LUISS Guido Carli

LIBERA UNIVERSITÀ INTERNAZIONALE DEGLI STUDI SOCIALI

Department of Economics and Finance

Chair of Econometric Theory

European equity portfolios and the effect of Sustainable and Responsible Investing

Supervisor: Cosupervisor:

Prof. Paolo Santucci De Magistris   Prof. Stefano Grassi

Candidate:

Arianna Luccini

706351

Academic year 2019/2020

Table of Contents

Executive summary

During the last decade Socially and Responsible Investing has become a widely accepted practice. The objective of this study is to find out if socially responsible stock investments guarantee an overperformance with respect to non-responsible ones. It replicates the United States-based study by Kempf and Osthoff (2007) for the European market, using an investible universe made by the EURO STOXX index's components as of 01/01/2009. The main strategy applied is buying portfolios with high ESG score and sell those with a low score built applying different incorporation techniques. The socially responsible portfolios are built using different incorporation criteria in order to also examine the impact on portfolio performance of the different socially responsible criteria. To measure performance we implement a multivariate regression study using Carhart's (1997) model.

The results of the analysis suggest that in Europe, in the period and for the portfolios analysed – consistently with what happens in the United States – there is an overperformance of the highly-rated stocks over the low-rated ones. Although, regarding the best ESG incorporation technique to exploit this overperformance, there is further investigation to be done.

Introduction

Among the challenges of our time, we find fighting climate change, production and consumption waste together with human capital related challenges like gender inequality. Public policy and public financing try to address these – and many other – challenges, for example the European Union committed 25% of its budget for 2021-2027 to contribute to the fight against climate change and to foster the transition to a low-carbon economy. Nonetheless, there is the need to underline that public financing alone will not be able to cope with these challenges. Private investments must contribute to the cause. This is why there is a need for investors – individuals and institutional investors – to engage in Sustainable and Responsible Investment (SRI) practices. The question to ask is whether these practices are profitable. The present work analyses if Sustainable and Responsible stock investments in Europe guarantee an overperformance with respect to traditional ones.

Sustainable and Responsible Investing consists in introducing environmental, social and governance considerations into the investment decision process. To do so, investors use ESG factors that are based on three pillars: the environmental pillar (E) to examine the company's interaction with the environment, the social pillar (S) to examine a company's interaction with society and its components and, finally, the corporate governance pillar (G) to analyse the quality of a company's management, culture, risk profile and other governance-related characteristics. Within each pillar, it is possible to define some key themes. Each key theme can be further investigated through some specific issues that are used to measure both the risks to which a company is exposed with respect to the three factors and the risks to which the company exposes the surrounding world producing negative externalities because of the way they use natural resources, human and intellectual capital. In order to include ESG considerations in the investment process, incorporation or active ownership can be used. Investors employ incorporation techniques when they incorporate ESG issues into existing investment practices and it can be done using three different kinds of strategies: integration, screening and thematic investing. Instead, active ownership is when investors are directly involved in the decisions made by a company they are already invested in, and their actions are designed to improve the company's risk management or develop more sustainable business practices. This can be done either through a direct engagement or proxy voting.

According to the Global Sustainable Investment Alliance (2018), the assets globally employed in sustainable investments in 2018 were $30.68 trillion and they have grown by 34% with respect to 2016. The same study assesses that 46% of these assets are from European countries. The main reasons for institutional investors (investors from now on) to choose to incorporate ESG considerations into their investment process are client demand and materiality in relation both to risk management and profitability (PRI, 2020). However, many are the obstacles to the incorporation of ESG factors in the investment-decision process. Among them we find the belief that incorporation means giving up returns and therefore violating the fiduciary duty, the time frame mismatch between the long-term impact of ESG-investing and investors' short-term performance evaluation, the absence of common standards on ESG valuation and the hardship of finding high-quality data. Not having common standards means that economic actors use many different ways to measure and incorporate ESG performance. Kotsantonis & Serafeim (2019) examining fifty random companies of the Fortune 500, find out that the firms analysed used twenty different ways to report employee health and safety data, "using a different terminology and, most importantly, different units of measure". These differences in reporting exist for any ESG issue hampering the comparability of companies' ESG performances. As a consequence, finding high-quality ESG data is still a hard task, despite the fact that new technologies improve data availability, accessibility and transparency. In fact, although an accurate non-financial company disclosure is desirable to enable investors to incorporate ESG issues in the investment process, Corporate Social Responsibility disclosure is often based on voluntary disclosure programmes and, being costly, firms are still sceptical about it. The European Union in 2013 issued a directive (amended in 2014) "on the annual financial statements, consolidated financial statements and related reports of certain types of undertaking". According to article 19a of the consolidated version of the directive, large entities of public interest, within the management report, must disclose "both financial and, where appropriate, non-financial key performance indicators relevant to the particular business". It must be noted that the European Union, in order to take into account all the different Member States' approaches and the fact that different companies have different characteristics such that a piece of information deemed as material for a sector/company may be irrelevant for another, made sure to introduce an extremely flexible framework

avoiding a one-size-fits-all approach through the indorsement of a set of national, international or EU-based frameworks like the Global Reporting Initiative or the International Integrated Reporting Framework. One of the consequences of this regulation – which is far from being strict – is that companies usually decide to disclose their ESG performance to satisfy the increasing demand for this kind of information, putting major emphasis on strengths diminishing weaknesses and gaining good publicity. It could also be that companies decide to greenwash or bluewash their reputation. Greenwashing happens when companies mislead investors and consumers about their environmental performance or about the environmental benefits of one of their products or services. Instead, bluewashing consists of "signing up to UN initiatives merely to appear aligned with the organisation which has a blue logo" (Johnson, 2010), without effectively fulfilling the obligations that the initiative entails. So, the Regulation on CSR disclosure is an important step to encourage companies, investors and consumers to take into consideration ESG issues, but there is still the need for improvements.

Given all these issues related to ESG data quality and information asymmetries, investors use ESG ratings to evaluate companies' ESG performances. ESG rating companies, therefore, have the function of information intermediaries between companies and investors (or other stakeholders), collecting and validating all the information publicly available and disclosed by companies. Although ratings may make investors' life easier, they have some shortcomings that need to be kept in mind when ratings are used and that can impede a further diffusion of their usage to make investment decisions. The main ones are the lack of uniformity among ratings form different providers and the lack of transparency on the criteria and assumptions used to compute them.

As mentioned before, this study focuses on testing whether companies' ESG rating influences portfolio returns and – in particular – if investors can gain extra-returns implementing Sustainable and Responsible Investment strategies. The literature review on Socially Responsible Investing's profitability highlights that many are not only the publications on this topic as but also the methodologies used to conduct them. Auer and Schuhmacher (2016) performing a portfolio study and using the Sharpe ratio as a portfolio performance indicator, find out that, in the European market there is no evidence of outperformance of high ESG rated portfolios with respect to the low rated ones and investors

often have to sacrifice returns for the sake of social responsibility. At the same time Kempf and Osthoff (2007), using a USA-based investible universe and performing a multivariate regression study with Carhart's model, find out that Socially Responsible Investing has a positive influence on US-based stock portfolio returns, therefore generating extra-returns with respect to non-SRI investing. Herremans, Akathaporn, and McInnes(1993), performing a portfolio study on manufacturing companies with a high corporate social responsibility reputation, find out that the latter outperform companies with a poorer reputation. Capelle-Blancard and Monjon (2014), performing an event study on French SRI mutual funds, conclude that they do not outperform the market or their non-SRI counterparts. Moreover, in Bauer, Koedijk, and Otten (2005) risk-adjusted returns of ethical and conventional funds have been compared without finding any significant evidence of an enhanced performance of ethical funds over the conventional ones. As can be noticed, there is still no agreement on whether financial and ESG performance have a positive, negative or neutral relationship. It must be noted that the majority of these studies are based on the American market. Therefore, this study replicates Kempf and Osthoff (2007) using a Europe-based investible universe in order to check if their findings are valid for the European market, trying to contribute to the clarification on what is the relationship between ESG issues and portfolio performance. The investible universe used in this study is the set of EURO STOXX's components as of 01/01/2009 and the period analysed goes from 01/12/2009 to 31/12/2019. Fifteen value weighted 10% cut-off portfolios are built; a Best, Worst and Long-Short portfolio for five different ESG screening methods. The ESG rating used in this study are the ones provided by Refinitiv-Asset4. The focus of the analysis is on the performance of the Long-Short portfolios – the result of going long on the Best portfolio and short on the Worst one – evaluated looking at their Sharpe ratios and at the Jensen's alpha generated from Carhart (1997) model performed using as dependent variable the monthly returns of the portfolios. The analysis shows that the Sharpe ratios are all positive, apart for the Long-Short portfolio obtained using the negative screening. Moreover, Carhart (1997) model's alphas are positive and are mostly statistically significant apart from a negative alpha for the negatively screened Long-Short portfolio. These results allow to infer that – as happens in Kempf and Osthoff (2007) for the United States – in the period analysed in Europe there is an overperformance of the high ESG-rated stocks over the low rated ones and the negative

screening technique is not suitable to gain an extra-return. This suggests that if European investment and asset managers – who prefer negative screening over the other incorporation techniques (Eurosif, 2018) – want to fully exploit the advantages of SRI, they should abandon negative screening for more complex and modern techniques. In order to be able to generalise the conclusions just drawn, some robustness checks are performed to check whether results are influenced by the portfolio weighting scheme: equally weighted 10% cut-off rate portfolios, instead of value weighted, are built. Moreover, it is checked if results are influenced by the cut-off rate: for every incorporation technique, a portfolio with 5%, 25% and 50% cut-off rates are created. The robustness check using equally weighted portfolios confirms that it can be inferred that there is the possibility to gain an extra-return investing in SRI portfolios with respect to non-SRI portfolios, but it also suggests the weighting scheme strongly influences the magnitude of the impact of the ESG performance on portfolios' returns. Performing the robustness check using different cut-off rates, we get contradicting results since – as in Lanza, Bernardini & Faiella (2020) – the choice of a particular threshold leads to a Best portfolio overperforming over the Worst one while with another cut-off rate the situation is reversed. Lanza, Bernardini & Faiella (2020) deal with this situation using a Machine Learning approach to determine which thresholds allow the Best portfolios to have superior risk-adjusted performances. So, it could be that using a more sophisticated investment decision process that allows to determine the optimal cut off rates is possible to get – in a robustness check like the one just performed – results more consistent with the rest of the analysis.

Two are the main limitations of this study: it is based on just one ESG rating provider and, given that ratings for the same company can be significantly different, the same analysis should be performed using at least another rating provider to be able to assess with certainty that the findings of this study mirror what happens in reality. Another limitation is that the Carhart (1997) model does not allow to determine a priori whether a significative alpha is due to market inefficiency or to a risk factor that is not captured by the model.

Overall, the results of this study suggest that companies' ESG performance represents valuable information for investors. Future research should try to assess whether the extra-return observed for companies with high ESG rating using Carhart (1997) model, is the result

of the compensation for an additional risk factor or is just the consequence of temporary mispricing.

The remainder of the work proceeds as follows. Chapter 1 gives further details on what means Socially Responsible Investing, analysing the reasons for investors to implement it together with the obstacles to its implementation. It also illustrates the European Regulation on companies' ESG disclosure. Chapter 2 dig into ESG ratings, their advantages and shortcomings. It also talks about the different ESG rating providers. Finally, it gives details about Refinitiv's ESG Combined scores that are used in the analysis conducted in chapter 3. Chapter 3 examines in depth the results of the analysis conducted.

1   ESG factors and equity portfolios

1.1   Sustainable and Responsible Investing

ESG factors are based on three pillars: the environmental pillar (E) to examine the company's interaction with the environment, the social pillar (S) to examine a company's interactions with society and its components and the corporate governance pillar (G) to analyse the quality of a company's management, culture, risk profile and other governance-related characteristics. Within each pillar is possible to define some key themes; each key theme can be further investigated through some more specific issues (Table 1).

Key Pillars	Key Themes		Key Issues
Environment	Climate change	Carbon footprint	Vulnerabilities from climate change events
	Natural resources	Energy efficiency Sourcing of raw materials	Water efficiency Usage of land
	Pollution and waste	Toxic emissions Wastewater management Hazardous materials management	Air quality Electronic waste management
	Opportunities and policy	Renewable energy Clean technology	Green buildings Environmental and biodiversity targets and investment
Social	Human capital	Workplace health and safety Development opportunities	Employee engagement, diversity, and inclusion Labor practices (e.g., wages, working conditions)
	Product responsibility	Product safety and quality Selling practices and product labeling	Customer privacy and data security Access to products
	Relations	Community Government	Civil society
Governance	Corporate governance	Board structure and accountability Accounting and disclosure practices	Executive compensation and management effectiveness Ownership and shareholder rights
	Corporate behavior	Management of corruption Systemic risk management Earnings quality	Competitive behavior Management of business environment (e.g., legal, regulations) Transparency on tax and related-party transactions

Table 1- The scope of ESG factors (Source: Global Sustainability Report IMF (2019))

Each ESG issue, theme and pillar can be used to measure the risks to which a company is exposed with respect to the three factors together with the risks to which the company exposes the surrounding world producing negative externalities because of the way they use natural resources, human and intellectual capital. In essence, these factors measure a company's level of Corporate Social Responsibility (CSR) which is "the voluntary integration of social and environmental concerns of companies in their business operations and in their relations with stakeholders" (European Commission, 2001). Given that, ESG pillars and issues can be evaluated in many different ways: ESG data are multifaceted and can be defined as alternative data, i.e. data that "do not necessarily show up in a company's quarterly or annual financial reports, but still have material long-term implications for company value" (In, Rook, & Monk, 2019). Alternative data's unique value stands in its power to explain investment performance when conventional financial data can't (In, Rook, & Monk, 2019). This kind of data is being increasingly included in the investment analysis process. Equity investors started using ESG information no sooner than three decades ago; but only in recent years it has become a widely spread practice among investors that look either for ways to avoid specific reputation-damaging risk exposures or for long-term-value-creating information (International Monetary Fund, 2019). To acknowledge the evolution of this phaenomenon, and today's commitment of the institutional-investment community, the evolution across time of the number of United Nations' Principles for Responsible Investment[1] (PRI from now on) signatories can be analysed. In the chart below (Figure 1) we can see that the number of signatories started increasing after the financial crisis in 2008, but it is from 2014 that an exponential increase can be observed: in 2014 the number of

---

[1] The United Nations' Principles for Responsible Investment (PRI) is a global organization that encourages and supports transparency on ESG issues together with the uptake of responsible investment practices in the investment industry. In partnership with UN Environment Programme Finance Initiative and the UN Global Compact, it works "to achieve this sustainable global financial system by encouraging adoption of the Principles and collaboration on their implementation; by fostering good governance, integrity and accountability; and by addressing obstacles to a sustainable financial system that lie within market practices, structures and regulation." Becoming a PRI signatory, an organization commit to implement six Principles for Responsible Investment. The signatories come from over 60 countries around the world. (https://www.unpri.org/)

signatories were 1,250 for \$45 trillion of assets under management, while in 2019 the signatories were over 2,250 (more than an 80% increase with respect to 2014) for \$85 trillion of assets under management.
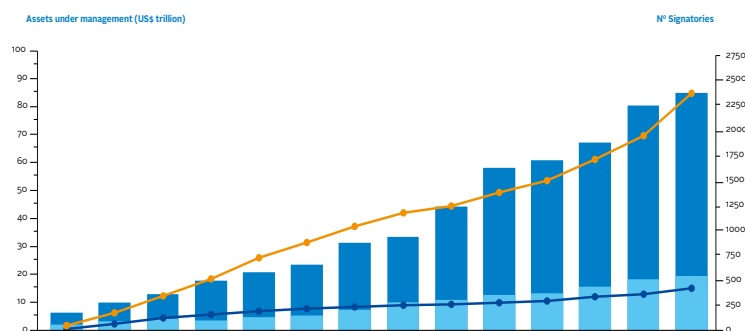

Figure 1 - Assets under management and number of signatories of UN's Principles of Responsible Investment between 2006 and 2019 (Source: UN-PRI website)

When adopting ESG considerations in the financial analysis and selection process of securities, equity investors are implementing a Sustainable and Responsible Investment (SRI) approach. A precise definition of SRI is yet to be found, academia's focus is on ESG factors' influence on performance, without worrying about giving specific definitions of what ESG or SRI are using interchangeably terms like SRI, Responsible Investing, Impact Investing or Ethical Investing. According to Eurosif (2018), Sustainable and Responsible Investing is a long-term investment approach that "combines fundamental analysis and engagement with an evaluation of ESG factors in order to better capture long term returns for investors and to benefit society by influencing the behaviour of companies".

