish data regarding their sustainability performance, due to the lack of uniformity in regulation, they can adopt different international standards generating misleading outputs. One of the consequences is the existence of the so-called “green washers”, basically firms that tend to mislead consumers about their environmental performance and practices for marketing and advertising purposes. Another consequence is that the lack of quality in sustainability data has a direct impact on several actors that interact or take decisions because of this data. What can happen is that a certain firm takes an investment decision based upon another company’s sustainability rating or report and if these information were compromised it would lead to a bad investment with terrible effects on the environment and from a financial point of view. This scenario is not that unlikely to happen also due to the fact that the vast majority of ESG-related information disclosed is **self-assessed and** often undergoes limited internal control. Not to mention that **no general auditing** procedures and practices have been drafted yet **thus** posing serious problems due to the incentives at stake.

As of today, the most reliable and credible information regarding ESG matters is provided by rating agencies, third party entities that are involved in the process of collecting and evaluating data from different sources in order to estimate companies’ performance under a sustainability lens. The general procedure involves collecting and analyzing data from several sources, like corporate reports, websites, CSR reports and others to then produce a final ESG score for the company. However, the heterogeneity in the methodologies used by the various agencies causes the presence of different scores that associated with the same company.

A very important concept to introduce when talking about ESG reporting is materiality, firstly introduced by Kahn in 2016. It is defined as the ability of a sustainability factor to have an influence on a company's financial performance. Financially material ESG factors are factors that have a significant effect – can be positive or negative – on a company's business model and value drivers (eg. revenue growth, required capital and risk). These factors vary across sector and industry, so two companies, in two different industries/sectors, could have similar performances for the same ESG factor, but if this element is more relevant in one of the two industries, then the weight and impact of that same factor on the rating will surely differ. For the purpose of this dissertation Thomson Reuters has been the choice to collect data regarding the ESG performance of the companies in our sample. Since our objective is to understand if the E, S and G components are priced in the U.S. market the scores we are interested in are:

1. ENSCORE (Environment)
2. SOCSCORE (Social),
3. CGSCORE (Governance)

Data regarding companies' performance under a sustainability point of view were initially collected by ASSET4, a leading provider of objective, comparable and auditable extra-financial information and a pioneer in this particular niche of market. In 2009, Thompson Reuters acquired the Swiss EGS data provider, enlarging its data availability and enhancing its quality. There are several reasons that make this instrument suitable for our analysis, one of which is that unlike other data providers 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 2002, year in which ASSET4 was founded, has been retrospectively acquired). Another reason that makes Thomson Reuters a good choice is their attention towards the concept of materiality. When evaluating a specific company Reuters captures and calculates over 400 company level ESG measures, of which a subset of the 178 most relevant data points contribute to the overall company assessment and scoring. The underlying measures are based on considerations around materiality, data availability, and industry relevance. They are grouped into 10 categories and a combination of these categories (following a particular weighting scheme) formulates the final ESG score, ranging from 0 to 100, where 0 is the minimum and 100 is the maximum. Another amazing aspect about their ESG scores is that because of how they are constructed, taking into consideration materiality, they already account for industry effect and no additional transformation is needed.

The time horizon over which the analysis will be carried out is the period after the dotcom bubble of 2001. Therefore, the time horizon is from 2002 to 2020. The choice is also conditioned by the significant availability of ESG data starting in December 2001, at the inception of ASSET4. This time horizon is particularly interesting since it includes periods of recessions during which the markets plummeted – mainly GFC of 2007 to 2009 – and it will be fascinating to understand how ESG has behaved during these tumultuous periods. Historically the belief and general consensus has been that high ESG performance can help in mitigating risk during time spans of increased volatility (David C. Broadstock 2021). In order to achieve the objective of the thesis, understanding whether the E, S and G components are priced in the U.S. market, factor mimicking portfolios for the Environmental, Social and Governance factors were needed. Unfortunately, as of today, no such an asset is present or available in the market, implying that we had to create them by ourselves. The stock sample of choice to construct those portfolios has been the S&P1500 and there are two main reasons:

1. Incorporating smaller companies provides a more holistic view of the U.S. economy.
2. Compared with other U.S. equity market indices, the S&P 1500 avoids relatively illiquid, lower priced, and lower quality stocks implying that there will be a higher probability to have information relatively to their sustainability profile.

The first filter applied to the sample has to do with price data availability. Given that our time horizon of interest begins in January 2002 until December 2020, all the stocks that did not have monthly price data for this period have been eliminated. Then, using each remaining stock's ISIN and RIC the respective scores for the Environmental, Social and Governance with a yearly frequency have been retrieved from Thomson Reuters from the end of 2001 to 2019. As has previously been discussed in this chapter there are still many issues related to the reporting of ESG data from companies. In the case companies did not report at all scores those were directly eliminated from the sample, instead firms were kept in the sample if they did not report scores for a maximum of two years. In order to correct and account for the years during which scores were missing average values from the scores of other years were used. This adjustment was performed to maintain as much as possible an unbiased sample without reducing sample size to an excessively small size. The following step involved ordering the companies, according to their score, from highest to lowest under each individual component. Then the final factor mimicking portfolios were built using a classical long-short approach, going long on the top performers for the year under a sustainability point of view and shorting the companies in the bottom quartile. Moreover, these portfolios were rebalanced yearly, according to the changes in the E, S and G scores. Important to mention that a time lag of six months between the publication of the scores and the investment in the long-short portfolios was assumed, this delay is needed to ensure that investors have enough time to gather, process and act upon more recent data (eg. Given the scores for E, S and G components at the end of December 2010 we construct the portfolios starting from June 2011)

6 Statistical adjustments

Before diving into the empirical analysis part of the thesis it is very important to perform some preliminary transformations and tests to ensure the correctness of our procedure. The first step involves applying logarithms to our returns time series. There are several reasons for applying this transformation, some of which include:

- To accommodate non-linearity and to reduce right skewness.
- To reduce heteroskedasticity.
- Log returns can be added across time periods.

6.1 *Stationarity check*

Often, one of the steps in any data analysis procedure involves performing regression analysis. Stationarity is an important notion in time series analysis to establish suitable inference and evade the situation of spurious / imbalanced regressions. More specifically, variables are required to exhibit certain stability over time for us to provide any meaningful regression analysis. If for every time index collection $1 \leq t_1 < t_2 < \cdots < t_m$ the joint distribution of $(x_{t_1}, x_{t_2}, \dots, x_{t_m})$ is the same as the joint distribution of $(x_{t_{1+h}}, x_{t_{2+h}}, \dots, x_{t_{m+h}})$ for all integers $h \geq 1$, then the process $\{x_t\colon t = 1,2,\dots\}$ is said to be stationary. In other words, as we move through time the joint probability distributions remain constant – they are not a function of time. The process is identically distributed. This assumption of strict stationarity is a necessary one to hold such that the Law of Large Numbers and the Central Limit Theorem can be used in regression analysis. A weaker form of stationarity, covariance stationarity, has three necessary conditions and only pertains to a process' first two moments (mean and variance):

1. $E[x_t] = \mu$ where μ is some constant; hence it does not vary over time.
2. $Var[x_t] = \sigma^2$ where σ^2 is the variance of the series; hence it does not vary over time.
3. $Cov[x_t, x_{t+h}] = f(h)$ for any $t, h \geq 1$; importantly, the covariance is a function of h and not a function of time – this means that the only relevant factor is the distance, *h*, between two terms and not the initial time period, *t*.

Another reason for trying to make a time series stationary is to be able to obtain meaningful sample statistics such as means, variances, and correlations with other variables. Such statistics are useful as descriptors of future behavior only if the series is stationary. For example, if the series is consistently increasing over time, the sample mean and variance will grow with the size of the sample, and they will always underestimate the mean and variance in future periods. And if the mean and variance of a series are not well-defined, then neither are its correlations with other variables. The most basic methods for stationarity detection rely on plotting the data, or functions of it, and determining visually whether they present some known property of stationary (or non-stationary) data. This allows us to get an initial clue and some "initial feeling" about the behavior of our time-series data. As a matter of fact, Chatfield (2004) – one of the most renowned experts in the field of time series analysis – noted that "Anyone who tries to analyze a time series without plotting it first is asking for trouble." The graphs of our returns time-series are depicted in Figure 2 in the appendix down below. From our preliminary analysis it seems like these time series should be stationary. However, a more formal test is needed. As such an Augmented Dickey-Fuller (ADF) test was run on the various variables using the AIC criteria to choose how many lags to include in each regression. The testing for the ADF test involves the estimation of this model:

$$\Delta \mathrm{yt} = \alpha + \lambda_{t-1} + \delta_1 \Delta \mathrm{y}_{t-1} + \delta_2 \Delta \mathrm{y}_{t-2} + \cdots + \delta_n \Delta \mathrm{y}_{t-n} \qquad \textit{with n = nr of lags}$$

The hypotheses for the test:

1. The null hypothesis for this test is that there is a unit root: $H0\colon \lambda = 0$
2. The alternate hypothesis differs slightly according to which equation you're using (depending on whether a time-trend was included in the ADF equation[4]. The basic alternate is that the time series is stationary (or trend-stationary): $H1\colon \lambda < 0$

Comparing each's t-statistic with the appropriate ADF at both 10%, 5% and 1% critical values we managed to reject the non-stationary null hypothesis in all cases. The results of the ADF tests are included in the table down below (Figure 3 in the appendix).

