POLITECNICO
MILANO 1863

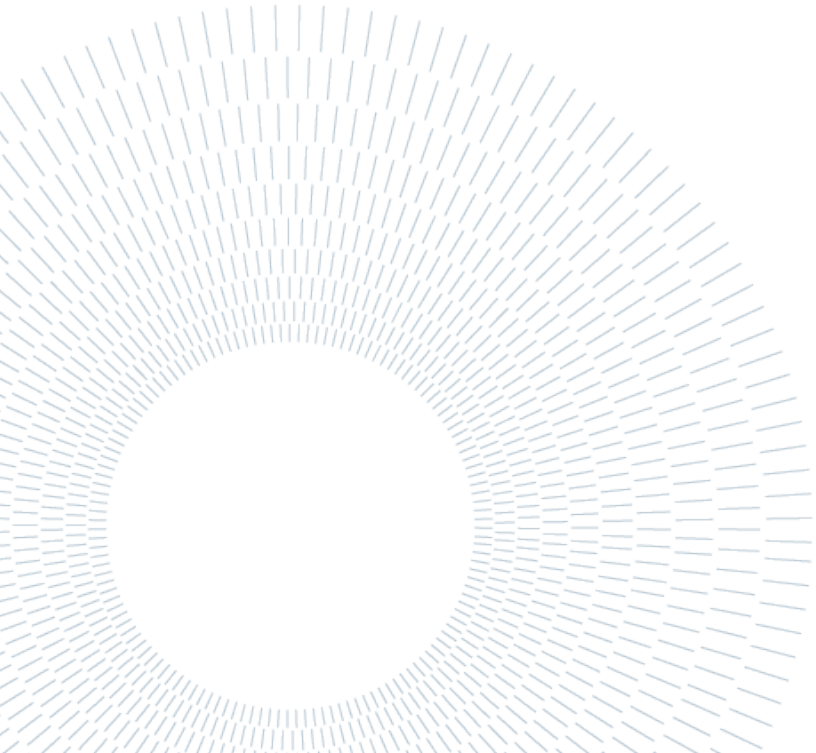
SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

The Impact of ESG Ratings on Stocks' Performance: An Empirical Analysis

TESI DI LAUREA MAGISTRALE IN
MANAGEMENT ENGINEERING
INGEGNERIA GESTIONALE

Authors: **Alessandro Illuzzi**
Simone Inguanta

Student ID: 953202, 942654
Advisor: Prof. Paolo Maccarrone
Academic Year: 2020-21



Abstract

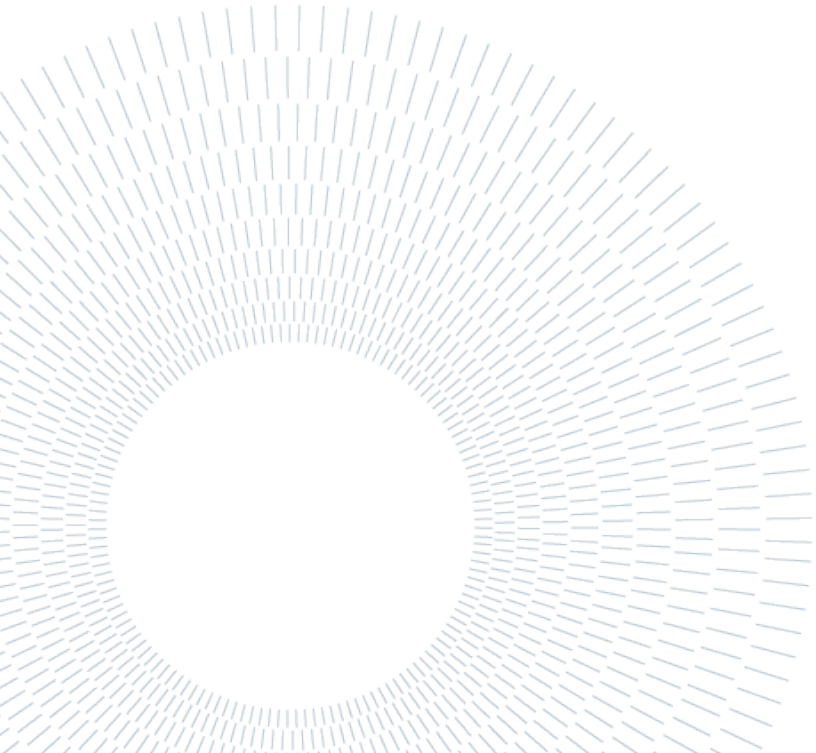
The present work comes to life in a context where sustainability-related topics are assuming in the last years more and more relevance in the financial world, through the ESG paradigm and its integration in the whole economic system. In this scenario, the analysis is inspired by two stylised facts: on the one hand, the positive correlation between the sustainability integration and the corporate financial performances, demonstrated by most of the literature. On the other hand, the existence of divergence in the sustainability evaluations provided by the rating agencies, which leads to the dispersion of the value associated to the ESG topics. Starting from these two drivers, this work aims at investigating the reaction of financial investors to the ESG ratings updates provided by MSCI and Refinitiv. In particular, the objective is to understand whether possible variations in the sustainability ratings of the corporations exert a direct impact on their financial value and, in addition, to figure out if this impact has changed in the last years. The methodology adopted is the Event Study, which is a model used in order to evaluate the impact of an economic event on the valuations of the corporations through the assessment of their Cumulated Average Abnormal Returns (CAARs). In the present work, the economic events are represented by the ESG evaluations updates provided in the period ranging from 2016 to 2020 by the two above-mentioned rating providers, namely MSCI and Refinitiv. More than 700 rating updates (events) are tested on a pool of 75 companies belonging to the main stock exchange markets, and heterogeneous from the belonging industry point of view. The value provided by this analysis can be found in the investigation of how the financial markets price the ESG ratings announcements. The results of the present work integrate the exiting literature, showing that nowadays the ESG finance is far from being mature, with most of the financial investors not looking at ESG ratings in their investment processes. As a result, the companies' market capitalizations are not impacted by changes in their sustainability performances. Moreover, this trend has not changed in the last years, demonstrating that the transition towards an ESG integrated finance still has a long way to go.

Key-words: Event Study, ESG, Ratings, Divergence, Performance, Sustainability.

Abstract in italiano

Il presente lavoro nasce in un contesto nel quale le tematiche legate alla sostenibilità stanno assumendo sempre più rilevanza nel mondo finanziario negli ultimi anni, attraverso il paradigma ESG e la sua integrazione nell'intero sistema economico. In questo scenario, l'analisi prende corpo da due evidenze: da un lato, la correlazione positiva tra investimenti sostenibili e performance finanziarie, dimostrata da gran parte della letteratura. Dall'altro lato, l'esistenza della divergenza nelle valutazioni di sostenibilità fornite dalle agenzie di rating, che conduce alla dispersione del valore associato alle tematiche ESG. Partendo da questi due drivers, il presente lavoro si pone l'obiettivo di investigare la reazione degli investitori agli aggiornamenti delle valutazioni ESG fornite da MSCI e Refinitiv. In particolare, l'obiettivo è quello di capire se possibili variazioni nei rating di sostenibilità delle compagnie esercitano un impatto diretto sul loro valore finanziario e, inoltre, è quello di capire se questo impatto è cambiato negli ultimi anni. La metodologia utilizzata è quella dell'Event Study, che è un modello utilizzato al fine di valutare l'impatto di un evento economico sulle valutazioni delle compagnie attraverso l'analisi dei loro Cumulated Average Abnormal Returns (CAARs). Nel presente lavoro, gli eventi economici sono rappresentati dai cambi nelle valutazioni ESG fornite nel periodo 2016-2020 dalle due agenzie di rating sopra citate, MSCI e Refinitiv. Più di 700 cambi di rating (eventi) sono stati testati su un campione di 75 compagnie appartenenti ai principali mercati finanziari, ed eterogenee dal punto di vista settoriale. Il valore apportato dall'analisi risiede nell'investigare come i mercati finanziari valutano gli annunci dei rating ESG. I risultati del lavoro integrano la letteratura esistente, mostrando che al giorno d'oggi la finanza ESG è lontana dall'essere matura, con la maggior parte degli investitori finanziari che non guarda ai rating ESG nei processi di investimento. Di conseguenza, la capitalizzazione di mercato delle compagnie non è impattata dai cambiamenti nelle loro performance di sostenibilità. Inoltre, questo trend non ha mostrato un cambiamento negli ultimi anni, dimostrando che la transizione verso una finanza sostenibile ha ancora molta strada da fare.

Parole chiave: Event Study, ESG, Ratings, Divergenza, Performance, Sostenibilità.



Contents

1 Introduction

The first chapter of the present work shows the increasing relevance assumed by sustainability nowadays and all the main stakeholders that are embraced by its pervasive impact: governments & policy makers, consumers and companies.

1.1. The Development of Sustainability

Sustainability or sustainable development was defined in the World Commission on Environment and Development's 1987 Brundtland report "*Our Common Future*" [1] as "development that meets the needs of the present without compromising the ability of future generations to meet their own needs". It aims at reconciling the economic development with the protection of social and environmental related aspects, trying to optimize the balance between these three elements.

According to the well-known Shareholder Theory by Friedman (1970) [2], the maximization of the shareholders' returns was perceived as the main objective of a company since the beginning of the industrial revolution. In those years, the sustainability contents were not considered important by most managers for two main reasons: on the one hand, sustainability matters were not perceived as positively correlated to the financial performances; on the other hand, they were

perceived as leading to an increase of firms' costs. However, in the last years an inversion of this trend has been observed due to a higher consciousness related to these themes. This stylized fact has its roots in the higher frequency of extreme weather events (for example the alarming increase of the sea level due to an unnatural melting of glaciers), social scandals affecting some of the biggest players in the market (for example the child labour scandal in Nike and the increase of suicides in Apple), and the financial crisis affecting the financial markets in 2008 (which had a huge negative impact on the wealth of the consumers).

All these events led to a higher attention towards sustainability themes by three main stakeholders: governments, consumers and corporations.

1.1.1. Governments & Policy Makers

For what concern governments and policy makers, in the last years there has been a huge development in global instruments and norms related to sustainability. In this regard, the most important global developments are:

- The *Paris Agreement*, which is an international treaty on climate change signed in 2016 by 195 parties at the United Nations Climate Change Conference (COP21). The goal of the agreement is to limit global warming to well below 2°C compared to pre-industrial levels, and aiming for 1.5°C. This agreement has inspired many legislations aimed at limiting greenhouse gas emissions and setting zero emission targets. In addition, investors, asset owners and asset managers are increasingly looking to align their portfolios to the goal of the Paris Agreement. Furthermore, in the days between 31 October and 13 November 2021, the Paris Agreement has been confirmed

and reinforced by the Conference of the Parties 26 (COP26), though the drafting of the *Glasgow Climate Pact*, which keeps the 2016 targets alive and finalises the outstanding elements of the Paris Agreement.

- The *Sustainable Development Goals (SDGs)*, adopted by all United Nations Member States in 2015, contain 17 globally set goals to achieve a more sustainable future by 2030. The Agenda 2030 contains 231 indicators and 169 targets, among which the most relevant are related to ending poverty and other deprivations, reducing inequality, spurring economic growth and tackling climate change.

Figure 1.1: Sustainable Development Goals

- The *Taskforce on Climate-related Financial Disclosures (TCFD)* was established by the Financial Stability Board in 2017, and it designed a set of Recommendations with the aim of helping companies to provide better information on how organizations disclose information on ESG factors.
- The *United Nations Environment Program Finance Initiative (UNEP FI)* Sustainable Financial Roadmap Initiative proposes an integrated approach to boost the transition towards a sustainable financial system.

In addition, following these guidelines, the governments of different regions and countries emanated several regulatory norms. Considering the most relevant

geographical areas in the sustainable investments arena, only European, American and Japanese cases are presented, since they together own 90% of the global sustainable investing assets.

1.1.1.1. Europe

The most important regulatory developments in Europe include:

- The European Union *Sustainable Finance Action Plan*, published in 2018 and, in particular, the Sustainable Finance Disclosure Regulation (SFDR). It requires institutional investors and asset managers to disclose how they integrate sustainability risks in their investment decisions, reporting their ESG products' sustainability risks and adverse impacts.
- *Corporate Sustainability Reporting Directive* requires large companies to publish regular reports on the social and environmental impacts of their activities. These non-financial reports are going to integrate the companies' sustainable strategies to their financial performances.
- *NextGenerationEu*, is a temporary expansive monetary policy of more than 800 billions of euros designed after the Covid-19 crisis and aimed at fixing the economic and social damages caused by the pandemic, leading to a more resilient, digital and green Europe. This is relevant because it will boost sustainable investment.

1.1.1.2. United States

In the United States, the regulatory policy environment for sustainable investing has completely changed during the current administration. The administration under President Donald Trump tried to limit sustainable investing through actions taken at the Department of Labor (DOL) and Securities and Exchange Commission (SEC). President Biden's administration, instead, has released measures to try to reverse those actions.

Some notable initiatives since the election of President Biden include:

- In March 2021, the SEC issued a request for information on climate risk and sustainability disclosures. This is expected to lead to a regulatory proposal for mandatory issuer disclosure on climate change and potentially a broader set of environmental, social and governance issues.
- In May 2021, the White House released an *Executive Order on Climate-Related Financial Risk*, which included measures aimed at boosting sustainability disclosure.

1.1.1.3. Japan

The most notable policy and regulatory drivers of the last years in Japan include:

- In 2021, the Ministry of Economy, Trade and Industry (METI) introduced a *Green Growth Strategy Through Achieving Carbon Neutrality* in 2050.

- The Financial Services Agency (FSA), METI and the Ministry of the Environment jointly released in 2021 the *Basic Guidelines on Climate Transition Finance*, aimed at strengthening the position of climate transition finance.
- The Tokyo Stock Exchange's *Corporate Governance Code* and the *Guidelines for Investors' and Companies' Dialogue* was amended in 2021 to make specific mention of sustainability topics, including specifically climate change, human rights, and fair and appropriate treatment of the workforce.

1.1.2. Consumers

Beside the higher government pressure on sustainability topics, also the consumers have focused their attention on the impact that products and services have on the environment and on the society. This higher consciousness has led the customers, especially the millennials, to make evaluations during their purchases not only about the functional aspects of the products and services, but also about the social and environmental impact that they have during the product lifecycle and along the entire supply chain delivering them.