1.1.1 Incorporation and active ownership

Including ESG considerations in the investment process can be done in two ways: incorporation and active ownership. Active ownership is when investors are directly involved in the decisions made by the companies they are already invested in and their actions are designed to improve the company's risk management or develop more sustainable business practices. This can be done either through a direct engagement or proxy voting. While direct engagement entails discussing environmental, social and governance

issues with a company's management, in order to improve its approach to such issues, proxy voting consists in voting against resolutions non ESG-friendly in the shareholders' meetings. Investors use incorporation technique when they incorporate ESG issues into existing investment practices. It can be done using three different kinds of strategies (PRI, 2020): integration, screening and thematic investing. Integration is when, in order to produce better returns and risk management, there is an explicit and systematic inclusion of ESG risks and opportunities in the investment analysis and decision process, that traditionally would be focused only on financial analysis. Integration, therefore, means modifying the research, security valuation and portfolio valuation processes to include ESG parameters. To do so, investors analyse both financial and ESG relevant information in order to identify which issue is material to a company's performance. Then, they create a "centralised research dashboard" (CFA institute, 2018) that collects both financial and ESG data. The data gathered allow investors to modify – in light of the possible impact of material ESG issues on the company's, sector's or country's performance – the already estimated financials (e.g. revenues), valuation model variables (e.g. discount rates) or valuation multiples. Moreover, ESG data influence both Value at Risk models and portfolio weightings, in order to "mitigate ESG risks exposures and avoid breaching risk limits" (CFA institute, 2018). There is the need to specify that the integration process does not entail any limitation to the investible universe, and – most of the times – offers the opportunity to decide between two companies, sectors or countries having identical characteristics apart from the ESG performance. Investors may also choose to apply some screens to the investible universe. It can be done in the form of negative screening, norm-based screening, positive screening and best-in-class approach. Negative screening, also defined as exclusion, happens when investors decide to leave out from the investible universe, following an ethical rule, all securities issued by companies of certain sectors and involved in certain activities considered as 'sinful'. It is the oldest strategy employed to incorporate ESG considerations in the investing decision. According to Eurosif (2018) the most commonly excluded sectors are: weapons, pornography, gambling and tobacco. The norm-based screening is when investments are screened on the basis of national or international norms covering ESG matters. Positive screening is when investors decide to invest only in the top ESG performers, without paying attention to which sector or industry businesses belong. The best-in-class approach, a

particular kind of positive screening that ensures an industry-balanced portfolio. It is implemented selecting the best ESG performers per each industry or sector. Finally, Thematic investing is when investors seek to combine attractive risk-returns with a specific social or environmental outcome. We include in this strategy two approaches: impact investing and sustainable themed investing. Impact investing is when investments are made with the objective to generate, not only enhanced returns, but also an environmental and social positive impact. Examples of impact investing are microfinance and community investing. Sustainable themed investing is implemented in order to build investment portfolios focusing on a specific ESG theme like climate change or issues like water efficiency. All the strategies above may be combined together. For example, sustainably themed investing, as well as negative screening, can be used as the basis for best-in-class approach if the investor wants to focus on a particular issue.

1.1.2   SRI diffusion statistics

As said above, there is no unanimity on a specific definition of SRI, so the statistics about its diffusion change depending on how SRI is defined. According to the Global Sustainable Investment Alliance (2018), the assets under management globally employed in sustainable investments in 2018 were \$30.68 trillions and they have grown by 34% with respect to 2016 (\$22.89 trillion). The same study assesses that 46% of these assets are from European countries and that the most used investment strategy at a global level is the negative screening. The SRI investment strategy with the highest growth rate in usage are the best-in-class approach (+125% AuM[2] with respect to 2016) and impact investing (+269% AuM with respect to 2016). Focusing on the European market – using the data provided by Eurosif (2018) – it can be stated that the region is in line with the global trend when talking about most used strategies. In fact, the first choice of European asset managers is negative screening, even if in 2015 this strategy started having a setback (its use decreased by 3% between 2015 and 2017). The second most used strategy is ESG integration, whose use – between 2015 and 2017 – experienced the highest growth rate (+27%); while the best in class approach is the third last strategy by usage, but it is starting to be more popular: only

---

[2] Assets under Management

in 2017 its use has increased by 19% (Global Sustainable Investment Alliance, 2018). Also the recurse to active ownership has increased throughout the years, this increase is the product of investors' increasing need of performing a more active management. The most significant decrease in use – between 2015 and 2017 – has been registered in the norm-based screening strategy: -21% (Eurosif, 2018). We can conclude that asset managers are starting to abandon more traditional strategies, as negative screening, for newer and more sophisticated ones such as integration and best in class approach. In fact, the only thing done with negative screening is to exclude certain companies from the investible universe. This restriction can create problems of scarce diversification that can be avoided using the best-in-class and integration approaches. In addition, Eurosif's (2018) study shows that Sustainability themed investing between 2016 and 2018 didn't experience any significant change but it is evident that investors are starting to focus especially on the themes of climate change and water management, while no preference was detected in the previous edition of the study. This means that investors are starting to identify the areas of sustainability they are more interested in. It can be seen as the result of increased frequency and popularity of discussions about climate change, both on media and at policy level.

1.1.3 Reasons for SRI implementation

There are several reasons why investors choose to incorporate ESG considerations into their investment process. The main ones are materiality, client demand and regulation (PRI, 2020). As far as materiality is concerned: many are the studies looking for evidences of the over- or under-performance of ESG-integrated investment portfolios with respect to traditionally built ones. Most of these studies (90% according to PRI) find out that taking into consideration ESG factors in the investment process influences investment returns in a positive way. Moreover, ESG factors are material also to risk management. In fact, excessive exposure to environmental and social risks[3] may be extremely detrimental to a firm's performance and therefore to its equity return. Note that materiality is not a static concept, but a dynamic one: it depends on investors' needs and constraints and on a company's characteristics. In fact, investors – in order to assess materiality– analyse the different issues

---

[3] See paragraph 1.1.3.1

looking at the country and sector in which a firm operates, paying attention to any possible regulatory or technological change that may take place in the sector itself, and to all the business' unique characteristics. Moreover, there is not only an increasing focus on environmental and social impact of companies' business decisions but also a growing awareness that ESG factors influence returns and reputation, hence investors are increasingly asking both investment managers and issuers to pay more attention to ESG matters. The final reason for investors to take into account ESG considerations is that, because of the regulators' realization that the financial sector can have a great role in meeting social and environmental challenges, a significant increase in responsible investment regulation is taking place (Figure 2). This increase started being relevant especially after the 2008 financial crisis (PRI, 2020). SRI is, in fact, sponsored by the European Union as a way to "support jobs, growth, competitiveness and a low carbon economy" together with the creation of a stronger and more resilient financial system (European Commission, 2016).
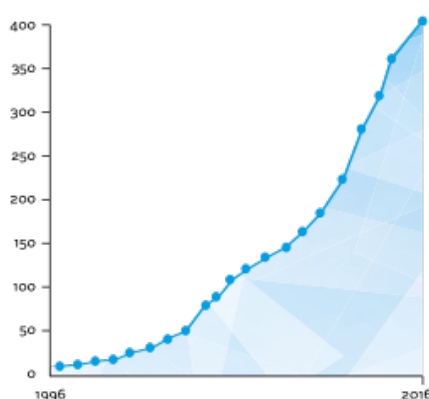
Figure 2- Cumulative number of policy intervention 1996-2016

(Source: United Nation Principles of Responsible Investment website)

According to Eccles, Kastrapeli and Potter's (2017) survey[4], investors' interest in ESG incorporation is mostly "market-driven": among the asset managers who responded to their survey, a great majority (38%) indicated their clients' demand as the principal reason for ESG incorporation, whilst only 18% of them indicated as principal driver the increasing

---

[4] A global survey directed to 582 institutional investors that "were or were planning to incorporate" ESG matters.

strictness of regulatory requirements, even if the number of policy interventions have been dramatically increasing in the past twenty years.

1.1.3.1 ESG and portfolio risks

According to PRI (2020) ignoring ESG factors means ignoring "risks and opportunities that have a material effect on the returns delivered to clients and beneficiaries". The risks mentioned by PRI are those related to the occurrence of value-destroying events like ESG incidents or of events out of a firm's control that result in negative externalities for the firm itself. ESG incidents are all those events resulting from "unethical or even illegal corporate behaviour" (Glossner, 2018) that generate a reputational damage for the company. Today, thanks to Internet and social media, the news about any kind of misbehaviour will spread as fast as ever. People will "use their greater knowledge to vote with their wallets" (Haefele, 2018) penalising a company's returns. Moreover, whenever there is a reputational damage, the ability of a firm to charge premium prices or attract talents is hampered with an even more severe consequence on a company's returns and valuation. A significant example of ESG incident is BP's oil spill in 2018, after which the company's stock price sunk. If an ESG incident happens, there is also the risk of incurring in regulatory actions. In fact, "firms that generate profits by imposing significant negative externalities on society are at a higher risk of regulatory action" (Ioannou & Serafeim, 2011). Hence, companies that do not take into account ESG issues in their business decisions, have a higher probability of being subject to sanctions and regulatory actions than those who decide to do so. The government's actions may, indeed, produce not only a negative effect on returns, but also a reputational damage. Negative externalities may come from both the environmental and social pillar, but most of the events out of a firm's control are related to the climate and the environment. Environmental risks can be divided into: physical and transition risks. Physical risk represents the risk that natural disasters and fatalities, as well as broader climate trends, may damage the property, land or infrastructure of a company; transition risk is the risk that some assets – because of evolving climate policies and technologies – in the transition to a lower carbon economy, may lose value as a result of divesting. In fact, "a growing number of asset owners have pledged to divest from fossil fuels" (International Monetary Fund, 2019).

Committing to Sustainable and Responsible Investing, therefore, means investing in companies whose management takes into account ESG issues and – as a consequence – the risks listed above, being able to reinvent itself in light of the transition to a low carbon economy and to invest in ESG-friendly projects avoiding, as much as possible, any potential source of reputational damage. Hence, including ESG consideration in the investment decision process helps managing the portfolio's risk exposure.

1.2 Obstacles to the implementation of SRI

Many are the obstacles to the incorporation of ESG factors in the investment-decision process. Among them we find: the absence of common standards on ESG valuation and the hardship of finding high quality data, the belief that incorporation means giving up returns therefore violating the fiduciary duty and, finally, the time frame mismatch between long-term impact of ESG-investing and investors' short-term performance evaluation.

1.2.1 Absence of common standards and the hardship of finding high quality data

The absence of a common ESG terminology and of an adequate and clear regulatory framework cause both the lack of common standards on ESG valuation and bad ESG data quality (European Commission, 2018). Not having common standards means that economic actors use many different ways to measure and incorporate ESG performance. Kotsantonis & Serafeim (2019) examine fifty random companies of the Fortune 500, finding out that the firms analysed used twenty different ways to report employee health and safety data, "using a different terminology and, most importantly, different units of measure". In fact, a consistent corporate disclosure system on ESG matters is missing both at a European and international level, so firms can – most of the times – decide by themselves what to disclose and which level of detail to use[5]. So, finding high quality ESG data is still an hard task in spite of new technology improving data availability, accessibility and transparency. These differences in reporting exist for any ESG issue, generating problems of data inconsistency both between companies and across time, hampering – therefore – the comparability of ESG performances. Consequently, a more transparent measurement and reporting could be

---

[5] See paragraph 1.3

fostered by the introduction of a common integration policy. However, it needs to be emphasized that investors' needs and constraints – on which investors base their data collection – cover an extremely wide range. It needs also to be noted that ESG data are both multifaced and context-dependent, so that every investor has its definition of high-quality data. There are some characteristics that ESG data must have (e.g. reliability, granularity, freshness, comprehensiveness, actionability and scarcity), but the relative importance of every one of these characteristics depends on the investor's priorities and constraints, and, most of the time, there is a trade-off between the various characteristics (In, Rook, & Monk, 2019).

In the European Union, in order let investors use common standards, the European Technical Expert Group on Sustainable Finance introduced a taxonomy on climate change (planning to define, in the future, other taxonomies for all the remaining environmental and social issues) with the objective to create a classification of what can be considered as 'green' (and 'sustainable'), to help investors to determine whether a business can be considered environmentally (and socially) sustainable and creating a "common language between investors, issuers, project promoters and policy makers" (EU Technical Expert Group on sustainable finance, 2020).

As will be illustrated more in depth in chapter 2, another consequence of the absence of consistent disclosure mechanisms is the inconsistency between the different ESG ratings.

1.2.2   Underperformance and violation of fiduciary duty

Many argue that environmental and social issues are not material for companies nor for investment portfolios and the belief that incorporating ESG factors means sacrificing returns is still widespread among practitioners. This idea is strengthened by the fact that research produces mixed evidences on whether ESG integration generates over- or under-performance. The belief of ESG-related underperformance fosters another obstacle to the diffusion of ESG investing: the conviction that these strategies imply the violation of fiduciary duty. Fiduciary duty is an ensemble of obligations that common law jurisdictions[6]

---

[6] While Fiduciary duty is explicitly mentioned in common law legal systems, in civil law jurisdictions (as the European Union and most of its Member States) there rarely is an explicit reference to it. Nonetheless, every

establish to manage information asymmetries in any relationship where there is someone "undertaking to exercise some discretionary power in the interests of another person" (Mugnier, Delerable, Tan, & Hélouin, 2014). The core duties comprehended in fiduciary duty are loyalty and prudence. The former states that the interests of the beneficiaries must be enacted in good faith, and any conflict of interest should be managed or eliminated. The latter consists in the fact that due care, skill and diligence should be adopted by fiduciaries in their actions "avoiding speculative and unduly risky investments " (Mugnier, Delerable, Tan, & Hélouin, 2014). The reason why many investors think of ESG integration as a violation of fiduciary duty, is that they interpret this concept in a strict and traditonal way. Traditionally fiduciary duty, in fact, has been interpreted as the duty to maximize short-term profits without any regard to other external factors. In reality, fiduciary duty is an organic and everchanging concept and its interpretation must change together with the everchanging world we live and invest in (UNEP FI, 2015; European Union; Union, 2014). So, using an extensive interpretation of fiduciary duty we can deem as compliant with the duty of loyalty considering in investment decisions not only future investment returns, but also other factors as ESG ones if they contrubute to maximizing long-term returns. Moreover, avoiding "unduly risky investments" can be interpreted as reducing a portfolio's exposure not only to traditional risks but also to the risk of be subjected to negative externalities generated by environmental and social matters or to the risk of ESG incidents. In fact, as previously pointed out, one of the consequences of the transition to a lower carbon economy is that some assets (i.e. fossil fuels) may be stranded and loose value and all the investments linked to that asset will have increasingly lower returns, while opprotunities for companies specialized in more environmentally friendly products will arise. Therefore it would be in the best interest of beneficiaries, avoiding investing in assets that are going to be stranded and focus on new businesses, or at least business that are trying to become more and more independent from the said asset. A real turning point on this matter has been the so called

---

EU Member State has similar specific obligations to institutional investors. For instance, in Italy the duty to act prudently has been introduced as the duty to act 'professionally'.