Another detail that deserves attention when performing regression analysis has to do with one of the assumptions of OLS, specifically the hypothesis that the data displays no autocorrelation. No serial correlation (autocorrelation) is a vital assumption for regression analysis. Formally it is defined as: $Corr(u_t u_s) = 0$ for all $t \neq s$. This means that the errors between two different instances in the series have a correlation equal to zero. If this is not the case, then we can speak of the presence of autocorrelation which suggests the errors are correlated across time. The OLS estimators will be inefficient and therefore no longer BLUE. The estimated variances of the regression coefficients will be biased and inconsistent and will be greater than the variances of estimate calculated by other methods, therefore, invalidating the results of hypothesis testing. In most of the cases, the R^2 will be overestimated (indicating a better fit than the one that truly exists). Moreover, the t- and F-statistics will tend to be higher. A correlogram is simply a plot of the autocorrelation function for

[4] In this case we opted for the ADF test equation without a time trend

sequential values of lag $s = 0, 1, 2, \ldots, n$ and represents a useful tool for detecting the presence of autocorrelation. As can be seen from Figure 4 in the appendix our time series data does not show signs of the presence of autocorrelation with all the lags – excluding lag at t = 0 – falling inside the confidence interval.

6.2 Winsorization

Finally, before moving on to the regression analysis section one last preliminary statistical technique was applied to ensure results' robustness. Winsorization is a technique used to minimize the influence of outliers in the data. Since the first and second moment of distributions are very susceptible to outliers, this technique is a powerful tool that can be used to cope with this problem by increasing robustness. Practically, winsorization consists in the process of sorting the data and replacing the smallest k values with the k+1 smallest values. In our case, this tool was applied on the top and bottom quartile of the E, S and G factor mimicking portfolios and top 1% and bottom 1% were "winsorized". This approach is more statistically robust than simply removing data points as it allows to keep the sample size constant. Nevertheless, there is a downside when applying this technique, which consists in the fact that bias is introduced into the results, even though this bias is less important than if the data point was simply removed.

7 Empirical findings

This section's objective is to present the empirical results of the study. We will start by focusing on the descriptive statistics of our sample in order to understand better the characteristics of the data, the next step will involve analyzing the regression results and finally our initial research question will be addressed.

7.1 Summary Stats

Before focusing the attention on the regressions and their results it is a good practice to deeply understand the main aspects that characterize our data. Therefore, descriptive statistics for all the factors were displayed in the table number 1 in the appendix. For each factor in the table the 1st until the 4th moment were computed, namely mean, standard deviation, skewness and kurtosis. Moreover, minimum and maximum value for each variable were presented. Important to remember that this sample comprehends observations with a monthly frequency from July 2002 up until June 2020. The first interesting aspects relates to the distribution of the returns time series. Due to the values of skewness and excess kurtosis a normal distribution has to be excluded, except for two factors. Skewness is a useful indicator that measures the degree of "symmetry" with respect to a perfect normal distribution. A general rule of thumb is that if skew is between -0,5 and 0,5 then the distribution is fairly symmetrical. As can be seen from the values in the table there are three factors with skewness values above the thresholds, respectively Mkt-RF, HML and RMW. Those three returns time series appear to be negatively implying that there is a higher probability of a negative return rather than positive. Instead, kurtosis is an indicator of "fat tails", the probability of extreme event happening. Normal distributions display kurtosis values of three, higher values are symptomatic of a leptokurtic distribution. All the factors, expect for SMB and CMA, present evidence of leptokurtic behavior. These results indicate that out of our nine factors there may be only two whose returns distribution resembles the once of a "normal" one (Figure 5 in Appendix).

A first glimpse at the performance of ESG investments is given by the average monthly return of our factor mimicking portfolios across our time horizon (July 2002 / June 2020). As can be seen from the bar chart down below (Figure 6 in the appendix), all the factors present a negative average monthly return, except for the market factor, quality minus junk and the profitability factor. According to this exhibit investors would have "earned" negative returns, on average, on a monthly basis (in excess of the risk-free rate) if they had invested in the E, S and G factor mimicking portfolios. Moreover, cumulative returns for the factors were computed and are displayed below (Figure 7).

The results show that the factor mimicking portfolios – calculated as top ESG scorers minus bottom ESG scorers – are amongst the worst performers, with a cumulative negative return over this eighteen-year period of respectively – 62%, -45% and -23%. This result despite being so drastic is still coherent with some pieces of literature. A first possibility could be that the market is implying a premium for firms with excellent ESG ratings implying that they are priced higher and will then return less, lowering considerably the returns of our portfolio. In order to really comprehend the reasons behind our negative cumulative returns further analysis was conducted. Focusing our attention specifically on the "long leg" of the factor-mimicking portfolios an interesting result manifested. As can be seen from the graph below **(**Figure 8 in Appendix), the "long legs" of the factors performed extremely well with cumulative returns over the same period of respectively 181% for the Environmental factor, 193% for the Social factor and 210% for the Corporate Governance factor, thus outperforming by a significant percentage all the other factors. This result is not surprising since the link between high ESG firms and good financial performance has already been explored by the literature considerably. Investors now more than ever are understanding that long term stability and profitability are related with how well companies are capable of managing their risks. There is a strong correlation between companies that are implementing ESG risk management and better preservation of investors' capital. Additionally, firms that incorporate ESG into their long-term strategic plan and business model are able to attract the best talent easier, amplify their customer bases etc. Lastly, the higher the ESG score, the higher the probability that these companies will be included into investment funds and vehicles. This creates a positive feedback loop effect, allowing continuous investment driving profitability. The more surprising result is related to the drastic performance of the "short-leg" of the factor mimicking portfolios. Shorting companies within the bottom quartile has led to negative returns of over 200% for all three factors. But what are possible reasons for this dramatic performance? One similar scenario in financial literature that comes to mind is the niche that focuses on the "sin stock" anomaly. Sin stocks, defined as "shares in companies involved in activities that are considered not ethical[5]" – that certainly display very low ESG scores – have earned consistently positive abnormal returns. Some of the explanations include the presence of a "reputation risk premium" that can be earned by investors that are willing to invest into those stocks that are usually neglected, particular market and the fact that sin industries are capable of benefitting from market structures that resemble monopolies. Even if our case of study is different, similar dynamics could be in play. In this first section we focused on returns but another aspect which is worth investigating is the factor mimicking portfolios' volatility profile. Interestingly, out of all the factors the E, S and G are the

[5] Fabozzi, Ma, and Oliphant (2008)

ones with the lowest average monthly standard deviation, as can be seen in the graph displayed below (Figure 9). This graph suggests that firms that value ESG are less volatile on average. This is a particularly important feat, especially during periods of recession and crashes. As a matter of fact, several studies testified how during market downturns there is a "flight to quality" that supports high ESG rated companies. Due to their more robust risk management practices, strong fundamentals and balance sheets those firms tend to perform better than their peers. Additionally, those companies tend to attract managers who display a certain ethic and have desirable values in terms of how they interact with co-workers, customers etc. These human factors, often under looked, can lead to better organizational performance, higher productivity and eventually higher revenues. Lastly, higher brand loyalty from customers who pay attention to sustainability practices, which in turn translates into better profitability reducing sensitivity of responsible business against market crashes. Furthermore, prior to moving on to the regression analysis section one last step was performed. The following graph (Figure 10) displays all the pairwise correlations between factors. Using correlation we can have a first impression about the degree of the relationships between two variables. A more in-depth analysis of what are the factors' drivers and what is the magnitude of influence of one on the other will be carried on later using OLS. As expected, there is evidence of mid to strong correlation between the E, S and G component. Unexpected findings include the absence of correlation between sustainability factors and the other ones.

7.2 Control Regressions

In this section the results of regression analysis on the E, S and G factor will be discussed. The focus and objective will be to understand if there are any significant relationships with other factors and in the case in which evidence should be found we will try to understand the direction, magnitude and rationale for these results. This first table presents the six-factor regressions of monthly returns from firm portfolios sorted by their respective ESG score in the United States divided into quartiles, with Q4 representing the top 25% companies ranked by E, S and G scores while Q1 comprises the companies with the lowest scores. Important to remember that these portfolios are rebalanced on an annual basis. The difference portfolio is the classic long-short that buys Q4 companies and assumes a short position in Q1 companies. Coefficients are estimated using OLS regressions. For this first section explanatory variables are Mkt-RF, SMB, HML, QMJ, RMW and CMA. The focus will lie on understanding whether it is possible to earn "abnormal returns", thus the significance and magnitude of the alpha of the regressions will be analyzed. As already stated, the regression intercept α is our variable of interest, as it can be interpreted as the "abnormal" return due to E, S and G activity in excess of the return from an investment into other risk factors.