More in detail, consumers are increasingly demanding for:

- Clearer information about the sustainability impact of the product and services they purchase.
- Products built using long-lasting and renewable materials that can be recycled at the end of the product lifecycle. In this regard, some relevant actions have been undertaken at the world level to protect the consumer rights: among them, it is interesting to mention on the one hand the "*Right to Repair*" and, on the other hand, the signaling mechanism "*Trashed Too Fast*",

which gives the consumers the possibility to signal those products that are suspected of having reach the end of their lifecycle too early.

- Reusable packaging, aimed at reducing unnecessary wastes. An important initiative boosted by this specific costumers' need is the "*New Plastics Economy*", aimed at eliminating all problematic and redundant plastic items, innovating to ensure the reusability of the unavoidable plastics and, eventually, the initiative has also the objective to circulate all the plastic items so to keep them inside the economy and outside the environment.

This attention to environmental and social issues is testified by the historical sentence issued by the Dutch court against Shell. In particular, the legal case was initiated in 2019 by Milieudefensie, which is part of the international organization "*Friends of the Earth*": more than seventeen thousand Dutch citizens, supported by six NGOs, sued the Dutch oil company for not undertaking sufficient and efficient actions in order to respect the Paris agreement. The Dutch court imposed Shell to reduce its net carbon emissions by 45% within 2030. This event represents an historical breakthrough, since it is the first time that a court forced a private corporation to change its strategy to reduce its carbon footprint. Furthermore, this sentence establishes a legal precedent in all the global courts, where legal causes against polluting multinationals are increasing in number.

1.1.3. Companies

From what has been written so far, the attention to sustainability has become a "must-have" and not anymore a "nice-to-have" for emerging public companies, which have to manage and disclose about the environmental, social and governance

topics if they want to maximize their value. Corporations must look at this forced alignment with emerging political directives not as a burden on their operational costs and on their Profit and Loss account but, on the contrary, they should look at this trend as an opportunity. Indeed, there are two main driving forces that push companies towards the adoption of a more sustainable management: first, investors are increasingly looking to allocate capital in ways that align with their values. Second, there is growing evidence that a strong focus on sustainability issues drives value for corporations. There are several ways through which a strong sustainability integrated strategy can lead to an increase of the enterprise value:

- *Access to capital*: companies caring about environmental, social and governance topics are more likely to attract capital from public and private investors focused on sustainability. Furthermore, these corporations have a privileged access to the debt capital markets.
- *Risk management*: having a strategic management of sustainability-related issues can help companies to lower operating risks, such as natural disasters leading to supply chain disruptions, and legal risks.
- *Cost of capital reduction*: according to the Capital Asset Pricing Model, the lower is the perceived risk of the corporations the smaller is the market premium required by the shareholders.
- *Brand and competitive advantage*: ethically driven companies can increase their market share by targeting segments that care about sustainability issues, such as Gen-Z and Millennial customers. 63% of customers declare to be engages with purpose-driven companies.
- *Talent attraction and retention*: according to the Fast Company survey, 70% of employees declare that a strong substantiality strategy affects their decision to remain at the company.

As reported by Morgan Stanley [3], in the US market large corporations are responding to this trend with improved sustainability transparency and disclosure aimed at capitalizing the advantages listed above. Looking at some numbers, 90% of S&P 500 companies published a sustainability report in 2019, up from 20% in 2011. Also in Europe, after the release of the *Corporate Sustainability Reporting Directive*, companies are obliged to disclose information related to their sustainability strategy.

In this regard, the main concern for the firms is related to how they could communicate their sustainability-related efforts to all the relevant stakeholders. This problem is amplified by the fact that, contrarily to what happens with traditional financial reporting, there is not a standardized way of drafting sustainability reports. Indeed, as shown by the section [1.1.1], the regulatory requirements and the methodology to disclose about environmental, social and governance performances are continuously evolving in the last years, making difficult for the companies to understand which is the best way to draft their non-financial statements. To solve this problem, in the last years some specialised organizations developed a set of guidelines which go in parallel with the evolving ones provided by the regulatory institutions. In this regard, the most important disclosure framework is the one designed and proposed by the Sustainability Accounting Standards Board (SASB), an independent non-profit organization whose mission is to develop sustainability accounting standards that help public corporations to disclose decision-useful information to financial investors. It is important to specify that SASB standards are intended for voluntary use by public companies. Going more in depth, the sustainability topics of the framework are organized under five broad sustainability dimensions (which are shown by the SASB materiality map reported below): environment, social capital, human capital, business model and innovation, leadership and governance. In building its

framework, the SASB identified the sustainability topics from an initial set of 30 broadly relevant sustainability issues organized under these five sustainability dimensions. This comprehensive set was refined through a series of steps designed to identify those issues reasonably likely to have material impact on companies, which means having the potential to affect corporate value: these are topics that can or do affect operational and financial performance through three channels of impact: (1) revenues and costs, (2) assets and liabilities, and (3) cost of capital or risk profile.

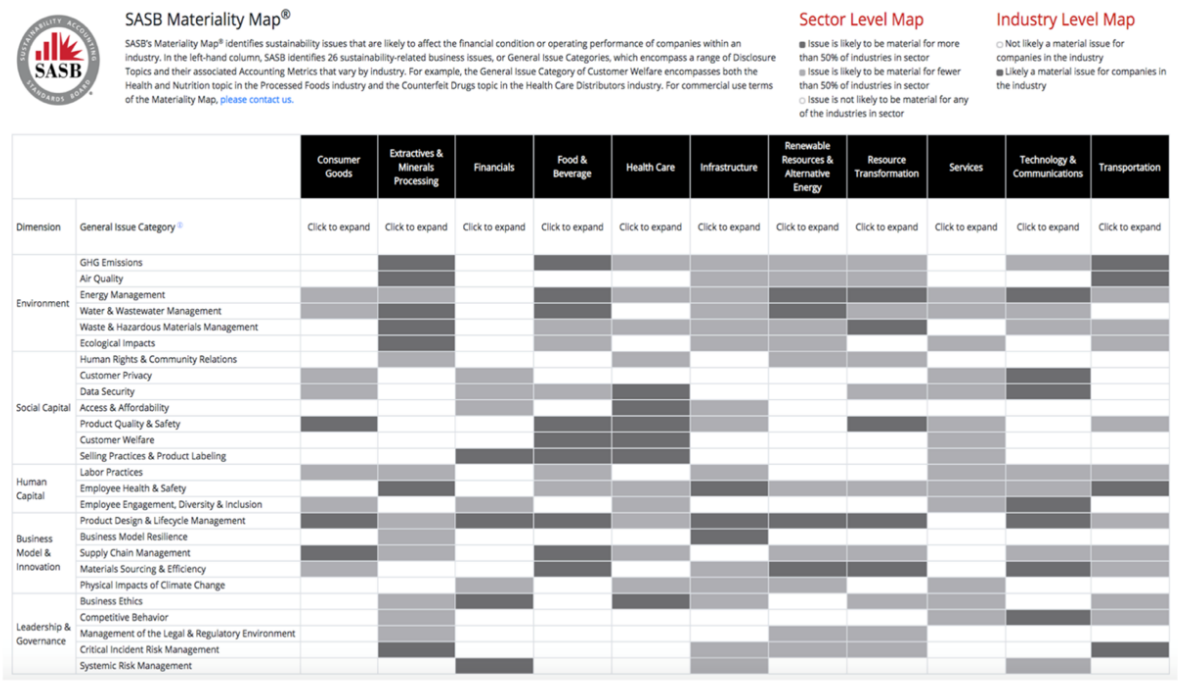


Figure 1.2: SASB Materiality Map, Source: SASB

2 ESG Finance

This chapter is dedicated to the presentation of how sustainability translates in the financial world through the definition of the ESG paradigm. After having introduced the ESG finance in the first section, the following ones describe, on the one hand, the ESG investing phenomena and, on the other, how ESG rating providers adopt their proprietary methodologies in order to assess the sustainable footprint of the corporations.

2.1. ESG Paradigm

In the financial world, sustainable development translates into the term ESG, which was coined in 2004 with the publication of the report "*Who Cares Wins*" by the UN Global Compact Initiative [4]. The term ESG is the acronym for Environmental, Social and Governance, which are the three broad areas containing the criteria utilized in order to assess the sustainable profile of a corporation. Furthermore, these three constituents represent the areas of interest for the so called socially responsible investors, who are the financial market agents that include the sustainability topics inside their investment decision-making process. The next paragraphs are dedicated to the description of each one of the ESG pillars.

2.1.1. Environmental Pillar

The environmental pillar evaluates the sustainability of those companies' activities carrying a direct and indirect impact on the surrounding environment. In particular, as shown by the Corporate Finance Institute (CFI), it includes criteria as the usage of renewable energy sources by the companies, their efforts in reducing the environmental footprints of their operations through waste management programs, and their ability to handle and react to potential issues such as air or water pollution, deforestation and climate change problems.

Other relevant themes related to this pillar could be the ones concerning the biodiversity practices adopted by the corporations on the land in which they run their processes, and also the way in which they conduct the raw material supply.

Enlarging the scope, the table below shows the criteria listed by the European Banking Authority (EBA) in its report [5] in order to assess the sustainable profile of a firm from the environment perspective.

Table 2.1: EBA environmental evaluation criteria. Source: EBA

Environmental Pillar	▪ GHG emissions ▪ Energy consumption and efficiency ▪ Exposure to fossil fuels ▪ Water, air, soil pollutants ▪ Water usage, recycling and management ▪ Land degradation, desertification, soil sealing ▪ Waste production and management (hazardous, nonrecycled) ▪ Raw materials consumption ▪ Biodiversity and protection of healthy ecosystems ▪ Deforestation

2.1.2. Social Pillar

The social pillar evaluates the sustainability of the corporations by looking at the way they are able to manage the social relationships and interactions inside and outside their boundaries. According to the CFI, some of the issues that have to be taken into account inside the social area are the following ones: the fairness level of the employees' salaries, the benefits that are guaranteed to them, the enforcement of policies related to inclusion, diversity and prevention of sexual discriminations, the presence of employee training and education programs and, eventually, the level of human capital turnover, More in general, this pillar includes all the aspects related to gender policies, protection of human rights, labour standards, workplace and product safety, public health and income distribution.

Even in this case the EBA provides a list of criteria [5] that can be used in order to evaluate the sustainable profile of a firm from the social perspective.

Table 2.2: EBA social evaluation criteria. Source: EBA

Social Pillar	▪ Implementation of fundamental ILO Conventions ▪ Violation of UN Global Compact Principles ▪ Inclusiveness/Inequality ▪ Exposure to controversial weapons ▪ Discrimination ▪ Insufficient whistleblower protection ▪ Rate of accidents and number of days lost to injuries, accidents, fatalities or illness ▪ Human rights policy ▪ Investment in human capital and communities ▪ Trafficking in human beings

2.1.3. Governance Pillar

The governance pillar is related to the way in which a firm is managed by its top management offices, and it can be seen as a measure of the alignment between the interests of the executive management and the ones of the company's stakeholders. In general, this pillar embraces aspects such as the independence of the board of administration, shareholders' rights, managers' remuneration, control procedures and anti-competitive practices, as well as the respect of the law. Moreover, the CFI enlarges the scope by including inside the governance themes like the financial and accounting reporting transparency, the avoidance of conflicts of interest from the board members and, eventually, the reduction of conspicuous bonuses for executives while the wages of all the other company's employees are frozen.

The table below shows the criteria provided by the EBA [5] in order to assess the sustainable profile of a firm from the corporate governance perspective.

Table 2.3: EBA governance evaluation criteria. Source: EBA

Governance Pillar	▪ Codes of conduct and business principles ▪ Accountability ▪ Transparency and disclosure ▪ Executive pay ▪ Board diversity and structure ▪ Bribery and corruption ▪ Stakeholder engagement ▪ Shareholder rights

2.2. ESG Investing

The increasing consciousness towards sustainability described in the section [1.1] has also affected the investing behaviours of retail and institutional investors, as shown by the 2020 Global Sustainable Investment Review [6]: this report maps the state of sustainable and responsible investment of major financial markets globally. In the Global Sustainable Investment Review, "Sustainable investment" is defined as investment approaches that consider environmental, social and governance (ESG) factors in portfolio selection and management, and it reveals a growth in sustainable investing of 55% moving from 2016 to 2020, reaching a total amount of investments of USD 35.3 trillion.

Investors integrate ESG factors into their financial decisions for several reasons. Among them, the most relevant are listed below:

- ESG data can help to have a broader and more comprehensive picture of a company's operating environment. Consequently, relying on ESG investing it is possible to identify and manage risks and opportunities which cannot easily be detected through standard financial analysis. ESG can be considered as a proxy for risk that is not priced in, and companies that better manage these risks can deliver returns with greater certainty. For example, reducing exposure to polluters can help mitigate regulatory risk.
- According to part of the literature – which is examined in depth in the dedicated section [3.1] - ESG investing could enhance the performances of the sustainability-driven portfolios, arriving to provide higher returns compared to the traditional ones.

- Other investors opt for ESG investing to meet their values (e.g., ethical, religious, political, cultural) or to promote specific environmental, social, or governance outcomes they would like to achieve. Investors, for instance, may integrate ESG factors into their financial decisions in order to identify and exclude companies adopting immoral practices.
- Institutional investors or financial advisors acting on behalf of a third party may rely on ESG criteria to satisfy specific legal requirements.

2.2.1. Strategies of Sustainable Investing

According to the Global Sustainable Investment Alliance (GSIA) [6], several strategies can be adopted by investors when approaching ESG investments, among which the main ones are the following:

- *Negative/exclusionary screening*. Exclusion of specific unacceptable, controversial and immoral sectors or companies whose activities may have a negative impact on the environment or on the society.
- *Positive/best-in-class screening*. Selection of the best ESG performing companies within a specific business sector and exclusion of companies not meeting certain performance thresholds.
- *Sustainability themed investing*. Targeted investments, including only activities related to the chosen theme (for example, clean energy, pollution, reduction, low carbon emissions, water resources management, sustainable agricultural activities).
- *Impact/community investing*. Private investments dedicated to specific projects solving social and environmental issues such as renewable energy use or social housing investments.