Freshfield Report[7] (2005) that concluded that ESG inclusion is obligatory when ESG issues are relevant from a financial point of view, hence when ESG is likely to have an impact on financial performance or evaluation. Moreover, according both to Sullivan, Martindale, Feller, & Bordon (2015) and to Mugnier, Delerable, Tan, & Hélouin (2014), there is no actual legal obstacle to ESG integration. And the High Level Expert Group (2018) says it is necessary to make clear that "in fulfilling their duties, investors should incorporate sustainability factors consistent with the broad interests, investment horizons and sustainability preferences of their clients and beneficiaries". Therefore we can conclude that integration is indeed compatible with fiduciary duty: it is sufficient to intepret it in a more modern way.

1.2.3 Time horizon

Another reason put forward by investors for not incorporating Environmental, Social and Governance factors is that it takes too long to see ESG investing advantages. There is indeed a time-frame mismatch: ESG factors have an impact on long term risk-return profile of investment portfolios while financial markets are mainly skewed towards the short term. In fact, sustainable investing entails large investments in long-lasting assets whose amortisation happens over time. Hence sustainable finance is "axiomatically linked to the long term" (EU High-Level Expert Group on Sustainable Finance, 2018). Eccles, Kastrapeli, & Potter (2017) survey supports this time-frame mismatch: 60% of the interviewees aknowledge the fact that ESG investing outperformance would come in five to ten years. At the same time 34% of asset managers interviewed have an investment horizon between three and five years (medium term), 12% have a time horizon btween five and seven years, and only 11% of them have a time horizon higher than 10 years. Moreover, the time frame of performance evaluation of portfolios in most cases is at best of 5 years. So, many institutional investors still worry about short term performance, therefor a shift from short-term profit evaluation model to a long-term profit one should be made for investors to fully appreciate the benefits of SRI. To that extent, it is important that not only trustees but also beneficiaries understand

---

[7] A report commissioned by UNEP FI's Asset Management Working Group in 2005 to the Freshfields Bruckhaus Deringer.

the potential economic benefits of SRI and its returns' long-term sustainability (European Commission, 2018). According to EU High-Level Expert Group on Sustainable Finance (2018) there is the need for a stable regulatory framework harmonized with the long-term European sustainable goals, allowing investment managers to align their time-horizons with the ones of end investors. However, it must be noted that short-termism is incentivised by existing regulation like the accounting mark-to-market rule for assets held in long term portfolios (EU High-Level Expert Group on Sustainable Finance, 2018) together with the fund and asset managers' compensation mechanisms. Nonetheless short termism is not a problem affecting financial markets only, also companies are too often focused more on the short term than on the long term. This attitude is detrimental to any sustainable investment that companies could undertake and therefore an obstacle to the diffusion of ESG corporate investments.

1.3   Corporate sustainability disclosure

While ESG investing has been gaining much popularity ESG performance disclosure isn't as popular. However, given that the source of ESG information on which investors base their investment is mainly company disclosure, an accurate non-financial company disclosure is desirable in order to enable investors to incorporate ESG issues in the investment process. However, ESG disclosure lacks standardization and is often based on voluntary disclosure programmes like the Global Reporting Standards[8] or the ones promoted by the International Integrated Reporting Council [9]. As a consequence, as already mentioned, there may be significant differences in the kind of information disclosed, the metrics used and the way in which it is disclosed. As far as the means to disclose are concerned, companies may opt either for more traditional disclosure methods like a report or for unconventional ones like

---

[8] GRS: a series of reporting standards in order to allow companies to provide for "an inclusive picture of an organization's material topics, their related impacts, and how they are managed" https://www.globalreporting.org/standards/gri-standards-download-center/

[9] IR is sponsored by the International Integrated Reporting Council (IIRC, https://integratedreporting.org/) that provides a framework to implement IR aiming at the improvement of information availability and firms' transparency to encourage efficiency and productivity in capital allocation.

publishing information on websites, social media or even releasing interviews to the press. Already in 2009 more than 75% of SP500's companies had a section in their website disclosing ESG policies and performances (Cuerel Burbano & Delmas, 2011). Reports disclosing ESG information for stakeholders and the general public are usually called sustainability reports. The latter can be either in a stand-alone or integrated format. The stand-alone sustainability report provides all the non-financial information separately from the financial report. This tool for disclosure, according to Eccles & Serafeim (2014), lacks of credibility, timeliness and relevance. Auditing would enhance the information's reliability and credibility, but it is still uncommon for sustainability reports to be audited. The main reason for this shortcoming is that a standardisation of ESG measures is missing and it is difficoult to create and establish auditing standard. Therefore, auditing firms would bear a very high legal risk auditing ESG information given the ambiguity on the topic. Moreover, the lack of timeliness comes from the fact that, while financial information is disclosed not much time after the end of the financial year, sustainability information is usually disclosed after several months and does not necessarily depict the firm's current ESG performance and commitment. Another relevant deficiency of the stand-alone sustainability reports is that the information contained is rarely put in the context of the firm's overall strategy and business model, so that it is extremely difficult for investors to understand the relationship of ESG performance with financial performance and revenues**.** These shortcomings hampers corporate reporting's information and internal transformation function[10] (Eccles & Serafeim, 2014). According to many researchers the integrated report (IR) would allow to overcome some of the limitations of the stand-alone one. In fact, in the IR non-financial information is reported together whith the financial one, therefore is disclosed without delay and is also more contextualized in the firm's overall strategy. As a matter of fact, the integrated report is "a holistic picture of the combination, interrelatedness and dependencies between the

---

[10] According to Eccles & Serafeim (2014), any kind of corporate reporting not only provides information to the different stakeholders and to the general public, but also gives to the company itself the possibility to transform in order to perform better in the future. The transformation function may be either internal or external. The first happens when the company's management decide to act on the basis of the report, the second is when shareholders force changes into the company by voting in shareholders' meetings.

factors that affect the organization's ability to create value over time" (International Integrated Reporting Council, 2013) that improves the quality and transparency of the info available with respect to the stand-alone report (Mervelskemper & Streit, 2016). Moreover, in Mervelskemper & Streit's (2016) study is shown that publishing an integrated report can help a firm enhancing its ESG performance evaluation more than a stand-alone report can do, and it allows to explain more thoroughly to what extent ESG and corporate governance performances influence market value. So, integrated reporting can be considered as the best option for ESG performance disclosure, given also that it "reduces the info-processing costs and mitigates informational asymmetries between corporate insiders external capital providers" (Mervelskemper & Streit, 2016).

1.3.1 European directives on corporate sustainability disclosure

To be able to monitor and manage undertakings' impact on society, to encourage Corporate Social Responsibility and its disclosure and to increase investors' and consumers' trust, the European Union in 2013 issued directive 2013/34/EU (amended in 2014 by directive 2014/95/EU[11]) "on the annual financial statements, consolidated financial statements and related reports of certain types of undertaking". According to Article 19a of the consolidated version of the directive, large entities of public interest (PIE)[12], within the management report, must disclose "both financial and, where appropriate, non-financial key performance indicators[13] relevant to the particular business". The article also establishes that "the management report shall, where appropriate, include references to, and additional

---

[11] Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups.

[12] We can consider as large entities of public interest: listed companies, credit institutions, insurance undertakings and any other entity that single EU Member States consider as such, with more than 500 employees (European Union, 2017).

[13] According to the European federation of financial analysts' societies (2009) KPIs are quantifiable measures, that allow for comparisons and can be "benchmarkable" from peer-to-peer. In addition, they need to "depict a correlation to risk or success factors of corporate business".

explanations of, amounts reported in the annual financial statements" therefore it refers to an integrated form of reporting non-financial information, instead of a stand-alone non-financial report. Nonetheless, Member States may allow for a separate report that needs to be disclosed no later than 2 months after the disclosure of the management report itself. Additionally, the directive assesses that non-financial information must include, at a minimum, information relating to environmental, human rights and employee issues together with anti-corruption and bribery matters. Moreover, a company's risks must be disclosed together with the actions taken to manage them and the due diligence process implemented. Therefore, a company has to reveal its approach to business sustainability and what ESG commitments are fundamental to achieve the goal set by the company itself, like the target level for some KPIs like CO_2 emissions. PIEs must also be transparent about their diversity policies: they need to disclose the level of diversity in administrative, supervisory and management bodies according to gender, age and background. Although non-financial reporting is stated as mandatory for PIE, companies may benefit of the *comply-or-explain* approach: they can decide not to disclose any of the required information conditional upon a thorough and clear explanation of why the company made this decision. However, it must be underlined that the *comply-or-explain* approach does not apply to the disclosure of the risks. Furthermore, article 19a assesses that "Member States shall ensure that the statutory auditor or audit firm checks whether the non-financial statement […] or the separate report […] has been provided". So, there is no obligation on auditing non-financial data. In 2017 on the Official Journal of the European Union the non-binding "Guidelines on non-financial reporting"[14] were published laying out a principle-based methodology for non-financial data disclosure in order to help companies in the implementation of directives 2013/34/EU and the following 2014/95/EU. The guidelines specify that the purpose of the disclosure is to provide the necessary information to understand the companies' "development, performance, position and the impact of their activity" in order to improve firms' performances both on a financial and non-financial level, together with making them more resilient. The first principle introduced is the one of materiality: it must be disclosed any

---

[14] Communication from the Commission "Guidelines on non-financial reporting (methodology for reporting non-financial information)" (2017/C 215/01).

non-financial information whose omission or misstatement would influence stakeholders' decisions. There is no specific indication of what information to disclose; materiality is a context-driven concept, hence an information deemed as material for a sector/company may be irrelevant for another. In fact, in order to take into account these characteristics of non-financial data and all the different Member States' approaches, the European Union made sure to introduce an extremely flexible framework avoiding a one-size-fits-all approach. There are, indeed, some examples of Key Performance Indicators (KPIs) that companies may disclose, that by no means represent an exhaustive list. For example, given that companies are expected to disclose their business' actual and potential impact on the environment, according to the guidelines, it could use: energy performance and improvements in energy performance, energy consumption from non-renewable sources and energy intensity, greenhouse gas emissions in metric tonnes of CO_2 equivalent and greenhouse gas intensity, emissions of other pollutants (measured in absolute value and as intensity), extraction of natural resources, impacts and dependences on natural capital and biodiversity or waste management (e.g. recycling rates). Moreover, the disclosure must include concise qualitative and quantitative data: the first providing context useful to fully understand the second, enhancing comparability and consistency of disclosed information. As far as the way in which non-financial information need to be disclosed, the European Union endorses a set of national, international or EU-based frameworks like: Global Reporting Initiative, the European Federation of Financial Analysts Societies' KPIs for ESG, the International Integrated Reporting Framework, the Sustainability Accounting Standards Board, the United Nations Global Compact (UNGC) and the UN Sustainable Development Goals. The European Directive 2013/34/EU and the guidelines on non-financial disclosure represent a "first significant step towards making businesses accountable to society" (European Coalition for Corporate Justice, 2014) and has the merit to enhance European business' transparency, nonetheless there are some shortcomings pointed out by European Coalition for Corporate Justice (2014) that are worth analysing in order to be aware of the possible improvements that may come from future provisions. First of all, the scope of the directive is limited to all listed companies having more than 500 employees to avoid putting an excessive burden on small and medium enterprises. These large companies represent just

a portion of the undertakings insisting in the European territory[15]. The hope is that, since these companies – in most cases – are leaders in their sector, smaller undertakings will try to emulate them, so that ESG disclosure can become as widespread as possible. The second major shortcoming is the fact that, in order to provide a flexible guidance for disclosing non-financial information, there are no clear instructions on how to report a company's risks and impact on society and environment. In fact, the Regulation only says that the risks reported must be "relevant and proportionate" without any further specification on what it means. Last but not least: undertakings are encouraged to use one (or more) framework(s) of their choice, with the only obligation to state in the report which one(s) they used, this provides flexibility but the comparability of reports may be hampered since the frameworks are significantly different in their purpose, content and definitions.

So, the Regulation on ESG disclosure is an important step to encourage companies, investors and consumers to take into consideration ESG issues, but there is still the need for improvements. Moreover, according to the IMF (2019) policymakers should not limit themselves to issuing directives or regulation on disclosure, but should also – in order to give a more convincing stimulus to ESG investing and disclosure – incorporate ESG principles and climate-related financial risks into financial stability monitoring and into micro-supervision of financial system.

1.3.2 European regulation on financial service sector's sustainability disclosure

As far as the financial service sector's sustainability disclosure is concerned, in 2019 the European Union introduced a regulation on "sustainability-related disclosures in the financial services sector"[16] which establishes the need for financial advisors and asset managers to disclose if they incorporate environmental, social and governance issues and the ways in which it is done, in order to reduce – as much as possible – the information asymmetry between beneficiary (end investors) and agents. So, financial service companies must disclose on their website information on how sustainability risks are taken into account,

---

[15] In 2014 PIE with more than 500 employees represented the 14% of the entrepreneurial system, according to European Coalition for Corporate Justice (2014).

[16] Regulation (EU) 2019/2088 of the european parliament and of the council of 27 November 2019.

and how this practice affects the internal processes like risk management and portfolio selection.

1.3.3   The decision to disclose Corporate Social Responsibility information

The consequence of a far from strict regulatory framework on Corporate Social Responsibility disclosure, as the one described above, is that the decision on what to disclose and how to do it is left to the single company. To explain the incentives for ESG disclosure and – more in general – non-financial disclosure, academics and scholars have introduced two theories: institutional theory and legitimacy theory that explain how disclosure is influenced respectively by country-specific and firm-specific characteristics. The institutional theory explores the influence of the structural characteristics of a country (e.g. political, labour and cultural systems) on disclosure. For example, Baldini, Dal Maso, Liberatore, Mazzi, & Terzani (2016) find out that in a country with high unemployment rate the disclosure of ESG information will be more frequent than in one with a low rate since "a high unemployment rate creates competition among firms to attract a higher number of skilled employees". In turn, a high corruption rate will hamper disclosure. The legitimacy theory, instead, explains how a firm's characteristics influence the disclosure of non-financial information. According to this theory, a company's survival in the market is subject to its social acceptance that depends, in turn, on how a firm's behaviour satisfies societal expectations. As said, sustainable and responsible investing has been gaining more and more popularity, hence with time investors' demand for ESG information has increased. It needs to be noted that the demand for ESG information does not come just from investors, but also from other stakeholders like employees or clients given the "growing public awareness of companies' roles in society and interest in social, environmental, and ethical issues " (Baldini, Dal Maso, Liberatore, Mazzi, & Terzani, 2016). Therefore firms "will prioritize their sustainability management activities that increase and secure legitimacy, whereas profit-orientation will be emphasized much less" (Baldini, Dal Maso, Liberatore, Mazzi, & Terzani, 2016). So, one of the main reasons for companies to disclose their ESG performance is to satisfy the increasing demand for this kind of information, in order to put major emphasis on strengths diminishing weaknesses and gaining good publicity or in order to

announce a variation in ESG policy putting remedy to (or preventing) a reputational damage caused by an ESG incident. It could also be just for "façade". Another possible reason for a company to decide to disclose ESG data could be that it may bring new opportunities (and customers) and the possibility to increase cost efficiency and to be able to "identify any future legislative initiative or opinion shaping process, which could lead to regulatory intervention" and additional costs (European federation of financial analysts societies, 2009). Moreover, various studies actually notice a positive relationship between non-financial performance indicators and ESG disclosure: disclosing ESG information would help a firm strengthening the relationship with its stakeholders and attracting more motivated and talented workforce; nowadays people don't want just to buy from sustainable companies, but they also want to work in them. Moreover, it could foster consumer loyalty and reduce regulatory burdens (Fatemi, Glaum, & Kaiser, 2018). However, it could also happen that companies may decide not to completely disclose their ESG activities, in order to avoid stakeholders thinking that the company is investing too much in ESG actions (that they deem as too costly), that the firm is "covering up for a lack of depth in ESG actions with 'too much talk' " (Fatemi, Glaum, & Kaiser, 2018) or even practicing *greenwashing or bluewashing*[17]. So, firms may find more fruitful explaining and justifying ESG concerns instead of praising themselves disclosing ESG strengths[18]. As a matter of fact, Fatemi, Glaum, & Kaiser (2018) find out that ESG disclosure tends to counteract the decrease in valuation due to an ESG concern, more than how much it increases the valuation because of ESG strengths. It actually seems that the positive relationship between firm value and ESG strengths is more evident when the firm decides to disclose less. Therefore, companies tend to put more efforts in legitimizing its operations and policies than to disclose positive information ESG-wise.