According to our analysis it does look like firms with higher E, S and G scores / ratings do not offer abnormal returns, in the U.S., after controlling for other variables. As a matter of fact, investing in a portfolio that has a long position in companies that highly value sustainability and a short position in firms with lower ESG scores earns significant negative returns on a monthly basis that range between 12 and 22 basis points approximately. Instead, had we invested in a portfolio with a long position in the companies within the first quartile, thus the companies with the highest sustainability scores we would have earned positive abnormal returns, respectively 0,93% for the Environmental factor, 0,90% for the Social factor and 1,1% for the Corporate Governance Factor. This result is coherent with our previous assumptions of a positive relationship between companies' scores and their returns. Lastly the Q1 which included the companies with the lowest E, S and G scores displayed highly abnormal returns, ranging from 1,2% for the Social factor to 1,3% for the Environmental factor, after controlling for other risk factors. A possible explanation for this significant outperformance is that those stocks are perceived by the public to be riskier due to their low attention to corporate social responsibility, justifying the presence of a risk premium.

Furthermore, in order to understand what are the dynamics in play that can influence returns of ESG portfolios additional regressions on control risk factors were computed. For each factor different regressions were carried out, with the dependent variable – the factor mimicking portfolio excess returns – remaining fixed. The first regression included only the classic Fama-French factors, namely Mkt-RF, SMB and HML, in the second one the Quality minus Junk factor of Frazzini, Asness (2013) was added as an explanatory variable, in the third one the Robust minus Weak and Conservative minus Aggressive factors were included and finally the "E, S and G" factors were also considered. The detailed results of our statistical analysis are displayed in the three tables below (Table 2, Table 3 and Table 4). The exhibits report the coefficient estimates in percentage terms for each variable in each regression and the t-statistics used to determine their significance. The results are mostly coherent with our previous correlation analysis, nevertheless some surprising relationships emerge which are worth to be further investigated. At first, we note how the different E, S and G components behave in a very similar way when it comes to the variables which they are being affected by. In all three cases, it looks like the classic Fama-French factors do not exhibit any degree of predictable power and thus are not able to explain the returns of the sustainability factor-mimicking portfolios.

This result could possibly suggest that firms that deeply value ESG, have a low market-Beta and are less prone to volatility in the broader market, consistent with the fact that those companies exhibited a low average monthly standard deviation. Moreover, the only significant factor that appears to have some degree of predictable power, in all three cases, is the Quality minus Junk factor, with a significant positive coefficient. This outcome could signal the presence of a positive relationship between high ESG ratings and returns during very volatile cycles, supporting the idea that investors tend to value a company's solidity and robustness, which is also indicative of its board and management quality. Intuitively this makes sense, since companies that value sustainability have been proven to be more resilient than their peers, especially during periods in which markets collapse. As expected, the E, S and G factors all show a strong positive relationship amongst themselves when included as explanatory variable. This should not come as a surprise as companies that invest time and resources in improving their sustainability profile tend to focus on all three components rather than pinpointing only one of them, also creating a "momentum" dynamic. One fascinating result when looking at the tables relates to the fact that when regressing the CG factor mimicking portfolio's returns a positive significant relationship was found for the RMW factor, despite it not being significant in the other two cases. It is difficult to precisely understand what the reasons for this result are, logically one possibility could be that firms with a solid corporate governance structure in place and a solid management team tend to be more profitable firms. To ensure the solidity and validity of the previous results additional regressions were computed. The only change that was made model was to the long-short portfolio returns which acted as our dependent variable. Previously the "sustainability" portfolios have been constructed using the top and bottom quartile of all the stocks ranked by E, S and G score. This time we are changing the size of the portfolios to check whether the regressions display coherent results in terms of both significance, magnitude and direction. The results of our new regression analysis are displayed in the three tables below (Table 4, 5 and 6).

The result of this additional series of regressions is consistent and coherent with our previous analysis. Using deciles instead of quartiles to build the long-short portfolios does not have an impact on the significance, magnitude and direction of the relationship between the variables. The Quality minus Junk factor remains the only one that is consistently able to explain the E, S and G factor mimicking portfolios' returns whilst the other control variables display close to zero predictable power. Overall this result confirms the "stability" and "robustness" of our model.

7.3 Fama-MacBeth procedure

In this section we will try to answer the original research question: "Can the sustainability factors, namely E, S and G explain the cross-sectional variation of American stocks?" In the previous chapters we conducted preliminary analysis in order to gain more information and have a better understanding of the sample we are working with and what are the relationships amongst all the variables. But now, we are going to explicitly address our initial question. To do so a Fama-MacBeth procedure was carried out. It is a classic tool used in the asset pricing field to estimate parameters for multi factor models. This approach consists of two main building steps that allow the estimation of certain specific parameters called "lambdas". These parameters indicate whether the factor loadings command a risk premium and whether this risk premium significantly differs from zero (to test if the lambdas are statistically significant we will need to compute t-stats or use p-values). As already mentioned, it involves two major steps:

1. Step 1: Time-series regressions -> we run N regressions, one for each asset in our portfolio, with the dependent variable being its excess returns and the regressors the different factors time series we want to investigate.
2. Step 2: Cross-sectional regressions -> we run regressions, one for each period, to estimate the risk premium related to each factor.

Once these two passages are completed, we can compute the risk-premia by taking the average and then assessing their significance checking for the value of their t-stat[6].Translating this to an equation form, for a portfolio with n assets and m factors in the first step we compute the factor exposures for each asset by taking the time series regressions with the form of:

$$r_{1,t} = \alpha_1 + \beta_{1,F1}F_{1,t} + \beta_{1,F2}F_{2,t} + \cdots + \beta_{1,Fm}F_{m,t} + u_{1,t}$$

$$r_{2,t} = \alpha_2 + \beta_{2,F1}F_{1,t} + \beta_{2,F2}F_{2,t} + \cdots + \beta_{2,Fm}F_{m,t} + u_{2,t}$$

with $r_{i,t}$ = return of asset i at time t, $F_{j,t}$ = factor j at time t, $\beta_{i,Fm}$ = the factor loadings. Having calculated the factor exposures for each stock in our portfolio we can now continue by running the cross-sectional regressions of the returns on the *m* estimates of the βs computed during the previous step. Each regression uses the same betas from the first step, since now the goal has become understanding the exposure of the *n* returns to the *m* factor loadings over time:

$$r_{i,1} = \gamma_{1,0} + \gamma_{1,1}\beta_{i,F1+}\gamma_{1,2}\beta_{i,F2} + \cdots + \gamma_{1,m}\beta_{i,Fm} + u_{i,1}$$

$$r_{i,2} = \gamma_{2,0} + \gamma_{2,1}\beta_{i,F1+}\gamma_{2,2}\beta_{i,F2} + \cdots + \gamma_{2,m}\beta_{i,Fm} + u_{i,2}$$

$$\vdots$$

$$r_{i,t} = \gamma_{t,0} + \gamma_{n,1}\beta_{i,F1+}\gamma_{n,2}\beta_{i,F2} + \cdots + \gamma_{n,m}\beta_{i,Fm} + u_{i,t}$$

with γ = regression coefficients that are later used to calculate the risk premium for each factor. In the end there are m + 1 series γ (including the constant in the second step) for every factor, each of length T. If the errors are assumed to be i.i.d, then to calculate the risk premium γ_m for factor

[6] This assumes that the risk-premia are independent over time

Fm we can take the average γ over t. As already stated, computing the lambdas is a crucial step for the Fama MacBeth procedure. Those coefficients represent the estimates for the risk-premium of the factors included in the model. But what does this mean practically? Applying a cross sectional regression at each point in time, in our case for each month starting from 2002/07 until 2020/06. If there is a relationship between our risk factors and our stock returns in the same period then we would find a statistically significant risk-premium. Instead, should the results indicate that there is no significant relationship – implying that the lambdas are not statistically significant – it would imply that our risk-factor is not able to explain the variation in cross sectional stock returns. The economic rationale and basic interpretation of the lambdas is the expected change in the stock returns, given an increase of one unit in the risk factor. In the exhibit above the results for the lambda estimation are displayed (Table 7). The results are more in less in line with our expectation and preliminary analysis. The three sustainability factors command a negative premium, implying a negative relationship between the stock returns and the risk factor. One option could be that the market is implying a premium for firms with excellent ESG ratings implying that they are priced higher and will then return less, implying a negative effect on stocks returns. Another possible explanation we can derive based on these results is that these negative relationships can be interpreted 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. However, it is important to notice that amongst those three only the Environmental component is statistically significant at the 10% level, with a t-stat value of -1,78. Thus in general the model does not a good job in explaining cross sectional variation of stock returns in the U.S. market.