- *ESG integration*. Systematic and explicit inclusion of ESG factors into financial analysis.
- *Corporate engagement and stock activism*. Exercise of the shareholders' rights aiming to influence corporate behaviour through direct dialogue with corporate management and proposal submissions.
- *Norm-based screening*. Investing only in stocks respecting minimum thresholds of ethical business practices.

2.2.2. Sustainable Investment Overview

After having introduced the ESG finance world, the motivations that lead financial investors to prefer sustainable investing and the strategies through which they can implement it, this section is aimed at showing the economic dimension of this trend. According to the 2020 Global Sustainable Investment Review [6], sustainable investment assets under management are continuing to increase worldwide except for Europe, which is experiencing a decline due to significant changes in the way sustainable investment is defined under EU legislation.

Table 2.4: Snapshot of global sustainable investing assets, 2016-2018-2020 (USD billions). Source: GSIA (2020) [6]

REGION	2016	2018	2020
Europe*	12,040	14,075	12,017
United States	8,723	11,995	17,081
Canada	1,086	1,699	2,423
Australasia*	516	734	906
Japan	474	2,180	2,874
Total (USD billions)	**22,839**	**30,683**	**35,301**

The largest increase (48% of sustainable assets growth) over the past two years has been observed in Canada, followed by The United States (42%) and Japan (34%). Europe reported a 13% decline in the growth of sustainable investment assets in the period from 2018 to 2020 due to a new set of revised definitions of sustainable investment that have become embedded into legislation in the European Union as part of the *European Sustainable Finance Action Plan*.

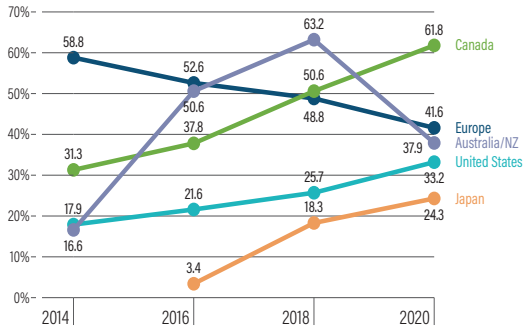


Figure 2.1: Proportion of sustainable investing assets relative to total managed assets 2014-2020. Source: GSIA (2020) [6]

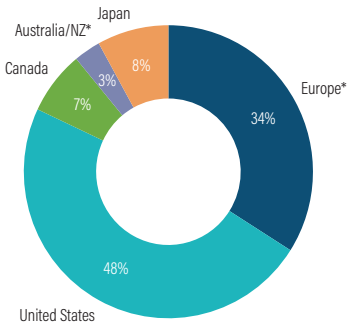


Figure 2.2: Proportion of global sustainable investing assets by region 2020. Source: GSIA (2020) [6]

Figure 2.1 shows the proportion of sustainable investing relative to total managed assets; while Canada, the United States and Japan experienced a continuous growth in this ratio, Australasia and Europe both reported a lower proportion of sustainable investing assets relative to total managed assets for the period 2018-2020.

Looking at the geographical distribution of the phenomena shown by Figure 2.2, more than 80% of global sustainable investing assets during 2018 to 2020 are represented by the two most important global financial hubs (United States and Europe), followed by Japan (8%), Canada (7%) and Australasia (3%).

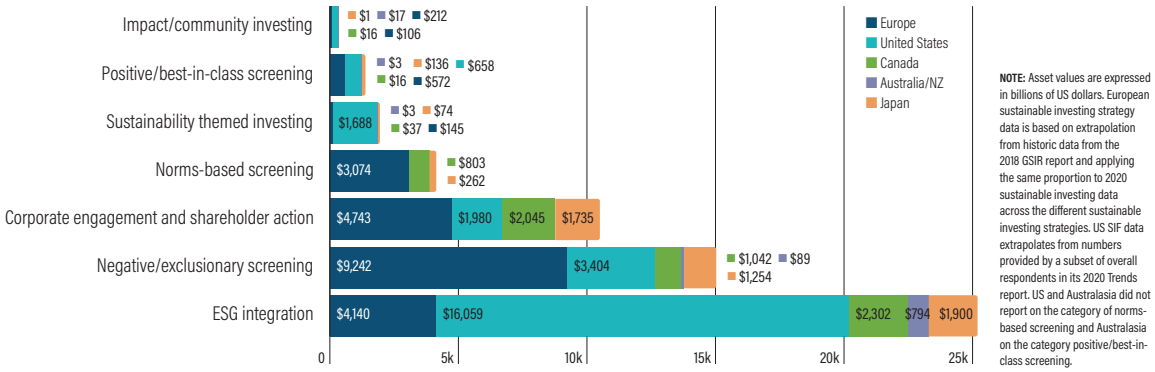


Figure 2.3: Sustainable investing assets by strategy & region 2020. Source: GSIA (2020) [6]

After having shown the relevance assumed by ESG investments also in quantitative terms, it could be interesting to provide some numbers about how the sustainable assets under management are partitioned among the different investment strategies.

Table 2.5: Sustainable investing assets by Strategy, 2016-2018-2020. Source: GSIA (2020) [6]

	2020	2018	2016	GROWTH 2016-2020	COMPOUND ANNUAL GROWTH RATE
Impact/community investing	$352	$444	$248	**42%**	**9%**
Positive/best-in-class screening	$1,384	$1,842	$818	**69%**	**14%**
Sustainability-themed investing	$1,948	$1,018	$276	**605%**	**63%**
Norms-based screening	$4,140	$4,679	$6,195	**-33%**	**-10%**
Corporate engagement and shareholder action	$10,504	$9,835	$8,385	**25%**	**6%**
Negative/exclusionary screening	$15,030	$19,771	$15,064	**0%**	**0%**
ESG integration	$25,195	$17,544	$10,353	**143%**	**25%**

The most adopted sustainable investment strategy globally is ESG integration, as shown in Figure 2.3, with USD 25.2 trillion in assets under management employing an ESG integration approach. This sustainable investment strategy is followed by negative/exclusionary screening (USD 15.9 trillion) and by corporate engagement/shareholder action (USD 10.5 trillion).

This stylized fact shows a clear change in the preferred strategy that, starting from 2018, shifted from negative/exclusionary screening to the ESG integration. Furthermore, an increasing number of reports shows that many investment organizations are adopting a combination of strategies, rather than solely relying on just one.

2.3. ESG Ratings

In the recent years, the integration of sustainability-related topics in the investment strategies has gained more relevance and has become a general preoccupation rather than a niche investment practice. At the basis of this kind of trend, there is the necessity to have a set of objective metrics able to properly provide a comprehensive picture of the sustainability effort undertaken by the target companies belonging to the investment universe. In response to this rising demand for reliable ESG data and grades, the sustainability rating market has grown noticeably and is becoming mature, with lot of agencies providing their sustainability evaluations. In this regard, the ESG risk ratings assess the extent to which a company's economic value is at risk driven by ESG factors or, in other words, the magnitude of a company's non-properly managed ESG risks. However, differently from credit ratings, ESG measurement and disclosure is somehow nebulous and more flexible given the lack of a common regulation, reporting standards and shared characteristics among each ESG component and across rating providers. For this reason, rating agencies are proposing several different metrics derived from alternative and competing definitions, making the sustainability performances of a company very difficult to assess.

This section is aimed at providing an overall picture of the methodologies adopted by the main rating agencies that give to the financial investors the relevant pieces of information necessary to build balanced and performing ESG portfolios.

Among all the players belonging to the ESG rating providers' arena, the present work shows the methodologies of the most relevant ones, namely: MSCI, Refinitiv, Sustainalytics, Vigeo Eiris and Standard & Poor.

2.3.1. MSCI

MSCI ESG Research, launched in 2010, is one of the largest third-party providers of ESG ratings. The data utilised by this company in order to draft its reports are mainly represented by:

- Company disclosures (i.e., sustainability reports)
- Macro data from academic and NGOs datasets
- Governments databases.

MSCI ESG Ratings aim at helping investors to understand relevant ESG risks and opportunities that could be useful to consider in their investment processes. Going more in dept, these evaluations seek to assess four key aspects about companies:

- the most relevant material ESG risks and opportunities facing a company and its industry.
- the exposure of a company to those key risks and/or opportunities.
- the quality and effectiveness of the company's risks/opportunities management system.
- the performance of the company relative to the ones of its global industry peers.

The MSCI ratings quantitative model [7] focuses on specific industries/sectors, since corporations in the same industry are likely to face the same major risks and opportunities. Furthermore, MSCI considers only those ESG risks and opportunities which are considered as material, which means that they are the factors likely to affect the ability of a company to create long-term value. In particular:

- A risk is material when companies in a specific industry are likely to incur substantial costs in connection with it.

- An opportunity is material when it is likely that companies in a specific industry could capitalize on it in order to make profit.

MSCI identifies which are the material risks and opportunities for each industry through a quantitative model [7]. The MSCI methodology is constituted by 35 ESG key issues, which are clustered in 10 themes belonging to the three pillars of sustainable thinking, namely environmental, social, and governance.

Table 2.6: MSCI methodology themes and key issues. Source: MSCI

3 Pillars	10 Themes	35 ESG Key Issues	
Environment	**Climate Change**	Carbon Emissions Product Carbon Footprint	Financing Environmental Impact Climate Change Vulnerability
	Natural Capital	Water Stress Biodiversity & Land Use	Raw Material Sourcing
	Pollution & Waste	Toxic Emissions & Waste Packaging Material & Waste	Electronic Waste
	Environmental Opportunities	Opportunities in Clean Tech Opportunities in Green Building	Opportunities in Renewable Energy
Social	**Human Capital**	Labor Management Health & Safety	Human Capital Development Supply Chain Labor Standards
	Product Liability	Product Safety & Quality Chemical Safety Financial Product Safety	Privacy & Data Security Responsible Investment Health & Demographic Risk
	Stakeholder Opposition	Controversial Sourcing Community Relations	
	Social Opportunities	Access to Communications Access to Finance	Access to Health Care Opportunities in Nutrition & Health
Governance*	**Corporate Governance**	Ownership & Control Board	Pay Accounting
	Corporate Behavior	Business Ethics Tax Transparency	

After having assigned a score to the key issues, normalizing them to the industry peers, their weighted average is computed in order to arrive to the final score of the rating. The next step consists in transforming the final score into the correspondent rating category, which could range from AAA (best-in-class) to CCC (laggard). The final evaluation is not absolute, but it is relative to the peers of a company's industry.

2.3.1.1. Evaluation Methodology

Going more in depth with the evaluation methodology [7], the first step consists in the identification of the key issues for each specific industry. Then, there is the setting of the weights, which reflect the contribution of each key issue to the final rating. Usually, the key issues' weights vary from 5% to 30% depending on the contribution of the industry to the negative or positive impact on environment and society, and depending on the expected time frame for the risk/opportunity to materialise in the industry.

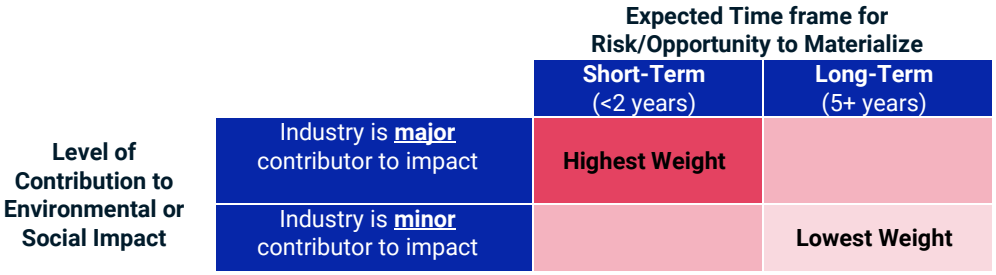


Figure 2.4: Framework for setting key issue weights. Source: MSCI

A key issue belonging to the "High Impact" and "Short-Term" quadrant is assigned a weight three times higher than a key issue categorised as "Low Impact" and "Long-Term".

At this point, there is the assessment of the key issues, which can be either risks or opportunities:

- *Risk assessment:* on the one hand, the MSCI rating evaluates the exposure of a company to a specific issue while, on the other hand, it assesses the effectiveness of the company's management strategy to manage this risk. Thanks to this process, two metrics are computed, namely risk exposure and risk management. To achieve a positive score on a specific key issue, the firm's management strategy needs to be commensurate with the level of exposure: for example, a company with high exposure must also have very strong management, while a modest approach could be sufficient for a corporation with limited exposure. Going more in depth, the risk exposure is assessed looking at the company's core products, its business segments and the location of its facilities. Risk exposure is scored on a 0-10 scale, where 0 represents no exposure and 10 represents very high exposure. On the other hand, risk management is assessed by looking at the strategies developed by the company to handle all the emerging criticalities. Management is scored on a 0-10 scale, where 0 represents no evidence of management efforts and 10 represents indications of very strong management and commitment. The risk exposure score and the risk management score are then combined to achieve the key issue scores, which are also on a 0-10 scale, where 0 is very poor and 10 is very good.

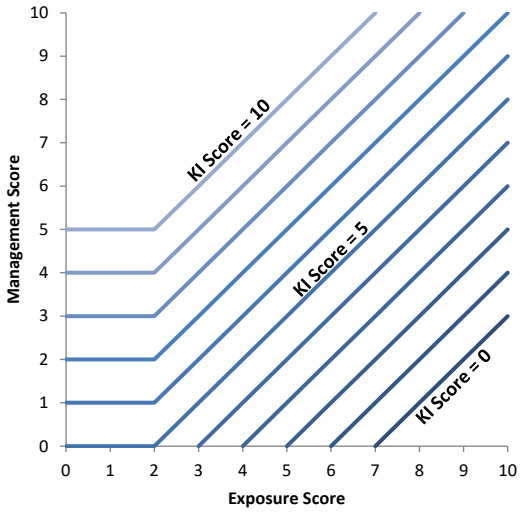


Figure 2.5: Combining exposure and management, risk key issues. Source: MSCI

- *Opportunities assessment:* the assessment of the opportunities works similarly to risks, but the model for combining exposure and management differs.

Exposure reflects the relevance of the opportunity to a given company according to its business and geographic segments. Management indicates the company's ability to exploit the opportunity and create value out of it.