---

[17] See paragraph 1.3.3.1

[18] In the ESG lexicon the opposition of strengths and concerns has been introduced by KLD Research and Analytics, an ESG rating provider (proprietary of MSCI) that measures the ESG performance of a company by counting every ESG incident or possible criticality (named concern) and ESG-friendly decisions and investments (named as strengths).

1.3.3.1 Greenwashing and bluewashing

Given the increasing pressure from customers and investors to disclose ESG information in accordance with the legitimacy theory, companies often incur in the practice of *greenwashing*. *Greenwashing* happens when companies mislead investors and consumers about their environmental performance (firm-level *greenwashing*) or about the environmental benefits of one of their products or services (product-level *greenwashing*). An example of firm-level *greenwashing* is when General Electrics was in 2005 simultaneously lobbying to fight new clean air requirements and extensively advertising the company's work to improve its environmental performance (the 'Ecomagination' campaign started). This practice has the effect of shrinking the socially responsible investment capital market because it undermines investors' confidence in environmentally friendly companies (Cuerel Burbano & Delmas, 2011). We can consider *greenwashing* as a spectrum: a firm can go from just using an extremely ambiguous and misleading language to completely lying about its environmental performance. In the middle we find the practice of 'selective disclosure'. It happens when a company operates a strategical and careful selection of the information to disclose, "disproportionately revealing beneficial or relatively benign performance indicators to obscure less impressive overall performance" (Marquis, Toffel, & Zhou, 2016). According to Marquis, Toffel, & Zhou (2016) the drivers for *greenwashing* are mostly 'external' and can be divided in market related and non-market related. In the first category we find the pressure coming from investors and consumers to be more environmentally friendly, so poor environmental performers may decide to disclose partial or even false information to comply with the raising investors' and customers' requests to know what is a company's environmental impact. Additional pressure can come from competitors: companies may feel compelled to disclose their 'green' performance for the fear of falling behind competitors. The non-market related drivers usually have the function to discourage greenwashing practices; examples of these drivers are regulation, NGOs and media pressure. However, as said, the regulatory framework does not specifically address *greenwashing* and the consequences for practicing it are uncertain and vague, in this way companies may be incentivized to greenwash by the absence of definite repercussions. Marquis, Toffel, & Zhou (2016) identify in activists', NGO's and media's pressure a possible deterrent. In fact, they can act as "informational monitors of *greenwashing*" and

hold companies accountable of their actions through campaigns against those who decide to greenwash, causing them serious reputational damage. The actions of activists, NGOs and media together with a more formalized and defined global norms would be the best combination to avoid companies disclosing misleading information about their environmental performance: the higher the scrutiny and pressure coming from these actors, the higher the compliance with institutional demand will be (Marquis, Toffel, & Zhou, 2016).

A practice similar to *greenwashing* is *bluewashing* which consists in "signing up to UN initiatives merely to appear aligned with the organisation which has a blue logo" (Johnson, 2010), without effectively fulfilling to the obligations that the initiative entails. Subscribing to one of these programs allows, in fact, to gain a boost in reputation and goodwill, a regulatory relief and fosters consumer loyalty (Berline & Prakash, 2005) in change of a serious commitment to change a firm's behaviours in order to help the community reaching the program's objective. However, complying with all the tasks of the programme – most of the times – is extremely costly, so a company could decide to simply act "symbolic low cost steps to convey the impression that they are fulfilling their obligations" (Berline & Prakash, 2005). This is possible because of the loose entry barriers and insufficient monitoring mechanisms.

2   ESG ratings

The first ESG rating agencies were being launched between twenty and twenty-five years ago, when sustainable and responsible investing started being popular among institutional investors (Windolph, 2011). As credit ratings help investors in the evaluation of the creditworthiness of an issuer, ESG ratings help them evaluating a firm's ESG performance. ESG ratings are also used by scholars in their research papers (e.g. Kempf and Osthoff, 2007). They systematically analyse the ESG performance of companies aggregating different indicators into a single score (either a number or a letter), allowing investors to compare the ESG performance of two or more companies. Because of information asymmetries and because of time constraints investors usually prefer to use ESG ratings instead of collecting and verifying by themselves information disclosed by companies and other available information. ESG rating companies, therefore, function as information

intermediaries between companies and investors (or other stakeholders) interested in ESG performances, collecting and validating all the information publicly available and disclosed by companies. Moreover, rating companies often integrate their dataset with companies' answers to surveys, that a single investor wouldn't be able to collect. So, ratings are the result of the combination of data that may come from different sources like government data (e.g. regulatory actions or toxic emissions) or information disclosed by the rated company itself through company sustainability reports, press releases, questionnaires or interviews.

2.1 Shortcomings of ESG ratings

Although, ratings may make investors' life easier, they have some shortcomings that can impede a further diffusion of their use to make investment decisions; the main ones are the lack of uniformity and the lack of transparency.

2.1.1 Lack of uniformity

The most popular rating providers are: KLD (MSCI Stats), Sustainalytics, Vigeo Eiris (Moody's), RobecoSAM (S&P Global), Refinitiv-Asset4. Their ratings often disagree: it could happen that a company is rated as AAA (the highest possible) by MSCI and at the same time assigned a middle grade from another provider, as happened to Tesla in 2017 (Kerber & Flaherty, 2017). There is no standardization for ESG ratings, therefore every company has its proprietary algorithm to compute ESG scores. As a consequence, as illustrated in Table 2, the correlation between the scores computed by the different providers is 0.54 on average, ranging from 0.38 to 0.76; not very high value considering that the correlation between the credit ratings from different providers is on average 0.99 (Berg,

	KL SA	KL VI	KL RS	KL A4	KL MS	SA VI	SA RS	SA A4	SA MS	VI RS	VI A4	VI MS	RS A4	RS MS	A4 MS	Average
ESG	0.53	0.49	0.44	0.42	0.53	0.71	0.67	0.67	0.46	0.7	0.69	0.42	0.62	0.38	0.38	0.54
E	0.59	0.55	0.54	0.54	0.37	0.68	0.66	0.64	0.37	0.73	0.66	0.35	0.7	0.29	0.23	0.53
S	0.31	0.33	0.21	0.22	0.41	0.58	0.55	0.55	0.27	0.68	0.66	0.28	0.65	0.26	0.27	0.42
G	0.02	0.01	-0.01	-0.05	0.16	0.54	0.51	0.49	0.16	0.76	0.76	0.14	0.79	0.11	0.07	0.30

Table 2- Correlations between ESG ratings from different providers at the aggregate rating level (ESG) and at the level of the environmental dimension (E), the social dimension (S), and the governance dimension (G).

(Source: Berg, Koelbel & Rigobon (2020))

Koelbel, & Rigobon, 2020). Note that in Table 2 KL stands for KLD, SA for Sustainalytics, VI for Vigeo Eiris, RS for RobecoSAM and A4 for Refinitiv-Asset4.

The discrepancies between the scores are mostly due to the high level of subjectivity that ESG evaluation entails, given also the lack of a strict disclosure regulation for companies. Three are the discretionary elements that can be identified as the main causes for the differences among raters: divergence on scope, on measurement and on weights. The divergence that has the biggest impact on the final score is the one on measurement, then there is the one on scope, and the least relevant difference is the one on weightings (Berg, Koelbel, & Rigobon, 2020). The divergence on scope is the divergence on the issues or attributes that the different rating providers take into account. This divergence happens if a rater includes an attribute (e.g. lobbying activities) while the others don't. It stems from the fact that, because of the lack of a precise definition of what socially responsible investing means, every rater has its own definition. Some, like Windolph (2011), state that this divergence is desirable, since in this way investors may 'shop' for the rating based on a definition of SRI as close as possible to theirs. The divergence on measurement attains to how the different attributes are measured. For example, gender equality can be measured using as proxies both the number of women on the board or the gender pay-gap between employees. The cause of this kind of divergence can be spotted in the difficulty of retrieving high quality information on certain topics and the loose regulation on ESG company disclosure[19]. Moreover, the divergence on weights happens because rating agencies assign a weight to each according to what is, in the raters' opinion, their relative importance. Therefore, each rating agency breaks down the ESG performance and decides the hierarchy of the components in its own way and most often there is a divergence on the weights the different rating providers assign to the same issue. For example, raters based in the USA have 71% of the issues analysed within the social pillar while the European ones only 47% (Chatterji, Durand, Levine, & Touboul, 2016). This depends, again, on what raters mean by being sustainable and responsible and on what are their major concerns.

An additional possible cause for the divergence between ratings may be spotted in the way raters select the benchmarks used to define the scores. In fact, the definition of what is a good or bad performance may be different across rating providers. Benchmarking can be

---

[19] See paragraph 1.3

done either through the comparison with a peer group or using absolute levels of performance based on a "predefined optimal level of performance" (Kotsantonis & Serafeim, 2019). A peer group can be universal and composed by companies of different industries and sectors or there may be one for each sector. The absolute levels of performance is a performance range based on a predefined rule (that most of the times – at least in the case of environmental measures – is based on scientific targets and data), assessing the real impact of a company's activity on society and the environment in which it insists. Therefore, the definition of what best performance means is crucial for the final assessment of a company and using a different method to do so will inevitably affect the final score.

How ESG rating agencies handle data gaps may be identified as another possible cause of the discrepancy between ratings. Whenever a data is missing, raters can use different techniques to fill the gap ('data imputation'), and this may, in fact, bring to significantly different ratings. The easiest way to perform data imputation is when the gap is filled by an arbitrary value that may be – for example – zero or the sector's industry's average. Another way to perform data imputation is using the input/output model: assigning the score for a single ESG metric computing a company's direct and indirect impact scaling industry-level data on the firm's size of operations (Kotsantonis & Serafeim, 2019).

According to Chatterji, Durand, Levine, & Touboul (2016) the divergence of the scores may represent the invalidity of the scores. Therefore, the authors suggest that ratings should be subject to a periodical assessment of their validity from practitioners and scholars, such as tests to verify if the companies with highest score actually have fewer major corporate scandals. Moreover, Berg, Koelbel, & Rigobon (2020) point out that, given that ESG ratings may convey mixed signals, there will be less ambition from companies to improve their ESG performance given that companies cannot be sure if – with a better behaviour – they will actually obtain a better ranking by every provider. Therefore, SRI investing wouldn't really serve, for companies, as a stimulus to improve their performance on sustainability issues and would result – in the worst-case scenario – in a misallocation of capitals.

2.1.2 Lack of transparency

Another criticism of ESG ratings is the lack of transparency: academics together with investors and companies see ratings as a black box that can't be opened. There is in fact, in most cases, a lack of disclosure on what assumptions are being made, what kind of data have been used and where they have been retrieved from and on which weightings and criteria have been used to compute the score. The demand for more transparency comes from investors (or any other actor interested in ESG rating) based on the need to assess the credibility and reliability of a specific score. The sustainable investing industry is still in its infancy, but if the situation does not change, investors would rapidly lose confidence in ratings.

A possible source of this lack of transparency is the fact that the methods used to synthetize a score are proprietary, hence considered as part of the intellectual property that raters need to protect in order to prevent competitors from copying it. Nonetheless, there still are some raters that are able to be more transparent than others because their rating process is continually updated – hence harder to replicate – on the basis of client's feedback, market analysis and industry analysis. Although, according to Regnan – an Australian rating provider – there will still be, for the time being, a level of opaqueness around ratings since the assessment and qualification of ESG factors is still at an experimental stage where change and innovation are the key. So, funders and clients will still require a certain level of confidentiality to continue to invest in them and buy their products (Rogers & Stubbs, 2013).

The easiest way to increase both uniformity and transparency would be to introduce a regulation of ESG rating agencies setting open and transparent disclosure and measurement standards. Stubbs & Rogers (2013) suggest the creation of an "independent standard setter body" similar to the already-existing one for the accounting profession. Nonetheless, it must be kept in mind that rigid standards hamper innovation, discouraging research and experimentation. In fact, as said, ESG evaluation is still in its infancy and innovation and flexibility are still essential to foster competitivity in the sector. So, the best solution would be finding the right balance between transparency and protection of intellectual property.

2.2 ESG scores by Refinitiv

The ratings used in chapter 3 are the ones computed by Refinitiv and retrieved through the Datastream platform.

Refinitiv is one of the biggest ESG providers: it assigns ratings to over 9000 companies (nearly 70% of the global market capitalisation) from all over the world, more than 2000 are European (Refinitiv, 2020). The provider's data and scores are "designed to transparently and objectively measure a company's relative performance" (Refinitiv, 2020). Refinitiv understands the importance of being transparent, accurate and of producing comparable ESG measures, hence the explanation of the methods they use to compute the scores is fairly detailed. Refinitiv uses a percentile rank scoring methodology: each score is based on the company's ESG performance compared to a peer group's one, allowing investors to easily compare companies' performances. To make it even easier to understand, the scores are also available in the form of a letter grade ranging from D- to A+, as we can see in the Table below.

Score range	Grade	Description	
0.0 <= score <= 0.083333	D -	"D" score indicates poor relative ESG performance and insufficient degree of transparency in reporting material ESG data publicly.	ESG Laggards
0.083333 < score <= 0.166666	D		
0.166666 < score <= 0.250000	D +		
0.250000 < score <= 0.333333	C -	"C" score indicates satisfactory relative ESG performance and moderate degree of transparency in reporting material ESG data publicly.	
0.333333 < score <= 0.416666	C		
0.416666 < score <= 0.500000	C +		
0.500000 < score <= 0.583333	B -	"B" score indicates good relative ESG performance and above average degree of transparency in reporting material ESG data publicly.	
0.583333 < score <= 0.666666	B		
0.666666 < score <= 0.750000	B +		
0.750000 < score <= 0.833333	A -	"A" score indicates excellent relative ESG performance and high degree of transparency in reporting material ESG data publicly.	
0.833333 < score <= 0.916666	A		ESG Leaders
0.916666 < score <= 1	A +		

Table 3- Score range and grades (Source: Refinitiv (2020))

Refinitiv's offers two types of scores: an overall ESG score and ESG Combined score.

2.2.1 Overall ESG score

The overall ESG score represents the evaluation of the ESG performance of a company, its commitment to an ESG-friendly business policy and the effectiveness of the said policy. This score is computed with as granular as possible data in order to be able to effectively

differentiate between non-transparent companies and those that set the example and can be considered as industry leaders, as far as ESG performance is concerned.

Refinitiv, in order to reach the highest possible data quality, combines algorithmic and human processes in the calculation of the scores. More than 450 company-level ESG measures are computed and made available to investors, but the rating company uses only the most comparable and material per each industry (186 in total) to actually compute the scores. Every measure used to compute a company's ESG score is grouped into themes that – in turn – are grouped into ten different categories. Categories are the three pillars' (Environmental, Social and Governance) building blocks: for the Environmental pillar Refinitiv analyses the resource use, emissions and innovation; for the Social pillar it analyses workforce, human rights, community and product responsibility; the Governance pillar's categories are management, shareholders and CSR strategy. Each category has its own weight ranging from 1 to 10, then normalized to a percentage to make calculation easier. Each industry has its own category weights for Social and Environmental pillars since on these topics the relevance of categories is comparable within the same industry. Every country has different weights in the Governance pillar's categories given that the governance practices are more consistent within countries. The rating company computes a score for each category, using the industry group as a benchmark for the categories belonging to the Social and Environmental pillars. For the Governance pillar the benchmark used is the country of incorporation. The formula used to compute the category scores is:

$$Cat.\,scores = \frac{no.\,of\ companies\ with\ a\ worse\ value + \frac{no.\,of\ companies\ with\ the\ same\ value\ included\ in\ the\ current\ on}{2}}{no.\,of\ companies\ with\ a\ value}$$

Then, the rating company computes a score for each pillar and, finally, for the overall ESG performance.