One of the conditions that needs to be respected in order to use OLS as an estimation method is that residuals have to be independent and identically distributed. If this assumption does not hold and the residuals are correlated across observations, there will be an issue related to estimation. OLS standard errors can be biased and either over or underestimate the true variability of the coefficient estimates, making them inefficient and no longer BLUE. The Fama MacBeth has been a staple in the asset pricing field for over a few decades now. It has been used by several important experts in all kinds of different applications within the financial literature landscape. Nonetheless, as the technique became more used by practitioners its limitations became more evident. Basically, this approach is affected by two major issues, that eventually threaten the validity and weaken many of the conclusions that may result from its practical application to financial data:

a) Errors-In-Variable Problem -> the cross-sectional regressions are based on the assumption that the betas resulting from the equations correspond to the true and unobservable market betas. The resulting and unavoidable errors in generating the needed beta-risk factor affect the precision by which the parameters of the cross-sectional regressions are estimated, hence the validity of the conclusions that may be derived from those estimates.
b) b) Cross-Sectional Independence Problem -> in estimating the cross-sectional model regressions using OLS we are implicitly making strong assumptions about the variance covariance matrix of the residuals at each point in time. This choice, when not adequately supported by actual data, makes the resulting estimates for the parameters consistent but not efficient. This may lead to false inference about the hypothesis under examination.

To account for these fallacies in the Fama MacBeth procedure the Newey-West correction has been applied. It is meant to adjust the covariance matrix of the parameters to account for autocorrelation and heteroskedasticity problems. Instead of using regular standard errors, HAC heteroskedasticity and autocorrelation consistent standard errors. These HAC standard errors allow us to account for several forms of serial correlation and the autocorrelation structure can be derived from the sample size. And in case of larger sample size we can be flexible in the amount of serial correlation. This implies that even if we use robust standard errors, we still will have to figure out and address the lag structure for the autocorrelation. Fortunately, Woolridge, one of the fathers of econometrics in one of his famous papers gives us some simple practical guide on how to deal with this issue:

- Annual data -> 1 or 2 lags
- Quarterly data -> 4 up to 8 lags
- Monthly data -> 12 up to 24 lags

Coherently with these indications, since our data has monthly frequency, we chose to use 12 lags. This correction is necessary in order to obtain more robust t-statistics and thus p-values to determine whether the independent variables are statistically significant or not. The new lambdas are calculated and displayed in the table below (Table 8). After the correction has been applied all the results stayed the same with the exception of the Environmental factor which before was significant at a 10% level and now is not statistically significant anymore. It appears thus that the model does a poor job in explaining the cross-sectional variation of stocks returns in the U.S and further dynamics are in play that the model is not able to capture properly. Coming back to our research question, we can conclude that our risk-factors, namely the Environmental, Social and Corporate Governance factor do not contribute to the explanation of the variation of cross sectional returns of stocks in the American market.

8 Conclusion

Nowadays investors are placing greater focus than ever on sustainability considerations when making investment decision. The positive or negative impacts of a company's activities have become an important criterion when pondering on whether to invest or not invest in set firm. Furthermore, ESG investing does not constitute anymore longer a "niche" concept within the financial markets landscape. Institutional investors with hundreds of billions, if not trillions have been shifting their focus on ESG and are approaching this area with more interest than ever. Given the rise of importance in the last decades of factor investing and more recently of ESG considerations, this dissertation's objective has been to try to understand what happens when these two different investing styles meet with each other.

In this research paper, I dedicated my attention on three different aspects, first of all to understand whether investing through a "sustainable lens" is profitable from a risk return point of view, secondly, using regression analysis what are the possible drivers that could have an impact on factor mimicking portfolios built using E, S and G scores. Lastly, I have investigated the possibility of the three sustainability factors being priced in the U.S. market, using a classic Fama MacBeth procedure. The study was carried out on a sample of companies drew out from the SP1500 index. Different screening "operations" were performed before starting our analysis to ensure to have a consistent and coherent sample. Since the dissertation's focus was the American market in the period from 2002 until 2020, companies whose price action was not available at any point in time during this timeframe have been eliminated. Moreover, firms that did not report E, S and G scores for more than three years have been removed from the sample, with the missing scores being replaced by the averages of the other years. This action was taken to maintain a high enough number of companies, since excessively small number of constituents would have harmed the validity of the analysis' results. The final sample included a total of 404 companies which were firstly ordered based on the E, S and G scores and the top and bottom quartile were used to build "sustainability" portfolios using a long-short approach.The main result of the first part of the dissertation highlights the statistically significant relationship between E, S and G scores and negative returns, as manifested by consistently negative alphas when regressing the sustainability factors on control variables. Upon further inspection it was possible to conclude that the negative relation between E, S and G scores and returns was caused by the excellent performance within our period of interest of the stocks in the bottom quartile, namely the stocks with exceptionally low sustainability considerations. Moreover when regressing the factor mimicking portfolios on different control variables, mainly the ones from the Fama French 5 factor models plus the Quality minus Junk factor, interesting results manifest. A common pattern stems from the ability of the QMJ factor to explain consistently the factor mimicking portfolios' returns, indicating a strong positive relation between companies that invest in sustainability with "high quality" firms. Instead, the other control variables appear to have close to zero predictable power and are not able to explain E, S and G returns. Lastly, to investigate whether these factors are priced in the American market, a two -step Fama MacBeth procedure was run to examine the cross-sectional variation of the stock's returns' with ESG factor mimicking portfolios.

In order to account for potential problems of autocorrelation and heteroskedasticity that are common when applying this approach, the Newey-West correction was implemented and consistent HAC standard errors were used.

The results of the Fama MacBeth regression indicate that the model does a poor job in explaining the cross-sectional variation of stock returns in the U.S. market and further dynamics are at play that the model was not able to capture properly. As a matter of fact, the computed lambdas all turned out to be statistically NOT significant. We can thus conclude, after our investigation that our risk-factors of interest, namely the Environmental, Social and Corporate Governance factor do not contribute to the explanation of the variation of cross sectional returns of stocks in the American market.

9 References

- Bolton, P., & Kacperczyk, M. (2020). *Do investors care about carbon risk?* National Bureau of Economic Research.
- Pastor, L., Stambaugh, R. F., & Taylor, L. A. (2019). *Sustainable investing in equilibrium*. National Bureau of Economic Research.
- Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2019). Responsible investing: The ESG-efficient frontier
- Gianfrate, G., Schoenmaker, D., & Wasama, S. (2015). Cost of capital and sustainability: a literature review. *Rotterdam School of Management, Erasmus University*.
- Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of financial economics*, *93*(1), 15-36.
- Henriksson, R., Livnat, J., Pfeifer, P., Stumpp, M., & Zeng, G. (2018). ESG literature review. *SSRN Electronic Journal*, 1-14.
- Skaife, H. A., Collins, D. W., & LaFond, R. (2004). Corporate governance and the cost of equity capital. *Available at SSRN 639681*.
- Klock, M. S., Mansi, S. A., & Maxwell, W. F. (2005). Does corporate governance matter to bondholders? *Journal of financial and quantitative analysis*, *40*(4), 693-719.
- Chollet, P., & Cellier, A. (2011, May). The impact of corporate social responsibility rating announcements on European stock prices. In *International conference of the French finance association (AFFI)*.
- Lee, D. D., & Faff, R. W. (2009). Corporate sustainability performance and idiosyncratic risk: A global perspective. Financial Review, 44, 213-237.
- Bohdanowisz, R. and Zientara, P. (2008). Hotel company's contribution to improving the quality of life of local communities and the well-being of their employees, Tourism and Hospitality Research, 9(2), 147-158.
- Bonilla-Priego, M.J., Font, X., and Pacheco-Olivares, M. (2014). Corporate sustainability reporting index and baseline data for the cruise industry, Tourism Management, 44, 149-160.
- Burke, L., Logsdon, J.M., Mitchell, W., Reiner, M., and Vogel., D. (1986). Corporate Community Involvement in the San Francisco Bay Area, California Management Review, 28(3), 122-141.
- Busch, T. and Friede, G. (2017). Corporate social and financial performance: A second-order metanalysis, Frankfurt, DE: Deutsche Asset Management; Carroll, A.B. (1979). A three-dimensional conceptual model of corporate social performance, Academy of Management Review, 4(4), 497–505
- Friend, Irwin & Blume, Marshall E, 1970. "Measurement of Portfolio Performance Under Uncertainty".
- Hsu J., Kalesnik V. 2014, "Finding Smart Beta in the Factor Zoo". Kent Baker H., Nofsinger John R. , (2012) Socially Responsible Finance and Investing, Wiley, Hoboken Ledoit O, Wolf M (2008) Robust performance hypothesis testing with the Sharpe ratio.
- OECD (2004). OECD Principles of Corporate Governance. In www.oecd.org PRI Association (2012),

- PRI Framework Pilot Main Definitions, in www.unipr.org Ross, S., 1976a.
- The arbitrage theory of capital asset pricing Sharpe,
- W. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk",
- mkriangkrai, 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
- 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.