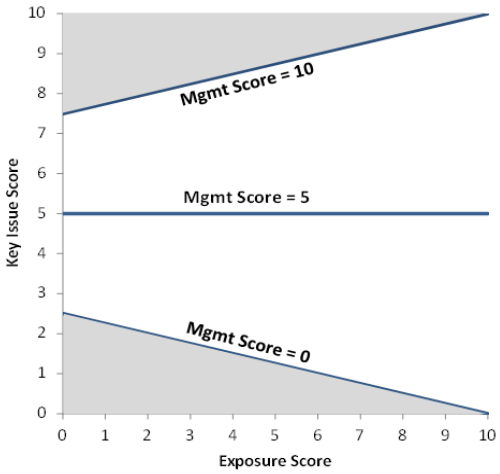


Figure 2.6:Combining exposure and management, opportunities key issues. Source: MSCI

After the assessments described above, MSCI determines the governance score. The governance pillar score is an absolute assessment of a company's governance, and it is done using a 0-10 scale. Each company starts by default with the maximum score, and then it is progressively decreased based on the assessment of the key metrics. The detailed process is shown by the image below.

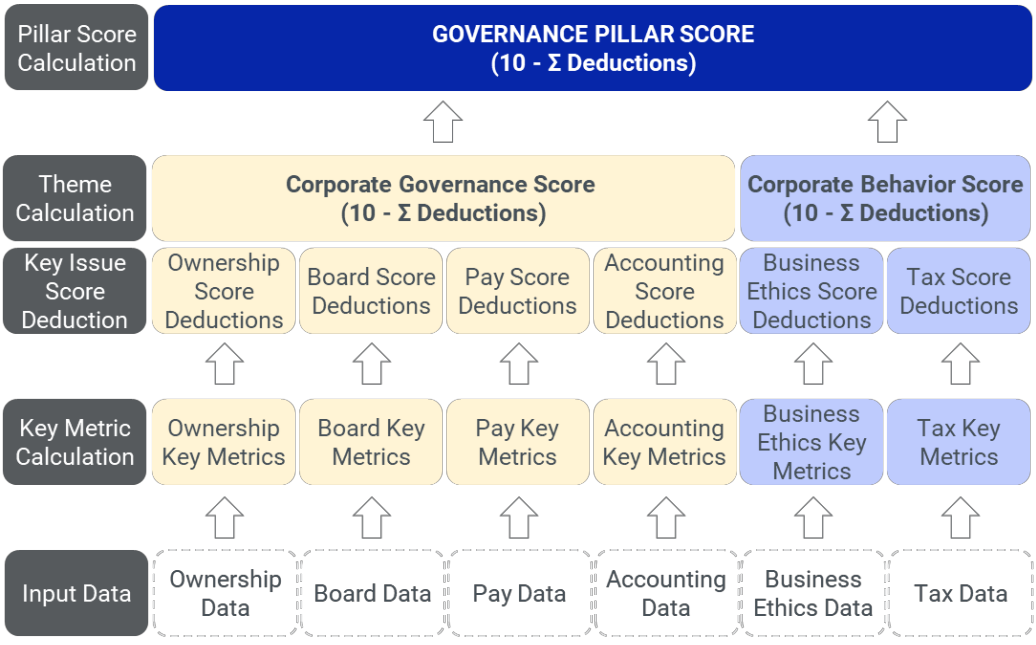


Figure 2.7: Governance pillar score computational process. Source: MSCI

2.3.1.2. Score Evaluation Process

The ESG Ratings model [7] is industry relative and uses a weighted average approach. In particular, key issue weights are assigned based on each industry's relative external impact and the time horizon associated with each risk. Corporate governance is always considered as material and therefore it is always weighted and analysed for all the firms. Finally, each company is assigned a weighted average key issue score according to the underlying key issue scores and weights, as shown by the Figure 2.8.

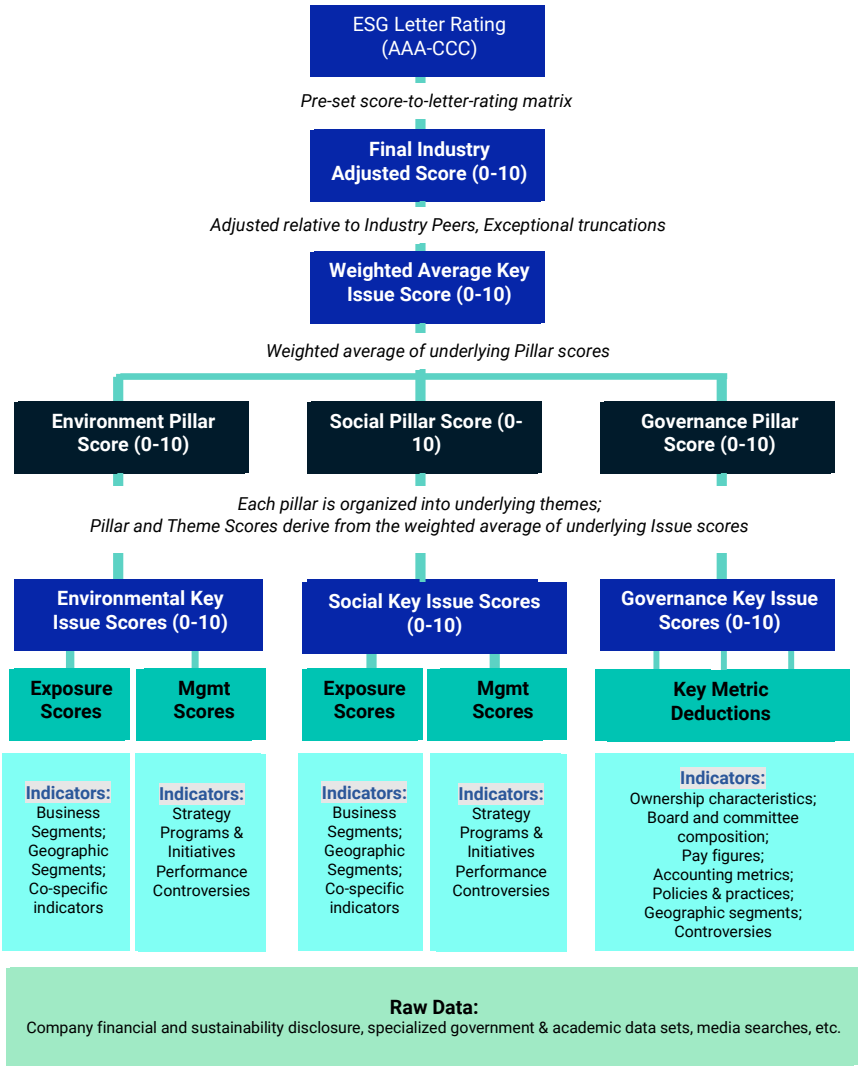


Figure 2.8: MSCI hierarchy of ESG scores. Source: MSCI

The last step of the process consists in the transformation of the weighted average key issue score into the final letter rating. To do so, the following rule is applied:

- The top benchmark value (industry maximum score, AAA) falls between the 95th and 100th percentile of modelled weighted average key issue scores within an ESG Rating Industry.
- The bottom benchmark value (industry minimum score, CCC) falls between the 0th and 5th percentile of modelled weighted average key issue scores within an ESG Rating Industry.

Table 2.7: MSCI letter rating transformation table. Source: MSCI

Letter Rating	Leader/Laggard	Final Industry-Adjusted Company Score
AAA	Leader	8.571* - 10.0
AA	Leader	7.143 – 8.571
A	Average	5.714 – 7.143
BBB	Average	4.286 – 5.714
BB	Average	2.857 – 4.286
B	Laggard	1.429 – 2.857
CCC	Laggard	0.0 – 1.429

**Appearance of overlap in the score ranges is due to rounding error. The 0 to 10 scale is divided into 7 equal parts, each corresponding to a letter rating.*

2.3.2. Refinitiv

The collection of information is at the basis of Refinitiv's assessment methodology [8], which is a transparent and data-driven process aimed at evaluating the companies' ESG performance. This rating agency relies on the expertise of more than 150 content research analysts, which are endowed with the responsibility of collecting ESG data. In particular, they process several publicly available information sources with the aim of providing up-to-date, objective and comprehensive coverage. The pieces of information collected are utilised to feed a set of more than 500 ESG measures that are at the basis of the Refinitiv's methodology. Each measure requires a specific process aimed at standardizing the information, ensuring its comparability across the entire range of firms.

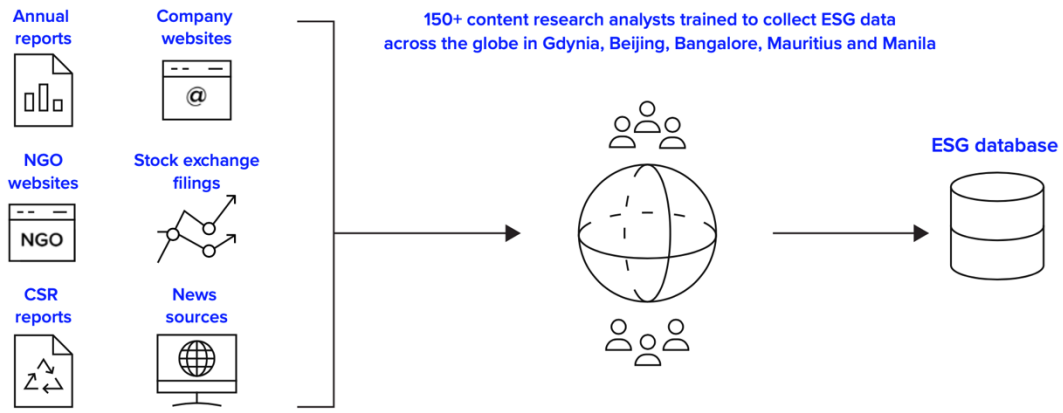


Figure 2.9: Refinitiv data sources. Source: Refinitiv

Going more in depth, the methodology adopted by Refinitiv leads to the computation of a comprehensive ESGC score, which adjusts the ESG score taking into account news controversies that materially impact the firms.

The model is made up by two overall ESG scores:

- ESG score, which measures the company's ESG performance according to data available in the public domain.
- ESGC score, which integrates the ESG score with ESG controversies in order to provide a comprehensive assessment of the company's sustainability impact and management over the years.

2.3.2.1. Evaluation Methodology

Refinitiv's assessment and scoring process [8] is based on a pool of more than 500 company-level ESG measures grouped into 10 categories, correlated to the three pillars of sustainability (i.e., environment, social and governance). Going more in depth, the category scores are aggregated into three pillar scores, namely environmental, social and corporate governance. The ESG pillar score is obtained through the relative sum of the category weights, which differ according to the specific sector for the environmental and social categories. For what concern the governance pillar, the weights remain unchanged across all industries. The pillar weights are normalized to percentages ranging between 0 and 100.

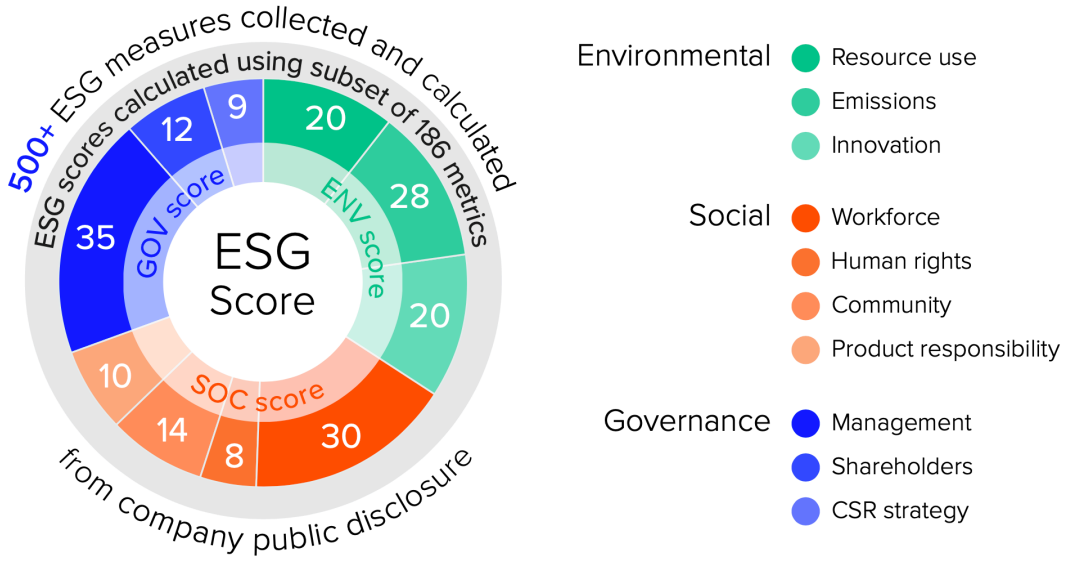


Figure 2.10: ESG score components. Source: Refinitiv.

The process does not stop to the calculation of the ESG score, but it goes on with the computation of the ESGC score. The ESGC score complements the previously computed ESG score by taking into account the reported information related to the ESG pillars regarding the ESG controversies captured from global media sources. This provides a more comprehensive scoring and picture of a company's ESG performance by discounting the ESG performance score based on negative media stories and incorporating the impact of material ESG controversies affecting a company. When a company is involved in ESG controversies, the ESGC score is computed by calculating the weighted average of the ESG score and ESG controversies score per fiscal period. When a firm is not involved in any ESG controversy, its ESGC score equals the ESG score. The ESG scores are returned in letter grades in order to show how companies are performing relative to their peers, and which are the company's areas presenting weaknesses and/or strengths.