The overall ESG score, along with the ESG combined score (see following paragraph), is updated once a year, in line with companies' disclosure frequency. We can observe how Refinitiv visually describes its calculation process looking at Figure 3 from the bottom to the top.
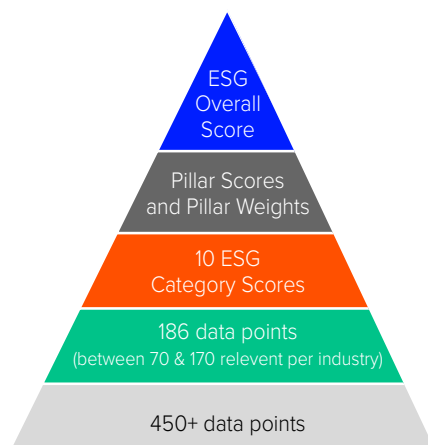

Figure 3 - The scoring process - bottom to top (Source Refinitiv (2020))

2.2.2 ESG Combined score

Throughout the year analysts examine media to check if a company has been involved in an ESG scandal. In doing so, they analyse in dept twenty-three ESG controversy topics (across all ten categories) that, according to the rating company, need to be taken into account when creating an ESG controversy score. The score range is from 0 to 100, where 100 means that the company had no controversy. The twenty-three topics include, for example, anti-competition controversies, public health controversies (relative to community), child labour controversies (relative to human rights), controversies on customer health, safety and privacy controversies (relative to product responsibility), insider dealings controversies (relative to shareholders), diversity, opportunity controversies (relative to workforce), environmental controversies (relative to resources) and management compensation controversies count (relative to management). To compute the ESG controversies score, analysts use industry peer groups as benchmark. To take into account the market cap bias that brings bigger companies to receive more media attention than small ones, severity weights are applied as in Table 4.

The assumption behind these weights is that a small firm's scandal – in order to be covered by media – must be extremely severe, while media will easily cover bigger companies', together with serious issues, also controversies of little relevance.

Global benchmark	Cap class	Severity rate*
>=10 billion	Large	0.33
>=2 billion	Mid	0.67
<2 billion	Small	1

* Logic to derive weights: Large=1/3= 0.33, Mid=.67, Small = 0.33+0.67=1.

Table 4- Severity weights (Source: Refinitiv (2020))

The ESG controversies score is used to discount the ESG score for news on controversies which may have a material impact both on the business overall performance and, more importantly, on the overall ESG performance. Hence, the ESG controversies score is incorporated into the ESG score creating the ESG combined score. This is done by computing the average of the ESG controversies score with the ESG score in case the controversies score is lower than the ESG score, otherwise the ESG combined score will be equal to the ESG score. This is to avoid inflating the ESG score whenever the company has behaved and wasn't involved in any controversy (or at least not many of them). Being the ESG combined score a "comprehensive evaluation of a company's sustainability impact and conduct in time" (Refinitiv, 2020), it can be considered more informative than the simple ESG score representing, therefore using it in this study seems the best choice possible. Here is an example of how Refinitiv incorporates the ESG controversies score in the ESG score:

Scenario	ESG controversies score	ESG score	ESG combined score
If controversies score is >=ESG score, then ESG score = ESG combined score	100	89	89
If controversies scores is <ESG score, then ESG combined score = Average of ESG & Controversies Score	48	49	48.5

Table 5- Combined score logics (Source: Refinitiv 2020)

3 The portfolio performance analysis

3.1 Literature review

The literature review on Socially Responsible Investing's profitability highlights that many are the studies on this topic as well as the methodologies used to conduct them. The latter can be divided into three categories: portfolio studies, multivariate regression studies and event studies (Derwall, Guenster, Bauer, & Koedijk, 2005). Portfolio studies analyse mutually exclusive portfolios based on ESG performance indicators (typically ESG scores) using data relative to a specific time horizon and study if there is any difference in portfolios' returns. Examples of portfolio performance measures used in these studies are Sharpe ratio and Jensen's alpha. Multivariate regression studies represent a variation of portfolio studies; they use multifactor models as Fama and French's (1993) or Carhart's (1997) ones. There is still no agreement on which model is the best to capture whether ESG brings an over-performance or not. Event studies, instead, study the difference in returns among SRI and non-SRI mutual funds. This kind of studies often find out that including ESG considerations in funds' investment process does not guarantee an extra return with respect to the funds who don't. This may be because mutual funds are actively managed funds and their performance is influenced, among other factors, by the portfolio manager's skills. Auer and Schuhmacher (2016) performing a portfolio study and using the Sharpe ratio as a portfolio performance indicator, find out that in the European market there is no evidence of an outperformance of high ESG rated portfolios with respect to the low rated ones and investors often have to sacrifice returns for the sake of social responsibility. At the same time Kempf & Osthoff (2007), using a USA-based investible universe and performing a multivariate regression study with Carhart's model, find out that Socially Responsible Investing has a positive influence on US-based stock portfolio returns, therefore generating extra-retrurns with respect to non-SRI investing. In particular, the highest possible extra return is achieved by those who implement the best-in-class integration approach[20]. Herremans, Akathaporn, and McInnes(1993), performing a portfolio study on manufacturing companies with a high

---

[20] see paragraph 1.1.1

corporate social responsibility reputation, find out that the latter outperform companies with a poorer reputation. Capelle-Blancard and Monjon (2014), performing an event study on French SRI mutual funds, conclude that they do not outperform the market or their non SRI counterparts. The authors, in fact, determine that the sectoral negative screens have a negative influence on the funds' performance, while transversal screens – like commitment to UN Global Compact Principles – do not have any significant impact on returns. In Bauer, Koedijk, and Otten (2005) risk-adjusted returns of ethical and conventional funds have been compared without finding any significant evidence of an enhanced performance of ethical funds over the conventional ones. As can be noticed, there is no unanimity on whether financial and ESG performance have a positive relationship and therefore Socially Responsible Investing allows to 'do good while doing well' enhancing both financial and non-financial utility[21] a negative relationship such that – including ESG considerations in the investment decision process – means giving up returns in favour of non-financial utility or there is no relationship between ESG performance and financial returns.

3.2   The empirical analysis

The objective of this study is to find out if socially responsible stock investments guarantee an overperformance with respect to non responsible ones. To do so, we replicate Kempf and Osthoff (2007) study for the North American market using a European investible universe in order to check if their findings are valid for the European market, trying to contribute to the clarification on what is the relationship between ESG issues and portfolio performance.

Fifteen portfolios are built: five socially responsible portfolios (the Best portfolios), five non socially responsible ones (the Worst portfolios) and, finally, five portfolios being the result of a long-short strategy (the Long-Short portfolios). These portfolios are built using five incorporation criteria in order to examine their impact on portfolio performance. To measure the portfolios' performances, a multivariate regression study is implemented using Carhart's (1997) model.

---

[21] With non-financial utility we mean the increase in society's wellbeing given that investors are cutting funds to polluting and socially unfair firms.

3.2.1 Data, model and methodology

3.2.1.1 Data

All data are retrieved through Reuters' Datastream service, apart from Carhart (1997) model's factors that are retrieved from Kenneth French's data library[22] which offers already computed factors for the European economy. However, Kenneth French computes the factors in USD and not in EUR, therefore we convert them using the USD-EUR spot exchange rate.

In order to include in the study only European stocks, we use as investible universe all EURO STOXX index's components. The EURO STOXX index is a value weighted index composed by a variable number of components (300 on average) and a subset of the STOXX Europe 600 Index. It represents large, mid and small capitalisation companies of eleven Eurozone countries[23].

To compute companies' stock returns we use monthly closing prices of each stock considering the period going from 01/01/2009 to 01/12/2019 (eleven years, 132 dates). This period has been chosen both to guarantee a time horizon as long as possible since sustainable finance is "axiomatically linked to the long term" (EU High-Level Expert Group on Sustainable Finance, 2018), and to avoid including periods of well-known economic financial distress (e.g. the financial crisis of 2008 and the more recent one due to the COVID-19 pandemics) that could strongly influence the final results. Following Lanza, Bernardini, & Faiella (2020), to avoid the survivorship bias, we use as investible universe the EURO STOXX index's components as of 01/01/2009. The survivorship bias would be, in fact, generated by having in our investible universe companies that we know will be part of the index in the future and therefore know that they will be successful. In practice, any investor knowing in 01/01/2009 which companies were going to be part of the index in 01/01/2019 would have a huge advantage. Hence the choice of using 01/01/2009's components instead of those as of 01/01/2019.

---

[22] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[23] Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain.

In order to clean the dataset from missing values the investible universe is reduced to all those companies having complete data, therefore all the companies having – for any reason – missing data are excluded. As a consequence, the investible universe counts 265 of the 315 EURO STOXX's components as of 01/01/2009. As mentioned in the previous chapter, the parameter measuring companies' ESG performance is Refinitiv's ESG Combined score.

3.2.1.2 Performance measurement: The Arbitrage Pricing Theory

The Arbitrage Pricing Theory (APT) was introduced to remedy to the Capital Asset Pricing Model's shortcomings, among which we find the fact that it has many stringent assumptions and that it doesn't capture the fact that value stocks outperform growth ones. So, the APT is based on less stringent and more realistic assumptions and allows to account for multiple factors. The main assumptions of the APT are:

- Markets are perfectly competitive

- Investors will always prefer more wealth to less wealth

- Security's i returns generating process is a linear function of k factors like the following:

$$R_i = \alpha_i + \beta_{i1}\delta_{1t} + \cdots + \beta_{ik}\delta_{kt} + \varepsilon_{it}$$

where R_i is the return of security i, α_i is the expected return without any risk, δ_{jt} represents a common factor (or index) with zero mean and that has an influence on assets' returns.

- No arbitrage is possible: according to the Law of One Price (LOOP) two assets with identical risk must have identical expected returns.

The equilibrium relation implied in these assumptions is the following:

$$E(R_i) = \lambda_0 + \lambda_1\beta_{i1} + \cdots + \lambda_k\beta_{ik}$$

where:

- $E(R_i)$ is the expected return for security i

- λ_0 is the expected return of a non-risky asset (r_f)

- λ_j represents the price of risk coming from factor j

- β_{ij} represents security i's sensibility to the common j-th factor, also called factor loadings

It must be noted that the theory does not identify the factors: both their number and their identity is unspecified. Therefore, is up to those who perform the analysis to specify the factors and decide how many of them are sufficient, usually empirical studies suggest that three or four is optimal choice. Throughout the years two are the main types of APT models that have been introduced: macroeconomics-based and microeconomics-based. The first assumes that risk premiums, therefore returns, are governed by 'broad economic influences' and its factors are macroeconomic variables like the monthly growth rate of a country's industrial production or the differential between expected and realized inflation levels (Reilly & Brown, 2012). Instead, the microeconomic-based models assume that it is possible to specify risk using factors based on specific stocks characteristics represented by proxy variables (Reilly & Brown, 2012). It can be also called a characteristic-based approach.

Many are the different microeconomic-based models that have been introduced and studied, among these we find Carhart's four factor model that will be used to investigate whether the companies' ESG performance has a positive influence on returns or not. Carhart (1997) model, also called Fama French four-factor model, is the following:

$$R_{it} - R_{ft} = \alpha_i + \beta_{1i}\left(R_{mt} - R_{ft}\right) + \beta_{2i}SMB_t + \beta_{3i}HML_t + \beta_{4i}MOM_t + \varepsilon_{it}$$

Where:

- R_{it} is month *t* return of the *i-th* portfolio.
- R_{ft} is month *t* risk free rate.
- $R_{mt} - R_{ft}$ is month *t* excess return of the market portfolio over the risk-free rate. Also called market factor.
- SMB_t is month *t* Small Minus Big factor. It represents the return of a portfolio resulting from going long on a portfolio composed by small capitalization stocks and short on a portfolio composed by large capitalization stocks. It is a factor designed to capture the elements of risk associated with firm size.
- HML_t is month *t* High Minus Low factor. It represents the return of a portfolio resulting from going long on a portfolio of stocks with high book-to-market ratio (value stocks) and short on a portfolio of stocks with low book-to-market ratio (growth stocks). It is a factor designed to capture the risk differential between growth and value stocks.

- MOM_t is month t momentum factor. It represents the result of a long short strategy such as going long on past year's winner and going short on past year's losers. It is a factor designed to capture firms' tendency of having high (low) returns when in the recent past they had high (low) returns as well.

The model showed above was built starting from Fama and French's (1993) three factor model which used just the market $(R_{mt} - R_{ft})$, SMB and HML factors, adding the Momentum factor.

In the following paragraphs we apply Carhart (1997) model using as independent variable the returns of the portfolios built (paragraph 3.2.1.3). To implement the model we perform a linear regression and looking at the factors' loadings (betas), the most significant determinants of returns can be established. For a multifactor model like this to fully explain a stock's (or portfolio's) returns' determinants, the alpha of the model should be non-statistically significant, otherwise it would mean that there are some other factors determining the portfolio's performance that the model does not capture or the model spotted a market inefficiency. So, in order to determine if the portfolios with a high ESG performance have an extra-return with respect to those having a low ESG performance we test the significance of the long-short portfolios' Jensen alphas: if they are positive and statistically significant – at least at 10% significance level – it means that the high-rated portfolios have returns higher than those of the low-rated portfolios.

3.2.1.3 The portfolio creation process

In this study five different SRI portfolio creation techniques are used: negative screening, positive screening, negative screening combined with positive screening, best-in-class approach and negative screening combined with best-in-class approach. Following Kempf & Osthoff (2007), for each technique two value weighted mutually exclusive portfolios are created: a Best portfolio and a Worst portfolio. Although the choice of building value weighted portfolios is sub-optimal, it seems appropriate in order to maintain as fair as possible the comparison with the results Kempf & Osthoff (2007) got for the United States. In addition, in order to better understand if companies' ESG performance really influences the portfolios' returns and lets investors gain an extra-return with respect to the portfolios built with the low rated stocks, the focus will be on the Long-Short portfolios. Each of the

latter is the result of going long on the Best portfolio built with an SRI portfolio creation technique and short on the Worst one built with the same technique. Ratings are usually published at the end of each year, therefore all the portfolios built – apart from the negatively screened ones – are rebalanced at the beginning of every year according to previous year's ratings. So, for example, to build 2009's portfolio the companies in the investible universe are ranked according to their 2008's rating.

The negative screens are based on the following controversial areas: Alcohol, Gambling, Nuclear and Tobacco. The data on whether the companies in the investible universe have revenues coming from the controversial areas, have been retrieved using the Datastream for Office template: "*ASSET4 Negative Screening Template*" that allows users to screen indices or custom portfolios against up to 35 exclusionary criteria. In this case the Best portfolio is composed by companies that among their sources of revenues do not have any of the above controversial areas and the Worst one is composed by companies that do have revenues coming from the same controversial areas. Both portfolios are created in 01/01/2009 and kept unchanged until 01/01/2019. The exclusionary criteria lead to a reduction of the investible universe by about 15% (39 companies are excluded).

To build the Best and Worst portfolio using the positive screening technique companies are ranked according to their overall ESG Combined score using a 10% cut-off rate. The first 10% is part of the Best portfolio, the bottom 10% is part of the Worst portfolio. In combining the negative and positive screens, a negative screen on the investible universe is performed so that it is possible to invest only in the companies that do not have revenues coming from the above-mentioned controversial areas. After this screen a positive screening is implemented on the reduced investible universe.