10 Appendix

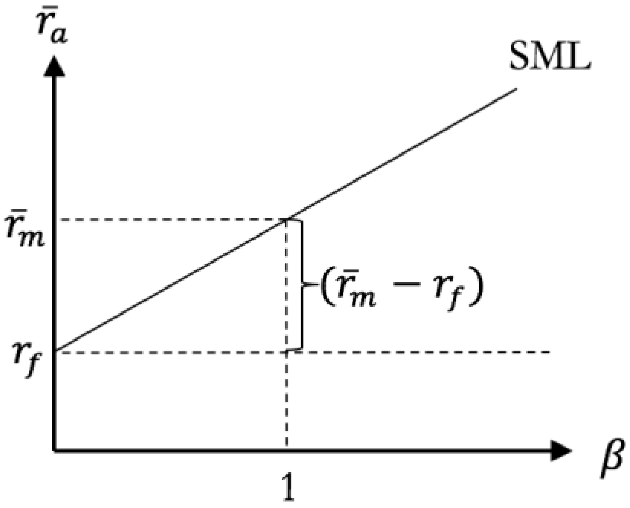
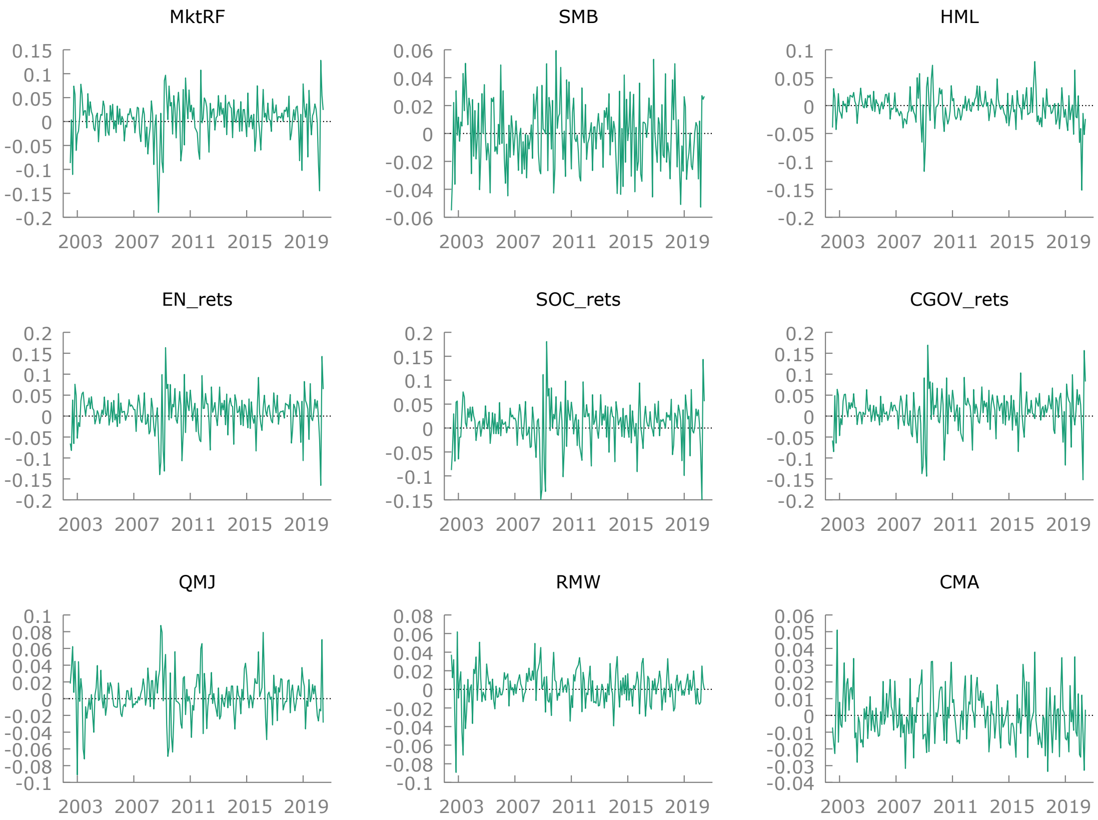


Figure number 1: Security Market Line (SML), Sharpe (1999)

Figure number 2: for each return time series of the different factors used in the analysis historical returns over the period from July 2002 until June 2020 have been plotted. Above each figure the title of the factor has been highlighted.



```
Augmented Dickey-Fuller test for MktRF
testing down from 14 lags, criterion AIC
sample size 215
unit-root null hypothesis: a = 1

  test with constant
  including 0 lags of (1-L) MktRF
  model: (1-L) y = b0 + (a-1) *y (-1) + e
  estimated value of (a - 1): -0.919277
  test statistic: tau_c (1) = -13.5956
  p-value 2.409e-023
  1st-order autocorrelation coeff. for e: 0.005
```

```
Augmented Dickey-Fuller test for SMB
testing down from 14 lags, criterion AIC
sample size 215
unit-root null hypothesis: a = 1

  test with constant
  including 0 lags of (1-L) SMB
  model: (1-L) y = b0 + (a-1) * y (-1) + e
  estimated value of (a - 1): -1.09487
  test statistic: tau_c (1) = -16.21
  p-value 2.379e-026
  1st-order autocorrelation coeff. for e: -0.010
```

```
Augmented Dickey-Fuller test for QMJ
testing down from 14 lags, criterion AIC
sample size 210
unit-root null hypothesis: a = 1

  test with constant
  including 5 lags of (1-L) QMJ
  model: (1-L) y = b0 + (a-1) *y (-1) + ... + e
  estimated value of (a - 1): -0.815197
  test statistic: tau_c (1) = -5.97783
  asymptotic p-value 1.354e-007
  1st-order autocorrelation coeff. for e: -0.020
```

```
Augmented Dickey-Fuller test for HML
testing down from 14 lags, criterion AIC
sample size 209
unit-root null hypothesis: a = 1

  test with constant
  including 6 lags of (1-L) HML
  model: (1-L) y = b0 + (a-1) *y (-1) + ... + e
  estimated value of (a - 1): -0.81785
  test statistic: tau_c (1) = -4.83143
  asymptotic p-value 4.46e-005
  1st-order autocorrelation coeff. for e: -0.010
```

```
Augmented Dickey-Fuller test for EN_rets
testing down from 14 lags, criterion AIC
sample size 214
unit-root null hypothesis: a = 1

  test with constant
  including one lag of (1-L) EN_rets
  model: (1-L) y = b0 + (a-1) * y (-1) + ... + e
  estimated value of (a - 1): -1.2314
  test statistic: tau_c (1) = -12.3218
  asymptotic p-value 9.358e-027
  1st-order autocorrelation coeff. for e: 0.005
```

```
Augmented Dickey-Fuller test for RMW
testing down from 14 lags, criterion AIC
sample size 215
unit-root null hypothesis: a = 1

  test with constant
  including 0 lags of (1-L) RMW
  model: (1-L) y = b0 + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.896542
  test statistic: tau_c (1) = -13.2662
  p-value 8.065e-023
  1st-order autocorrelation coeff. for e: 0.001
```

```
Augmented Dickey-Fuller test for SOC_rets
testing down from 14 lags, criterion AIC
sample size 208
unit-root null hypothesis: a = 1

  test with constant
  including 7 lags of (1-L) SOC_rets
  model: (1-L) y = b0 + (a-1) * y (-1) + ... + e
  estimated value of (a - 1): -0.993821
  test statistic: tau_c (1) = -4.60953
  asymptotic p-value 0.0001
  1st-order autocorrelation coeff. for e: 0.001
```

```
Augmented Dickey-Fuller test for CGOV_rets
testing down from 14 lags, criterion AIC
sample size 214
unit-root null hypothesis: a = 1

  test with constant
  including one lag of (1-L) CGOV_rets
  model: (1-L) y = b0 + (a-1) * y (-1) + ... + e
  estimated value of (a - 1): -0.796227
  test statistic: tau_c (1) = -8.67947
  asymptotic p-value 4.623e-015
  1st-order autocorrelation coeff. for e: 0.002
```

```
Augmented Dickey-Fuller test for CMA
testing down from 14 lags, criterion AIC
sample size 215
unit-root null hypothesis: a = 1

  test with constant
  including 0 lags of (1-L) CMA
  model: (1-L) y = b0 + (a-1) *y (-1) + e
  estimated value of (a - 1): -0.899664
  test statistic: tau_c (1) = -13.2023
  p-value 1.028e-022
  1st-order autocorrelation coeff. for e: -0.006
```

Figure number 3: for each return time series of the different factors used in the analysis an Augmented Dickey Fuller test was carried out to test for stationarity. The null hypothesis for the test is that a unit root is present, implying that we are in the context of spurious regressions, thus non-stationarity. The metric of relevance here is the p-value which allows us to determine if we are able to reject the null hypothesis. The general rule of thumb is that if the p-value is lower than 0,05 then we can confidently reject H0.