Table 2.8: Refinitiv letter rating transformation table. Source: Refinitiv

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.
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	
0.916666 < score <= 1	A +	

ESG laggards

ESG leaders

2.3.2.2. Score Evaluation Process

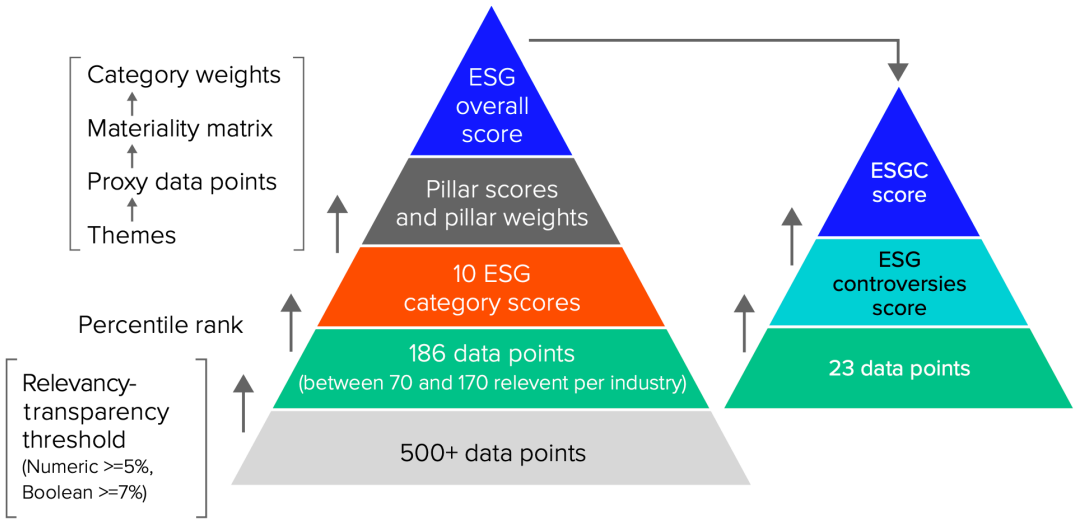


Figure 2.11:ESG score evaluation methodology. Source: Refinitiv

As shown by the Figure 2.11, the Refinitiv ESG scoring methodology [8] consists of a five-step process:

- *Step 1: ESG category scores.* In order to compute the 10 category scores and the ESG controversies score, it is adopted the percentile rank scoring methodology, which is based on the absolute value of the company and on its relative value compared with the industry peers.
- *Step 2: Materiality matrix.* Refinitiv defines materiality in the form of category weights, which are calculated based on an objective and data-driven approach aimed at determining the relative importance of each theme to each individual industry. The table below shows a detailed view on the ESG themes present in each category.

Table 2.9: Refinitiv ESG themes. Source: Refinitiv

Pillars	Catagories	Themes	Data points	Weight method
Environmental	Emmission	Emissions	TR.AnalyticCO2	Quant industry median
		Waste	TR.AnalyticTotalWaste	Quant industry median
		Biodiversity *		
		Environmental management systems *		
	Innovation	Product innovation	TR.EnvProducts	Transparency weights
		Green revenues, research and development (R&D), and capital expenditures (CapEx)	TR.AnalyticEnvRD	Quant industry median
	Resource use	Water	TR.AnalyticWaterUse	Quant industry median
		Energy	TR.AnalyticEnergyUse	Quant industry median
		Sustainable packaging *		
		Environmental supply chain *		
Governance	CSR strategy	CSR strategy	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		ESG reporting and transparency		
	Management	Structure (independence, diversity, committees)	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Compensation		
	Shareholders	Shareholder rights	Data points in governance category and governance pillar	Count of data points in each governance category/all data points in governance pillar
		Takeover defenses		
Social	Community	Equally important to all industry groups, hence a median weight of five is assigned to all		Equally important to all industry groups
	Human rights	Human rights	TR.PolicyHumanRights	Transparency weights
	Product responsibility	Responsible marketing	TR.PolicyResponsibleMarketing	Transparency weights
		Product quality	TR.ProductQualityMonitoring	Transparency weights
		Data privacy	TR.PolicyDataPrivacy	Transparency weights
	Workforce	Diversity and inclusion	TR.WomenEmployees	Quant industry median
		Career development and training	TR.AvgTrainingHours	Transparency weights
		Working conditions	TR.TradeUnionRep	Quant industry median
		Health and safety	TR.AnalyticLostDays	Transparency weights

* No data points available that may be used as a proxy for ESG magnitude/materiality

The magnitude matrix is calculated adopting the following two methods: the industry median, which is the relative weight determined by computing the relative median value with respect to the industry group, and the transparency weights, which are the relative weights determined based on the disclosure level relative to industry group. After that, the magnitude weights of all the 10 categories are summed up for each industry group. Each category's magnitude weight is divided by the sum of the magnitude weights of the respective industry group in order to get to the category weight.

$$Category\ weight\ of\ an\ industry\ group = \frac{Magnitude\ weight\ of\ a\ category}{Sum\ of\ magnitudes\ of\ all\ categories} \quad (2.1)$$

- *Step 3: Overall ESG score calculation and pillar score. ESG scores* are calculated by aggregating the 10 category weights, which are computed using the Refinitiv magnitude matrix.

Table 2.10: Refinitiv category weights. Source: Refinitiv

	Environmental			Social				Governance			
Industry group	Emission	Innovation	Resource use	Human rights	Product responsibility	Workforce	Community	Management	Shareholders	CSR strategy	ESG scores
Water and related utilities	0.15	0.13	0.15	0.05	0.04	0.13	0.09	0.17	0.05	0.03	
ABC	0.66	0.00	0.44	0.05	0.58	0.89	0.34	0.99	0.84	0.56	0.571146184
CBD	0.71	0.96	0.38	0.00	0.69	0.66	0.70	0.37	0.01	0.56	0.547913483
DEF	0.03	0.00	0.00	0.00	0.00	0.57	0.11	0.21	0.14	0.54	0.150536652
EFG	0.00	0.31	0.03	0.00	0.00	0.25	0.59	0.89	0.94	0.00	0.327824384
EMJ	0.87	0.31	0.68	0.20	0.86	0.84	0.98	0.33	0.87	0.68	0.639400132
EMQ	0.00	0.00	0.00	0.00	0.00	0.30	0.02	0.88	0.08	0.01	0.194782046
ENR	0.92	0.81	0.85	0.75	0.97	0.93	0.66	0.40	0.49	0.86	0.756319427
GPQ	0.24	0.31	0.00	0.00	0.17	0.02	0.16	0.56	0.56	0.00	0.223443757
HIJ	0.61	0.31	0.50	0.65	0.42	0.80	0.80	0.48	0.27	0.37	0.54145808
IBD	0.00	0.00	0.00	0.00	0.00	0.07	0.30	0.51	0.49	0.00	0.145398367
JKL	0.50	0.73	0.74	0.00	0.78	0.43	0.93	0.62	0.89	0.26	0.611504799
LMN	0.76	0.31	0.56	0.00	0.47	0.48	0.48	0.17	0.24	0.26	0.415151441
MNO	0.82	0.31	0.91	0.40	0.58	0.61	0.07	0.33	0.52	0.63	0.539888776
MSE	0.55	0.00	0.62	0.85	0.17	0.75	0.84	0.77	0.35	0.91	0.581805891
OPQ	0.29	0.00	0.32	0.00	0.17	0.16	0.48	0.15	0.42	0.08	0.212906948
PQR	0.45	0.65	0.79	0.55	0.78	0.52	0.75	0.76	0.76	0.16	0.640379494
PSF	0.97	0.88	0.97	0.95	0.92	0.98	0.89	0.15	0.73	0.34	0.776142465
RST	0.08	0.31	0.00	0.00	0.17	0.20	0.59	0.42	0.42	0.00	0.228111754
UVW	0.34	0.00	0.26	0.20	0.58	0.70	0.39	0.26	0.16	0.31	0.316400123
VPF	0.16	0.31	0.15	0.00	0.17	0.11	0.25	0.88	0.90	0.00	0.325828115
XYZ	0.39	0.00	0.21	0.40	0.17	0.39	0.48	0.95	0.73	0.51	0.429105164
YQM	0.16	0.00	0.09	0.00	0.36	0.34	0.20	0.69	0.34	0.00	0.25005416

ESG pillar scores are represented by the relative sum of the category weights. The table below shows the calculations needed in order to derive the pillar scores.

Table 2.11: Pillars' score computation process. Source: Refinitiv

Pillar	Category	Category scores*	Category weights	Sum of category weights	Formula: sum of category weights	New category weights*	Formula: new category weights	Pillar scores	Formula: pillar scores
Environmental	Emissions	0.98	0.15			0.35	(0.15/0.44)		
Environmental	Resource use	0.97	0.15	0.44	(0.15+0.15+0.13)	0.35	(0.15/0.44)	0.94	(0.98*0.35)+ (0.97*0.35)+ (0.85*0.29)
Environmental	Innovation	0.85	0.13			0.29	(0.13/0.44)		
Social	Community	0.89	0.09			0.28	(0.09/0.31)		
Social	Human rights	0.95	0.05		(0.09+0.05+0.04+0.13)	0.17	(0.05/0.31)	094	(0.89*0.28)+ (0.95*0.17)+ (0.92*0.13)+ (0.98*0.43)
Social	Product responsibility	0.92	0.04	0.31		0.13	(0.04/0.31)		
Social	Workforce	0.98	0.13			0.43	(0.13/0.31)		
Corporate governance	Shareholders	0.73	0.05			0.20	(0.05/0.26)		
Corporate governance	CSR strategy	0.34	0.03	0.26	(0.05+0.03+0.17)	0.13	(0.03/0.26)	0.32	(0.73*0.20)+ (0.34* 0.13)+ (0.19*0.67)
Corporate governance	Management	0.19	0.17			0.67	(0.17/0.26)		

*decimal places to be considered

- *Step 4: Controversies scores calculation. ESG* controversies score is computed based on 23 ESG controversy topics. The default value assigned to each controversy is set equal to 0. Companies that do not present any of these issues get the maximum score of 100. In the computation of the controversy score, the severity weights are used in order to consider the market cap bias that is suffered by large-cap companies, since they are more likely to attract media attention than smaller firms.
- *Step 5: ESGC score.* At the end of this process, the ESGC score is computed by calculating the average of the ESG score and ESG controversies score when there are controversies during the fiscal year. In the case in which the controversies score is greater than the ESG score, then the ESG score is equal to the ESGC score.

Table 2.12: ESGC score computation process. Source: Refinitiv

Scenario	ESG controversies score	ESG score	ESGC score
If controversies score is >=ESG score, then ESG score = ESGC score	100	89	89
If controversies scores is <ESG score, then ESGC score = average of ESG and controversies score	48	49	48.5

2.3.3. Sustainalytics

Sustainalytics is a leading ESG and corporate governance research, ratings and analytics company that helps global investors in incorporating ESG information into their investment processes, guiding them in the building and management of their investment portfolios and strategies. At the basis of the Sustainalytics' methodology [9], there is the concept of materiality, as in the case of MSCI.

The ESG Risk Rating developed by this provider is made up by three building blocks: corporate governance, material ESG issues (MEIs), and idiosyncratic ESG issues.

- *Corporate governance:* it is a fundamental building block of the ESG Risk Rating, since poor corporate governance may generate material risks for companies. On average, unmanaged corporate governance risk constitutes around the 20% of the overall unmanaged risk score of a firm.
- *Material ESG issues*: they are a set of related topics that require a common set of management initiatives in order to be effectively treated. For example, the

topics of employee recruitment, development, diversity, engagement and labour relations are all encompassed by the material ESG issue of Human Capital because they are all employee related. The evaluation of the material ESG issues occurs at the subindustry level and, at company level, material ESG issues can be eliminated from the rating if they are not relevant to the specific company's strategy. This block is the core of Sustainalytics' methodology; it starts from the assumption that material ESG issues have a predictable impact on the economic value of a corporation belonging to a particular industry. Sustainalytics tries to identify these material ESG issues looking at the business environment in which a company is operating. However, there are issues that may become material in an unpredictable manner; they are called 'Idiosyncratic Issues' and they constitute the third building block.

- *Idiosyncratic Issues*: they are 'unpredictable' or unexpected issues, since they are unrelated to the specific subindustry a company belongs to. It is important to note that idiosyncratic issues turn out to be material issues only for the specific company in question, not for the entire subindustry that the company is part of.

2.3.3.1. Evaluation Methodology

Sustainalytics adopts a two-dimensional architecture with the first dimension, exposure, reflecting the extent to which a company is exposed to material ESG risks, and the second one, management, reflecting how effectively a company is managing its exposure.

This framework starts from the concept of exposure, which is defined as the company's sensitivity or vulnerability to ESG risks. Material ESG issues' exposure scores are evaluated at the subindustry level and then investigated at the company level. The first step consists in the computation of the exposure of companies that operate in the same subindustry with respect to a set of relevant ESG issues. In order to make this assessment, Sustainalytics relies on the pieces of information coming from company reporting, structured external data (e.g., CO_2 emissions) and third-party research. The next step consists in moving from the industry perspective to the single company one, by computing the Beta indicator. The Beta indicator is used in order to make the ESG Risk Rating company specific, since it reflects how much a company's exposure to a material ESG issues deviates from the average exposure of the subindustry's peers.

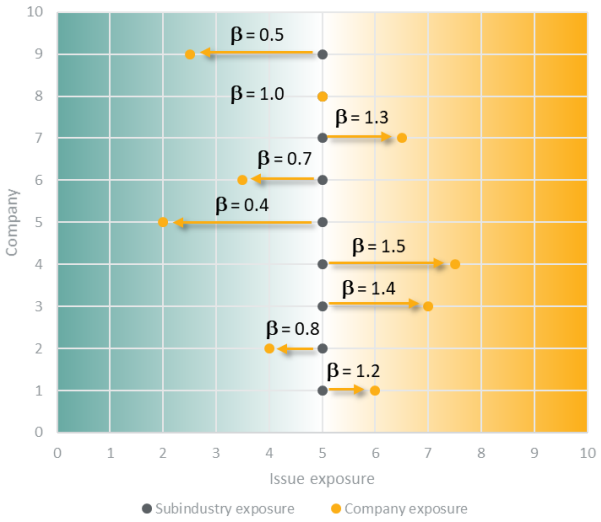


Figure 2.12: Beta indicator determinantion. Source: Sustainalytics

The subindustry exposure score is multiplied by the company's issue Beta in order to obtain the company's exposure score for a particular ESG issue. The Beta for a company with respect to an ESG issue is calculated in a three-stage process. As

starting point, Beta indicators have been created for four thematic areas: product & production, financials, events, and geographic. In a second step, Sustainalytics' analysts refine the indicator by taking into account company specific factors that are not reflected in the standard model. At the end, a technical correction factor is applied to assure that the average Beta within a subindustry is equal to 1.