To implement the best-in-class approach, companies in the whole investible universe are divided according to their industry classes, using the Industry Classification Benchmark (ICB)[24]. Within each industry, companies are ranked according to their overall ESG Combined score. A Best and Worst portfolio for every industry has been created choosing,

---

[24] ICB: identifies eleven different industry classes: basic materials, consumer discretionary, consumer staples, energy, financials, health care, industrials, real estate, technology, telecommunications, utilities. https://www.ftserussell.com/data/industry-classification-benchmark-icb

respectively, the best and worst rated 10% for each industry. These portfolios are used to create two final Best and Worst portfolios with a homogeneous representation of every industry. There are very few companies belonging to some industries (often less than 10) and – in order to let every industry to be represented in the final Best and Worst portfolios – we establish that there must be at least one company for each industry. To create the final Best and Worst portfolios the EURO STOXX's industry composition is used.

Finally, to combine the negative and best-in-class approach, a negative screen on the investible universe is performed so that it is possible to invest only in the companies that do not have revenues coming from the already mentioned controversial areas. After this screen a best-in-class approach, as described above, is performed.

3.2.2 Limitations of the study

Before proceeding to the analysis of the data, there is the need to acknowledge the limitations of this study. The first limitation is the fact that, in cleaning, the dataset has been necessary to exclude all the companies with missing values. In doing so, the investible universe went from 315 to 265 companies: a reduction of 16%. As a matter of fact, the sample has been reduced and it should be further investigated if this had major influences on the results. Moreover this study uses Jensen's alphas generated by the Carhart (1997) model, to investigate on whether there is an overperformance of high ESG-rated portfolios over the low rated ones. The limitation of this model is that it does not allow to determine a priori whether a significative alpha is due to a risk factor that is not captured by the model itself or to a market inefficiency. Another limitation is relative to the level of sophistication of the portfolio allocation process. In order to select the Best and Worst portfolios the only criteria used in this study is the companies' ESG rating. This means that the portfolio creation process is tied to a measure that may be biased towards one of the ESG pillars and, as illustrated in chapter 2, has a certain degree of subjectivity. A more sophisticated portfolio allocation process, perhaps based on multiple measures, would help overcoming this bias eliminating – at least in part – the subjectivity of the ESG scores. For example. Lanza, Bernardini, & Faiella (2020) use a Machine Learning approach "to better spot the most material E, S or G metrics for sustainable investing" and "overcome the current inconsistencies in the ESG scores". The last, but not least important, limitation is that only

one ESG rating provider has been used. Chapter 2 mentioned how different rating agecies can provide different scores for the same company. In order to be able to assess with certainty that the findings of this study mirror what happens in reality, the same analysis should be performed using at least another rating provider.

3.2.3   ESG Combined scores description

Figure 4 shows that the sample's ESG scores – in the period analysed – have been experiencing a growing trend. It can be, in fact, noticed that from 2009 to 2019 the average overall ESG score went from 50 to 65, following a linear trend.
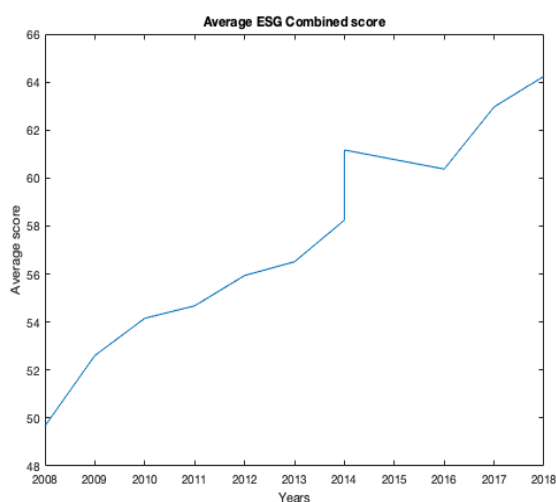

Figure 4 - ESG Combined score average 2008-2018

Figure 5 depicts the distribution of the whole sample's ESG Combined score (top left) together with those of the ESGC at the beginning of the sample (top right), in the middle and at the end of the sample (bottom left and right, respectively). The first plot shows that the whole sample's ESG Combined score distribution is skewed to the right: they are concentrated on the higher values and less companies are rated with the lowest possible values. The other three plots confirm that the ESG scores, during the period considered, have been experiencing a growing trend. Supporting this thesis there is also the fact that the medians in 2009, 2014 and 2019 are 51.2, 58.14 and 66 respectively. This phaenomenon is certainly due to the fact that, with time, having a high level of corporate social responsibility has become increasingly popular, but it may also be due to the fact that the scoring

methodology is everchanging and with time some issues may have become more important than others with time, affecting the overall score.
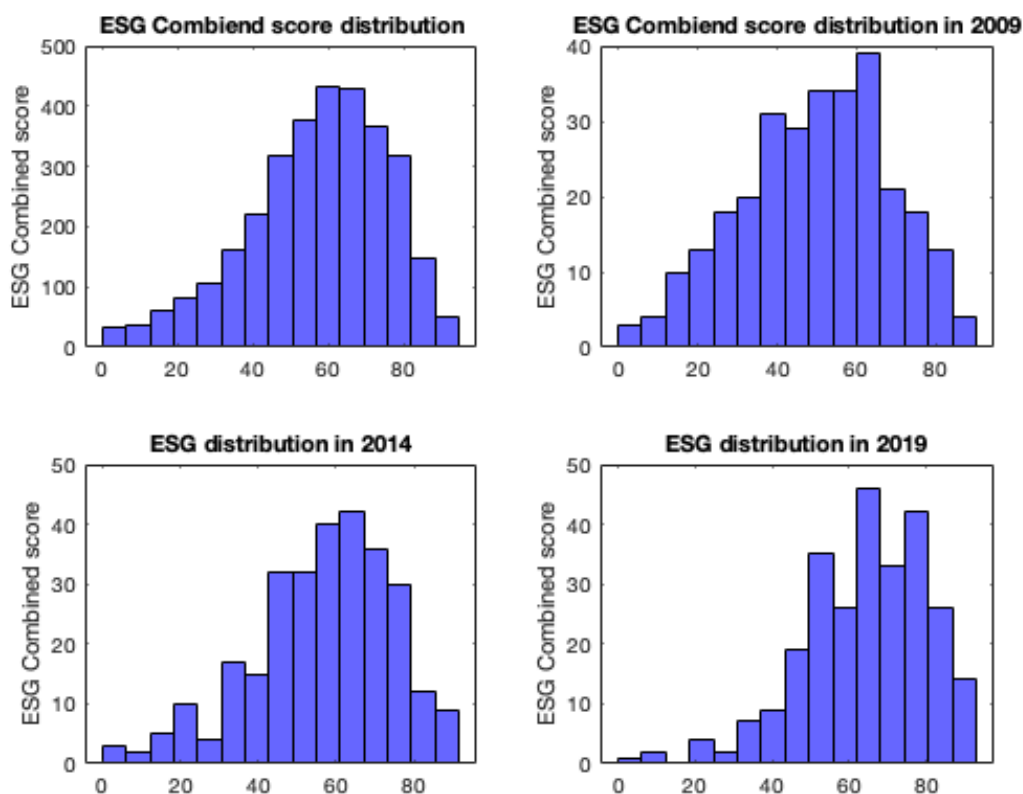

Figure 5- ESG Combined score distribution

3.2.4 The portfolios' Sharpe ratios

The Sharpe Ratio (SR) for portfolio *i* is a risk-adjusted performance measure that is computed as follows:

$$SR_i = \frac{\bar{R}_i - \bar{R}_f}{\sigma_i}$$

Where:

- $\bar{R}_i$ is portfolio the *ith*'s portfolio historical average return
- $\bar{R}_f$ is the average risk-free rate
- σ_i is *ith*'s portfolio standard deviation

As one can see, the Sharpe Ratio compares the mean excess return of a portfolio with its standard deviation, showing the risk premium earned for every additional unit of total risk

(systematic and idiosyncratic). Therefore, the higher the Sharpe Ratio, the higher the portfolio's performance. Being a risk-adjusted performance measure, it allows for the comparison between two or more portfolios, helping investors choosing what to invest in.

The Long-Short portfolio built using the negative screening is the one having the lowest Sharpe Ratio, with a value of -4%; the highest Sharpe ratio is the one built combining the negative screening with the best-in-class approach (20%). The Sharpe ratios for the other portfolios are as follows: 7% for the Long-short portfolio built using the positive screening and 5% for the other two Long-Short portfolios built, respectively, with the positive screening and combining the negative screening with the best in class approach. So, all the Long-short portfolios – apart from the negative screened one – have a positive Sharpe ratio. This is a first hint that the best ESG rated companies have a superior portfolio performance with respect to those with a low ESG rating. These results seem compatible with the analysis in Kempf and Osthoff (2007). In fact, they find not only that the Long-Short portfolio resulting from the combination of the negative and best-in-class approaches is the one allowing for the highest extra-return but also that the negative screening approach is the most disadvantageous one, having the Long-Short portfolio a negative alpha (therefore a negative extra-return). However, it must be noted that a Sharpe ratio of 20% seems excessively high especially with respect to the others. It is therefore needed further investigation to check whether we got results like these by chance or they really reflect what happens in reality. To do so, a Carhart (1997) model will be estimated in the following paragraphs.

In the appendix we analyse, and compare, the riskiness of the portfolios built using the Value at Risk.

3.2.5 Results of Carhart model

Table 6 summarizes the results of Carhart's model for all the portfolios built on the basis of the ESG integration techniques listed in paragraph 3.2.1.3. It shows the coefficients of the different factors along with the alphas, their p-values and the regressions' R-squared.

The different regressions' R-squared are fairly high: the range is 46%-82% with a high concentration on its upper end. This means that Carhart's model explains well the return generation process of the portfolios analysed. Therefore, using this model to examine whether there is the possibility to gain extra-returns including ESG considerations in the

investment process was a good choice. It can also be observed that – along with the market risk – the size and book-to-market factors have, in most cases, a statistically significant impact on portfolios' returns. As far as the momentum factor is concerned, the beta is significant (at most at 5% confidence level) for all the Best portfolios built using negative, positive and the two screens combined, together with all the portfolios built using the best-in-class approach and the combination between the negative and best-in-class approaches.

As anticipated, this study focuses on the sign and level of significance of the different regressions' alphas. In order to better understand if the Best portfolios really overperform the Worst ones, and therefore whether ESG investing can be decisive in determining a portfolio return performance, the primary focus is on the Long-Short portfolios' alphas. In fact, if the alpha of a Long-Short portfolio is positive and statistically significant, the Best portfolio outperforms the Worst one and it can be concluded that introducing ESG considerations into the investment process gives an advantage on portfolio performance.

We first look at the portfolios built using negative screens. The Best portfolio, in this case, is built excluding all the companies that have revenues coming from controversial business areas. As Table 6 shows, the negatively screened Long-Short portfolio has a negative although non-significant alpha (-0.05%), as happens in Kempf and Osthoff (2007) study. A possible explanation is that the companies that the businesses of Alcohol, Gambling, Nuclear and Tobacco are highly profitable and completely excluding them from the Best portfolio and putting them all in the Worst one ensures an overperformance of the latter over the Best portfolio.

We now turn to the results of the positively screened portfolios' regressions. The Best and Worst portfolios are built picking respectively the best and worst rated 10% stocks of our investible universe. In this case, the Long-Short portfolio has a positive (2.66%) and statistically significant alpha (at 5% confidence level). The Long-Short portfolio resulting from the combination of negative and positive screening has small (0.21%), positive and non-statistically significant alpha.

Lastly, we look at the portfolios built using the best-in-class approach. In this case, companies in the whole investible universe are divided into industry classes and, within each industry class, companies are ranked according to their overall ESG score in order to pick the best and worst 10% for the Best and Worst portfolio respectively. We can see that the

Long-Short portfolio has an alpha which is positive (1.24%) and significant with 10% confidence level. The Long-Short portfolio resulting from the combination of negative and best-in-class screenings has a positive alpha (1.88%) which is statistically significant at 1% confidence level.

The fact that the alphas of the Long-Short portfolios are all positive – excluding the ones resulting from the negative screening – suggests that, in the period analysed and for the investible universe considered, companies' ESG performances positively influence portfolio returns. However, as already mentioned, it must be noted that the model does not allow to assess if the extra-return of sustainable and responsible portfolios over the non-sustainable and responsible investing, can be attributed to an additional risk factor that Carhart's (1997) model doesn't capture or is the consequence of a market inefficiency. Moreover, although the signs and levels of significance of the alphas are substantially in line with what found by Kempf and Osthoff (2007) study (K&O from now on), their absolute values are much lower than K&O's ones, especially for the portfolios built using the best-in-class approach together with the negative screen which, instead, K&O identify as the best incorporation technique. This phaenomenon could be due to many different reasons. The first possible explanation is that the different behaviours are due to the period the two studies analyse: this study uses data from 2009 to 2019, while the Authors' analysis goes from 1992 to 2004. If the extra-return of SRI portfolios can be attributed to the presence of an additional risk factor, it could be that – with time – the risk has been reduced; if the extra-return of SRI portfolios is due to a market inefficiency, it could be that the inefficiency itself has been already exploited by investors and it is on its way to extinction. The second possible explanation for the difference in alphas' magnitude, is that the European market and the American market react differently with respect to ESG considerations: if the extra-return is due to a market inefficiency the European market, perhaps, is more efficient than the American one; if the extra-return is due to an additional risk factor, it may be that European investors deem this risk as less impacting than what their American counterparts do.

Carhart's model results						
	Alpha	MKT	SMB	HML	MOM	R2
Best Negative screening	-0.15	1.59	-1.09	0.44	-0.28	0.81
P-value	(0.45)	(0.00)	(0.00)	(0.04)	(0.03)	
Worst Negative screening	-0.14	1.55	-1.37	-0.12	-0.19	0.71
P-value	(0.56)	(0.00)	(0.00)	(0.65)	(0.23)	
Long-Short Negative screening	**-0.05**	**0.04**	**0.29**	**0.55**	**-0.09**	**0.14**
P-value	(0.76)	(0.63)	(0.17)	(0.00)	(0.45)	
Best Positive screening	-0.03	1.59	-1.23	0.52	-0.3	0.77
P-value	(0.89)	(0.00)	(0.00)	(0.04)	(0.05)	
Worst Positive screening	-2.73	6.06	2.00	4.07	-0.42	0.66
P-value	(0.01)	(0.00)	(0.14)	(0.00)	(0.58)	
Long-Short Positive screening	**2.66**	**-4.46**	**-3.21**	**-3.55**	**0.12**	**0.54**
P-value	(0.01)	(0.00)	(0.01)	(0.00)	(0.87)	
Best Negative screening & Positive screening	-0.09	1.59	-1.19	0.38	-0.34	0.78
P-value	(0.66)	(0.00)	(0.00)	(0.11)	(0.02)	
Worst Negative screening & Positive screening	-0.35	1.33	0.25	0.96	-0.35	0.47
P-value	(0.38)	(0.00)	(0.61)	(0.03)	(0.19)	
Long-Short Negative screening & Positive screening	**0.21**	**0.27**	**-1.42**	**-0.57**	**0.01**	**0.1**
P-value	(0.57)	(0.17)	(0.00)	(0.17)	(0.96)	
Best Best-in-class approach	0.64	5.65	-2.61	2.66	-1.35	0.8
P-value	(0.39)	(0.00)	(0.01)	(0.00)	(0.01)	
Worst Best-in-class approach	-0.64	6.16	2.83	1.56	-1.12	0.73
P-value	(0.47)	(0.00)	(0.01)	(0.11)	(0.06)	
Long-Short PORT Best-in-class approach	1.24	**-0.51**	**-5.43**	1.11	-0.23	**0.23**
P-value	(0.1)	(0.19)	(0.00)	(0.18)	(0.64)	
Best Negative screening & Best-in-class approach	0.77	4.16	-2.44	0.56	-0.54	0.7
P-value	(0.24)	(0.00)	(0.00)	(0.43)	(0.22)	
Worst Negative screening & Best-in-class approach	**-1.16**	4.39	0.48	3.02	-0.55	0.66
P-value	(0.17)	(0.00)	(0.64)	(0.00)	(0.33)	
Long-Short Negative screening & Best-in-class approach	**1.88**	**-0.23**	**-2.91**	**-2.46**	**0.01**	**0.15**
P-value	(0.01)	(0.56)	(0.00)	(0.00)	(0.98)	

Table 6 – Carhart (1997) model's regression coefficients and R-squared

3.2.6 Robustness checks

In order to be able to generalise the conclusions above, some robustness checks need to be performed. In the next paragraphs it is checked whether results are influenced by the portfolio weighting scheme: equally weighted – instead of value weighted – portfolios are built. Moreover, it is checked if results are influenced by the cut-off rate: thus far the portfolios built using the positive or best-in-class screening were based on a 10% cut-off rate: for every incorporation technique, a portfolio with 5%,25% and 50% cut off rates are now created.