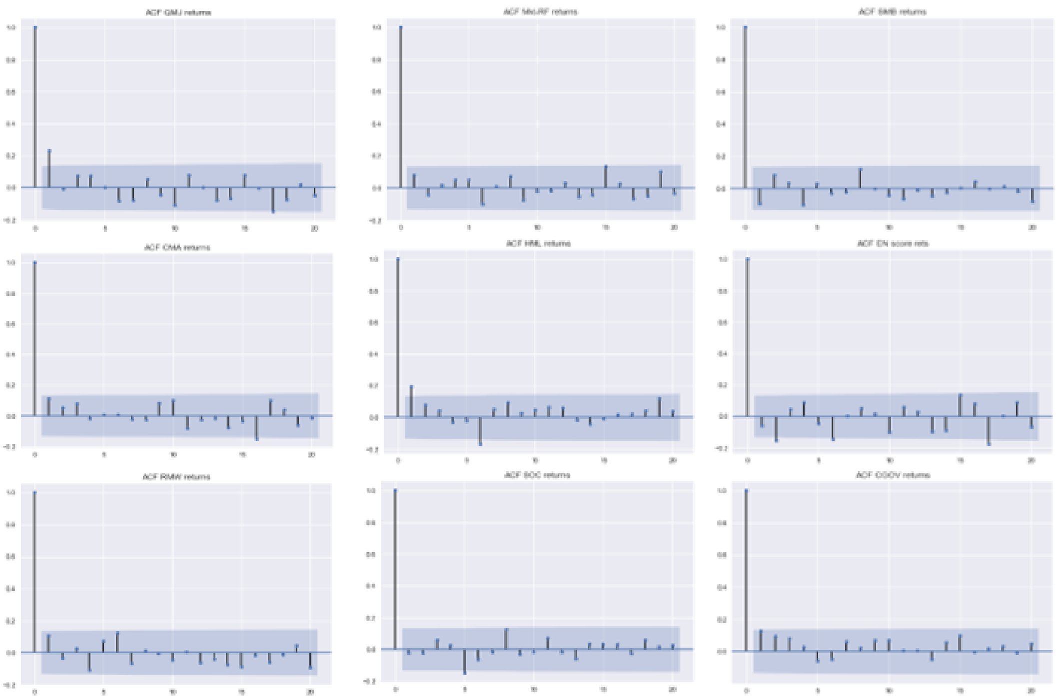


Figure number 4: for each return time series of the different factors used in the analysis the autocorrelation function has been displayed. The ACF is a very useful graph as it helps in determining how the correlation between any two values of the series behaves and it gives us information about the memory and persistence of the process.

	Mean	Median	Minimum	Maximum
MktRF	0.0052559	0.010917	−0.18990	0.12795
SMB	−0.00029291	0.00084884	−0.055111	0.059395
HML	−0.0036265	−0.0034977	−0.15156	0.078896
EN_rets	−0.0019124	−0.0023650	−0.034450	0.030249
SOC_rets	−0.00091176	−0.00088016	−0.029787	0.039414
CGOV_rets	1.3646e-005	0.00074216	−0.041039	0.044132
QMJ	0.0029602	0.0022204	−0.090987	0.087428
RMW	0.0018153	0.0028000	−0.089300	0.061800
CMA	-4.9074e-005	−0.00070000	−0.033500	0.051000

	Std. Dev.	C.V.	Skewness	Ex. kurtosis
MktRF	0.043941	8.3603	−0.86748	2.1457
SMB	0.023572	80.478	0.061247	−0.46217
HML	0.027174	7.4931	−0.82179	4.9487
EN_rets	0.0091359	4.7771	0.13940	1.3179
SOC_rets	0.0097360	10.678	0.36225	2.2823
CGOV_rets	0.010269	752.58	−0.20272	1.8922
QMJ	0.025819	8.7220	0.011799	1.8180
RMW	0.018819	10.367	−0.57809	3.1340
CMA	0.014930	304.24	0.39567	0.11406

Table number 1: for each return time series of the different factors used in the analysis summary statistics were calculated. These statistics include the mean, medium, minimum, maximum, average standard deviation lastly skewness and kurtosis.

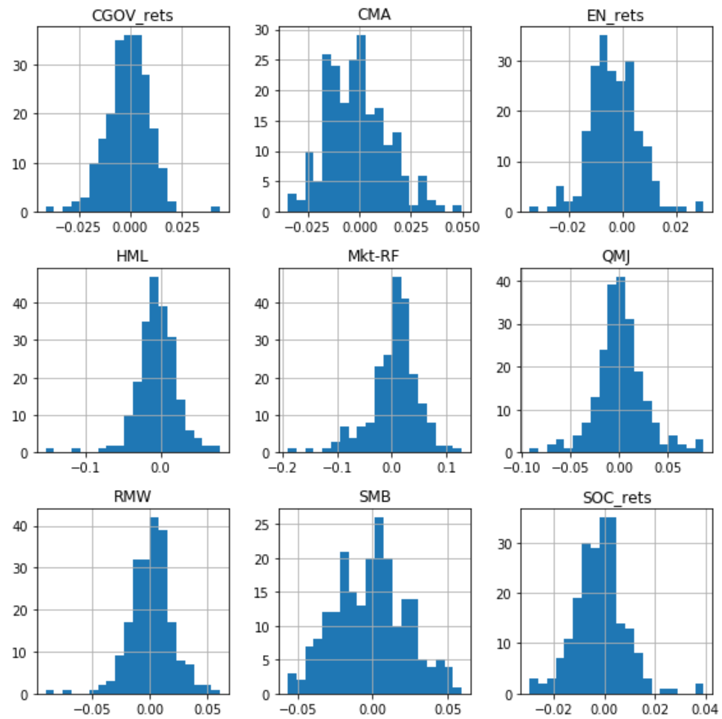


Figure number 5: for each return time series of the different factors used in the analysis a histogram to visualize the returns frequency distribution was plotted. This graph is useful as it allows to have a first impression about the 3rd and 4th moment, respectively skewness and kurtosis.

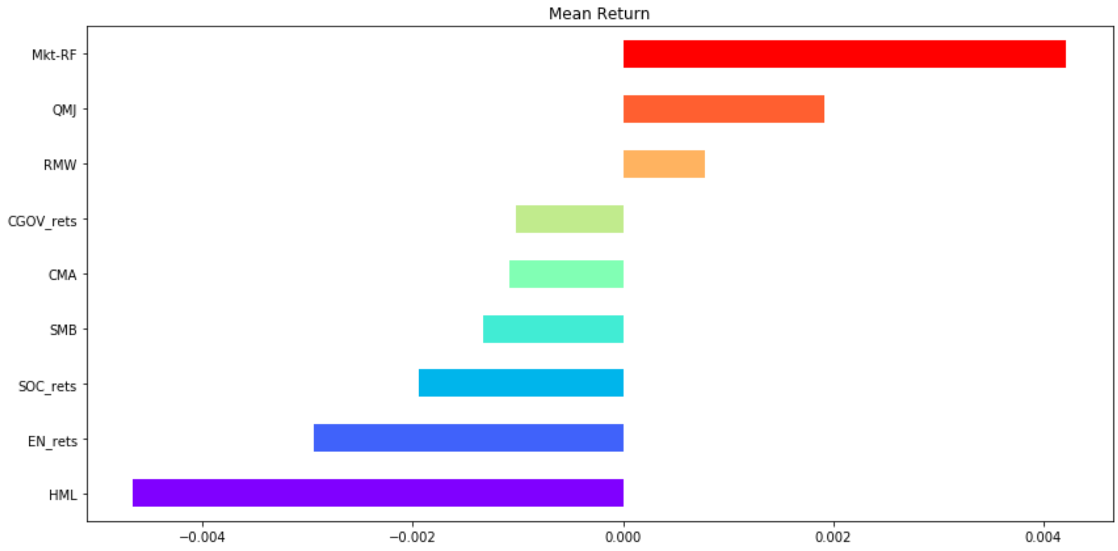


Figure number 6: for each return time series of the different factors used in the analysis the average monthly returns has been reported in this horizontal bar chart. It can be seen that of our nine time-series only three of them, Mkt-Rf, QMJ and RMW reported an average positive return, whilst the other factors consistently reported negative average returns on a monthly basis.