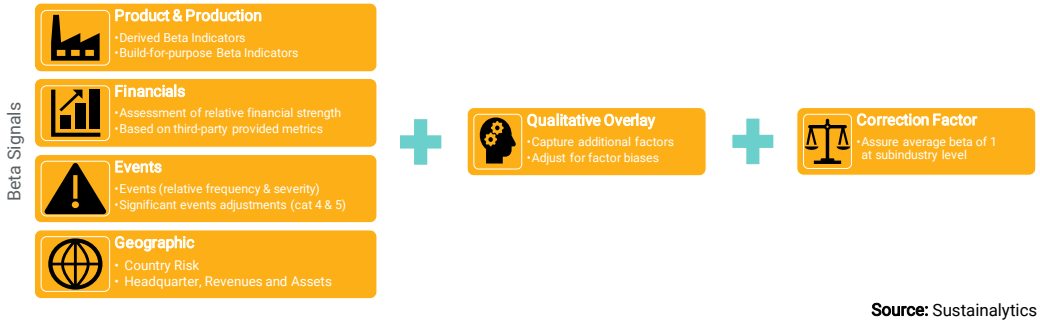


Figure 2.13: Model for calculating issue betas. Source: Sustainalytics

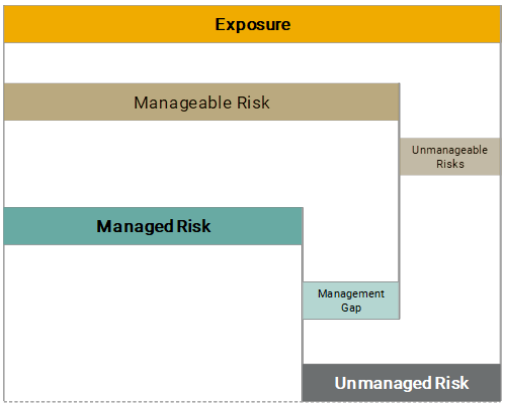
The ESG Risk Rating' second dimension is represented by management. It is defined as the set of company commitments, actions and strategies utilised by the company in order to deal with the risks it is exposed to. The management score of a company is derived from a set of management indicators and outcome-focused indicators. In particular, outcome-focused indicators evaluate management performance either in quantitative terms (e.g., CO_2 emissions) or through a company's involvement in controversies (represented by the company's event indicators). These indicators are then weighted in order to reach the comprehensive description of the management system of the company.

2.3.3.2. Score Evaluation Process

In order to arrive to the ESG Risk Rating scoring system for a company, three stages have to be undergone [9]. The starting point consists in determining the exposure. The next stage is assessing management and the degree to which the risk is managed by the company. The last step is calculating the unmanaged risk.

In particular, unmanaged risk is defined as material ESG risk that has not been managed by a company. It is made up by two different types of risk:

- Unmanageable risk, which cannot be treated with the company's actions.
- The management gap, which reflects the risks that could potentially be managed by a company but that are not sufficiently managed.



Company Exposure	=	Subindustry Exposure	*	Issue Beta		
	=	8	*	1.5	=	**12**
Manageable Risk	=	Company Exposure	*	MRF		
	=	12	*	90%	=	**10.8**
Managed Risk	=	Manageable Risk	*	Management score (as %)		
	=	10.8	*	75%	=	**8.1**
Unmanaged Risk	=	Company Exposure	-	Managed Risk		
	=	12	-	8.1	=	**3.9**

Source: Sustainalytics

Figure 2.14: Score computational process. Source: Sustainalytics

The final ESG Risk Rating score is calculated by summing the individual material ESG issues' unmanaged risk scores, and it represents the overall unmanaged risk of a company. Alternatively, it can be computed by considering the difference between a company's exposure and its managed risk.

2.3.4. Vigeo Eiris

Vigeo Eiris is an international provider of ESG research and services for investors and organisations, founded in France in 2002. In order to generate its ESG scores, Vigeo evaluates 38 distinct ESG criteria that are framed within 40 industry specific models [10]. These criteria are assessed by analysing the data coming from the following sources:

- *Corporate reporting*
- *Direct company contacts*
- *Stakeholders' websites*
- *Factiva press database*

The Figure 2.15 sums up the entire methodology adopted by Vigeo in order to generate its sustainability assessment.

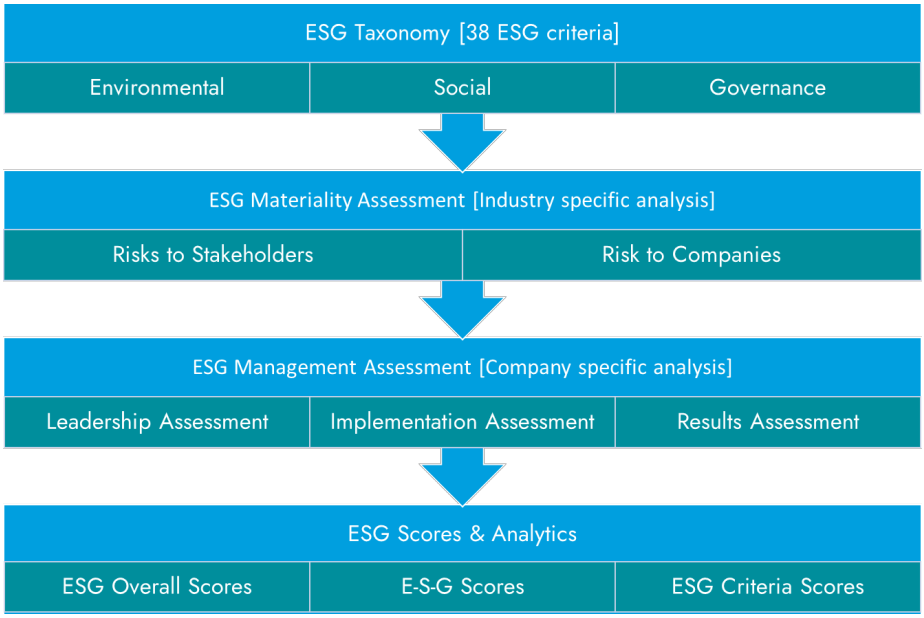


Figure 2.15: Vigeo Eiris scoring framework. Source: Vigeo Eiris

Going more in depth with the details, during the assessment of each corporation Vigeo focuses on the analysis of 38 ESG criteria, which belong to the three pillars of sustainability, namely environment, social and governance. Each one of these ESG criteria is assigned a score from 0-100 according to the quality and effectiveness of the management approaches adopted by the company, reflecting in this way its exposure to different types of risk.

Table 2.13: ESG pillars evaluation criteria. Source: Vigeo Eiris

ESG Pillars and the associated ESG Criteria		
Environment	Social	Governance
Environmental Strategy	Social Dialogue	Anti-Corruption
Accidental Pollution	Employee Participation	Anti-Competition
Green Products	Reorganization	Lobbying
Biodiversity	Career Management	Board of Directors
Animal Testing	Remuneration	Audit & Internal Controls
Water	Health & Safety	Shareholders
Energy	Working Hours	Executive Remuneration
Atmospheric Emissions	Information to Customers	Product Safety (G)
Waste	Customer Relations	
Local Pollution	Suppler Relations	
Transportation	Social Standards on the Supply Chain,	
Use & Disposal of Products	Social & Economic Development	
Environmental Standards in the Supply Chain	Societal Impacts of Products & Services	
	Philanthropy	
	Fundamental Human Rights	
	Fundamental Labour Rights	
	Non-Discrimination	
	Child & Forced Labour	
	Product Safety (S)	

Each ESG criteria includes a set of "Principles of Actions", which represent the expected minimum requirements that a high-performing company should satisfy in each dimension in order to get a positive score.

2.3.4.1. Evaluation Methodology

Vigeo's evaluation methodology [10] is based on two types of assessments:

- *ESG materiality assessment*
- *ESG management assessment*

Starting from the first one, since the challenges affecting companies are not uniform across the different industries, Vigeo customizes its assessment model in order to make it fit with the peculiarities of the specific sectors. In particular, in each industry framework the 38 generic ESG criteria are assigned a weight from w0 (not relevant to the sector) to w3 (highly material to the sector). In order to assign the weights to the criteria, Vigeo evaluates three different aspects: the nature of rights, the company's risks and the stakeholders' risks. On average, a specific sector counts about 25 criteria considered relevant to it, with an industry-specific materiality weight assigned to each criterion.

Table 2.14: ESG weighting criteria. Source: Vigeo Eiris

Weighting Criteria			
Nature of stakeholders' rights, and expectations	Risks to stakeholders	Risks to companies	ESG Criterion Weight
Framed as fundamental stakeholder rights in international reference texts. *For example – Human Rights Labour Rights*	Stakeholders in the sector are highly exposed if companies do not manage their responsibilities. Companies are using high volumes of raw materials or emitting high volumes from an environmental perspective (high environmental footprint)	High risk to a company's reputation, human capital, operational efficiency or exposure to legal risk.	W3 Highly Material
Framed as important in international reference texts *For example - Anti-Competition, Responsible Lobbying*	Stakeholders in the sector are moderately exposed if companies do not manage their responsibilities. Companies are using moderate volumes of raw materials or emitting moderate volumes from an environmental perspective (moderate environmental footprint)	Moderate risk to a company's reputation, human capital, operational efficiency or exposure to legal risk.	W2 Moderately Material
Minor interests and expectations from society *For example -Philanthropy*	Stakeholders are marginally exposed. There is a low environmental footprint.	Low risk to a company's reputation, human capital, operational efficiency or legal risk.	W1/0 Low Materiality

The weights are fundamental to the methodology since the company's ESG overall score is obtained through the weighted average of the scores of all the criteria.

Moving to the ESG management assessment, a three-pillar managerial questioning framework is used to evaluate each ESG criterion. The main areas touched by this framework are listed below:

- *Quality of leadership:*
  - Visibility of the commitments: existence of defined, understandable and accessible policies related to the criterion.

- Exhaustiveness of the commitment: the degree of alignment between the company policies and the principles of action expected for a high performing company.
- Ownership of the commitment: assignment of accountability to a person or a department for the realization of the stated objectives.

- *Extent of implementation:*
- Means allocated: sufficiency of processes and measures put in place to ensure that the organization can achieve its stated objectives.
- Geographical coverage: comprehensive coverage of all business locations.
- Scope: the extent to which processes and measures put in place cover all the relevant principles of action expected for a high performing company.

- *Results (measures of effectiveness):*
- KPI indicators: objective assessment of company's performance relative to its stated objectives and its sector.
- Stakeholder feedback: occurrence of controversies related to the principles of action under review.
- Controversy management: nature of company's response to any allegations (e.g., non-communicative, reactive, proactive).

Table 2.15: ESG criterion score framework. Source: Vigeo Eiris

Criterion Score x/100	Leadership Assessment	- Visibility - Exhaustiveness - Ownership	33% of the criterion score
	Implementation Assessment	- Means - Coverage - Scope	33% of the criterion score
	Results Assessment	- KPIs - Stakeholder Feedback - Controversy Management	33% of the criterion score

2.3.4.2. Score Generation Process

The generation of the ESG scores follows a bottom-up approach made up by three main steps: Criterion Level Scoring, E-S-G Level Scoring and, finally, ESG Overall Score.

- *Criterion Level Scoring*: the criteria level score is the average of the scores of the three pillars illustrated above (Leadership, Implementation, Results). Each of the pillars is assigned a score ranging from 0 to 100; then the average of the pillars' scores is computed in order to arrive to the score of the single criteria. The table below shows an example of this step of the process.

Table 2.16: ESG scoring process. Source: Vigeo Eiris

Criterion Score = 61/100		
Leadership Pillar [72/100] 33% of criterion score	Implementation Pillar [76/100] 33% of criterion score	Results Pillar [34/100] 33% of criterion score
Visibility score (65/100) 20% of pillar score	Means score (65/100) 40% of pillar score	KPI Trends score (0/100) 30% of pillar score
Exhaustiveness score (65/100) 60% of pillar score	Coverage score (65/100) 30% of pillar score	Stakeholder feedback score (30/100) 35% of pillar score
Ownership score (100/100) 20% of pillar score	Scope score (100/100) 30% of pillar score	Controversy management score (65/100) 35% of pillar score

- *E-S-G Level Scoring*: in the second phase, the scores of the environmental, social and governance factors are computed starting from the weighted average of the criteria's scores belonging to each specific pillar.

Table 2.17: E-S-G level scoring. Source: Vigeo Eiris

E SCORE - 58/100	S SCORE - 47/100	G SCORE - 28/100
Environmental Management [50/100] (w3)	Labour Rights [30/100] (w2)	Board [10/100] (w3)
Water [62/100] (w2)	Non-Discrimination [45/100] (w3)	Audit & Internal Controls [10/100] (w3)
Energy [62/100] (w3)	Reorganisations [65/100] (w2)	Shareholders [50/100] (w3)
Environmental Supply Chain [62/100] (w1)	Economic Development [50/100] (w1)	Corruption [75/100] (w2)

- *ESG Overall Score*: at the end of the process, the ESG overall score is the weighted average of the grades assigned to the three pillars of sustainability.

Table 2.18: ESG overall score. Source: Vigeo Eiris

ESG Overall Score 45/100		
Environmental Management [50/100] (w3)	Labour Rights [30/100] (w2)	Board [10/100] (w3)
Water [62/100] (w2)	Non-Discrimination [45/100] (w3)	Audit & Internal Controls [10/100] (w3)
Energy [62/100] (w3)	Reorganisations [65/100] (w2)	Shareholders [50/100] (w3)
Environmental Supply Chain [62/100] (w1)	Economic Development [50/100] (w1)	Corruption [75/100] (w2)

The ESG overall score is adjusted taking into account all the possible controversies that may affect the business of a company. In particular, these controversies are systematically monitored by a dedicated team, which assesses them according to the following three elements:

- *Severity of the controversy*
- *Frequency of controversies on the particular ESG issue*
- *The responsiveness of the company to this controversy*

The final output is an evaluation of the controversy risk mitigation capacity of the company. This evaluation can lead to three different adjusting outcomes on the ESG overall score of a company, namely Confirmation, Upgrading or Downgrading of the previously assigned score.