3.2.6.1 Equal weights

As far as portfolios' weighting schemes are concerned, many are the alternatives a portfolio manager has. In this study a value weighting scheme has been used, both to allow for a fair comparison with K&O's findings. However, if the weighting scheme influences portfolio returns, it must be checked if it also influences the relationship between the Best and Worst portfolios. In order to do so, equally weighted portfolios are built. The focus will, again, be on the Long-Short portfolios.

Table 7 shows the comparison of the alphas for the equally weighted Long-Short portfolios with 10% cut-off rate and their value weighted counterparts, together with the p-values. One can note that, in line with what observed until now, all the portfolios – apart from the negatively screened one – have positive alphas. As can be seen, the alphas with the highest level of significance are those built using the simple positive screen and the combination of the latter with negative screening. In this case the portfolios built using the best-in-class approach (simple or combined with the negative screen) yield positive, although non-statistically significant, alpha. It must be noted that the magnitude of the alphas is not comparable to those found using the value weighted portfolios, therefore using equal weights would result in lower extra-returns for the Best portfolios.

Equally (EW) and value (VW) weighted portfolios' alphas		
	EW	VW
Long-Short Negative screening	-0.23	-0.05
P-value	(0.08)	(0.76)
Long-Short Positive screening	0.63	2.66
P-value	(0.00)	(0.01)
Long-Short Negative screening & Positive screening	0.37	0.21
P-value	(0.07)	(0.57)
Long-Short best-in-class	0.22	1.24
P-value	(0.26)	(0.10)
Long-Short negative screening & Best-in-class	0.24	1.88
P-value	(0.30)	(0.01)

Table 7 - Carhart's alpha for equally weighted and value weighted 10% cut off Long-Short portfolios

After this robustness check, it can still be inferred that there is an over performance of the Best portfolios over the Worst ones, and that there is indeed the possibility to gain an extra-return investing in SRI portfolios with respect to non-SRI portfolios. The best incorporation technique is still the simple positive screen, while the best-in-class approach – in this case – does not seem to guarantee an overperformance with respect to the market of the Long-Short portfolio. Finally, it can be concluded that the weighting scheme strongly influences the magnitude of the impact of the ESG performance on portfolios' returns since equally weighted portfolios' returns seem less prone to be influenced by the different companies' ESG performance. It should be further investigated if, using other weighting schemes, the results are more similar to the ones obtained using the value or equal weighting schemes to assess whether the results just illustrated are due to the specific choice of weighting scheme or the over-performance of the Best portfolios over the Worst ones is not consistent across different weighting schemes.

3.2.6.2 Alternative cut offs

Until here, the analysis was based on portfolios built using a cut-off rate of 10%. This means that the Best portfolios are built with the 10% best ESG-rated of the investible

universe and the Worst ones are built using the 10% worst rated. This choice influences which companies a hypothetical investor could invest in. It must be therefore checked what happens if the cut-off rate changes. To do so, three other cut-off rates are used: 5%,25% and 50%. Also in this case, the focus is on the Long-Short portfolios in order to easily assess – for each ESG incorporation strategy – if the Best portfolio over-performs with respect to the Worst one.

Table 8 shows Carhart's model results for Long-Short portfolios built with the different cut off rates. The alphas of the Long-Short portfolios just built are mostly positive, but some are negative like the one of the positively screened 25% cut-off portfolio and the ones created using the negative screening in addition to the positive screening with 5% and 25% cut off rates. Until now negative alphas were observed only for negatively screened portfolios. Moreover, it can be noticed that for positive screening the portfolios built using 5% and 10% cut off rates are the only ones with a statistically significant alpha, leading to the conclusion that investors should concentrate on the very best stocks in order to take advantage of good ESG performances. The same happens for the portfolio resulting from the combination of negative screening and the best-in-class approach. Instead, as far as the simple best-in-class approach is concerned, it seems that it would be best for an investor to focus on higher cut off rates, being the alphas of the 25% and 50% cut off rates statistically significant and sensibly higher than the one obtained with the 10% cut off rate. This appears to be in contradiction with what concluded for the positively screened portfolios and those built using the combination of negative screening and best-in-class approach.

Alphas for different cut-offs				
	5%	10%	25%	50%
Long-Short Positive screening	1.02	2.66	-0.37	0.26
P-value	(0.05)	(0.01)	(0.70)	(0.48)
Long-Short Negative screening & Positive screening	-0.12	0.21	-0.04	0.08
P-value	(0.80)	(0.57)	(0.84)	(0.50)
Long-Short Best-in-class approach	0.59	1.24	1.29	2.27
P-value	(0.21)	(0.10)	(0.02)	(0.03)
Long-Short negative screening & Best-in-class	1.40	1.88	0.20	0.58
P-value	(0.03)	(0.01)	(0.60)	(0.30)

Table 8 - Alphas of value-weighted Long-Short portfolios at different cut-off rates

It seems difficult to spot a pattern in these results and this raises the question on whether the results got until here were only by chance supporting the thesis of an extra-return from SRI portfolios over non-SRI portfolios. Although also Lanza, Bernardini, & Faiella (2020), using the same investible universe of the present study and a sophisticated Machine Learning approach both to portfolio allocation and to the choice of the best indicators to use to measure corporate social responsibility, find that the choice of a particular threshold can lead to a Best portfolio overperforming over the Worst one while with another cut-off rate the situation was reversed. The authors deal with this situation using a Machine Learning approach to determine which thresholds allow a Best portfolio to have a superior risk-adjusted performance. So, it could be that using a more sophisticated investment decision process could generate more consistent results in a robustness check like the one just performed, allowing to spot a more clear pattern and helping investors choose the optimal cut-off rate.

Conclusions and future developments

This work – after an overview of what Socially Responsible Investing, ESG factors and ESG ratings are – shows that in Europe in the period that goes from 2009 to 2019 the high ESG-rated portfolios have an over-performance with respect to the low rated ones.
The analysis started in chapter 1 pointing out that introducing ESG considerations in the investment decision process, and therefore implementing a Sustainable and Responsible

Investment approach during the last decade has become an increasingly popular practice. The study also illustrated that ESG considerations can be included in the investment process either through incorporation or trough active ownership and the reasons for choosing to do so stand in: materiality of ESG information with respect to portfolio performance and to risk management; end-investors asking to investment managers and issuers to pay more attention to ESG matters, given their increased interest and awareness of the influence of ESG matters on a company's returns and reputation; finally, another reason for using ESG information in the investment process is the fact that regulators, after realizing that the financial sector can have a great role in helping dealing with social and environmental challenges, generated a significant increase in responsible investment regulation in the last twenty years, especially after the 2008 financial crisis. However, the incorporation of ESG information in the investment decision process faces many challenges among which we can find: the hardship of finding high quality data, the widespread belief that incorporation means giving up returns and therefore violating the fiduciary duty together with the time frame mismatch between ESG-investing's impact horizon and investors performance evaluation time frame. It was also shown that the scepticism over the importance of ESG information in investment decisions is fuelled also by the fact that companies' ESG information disclosure lacks standardization and is often based on voluntary disclosure. This study concluded, in fact, that although the European Union has introduced some important directives on the matter, but there is still room for improvements.

The second chapter of this work shows how – given the hardship of finding high quality data – investors often rely on ESG ratings provided by many different rating agencies that help evaluating a firm's ESG performance. These ratings have two major shortcomings: the lack of uniformity and transparency. Investors, when using ESG ratings, must keep in mind these limitations and perhaps find a way to overcome them before using the ratings in investment decisions.

In the third chapter an empirical analysis is performed in order to test whether companies' ESG rating influences portfolio returns, and – in particular – if high ESG-rated portfolios overperform with respect to the low rated ones. The investible universe used in this study is the set of EURO STOXX's components as of 01/01/2009 and the period analysed goes from 01/12/2009 to 31/12/2019. Fifteen are the value weighted 10% cut-off portfolios built; a

Best, Worst and Long-Short portfolios for five different ESG incorporation methods: negative screening, positive screening, negative screening combined with positive screening, best-in-class approach and negative screening combined with the best-in-class approach. The focus of the analysis is on the risk adjusted performance of the Long-Short portfolios that is evaluated computing the Sharpe ratio and estimating the Jensen's alpha generated from Carhart (1997) model. The Sharpe ratios are all positive, apart from the one of the Long-Short portfolio generated by implementing the negative screen. The alphas of the model estimated are positive and mostly statistically significant apart from a negative alpha for the negatively screened Long-Short portfolio. These results are backed by the alphas obtained performing a robustness check that consisted in applying Carhart (1997) model to equally weighted portfolios instead of value weighted ones. These results allow to infer – as happens in the United States according to Kempf and Osthoff (2007) – that in the period analysed in Europe there is an overperformance of the high ESG-rated stocks over the low rated ones and the negative screening technique is not suitable to gain an extra-return since it doesn't allow to 'do well while doing good' but only to give up returns in favour of non-financial utility. This suggests that if European investment and asset managers – who prefer negative screening over the other incorporation techniques (Eurosif, 2018) – want to fully exploit SRI advantages should abandon negative screening for more complex and modern techniques. At the same time the results of the analysis make it difficult to define which is the best technique to exploit companies' ESG performance in portfolio construction. In fact, while the analysis of the different Sharpe ratios suggests that the best incorporation method is the negative screening combined with the best-in-class approach – as Kempf and Osthoff (2007) observe for the United States – Carhart (1997) model's alphas suggests that the method with which investors can gain the highest extra-return is positive screening. As a consequence, there is the need for further investigation on this matter. It must be noted that the alphas registered from the application of Carhart (1997) model to equally weighted portfolios as a robustness check, resulted to be much smaller than the ones obtained using the value weighted portfolios. This draws to the conclusion that the weighting scheme may strongly influence the level of extra-return that investors can get by incorporating ESG considerations in the investment decision process. This issue must be further investigated as well, perhaps replicating the analysis performed using more

weighting schemes. Another robustness check was performed creating other five Long-Short portfolios using a 5%, 25% and 50% cut-off rates, in order to check how the cut off rates influence results. It is quite difficult to draw a conclusion, this is because the choice of a particular cut-off rate can lead a Best portfolio to overperform with respect to the Worst one while with another cut-off rate the situation is reversed (Lanza, Bernardini, & Faiella, 2020). Hence investors and portfolio managers will have to carefully choose the cut off rate, possibly using a Machine Learning algorithm that helps finding the threshold that maximises the performance of high ESG-rated portfolios. Talking about magnitude, it is necessary to point out that the absolute values of the alphas obtained regressing the value weighted 10% cut-off portfolio on the Carhart (1997) model's factors, are far lower than the ones obtained by Kempf and Osthoff (2007) study. This difference may be due to the different time period analysed in this study and in Kempf and Osthoff one as well as to differences in market structure and investors' behaviour between the American and European markets.

Two are the main limitations of this study: it is based on just one ESG rating provider and, given that ratings for the same company can be significantly different, the same analysis should be performed using at least another rating provider to be able to assess with certainty that the findings of this study mirror what happens in reality. Another limitation is that the Carhart (1997) model does not allow to determine a priori whether a significative alpha is due to market inefficiency or to a risk factor that is not captured by the model.

Overall, the results of this study suggest that companies' ESG performance represents valuable information for European investors as well as for American ones. Future research should try to assess whether the extra-return observed for companies with high ESG rating using Carhart (1997) model, is the result of the compensation for an additional risk factor or is just the consequence of a temporary mispricing.

Appendix: Value at Risk

In this appendix we investigate the riskiness of the portfolios. To do so the Value at Risk (VaR) is used. The latter is a risk measure that allows for comparison between different securities, representing the maximum loss that an investor can incur with a certain level of confidence. Assuming that the portfolios' returns (R_i) are normally distributed such that $R_i \sim N(\mu, \sigma_{it}^2)$ where $\mu = 0$. The monthly VaR at 95% confidence level of the i-th portfolio is as follows:

$$VaR_{it} = 1.64 * \sigma_{it}$$

where 1.64 is the 95-th percentile of the standard normal distribution and σ_{it} is the time varying variance of portfolio i. In order to estimate the latter, the Generalised Autoregressive Conditional Heterostkedasticity (GARCH) model is used. This model estimates the variance based on its past value allowing to account for volatility clustering, in particular we use the GARCH (1,1) which is as follows:

$$\sigma_t^2 = \omega + \alpha \, \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$$

where: ω, α, β are unknown coefficients to be estimated through maximum likelihood; ε_{t-1}^2 represents the past return innovation; σ_{t-1}^2 is the t-1 variance.

In Table 9 we find the average monthly 95% confidence level Value at Risk for all the Best and Worst portfolios. As can be noticed, the Worst portfolios always have a higher average monthly VaR with respect to the best ones. This shows that the Best portfolios we created, not only have higher returns, but are less risky too.

Average monthly Value at Risk (%)	
Best Negative screening	7.54
Worst Negative screening	7.32
Best Positive screening	8.08
Worst Positive screening	29.25
Best Negative screening & Positive screening	7.93
Worst Negative screening & Positive screening	8.99
Best Best-in-class	28.18
Worst Best-in-class	27.48
Best Negative screening & Best-in-class	20.53
Worst Negative screening & Best-in-class	23.47

Table 9 - Average Best and Worst portfolios' VaR

Carhart's (1997) model suggests that the positive screening incorporation technique stands out as the best technique to gain an extra-return in order to exploit companies' ESG performances. Therefore, we investigate if the VaR of the Best positively screened portfolio has a lower Value at Risk than the Worst one for the whole period analysed, being the VaR a time varying variable. The results of this analysis are shown by Figure 6, depicting the two portfolios' VaR. As can be noticed, the Best portfolio's Value at Risk is always lower than the one of the Worst portfolios.
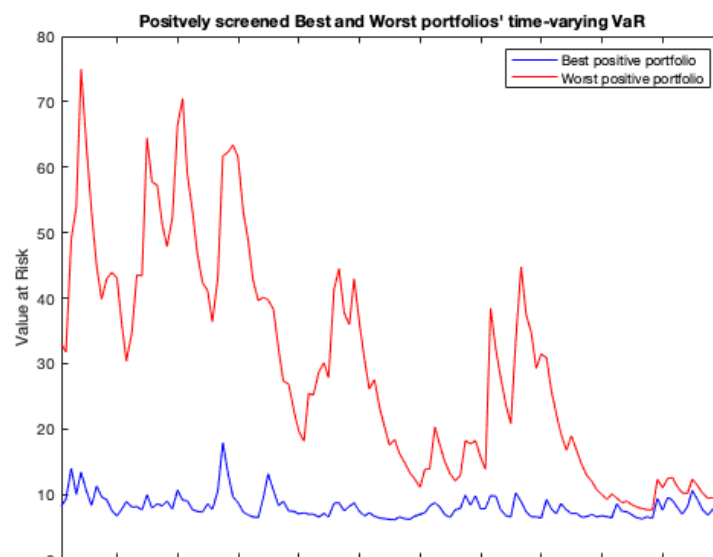
Figure 6- Positively screened Best and Worst portfolios' time-varying VaR

3.2.6.2.1 VaR backtesting

In order to assess the level of statistical accuracy of our VaR estimates, we use three different tests: Kupiec, Christoffersen and Berkowitz's test.