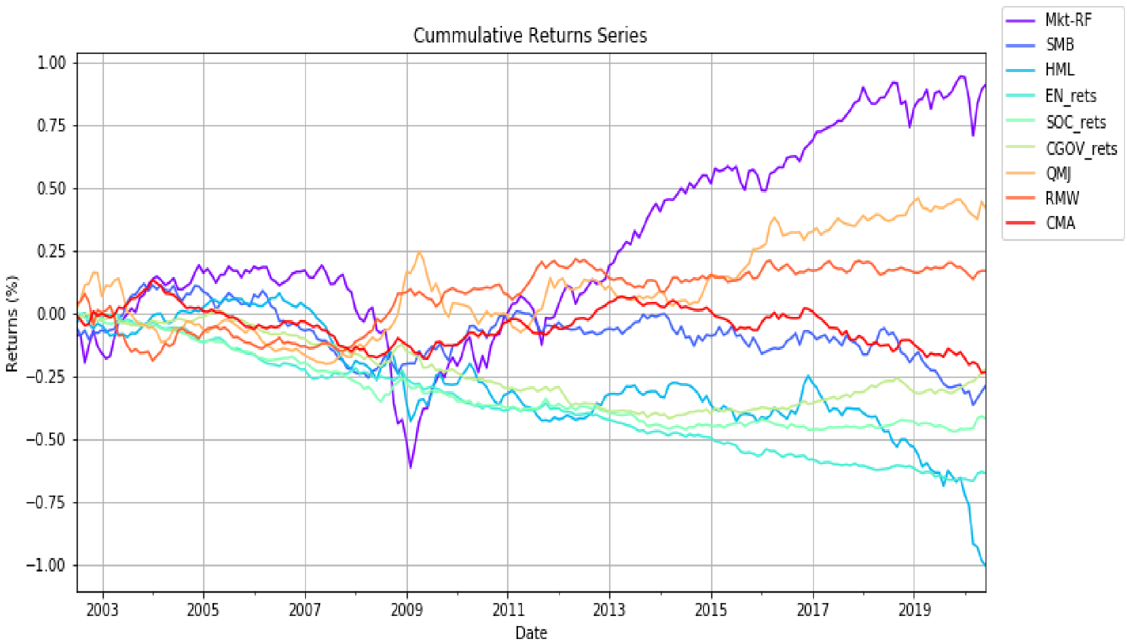


Figure number 7*:* for each return time series of the different factors used in the analysis the cumulative return in percentage terms has been reported in this horizontal line chart. The graph shows that the sustainability factors returned significant negative total returns over the period from 2002 to 2020.

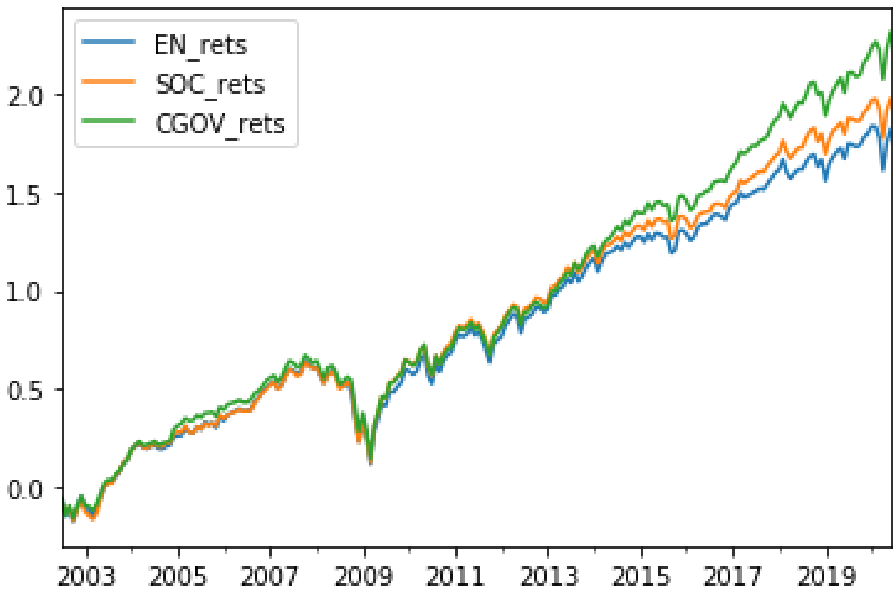


Figure number 8: for the E, S and G long component the cumulative return in percentage terms has been reported in this horizontal line chart. The graph shows that the long only portfolio for the sustainability factors returned significant positive total returns over the period from 2002 to 2020.

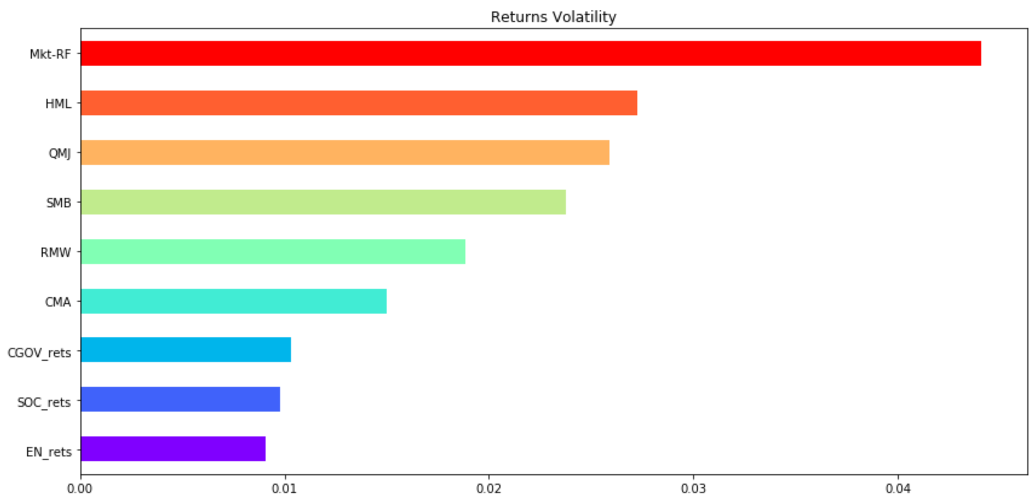


Figure number 9*:* for each return time series of the different factors used in the analysis the average monthly standard deviation has been reported. This graph shows us that the sustainability factors were the three with the lowest average volatility over our period of analysis.

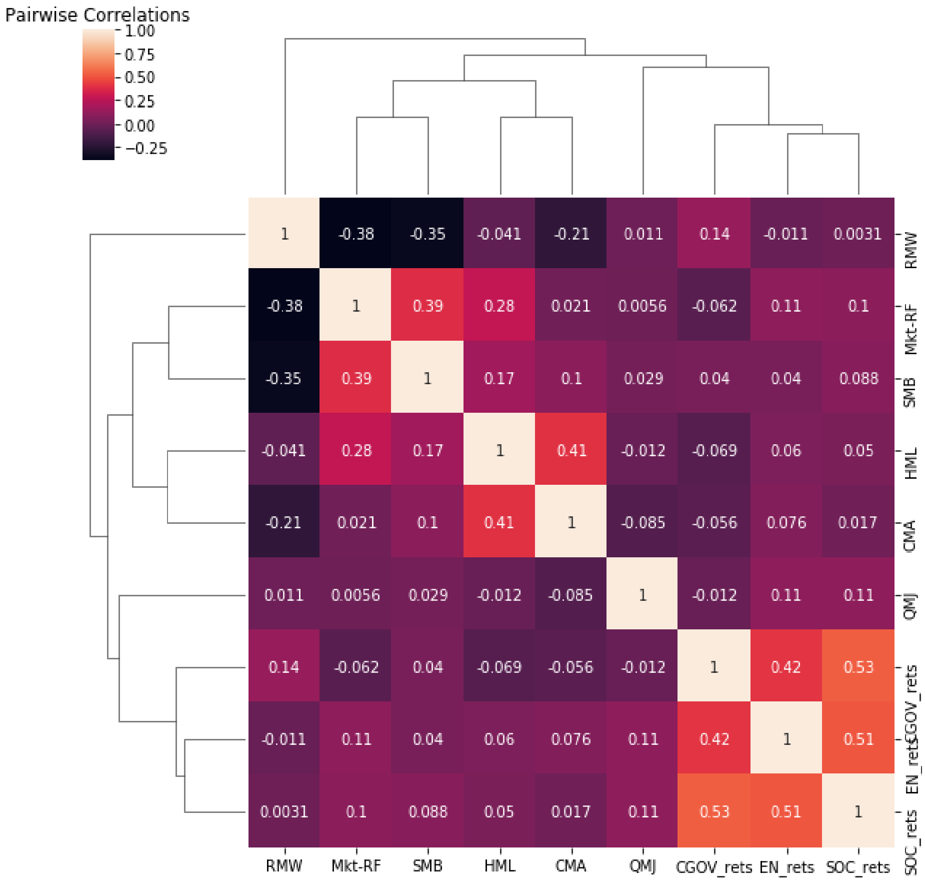


Exhibit 11: factors heatmap

Figure number 10*:* This graph depicts the correlation matrix between all the variables used in the analysis. It is a very useful tool of exploratory data analysis. They allow us to understand in a glance which variables are correlated, to what degree, in which direction, and alerts us to potential multicollinearity problems. As can be seen from the graph significant correlation values are only registered between the E, S and G components.