2.3.5. Standard & Poor

Standard & Poor's annual Corporate Sustainability Assessment (CSA) [11] was developed in 1999 with the aim of spotting the corporations that are leaders in recognizing, managing and reacting to new waves of sustainability trends and opportunities. In particular, a multidisciplinary team of analysts is endowed with the responsibility of designing, managing and continuously improving the CSA, with the aim of offering an integrated approach oriented to the analysis of a company's sustainability performance. The assessment is based on those sustainability criteria that are considered as material, meaning that they are potentially impacting the financial performances of a firm. The CSA's approach is

also unique: indeed, it is based on the pieces of information provided by the corporations directly through the online questionnaire. In this way, S&P has the possibility to analyse a company's sustainability performances with a higher granularity than the one obtainable by using the frameworks based on public disclosure alone. S&P Global ESG Research invites 5000 of the world's largest publicly traded companies to take part in the CSA. The starting point for the annual corporate assessment is represented by an industry-specific questionnaire which focuses on material environmental and social criteria.

2.3.5.1. Evaluation Methodology

In order to get to a company's ESG Score, S&P relies on a hierarchical process [11] that is based on sub-level scores that are progressively weighted and summed until a final aggregated score is reached. In particular, the first step of this process consists of individual questions, whose values are weighted, summed and clustered into broader areas called criteria. In a similar way, criteria scores are weighted, summed and aggregated into even broader areas called dimensions. Finally, dimensions values are weighted and summed to get to the maximum sustainability score. This score is complemented with the Media and Stakeholder Analysis (MSA), which examines all the relevant events linked to a company that have been reported via the media and other channels. The MSA is utilised in order to adjust the criteria score by taking into account all the negative company-related controversies. The Figure 2.16 shows a graphical representation of this process.

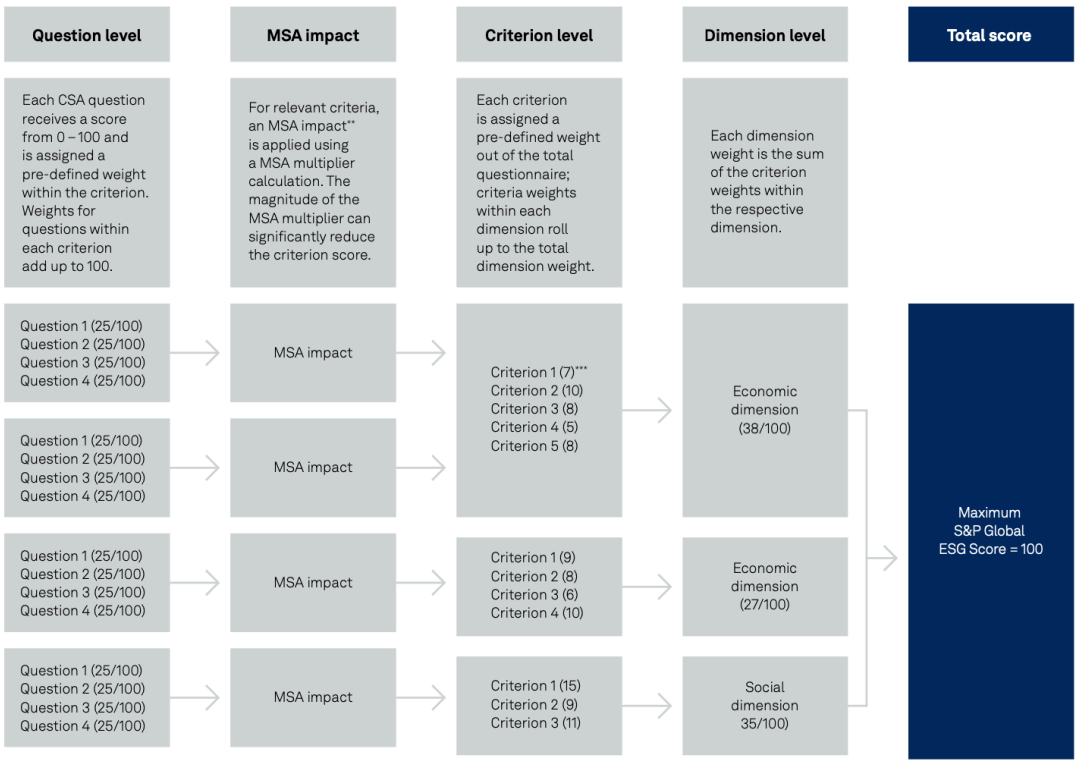


Figure 2.16: S&P ESG score generation process. Source: S&P Global

Based on the principal sustainability challenges identified by the team of analysts, the general criteria relating to standard management practices and performance measures such as Corporate Governance, Human Capital Development and Risk and Crisis Management are developed for all the 61 industries. These criteria contribute on average to about 50% of the total assessment. The remaining part of the CSA is constituted by industry specific ESG risks and opportunities that are relevant to the companies belonging to a particular industry. The relative weights of the economic, environmental and social dimensions of the questionnaire are different depending on the industry. For example, as shown in Figure 2.17, the environmental dimension presents a higher weighting in the Electric Utilities industry with respect to Banking or Pharmaceuticals industries.

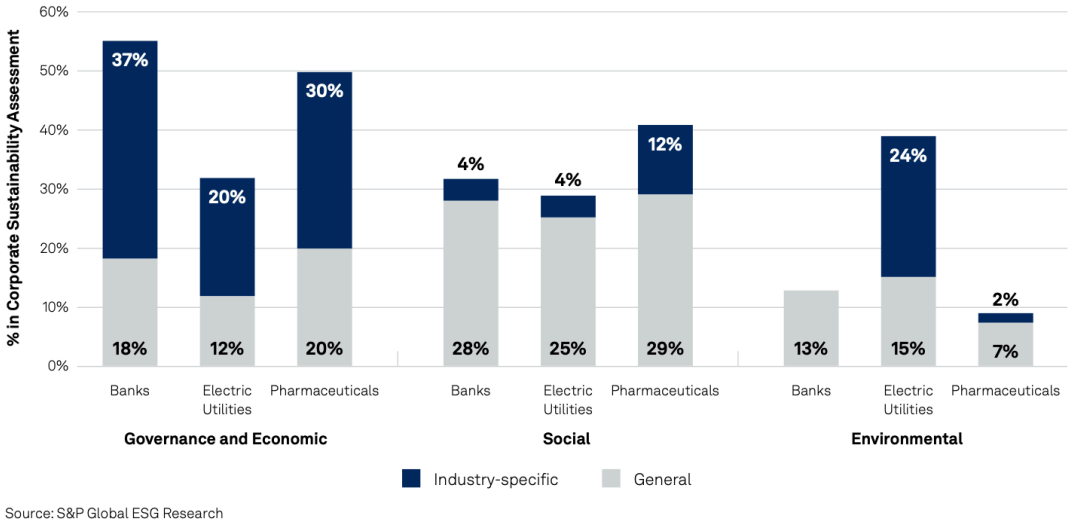


Figure 2.17: General versus industry specific weights by dimension. Source: S&P Global

Three important annotations are worthy to be mentioned at this stage:

- Criteria within the questionnaire can differ from industry to industry to reflect industry-specific drivers.
- Even though some criteria are applied to more than one industry, they could have different weights within the CSA.
- The same criterion, when applied to different industries, may contain a slightly different pool of questions.

2.3.5.2. Score Generation Process

The S&P sustainability evaluation process [11] is made up of three building blocks: The Materiality Matrix, the Corporate Sustainability Assessment and the Media & Stakeholders analysis.

The starting point of the CSA assessment is constituted by the concept of financial materiality. In particular, for each of the 61 industries analysed, the interdisciplinary team of analysts identify which are the sustainability topics that have an impact on the long-term value creation of a company. This analysis returns as output a materiality matrix for each industry, which is used as starting point for determining the applicability and weights of the various sustainability criteria in the CSA. The materiality matrix is obtained through the financial materiality analysis, which focuses its attention on industry-specific value drivers that contribute to company performance. More in depth, it consists of quantitative research aimed at identifying those intangible factors that have shown meaningful correlations to past financial performance. Each factor is then ranked according to the magnitude and likelihood of its impact on the company's financial performance over time. At the end of this process, the factors that are classified in the top of this ranking are assigned the highest weighting. An example of the outcome of the financial materiality analysis is shown by the image below.

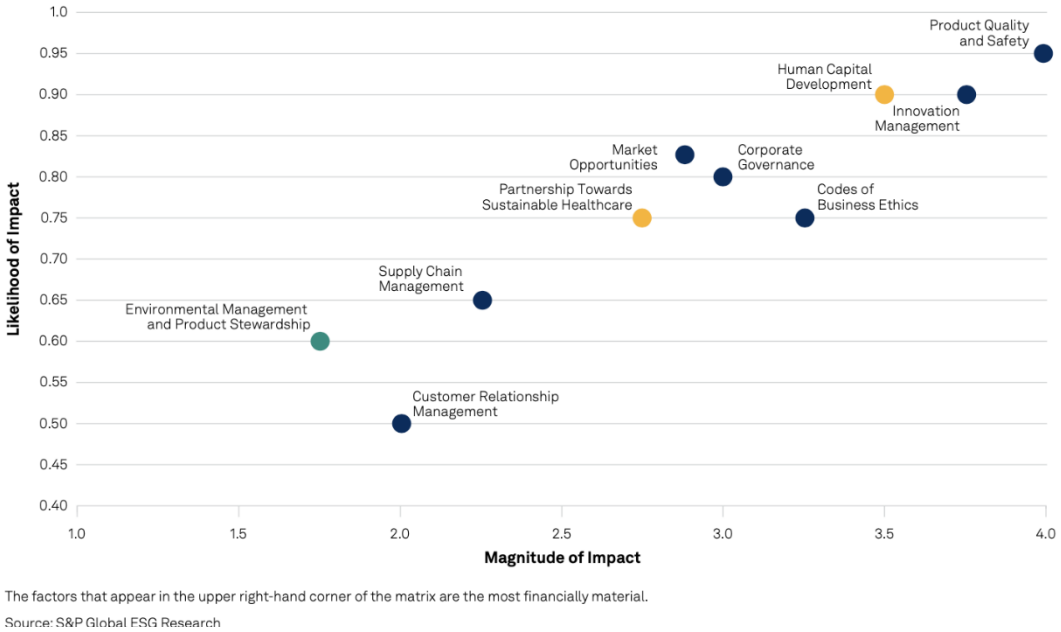


Figure 2.18: Financial materiality matrix for the pharmaceutical industry. Source: S&P Global

Once the materiality matrix has been determined, the Corporate Sustainability Assessment provide a score to the previously identified criteria through a questionnaire compiled by the companies under analysis. In particular, the questionnaire is designed in order to be objective and it adopts predefined scoring approaches in which each given answer is attributed a number of points between 0 and 100. For questions that could admit qualitative answers, S&P analysts codify the response by using a predefined method, which then translate the qualitative answer in a quantitative score. Furthermore, several questions require corporations to support the answers they have provided through the disclosure of some specific documents. In the cases in which companies are able to provide adequate supporting materials, they receive the maximum score for the question under inquiry. For some questions, points are only awarded if information covering the question requirements is publicly available.

Going more in depth, the questions belonging to each criterion are structured to capture and assess the following elements:

- Company's awareness of the importance of these factors to its financial success.
- The potential financial impact (i.e., materiality) of its exposure to sustainability factors.
- Implementation of management strategies to deal with these sustainability risks and/or opportunities in order to mitigate or capitalize on them respectively.
- Measurement of the effectiveness of a company's sustainability strategy through a set of Key Performance Indicators (KPI).
- Transparent disclosure of the company's corporate sustainability strategies and the results they have led to.

The Figure 2.19 provides an example of specific questions from the Banking industry, and it shows how a company's response to these questions could have an impact on the overall ESG Score.

Question: Customer Data Security and Data Privacy	**Which of the following qualitative and assurance aspects does your company's on-line financial service/system platform cover? Please provide supporting documents.**
Question Points	0–100
Question weight within criterion	15%
Criterion	Customer Relationship Management
Dimension	Economic
CSA Rationale	New lifestyles such as flexible working hours, increased mobility, and working from home are shifting consumer attitudes towards online services. By adopting a multi-channel strategy that includes online services, companies can further enhance their product offerings, service availability and standardization while improving customer loyalty and lowering costs. S&P Global ESG Research assesses what type of online services banks offer their customers. Networked data and globalized corporate activities require the diligent handling of information. Therefore, not only must companies have a comprehensive (online) privacy policy in place, they must also have the mechanisms to ensure the proper implementation of their policy. Over the past decade, the number of data breaches has grown exponentially. Therefore, this question asks companies if they have the necessary security systems in place and the ability to evaluate potential costs associated with such data breaches.

Possible Answers	**Number of Points Awarded**
A. List of potential approaches (company can check all that apply)	0–100 (depending on which approaches have been selected)
B. Not applicable	A question that has been marked "Not Applicable" will not be scored and the weight of the question will be equally redistributed across the other questions within the same criterion, only if the analyst agrees that the question does not apply to the company's business model. This option is only granted in exceptional cases.
C. Not known	0

Assuming the company receives 67 points for its response to this question, its score will be calculated as follows:

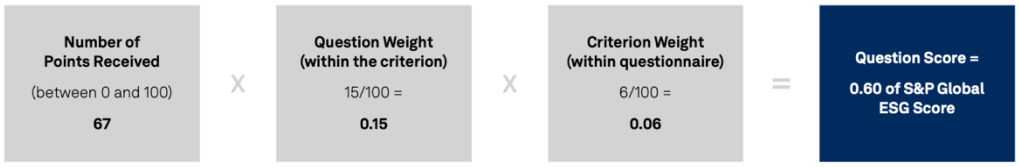


Figure 2.19: Example of questions for the Banking industry. Source: S&P Global

Finally, the company's S&P Global ESG Score is constituted by the sum of all the Question Scores. Each company is assigned an ESG Score ranging from 0–100. Once the S&P Global ESG Score have been computed, firms belonging to the same industry are ranked against their peers in order to determine which companies are eligible for inclusion in the Dow Jones Sustainability Indices (DJSI).

$$S\&P\ Global\ ESG\ Score = \sum (Number\ of\ Question\ points\ received \times Question\ Weight \times Criterion\ Weight) \quad (2.2)$$

The last building block of the S&P's evaluation methodology [11] is represented by the Media and Stakeholder Analysis (MSA). Indeed, beside the previously described scoring process, another relevant part of the CSA process is the continuous monitoring of the publicly available pieces of information coming from print and online media, government bodies, regulators and other sources. This continuous monitoring is aimed at identifying companies' involvement and subsequent response to environmental, economic and social incidents that may have a detrimental impact on their reputation and main business. In particular, a MSA case is opened in three main situations:

- If a company has been involved in a negative event for which it is considered to be responsible.
- In the case in which a company is found to adopt actions that are inconsistent with its stated policies and goals.
- In the case in which a negative incident reveals the failure of the company's management, systems and/or processes.