Kupiec (1995) introduced a test – also called proportion of failures test – that has the objective to test whether the probability of VaR being exceeded is 5% (as implied in VaR's level of confidence) or not. This hypothesis is tested using a likelihood ratio built as follows:

$$LR_{POF} = -2 \ln \left[\frac{\alpha^x (1-\alpha)^{n-x}}{\pi^x (1-\pi)^{n-x}} \right]$$

where:

- α is the probability of VaR being exceeded implied in VaR's level of confidence
- n is the number of observations in the sample.
- x is the number of times the VaR is exceeded, also called exceptions.
- π is the proportions of exceptions with respect to the total number of observations.

The LR is asymptotically distributed as a Chi-squared with one degree of freedom and POF stands for probability of failure.

Christoffersen (1998) introduce a test to verify the independence of exceptions. The main hypothesis behind it is: if returns are independent, exceptions must be independent as well and evenly spread during the period analysed without being clustered. The null hypothesis, therefore, is that the probability for an exception to happen is independent with the fact that an exception happened the day before. Also in this case a Likelihood Ratio asymptotically distributed as Chi-squared with one degree of freedom is used to test the hypothesis built in the following way:

$$LR_{CCI} = -2 \ln \left[\frac{(1-\pi)^{n00+n10} \pi^{n01+n11}}{(1-\pi_0)^{n00} \pi_0^{n01} (1-\pi_1)^{n10} \pi_1^{n11}} \right]$$

where:

- n00: represents the number of periods without an exception followed by a period without an exception as well
- n10: represents the number of periods with at least an exception followed by a period without an exception
- n01: represents the number of periods without an exception followed by a period with at least an exception

- n11: represents the number of periods with at least an exception followed by a period with at least an exception as well
- π is the probability for an exception to happen in period t
- π_0 is the probability for an exception to happen in period t given that no exception happened in the period before
- π_1 is the probability for an exception to happen in period t given that at least one exception happened the period before

As can be noticed looking at tables 10 and 11, Kupiec's and Christoffersen's test do not reject the null hypothesis for the VaR of all the portfolios, therefore the proportion of exceptions does not exceed 5% and the exceptions are independently distributed.

VaR backtesting: Kupiec test				
	LRatioPOF	PValuePOF	Number of failures	Outcome
Best Negative screening	1.60	0.21	10	accept
Worst Negative screeninig	0.83	0.36	9	accept
Long-Short Negative screening	0.03	0.87	7	accept
Best Positive screening	1.60	0.21	10	accept
Worst Positive screening	0.83	0.36	9	accept
Long-Short Positive screening	1.25	0.26	4	accept
Best Negative screening Positive screening	2.59	0.11	11	accept
Worst Negative screening Positive screen	0.29	0.59	8	accept
Long-Short Negative screening Positive screen	0.06	0.81	6	accept
Best Best-in-class	3.78	0.05	12	accept
Worst Best-in-class	0.83	0.36	9	accept
Long-Short Best-in-class	0.03	0.87	7	accept
Best Negative screening Best-in-class	0.03	0.87	7	accept
Worst Negative screening Best-in-class	0.29	0.59	8	accept
Long-Short Negative screening Best-in-class	1.25	0.26	4	accept

Table 10 - Kupiec test results

VaR backtesting: Christoffersen test								
	LRatioCCI	PValueCCI	Failures	N00	N10	N01	N11	Outcome
Best Negative screening	1.65	0.20	10	111	10	10	0	accept
Worst Negative screeninig	0.35	0.55	9	115	8	7	1	accept
Long-Short Negative screening	0.79	0.37	7	117	7	7	0	accept
Best Positive screening	0.08	0.78	10	112	9	9	1	accept
Worst Positive screening	1.33	0.25	9	113	9	9	0	accept
Long-Short Positive screening	0.25	0.62	4	123	4	4	0	accept
Best Negative screening Positive screening	0.01	0.93	11	110	10	10	1	accept
Worst Negative screening Positive screen	0.48	0.49	8	116	7	7	1	accept
Long-Short Negative screening Positive screen	0.58	0.45	6	119	6	6	0	accept
Best Best-in-class	0.75	0.39	12	109	10	10	2	accept
Worst Best-in-class	2.43	0.12	9	115	7	7	2	accept
Long-Short Best-in-class	0.79	0.37	7	117	7	7	0	accept
Best Negative screening Best-in-class	0.84	0.36	7	118	6	6	1	accept
Worst Negative screening Best-in-class	0.48	0.49	8	116	7	7	1	accept
Long-Short Negative screening Best-in-class	0.25	0.62	4	123	4	4	0	accept

Table 11 - Christoffersen test results

Figures index

Tables index

Bibliography

Auer, B. R. (2016). Do Socially Responsible Investment Policies Add or Destroy European Stock Portfolio Value? *Journal of Business Ethics*(235), 381-397.

Auer, B. R., & Schuhmacher, F. (2016). Do socially (ir)responsible investments pay? New evidence from international ESG data. *The Quarterly Review of Economics and Finance, 52*, 51-62.

Baldini, M., Dal Maso, L., Liberatore, G., Mazzi, F., & Terzani, S. (2016). Role of Country- and Firm-Level Determinants in Environmental, Social, and Governance Disclosure. *Journal of Business Ethics, 150*, 79-98.

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance, 29*(7), 1751-1767.

Berg, F., Koelbel, J. F., & Rigobon, R. (2020). *Aggregate confusion: the divergence of ESG ratings - MIT Sloan School Working paper.* Retrieved from SSRN: https://ssrn.com/abstract=3438533

Berline, D., & Prakash, A. (2005). "Bluewashing" the Firm? Voluntary Regulations, Program Design, and Member Compliance with the United Nations Global Compact. *Policy Studies Journal, 43*(1), 115-138.

Bernardi, C., Demartini, P., Dumay, J., & Guthrie, J. (2016). Integrated reporting: A structured literature review. *Accounting Forum, 40*(3), 166-185.

Capelle-Blancard, G., & Monhon, S. (2014). The performance of Socially Responsible Funds: does the screening process matter? *European Financial Management, 20*(3), 494-520.

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *Journal of finance, 52*(1), 57-82.

CFA institute. (2018). *Guidance and case studies for ESG integration: equities and fixed income.* Retrieved from UNPRI: https://www.unpri.org/investor-tools/guidance-and-case-studies-for-esg-integration-equities-and-fixed-income/3622.article

Chatterji, A. K., Durand, R., Levine, D. I., & Touboul, S. (2016). Do ratings of firms converge? Implications form managers, investors and strategy researchers. *Strategic Management Journal, 37*, 1597-1614.

Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How Well Do Social Ratings Actually Measure Corporate Social Responsibility? *Journal of Economics & Management Strategy, 18*(1), 125–169.

Christoffersen, P. (1998). Evaluating Interval Forecasts. *International Economic Review, 39*(4), 841–862.

Cuerel Burbano, V., & Delmas, M. A. (2011). The Drivers of Greenwashing. *California management review, 54*(1), 64-87.

Derwall, J., Guenster, N., Bauer, R., & Koedijk, K. (2005). The Eco-Efficiency Premium Puzzle. *Financial Analysts Journal, 61*(2), 51-63.

Eccles, R. G., & Serafeim, G. (2014). Corporate and Integrated Reporting: A Functional Perspective. *(Working Paper) Harvard Business School* , 1-22.

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science, 60*(11), 2835-2857.

Eccles, R. G., Kastrapeli, M. D., & Potter, S. J. (2017). How to Integrate ESG into Investment Decision-Making: Results of a Global Survey of Institutional Investors. *Journal of applied corporate finance, Volume 29*(Fall 2017, issue 4), Pages 125-133.

EU High-Level Expert Group on Sustainable Finance. (2018). *Final Report 2018 by the High-Level Expert Group on Sustainable Finance.* Retrieved from https://ec.europa.eu/info/publications/sustainable-finance-high-level-expert-group_it

EU Technical Expert Group on sustainable finance. (2020). *TEG final report on the EU taxonomy*. Retrieved from European Commission: https://ec.europa.eu/info/files/200309-sustainable-finance-teg-final-report-taxonomy_en

European Coalition for Corporate Justice. (2014). *Assessment of the EU Directive on the disclosure of non-financial information by certain large companies.* Retrieved from European Coalition for Corporate Justice: https://corporatejustice.org/eccj-publications/391-assessment-of-the-eu-directive-on-the-disclosure-of-non-financial-information-by-certain-large-companies

European Commission. (2001). *Green Paper.* Retrieved from https://ec.europa.eu/transparency/regdoc/rep/1/2001/EN/1-2001-366-EN-1-0.Pdf

European Commission. (2016). *Communication from the Commission to the European Parliament, the Council, the European Central Bank, the European economic and social committee and the committee of the Regions.* Retrieved from Eur-Lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1490966616011&uri=CELEX:52016DC0601

European Commission. (2017). *Guidelines on non-financial reporting (methodology for reporting non-financial information).* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017XC0705%2801%29

European Commission. (2018). *Feedback Statement - Public Consultation on Institutional Investors' and Asset Managers' Duties regarding Sustainability.* Retrieved from https://ec.europa.eu/info/sites/info/files/2017-investors-duties-sustainability-feedback-statement_en.pdf

European Commission. *Paris Agreement.* Retrieved from European Commission: https://ec.europa.eu/clima/policies/international/negotiations/paris_en#:~:text=The%20Paris%20Agreement%20sets%20out,support%20them%20in%20their%20efforts.

European federation of financial analysts societies. (2009). *KPIs for ESG - A Guideline for the Integration of ESG into Financial Analysis and Corporate Valuation.* Retrieved from EFFAS: https://effas.net/papers/effas-standard-setter.html

European Union. (2013). *Corporate social responsibility: promoting society's interests and a route to sustainable and inclusive recovery.* Retrieved from Eur-Lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2016.024.01.0033.01.ENG&toc=OJ:C:2016:024:FULL

European Union. (2013). Directive 2013/34/EU of the European Parliament *(Consolidated text).* Retrieved from Eur-Lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02013L0034-20141211

European Union. (2014). *Directive 2014/95/EU of the European Parliament and of the Council.* Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014L0095

European Union. (2017). *Guidelines on non-financial reporting (methodology for reporting non-financial information).* Retrieved from Eur-Lex: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52017XC0705%2801%29

European Union. (2019). *Regulation (EU) 2019/2088 of the European Parliament and of the Council of 27 November 2019 on sustainability-related disclosures in the financial services sector.* Retrieved from Eur-lex: https://eur-lex.europa.eu/eli/reg/2019/2088/oj

Eurosif. (2018). *European SRI study 2018.* Retrieved from http://www.eurosif.org/sri-study-2018/

Eurosif. *Responsible Investment Strategies.* Retrieved from www.eurosif.org/responsible-investment-strategies/

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journa of Financial Economics, 33*(1), 3-56.

Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal, 38*, 45-64.

Freshfields Bruckhaus Deringer. (2005). *A legal framework for the integration of environmenta, social and governance issues into institutional investment.* UNEP Finance Initiative.

Giannarakis, G., Konteos, G., & Sariannidis, N. (2014). Financial, governance and environmental determinants of corporate social responsible disclosure. *Management decision, 32*(10), 1928-1951.

Global Sustainable Investment Alliance. (2018). *Global Sustainable Investment Review 2018.* GSI Alliance.

Glossner, S. (2018, May 18). *The Price of Ignoring ESG Risks.* Retrieved from SSRN: http://dx.doi.org/10.2139/ssrn.3004689

Haefele, M. (2018). *Sustainable investing can propel long-term returns*. Retrieved from Financial Times: https://www.ft.com/content/292ecaa7-294c-3a4b-bde6-a7a744cb85a9

Herremans, I. M., Akathaporn, P., & McInnes, M. (1993). An investigation of corporate social responsibility reputation and economic performance. *Accounting, Organizations and Society, 18*(7/8), 587-604.

Hull, J. C. (2018). *Risk management and financial institutions - Fifth edition.* Wiley.

In, S. Y., Rook, D., & Monk, A. (2019). *Integrating Alternative Data (Also Known as ESG Data) in Investment Decision Making.* Global Economic Review.

International Integrated Reporting Council. (2013). *International <IR> Framework.* Retrieved from https://integratedreporting.org/resource/international-ir-framework/

International Monetary Fund. (2019). *Global financial stability report: lower for longer.* Retrieved from https://www.imf.org/en/Publications/GFSR/Issues/2019/10/01/global-financial-stability-report-october-2019

Ioannou, I., & Serafeim, G. (2011). The Consequences of Mandatory Corporate Sustainability Reporting. *(Working Paper) Harvard Business school* .

Johnson, S. (2010). *Companies fail UN's Global Compact.* Retrieved from Financial Times: https://www.ft.com/content/8b19f9ee-1806-11df-91d2-00144feab49a

Kempf, A., & Osthoff, P. (2007). The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management, 13*(5), 908-922.

Kerber, R., & Flaherty, M. (2017). *Investing with 'green' ratings? It's a gray area*. Retrieved from Reuters: https://www.reuters.com/article/us-climate-ratings-analysis-idUSKBN19H0DM

Kotsantonis, S., & Serafeim, G. (2019). Four Things No One Will Tell You About ESG Data. *Journal of Applied Corporate Finance, 31*(2), 50-58.

KPMG. (2017). *The KPMG Survey of Corporate Responsibility Reporting 2017.* Retrieved from KPMG: https://home.kpmg/xx/en/home/insights/2017/10/the-kpmg-survey-of-corporate-responsibility-reporting-2017.html

Kupiec, P. (1995). Techniques for Verifying the Accuracy of Risk Management Models. *The Journal of Derivatives, 3*(2), 73–84.

Lanza, A., Bernardini, E., & Faiella, I. (2020, June). Mind the gap! Machine learning, ESG metrics and sustainable investment. *Banca d'Italia Questioni di Economia e Finanza (Occasional Papers), 561*.

Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, Norms, and Selective Disclosure: A Global Study of Greenwashing. *Organization Science, 27*(2), 483-504.

Mervelskemper, L., & Streit, D. (2016). Enhancing Market Valuation of ESG Performance: Is Integrated Reporting Keeping its Promise? *Business Strategy and the Environment, 26*, 536-549.

Mugnier, E., Delerable, C., Tan, A., & Hélouin, A. (2014). *Resource Efficiency and Fiduciary Duties of Investors - Final Report. European Commission – DG Environment* . Retrieved from https://op.europa.eu/en/publication-detail/-/publication/439d9fcf-9fd2-11e5-8781-01aa75ed71a1

PRI. (2020). *Principles of responsible investing.* Retrieved from UN PRI Web site: https://www.unpri.org/pri/an-introduction-to-responsible-investment/what-are-the-principles-for-responsible-investment

Refinitiv. (2020). *Environmental, Social and Govenrance (ESG) scores from Refinitiv.* Retrieved from Refinitiv: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/esg-scores-methodology.pdf

Reilly, F. K., & Brown, K. C. (2012). Investment analysis & portfolio management. 250-260.

Rogers, P., & Stubbs, W. (2013). Lifting the veil on environment-social- governance rating methods. *Social responsibility journal, 9*(4), 622-640.

UNEP FI. (2015). *Fiduciary duty in the 21st century - Final Report.* Retrieved from https://www.unepfi.org/publications/investment-publications/fiduciary-duty-in-the-21st-century/

Windolph, S. E. (2011). Assessing Corporate Sustainability Through Ratings: Challenges and Their Causes. *Journal of Environmental Sustainability, 1*(1), 37-57.