Environmental factor	α	Mkt-RF	SMB	HML	QMJ	RMW	CMA
difference Q4 -Q1	-0,20%	2,60%	-0,60%	-0,43%	4,11%	2,66%	6,00%
	-3,97	1,31	-0,25	-0,12	**1,94**	0,6186	1,26
Q4	0,93%	0,98%	4,50%	14,04%	9,15%	-0,25%	-0,05%
	3,75	0,11	0,28	1,18	0,66	-1,16	-0,21
Q1	1,30%	-4,12%	5,35%	15,40%	-0,51%	-28,08%	-18,12%
	4,820	-0,43	0,38	0,99	0,035	-1,125	-0,712

Social factor	α	Mkt-RF	SMB	HML	QMJ	RMW	CMA
difference Q4 -Q1	-0,22%	2,24%	2,67%	0,35%	3,78%	2,89%	1,07%
	-3,362	0,84	0,87	0,11	**1,39**	0,57	0,189
Q4	0,90%	-0,50%	6,00%	1,60%	7,00%	-23,00%	-13,00%
	4,028	-0,066	0,39	1,37	0,55	-1,05	-0,45
Q1	1,22%	-0,04%	0,02%	0,15%	0,04%	-0,28%	-0,12%
	3,370	-0,44	0,11	1,069	0,035	-1,32	-0,46

Governance factor	α	Mkt-RF	SMB	HML	QMJ	RMW	CMA
difference Q4 -Q1	-0,12%	-7,00%	5,50%	-2,60%	-1,10%	8,30%	-0,99%
	-1,65	-0,44	1,65	-0,87	-0,42	1,90	-0,18
Q4	1,10%	-1,30%	8,80%	11,00%	4,00%	-25,00%	-11,00%
	4,548	-0,1661	0,54	1,028	0,34	-1,07	-0,46
Q1	1,22%	-0,74%	-0,92%	1,74%	0,74%	-4,10%	-0,90%
	4,092	0,073	-0,059	1,145	0,048	**-1,680**	-0,32

Table number 2*:* The table reports the first tranche of time series regressions. For each E, S and G factor three different regressions were calculated, the first one with the dependent variable being the long-short factor mimicking portfolio returns (Q4-Q1), the second one used the returns of the long only component of the factor mimicking portfolio (Q4), whilst the last one used the returns of the short only component (Q1). The independent variables are the different return time series of the factors. The t-statistic were reported below the coefficient. In bold we highlighted the t-stats that imply statistically significant coefficients.

	EN rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,02%	0,20%	-0,20%	0,18%
	-3,343	**-3,749**	**-3,970**	**-3,230**
Mkt-RF	0,18%	0,70%	2,60%	2,10%
	1,1460	1,148	1,310	1,440
SMB	-8,00%	-2,20%	-0,60%	-22,00%
	-0,360	-0,400	-0,250	-0,810
HML	-0,10%	-0,19%	-0,43%	0,00%
	0,370	0,380	-0,120	0,034
QMJ		3,80%	4,11%	3,00%
		1,820	**1,940**	1,451
RMW			2,7%	0,0%
			0,619	-0,013
CMA			6,00%	5,90%
			1,260	1,456
EN				
SOC				34,00%
				5,306
CG				21,00%
				3,422

Table number 3*:* The table reports the second tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Environmental factor – remaining always fixed. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate the coefficients that are statistically significant)

	SOC rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,09%	-0,21%	-0,22%	-0,09%
	-1,300	**-3,004**	**-3,362**	-1,380
Mkt-RF	1,40%	1,80%	2,24%	2,20%
	0,520	0,640	0,840	1,234
SMB	1,50%	2,00%	2,67%	0,98%
	0,550	0,770	0,870	0,040
HML	0,54%	0,70%	0,35%	2,10%
	0,830	0,260	0,110	1,230
QMJ		3,70%	3,78%	4,40%
		1,480	1,390	**1,909**
RMW			2,9%	-2,4%
			0,570	-0,720
CMA			1,07%	0,54%
			0,189	0,130
EN				13,00%
				1,780
SOC				
CG				46,30%
				5,221

Table number 4*:* The table reports the second tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Social factor – remaining always fixed. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate that the coefficient are statistically significant)

	CG rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,10%	-0,10%	-0,12%	0,05%
	-1,230	-1,290	**-1,644**	0,820
Mkt-RF	-1,20%	-1,80%	-7,00%	-2,20%
	-0,704	-0,690	-0,440	-1,320
SMB	3,77%	3,00%	5,50%	4,20%
	1,130	1,151	**1,648**	1,460
HML	-2,37%	-2,00%	-2,60%	-2,30%
	-0,750	-0,720	-0,870	-1,080
QMJ		-1,00%	-1,10%	-3,40%
		0,420	-0,421	1,450
RMW			8,3%	5,5%
			1,904	**1,705**
CMA			-0,99%	-2,44%
			-0,182	-0,610
EN				11,00%
				1,924
SOC				51,00%
				6,230
CG				

Table number 5: The table reports the second tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Corporate Governance factor – remaining always fixed. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate the coefficients that are statistically significant)

	EN rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,02%	0,20%	-0,20%	0,18%
	-3,343	**-3,749**	**-3,970**	**-3,230**
Mkt-RF	0,18%	0,70%	2,60%	2,10%
	1,1460	1,148	1,310	1,440
SMB	-8,00%	-2,20%	-0,60%	-22,00%
	-0,360	-0,400	-0,250	-0,810
HML	-0,10%	-0,19%	-0,43%	0,00%
	0,370	0,380	-0,120	0,034
QMJ		3,80%	4,11%	3,00%
		1,820	**1,940**	1,451
RMW			2,7%	0,0%
			0,619	-0,013
CMA			6,00%	5,90%
			1,260	1,456
EN				
SOC				34,00%
				5,306
CG				21,00%
				3,422

Table number 6*:* The table reports an additional tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Environmental factor – remaining always fixed. In this case the long – short portfolio returns were calculated as the top decile minus the bottom decile. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate the coefficients that are statistically significant)

	SOC rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,09%	-0,21%	-0,22%	-0,09%
	-1,300	**-3,004**	**-3,362**	-1,380
Mkt-RF	1,40%	1,80%	2,24%	2,20%
	0,520	0,640	0,840	1,234
SMB	1,50%	2,00%	2,67%	0,98%
	0,550	0,770	0,870	0,040
HML	0,54%	0,70%	0,35%	2,10%
	0,830	0,260	0,110	1,230
QMJ		3,70%	3,78%	4,40%
		1,480	1,390	**1,909**
RMW			2,9%	-2,4%
			0,570	-0,720
CMA			1,07%	0,54%
			0,189	0,130
EN				13,00%
				1,780
SOC				
CG				46,30%
				5,221

Table number 7*:* The table reports an additional tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Social factor – remaining always fixed. In this case the long – short portfolio returns were calculated as the top decile minus the bottom decile. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate the coefficients that are statistically significant)

	CG rets dependent variable (2002 - 07 / 2020 - 06)			
Alpha	-0,10%	-0,10%	-0,12%	0,05%
	-1,230	-1,290	**-1,644**	0,820
Mkt-RF	-1,20%	-1,80%	-7,00%	-2,20%
	-0,704	-0,690	-0,440	-1,320
SMB	3,77%	3,00%	5,50%	4,20%
	1,130	1,151	**1,648**	1,460
HML	-2,37%	-2,00%	-2,60%	-2,30%
	-0,750	-0,720	-0,870	-1,080
QMJ		-1,00%	-1,10%	3,40%
		0,420	-0,421	**1,647**
RMW			8,3%	5,5%
			1,904	**1,705**
CMA			-0,99%	-2,44%
			-0,182	-0,610
EN				11,00%
				1,924
SOC				51,00%
				6,230
CG				

Table number 8*:* The table reports an additional tranche of time series regressions. Four different regressions were computed with the dependent variable – returns of the long short portfolio for the Corporate Governance factor – remaining always fixed. In this case the long – short portfolio returns were calculated as the top decile minus the bottom decile. In the first regression the independent variables are the classic Fama-French variables, in the second one the Quality minus Junk developed by Frazzini, Asness was added, in the third one we included also the Robust minus Weak and Conservative minus Aggressive factors and lastly all the previous variables as well as the other sustainability factors were included. The table reports the estimated regression coefficients with the respective values for the t-statistics (the ones highlighted in bold indicate the coefficients that are statistically significant).

	Lambda estimation without Newey West correction								
	Mkt-RF	SMB	HML	QMJ	RMW	CMA	EN	SOC	CG
λ	8,40%	0,89%	0,30%	-5,30%	-0,24%	0,48%	-0,26%	-0,21%	-0,08%
	0,47	1,07	0,301	-0,68	-0,35	0,81	**-1,78**	-1,102	-0,33

Table number 9*:* the results from the second step of the Fama MacBeth procedure, the cross-sectional regressions, are reported in the table. The lambdas represent the estimates of the risk-premium for the factors included in the model specification. Below the lambda values we reported, as already done before, the t-statistics in order to assess the statistical significance of these parameters. In this case no correction to account for autocorrelation and heteroskedasticity problems was included.

	Lambda estimation with Newey West correction								
	Mkt-RF	SMB	HML	QMJ	RMW	CMA	EN	SOC	CG
λ	1,10%	0,90%	0,01%	-0,14%	-0,30%	-0,14%	-0,18%	-0,14%	-0,04%
	0,35	0,56	0,013	-0,08	-0,26	-0,14	-0,78	-0,52	-0,102

Table number 10*:* the results from the second step of the Fama MacBeth procedure, the cross-sectional regressions, are reported in the table. The lambdas represent the estimates of the risk-premium for the factors included in the model specification. Below the lambda values we reported, as already done before, the t-statistics in order to assess the statistical significance of these parameters. In this case we accounted for autocorrelation and heteroskedasticity problems using the Newey West procedure with HAC estimators.