The necessary condition that an MSA case must satisfy in order to be created is that it has to be material, meaning that it could potentially lead to relevant damages in reputation and financials. After that the MSA case is opened, the company is expected to fix the issue by undertaking countermeasures in order to reduce its detrimental impact as well as the likelihood of possible future incidents. S&P does not stop at this stage, but it takes the responsibility of contacting the companies for which an MSA case has been created and it continuously monitors the case until it has been resolved. To measuring the MSA impact, S&P adopts a process that starts with the identification of an MSA case. Then, the MSA case is assigned a score on the basis of its impact and according to the subsequent response of the company involved. Using the MSA score as input, the "MSA multiplier" is defined; this

coefficient is used to adjust relevant CSA criteria in proportion to the negative impact of the MSA cases. The MSA multiplier amplifies the negative impact of mediocre MSA scores on the final criteria scores. In particular, the larger the negative impact, the larger the downward adjustment of criterion scores. The image below shows the process through which a specific MSA case is identified, assessed and incorporated into the CSA.

In the case in which a corporation does not present any MSA case during the accounting period, the criterion score is not subject to variations.

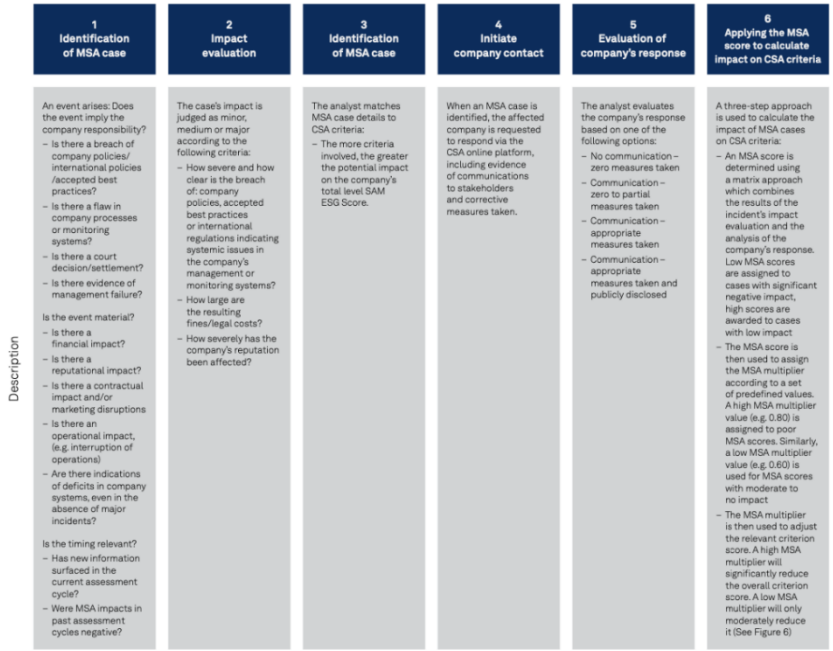


Figure 2.20: Overview of the MSA process. Source: S&P Global

3 Literature Review

This chapter describes the state of the art of the literature concerning, on the one hand, the relationship between the sustainability and the financial performances and, on the other hand, the ESG ratings divergence phenomenon, the financial investors' confusion deriving from it and, eventually, its possible consequences in terms of impact on corporations' value.

3.1. The Relationship Between Sustainability and Performance

In section [2.2], the motivations leading financial investors to integrate sustainability factors into their investment decision were listed. Among them, one of the main drivers for ESG investing is the fact that, according to the major part of the literature, it could increase the performances of the sustainability-driven portfolios, arriving to provide higher returns compared to the traditional ones. In this regard, this section has the aim to provide a comprehensive overview of the literature concerning the causal relationship between sustainability and financial performances and, therefore, financial value. In doing so, the starting point consists in analysing the studies that relate firms' value to the three main constituents of the ESG finance world, namely environment, society and governance.

3.1.1. Environmental Pillar and Performance

The environmental pillar assesses the sustainability of those companies' activities having a direct and indirect impact on the surrounding environment. In the financial literature there are several studies demonstrating the positive correlation between this pillar and the financial performances of the corporations. Among them, the most significant are the ones by Derwall et al. (2004) [12] and Manrique and Martí-Ballester (2017) [13]. More in depth, the work by Derwall et al. (2004) [12] investigates whether socially responsible investing (SRI) is able to provide superior portfolio performance. In particular, it focuses on the "eco-efficiency" concept, which can be interpretated as a proxy of the economic value a company is able to generate compared to the waste it produces. In order to accomplish the objective of the analysis, the authors construct and evaluate two equity portfolios that differed in eco-efficiency, based on Innovest Strategic Value Advisors' corporate eco-efficiency scores. What has been found is that the sustainable portfolio is the one experiencing higher financial performance, providing substantially higher average returns over the 1995–2003 period with respect to the low-ranked portfolio. Instead, the main objective of the study by Manrique and Martí-Ballester (2017) [13] was to explore the effect of environmental performance on financial performance during a global financial crisis, distinguishing the impact according to the economic development level of the firm's belonging country. In order to do so, the two authors start from a data set composed by 2982 corporations, gathering their pieces of information from 2008 to 2015. Then, this data sample is analysed trough the Petersen's approach, finding that the usage of environmental practices leads to significant financial performance benefits to both developed and developing countries. Nevertheless, this effect is surprisingly stronger for the companies

situated in the developing countries with respect to the ones located in more developed countries. Indeed, in the developing areas the financial performance improvement appears in the short term and remains in the long one, while in the developed regions the improvement effect is constrained only to the short term.

3.1.2. Social Pillar and Performance

The social pillar assesses the corporations' sustainability by looking at the way they are able to handle the social relationships and interactions inside and outside their boundaries. For what concerns this pillar, the most relevant study to be cited is the one by Edmans (2011) [14]. This paper investigates the impact that a high level of employee satisfaction could have on long-run stock returns. More in depth, the author starts by building a value-weighted portfolio composed by the stocks of those companies belonging to the ''*100 Best Companies to Work for in America*'', and he evaluates the extra-return of the portfolio with respect to the return of the four-factor model in the period between 1984 and 2009. What emerges from the study is that the socially sustainable portfolio shows an annual four-factor alpha of 3.5%. It is important to note that this result is robust independently from the company-specific characteristics, the weighting methodologies and also from the removal of outliers. Two main considerations emerge from what explained above. The first one is that socially responsible investing (SRI) screens may lead to investment returns improvements. The latter is that, consistently with human capital-centered theories, the employee satisfaction is positively correlated with corporations' value.

3.1.3. Governance Pillar and Performance

The governance pillar refers to the way in which a firm is managed by its top management offices, and it can be seen as a measure of the alignment between the executive management's interests and the ones of the company's stakeholders. Among the literature studies demonstrating the positive correlation between the corporate governance pillar and the financial performances of firms, the two papers cited in this work are the ones by Gompers et al. (2003) [15] and Velte (2017) [16]. In particular, the paper by Gompers et al. (2003) [17] is based on the observation that the rights of shareholders could change inside different companies. From this consideration and exploiting 24 governance rules, the authors developed a "Governance Index" for a sample of 1500 corporations evaluated during the 1990s; this indicator can be interpreted as a measure of the shareholders rights' level. The next step consists in the formulation of the investment strategy: in this regard, the authors decide to have a long position on the companies in the best decile of the index (which are the ones experiencing the strongest level of shareholder rights) and, instead, to sell the stocks of the firms belonging to the worst decile. From this analysis, it emerges that the portfolio would have earned abnormal returns equal to 8.5% per year during the considered period of the analysis. Looking at this result, it is possible to conclude that the companies with a higher level of shareholder rights benefit from higher firm value and greater profits. Moving to the study conducted by Velte (2017) [16], it focuses on environmental, social and governance performance in total, and then it breaks downs the ESG performance into the three constituents, with the aim of assessing their impact on financial performance. By looking at the methodology, the author started from the creation of a pool of companies listed on the German Prime Standard (DAX30, TecDAX, MDAX), and he

analyses them in the period going from 2010 to 2014. More in depth, the author conducts a correlation and regression analysis with the aim of assessing possible relationships between ESG performance and financial performance (which, in this case, are measured through the ROA). What emerges from the study is that ESG performance exert a positive impact on ROA. In particular, it is interesting to report that the corporate governance pillar is the one having the strongest impact on the financial performance of the companies with respect to the environmental and the social constituents.

3.1.4. ESG and Performance

After having explored the literature concerning each specific pillar of the ESG paradigm, the next paragraph aims at investigating the most relevant papers that study in a comprehensive way the benefits coming from considering the environmental, social and governance dimensions at the same time.

In doing so, part of the literature is focused on studying the disclosure about ESG activities and the firms' benefits arising from it. In this regard, the paper by Wong et al (2020) [17] investigates the effect of ESG certification on Malaysian firms, finding that corporations' commitment toward sustainability reduces their cost of capital. These results are strengthened by Fatemi et al (2017) [18]: this study examines how companies' value is impacted by ESG activities and the disclosure about them, finding that ESG strengths increase firm value and that weaknesses decrease it. In this contest, it is relevant to cite the paper by Brogi and Lagasio (2018) [19], which has the scope to indagate the relationship between ESG disclosure and the firm's profitability captured through the Return on Asset. The authors gather the sustainability scores about a set of listed American companies from MSCI, for a

period ranging from 2000 and 2016. Exploiting a statistical model, the connection between ROA and the environmental, social and governance pillars is investigated, finding a positive and significant correlation between ESG and the companies' profitability. The authors find heterogeneous results moving from industrial to financial firms. In particular, the latter are the ones experiencing the strongest bond between the two focal performances under analysis.

Another relevant part of the literature is composed by the studies analysing the impact that ESG exerts on the firms' performances. Starting from the study by Sila and Cek (2018) [20], it investigates the impact of the three dimensions of ESG finance on the corporations' performances, but focusing on the economic rather than the financial ones. For what concern the methodology adopted in the study, the authors run a regression analysis on the data regarding a sample of Australian firms, covering the period 2010-2016. From this analysis it is found that, among the three ESG pillars, the social dimension is the one leading to a higher improvement of the economic performance. The environmental performance exerts a positive influence on the economic one too, but with a lower magnitude of the impact with respect to the social dimension. For what concern the last ESG dimension, namely the governance, the authors find a positive but mostly not significant relationship between this pillar and the companies' economic performance. Moving to the paper by Aouadi and Marsat (2016) [21], it inspects the relationship between environmental, social, and governance controversies and firm market value. In particular, the authors start from a set of more than 4000 companies belonging to 58 countries, in the period ranging from 2002 to 2011. The firms of the sample are then split in two clusters: the high-attention firms versus the low-attention ones. High-attention companies are the larger ones, the more searched on the Internet, the more followed by analysts and are located in countries with greater press freedom. From the analysis conducted in this study, it emerges that higher Corporate Social

Performances score has an impact only on the market value of high-attention firms. Another relevant study in this field is the one by Giudici et al (2020) [22], which investigates the performance of 799 bonds listed on European stock markets in the period 2014-2018, according to the quarterly improvement of their ESG rating. In particular, the authors adopt a matching procedure, with the aim of extrapolating the differences not related to the variety of the bonds in terms of maturity and other factors. What emerges from the paper is that high-yield securities are the ones experiencing a positive and significant influence from sustainability, presenting a yield spread of around 9% over the period. Moreover, this positive impact on financial performance is mainly attributable to environmental and governance factors. Among all the studies which link the corporate performance to the ESG ones, the paper by Friede et al. (2015) [23] is by far the most exhaustive. This research starts from the observation that this relationship has been investigated since the 1970s, counting more than 2000 empirical studies. Hence, the knowledge on the causal effects of the ESG constituents on financial performance result to be fragmented. It is in this regard that the authors of this paper try to gather all the relevant pieces of information coming from the existing literature, with the aim of providing a comprehensive overview of all the studies' findings. From this recapitulatory work, it emerges that around the 90% of the analysed papers report a non-negative relationship between ESG and corporate performances. As final result, the authors highlight on the one hand that the majority of the papers find a positive relationship and, on the other hand, that this outcome appears stable over time.

Beside the literature exploring the stylised fact that ESG topics integration leads to better economic and financial performance for both firms and investment portfolios, another relevant part of the literature shows that this positive relationship is enhanced during the crisis periods. This evidence is due to a systematic reduction

of the risk, which is caused by the inclusion of sustainability pillars inside the decision-making process. Among all the papers treating this topic, the three most relevant are the ones by Nofsinger and Varma (2014) [24], Henke (2016) [25] and by Giese et al. (2019) [26]. Starting from the paper by Nofsinger and Varma (2014) [24], it compares the performance of socially responsible mutual funds with respect to the conventional ones. What emerges from the study is that the sustainable funds experience better performances during the periods of financial crisis due to a lower downside risk; however, these funds show lower performances during the market non-crisis periods. The result of this study is especially true for the ESG funds that adopt positive screening techniques when building their investment portfolio. Eventually, the authors highlight that these observed patterns are exclusively due to the socially responsible attributes and not to the different ways of managing the funds or to the specific firms' characteristics. The evidence found by Nofsinger and Varma (2014) [24] are confirmed in the study by Henke (2016) [25]. In particular, it assesses what it the financial impact of the inclusion of ESG criteria in the bond portfolio construction methodology. More in depth, the analysis is focused on the risk-adjusted financial performance of 103 sustainable funds in the US and the Eurozone, which are compared with conventional portfolios during the time horizon going from 2001 to 2014. What emerges from the analysis is that the sustainable bond funds experience higher annual performances, which are attributable to the mitigation of the ESG-related risks. Moreover, the author splits the analysis period into two time buckets, separating the crisis periods from the non-crisis ones. This further step of the analysis highlights that the outperformance of the ESG integrated funds is more likely to be observed during the crisis periods. The last relevant contribution to this topic is provided by the study conducted by Giese and Lee (2019) [26]. The authors start from the observation that the adoption of MSCI ESG ratings leads to a reduction of the risks in investment portfolios. This

fact can be explained by looking at the MSCI's risk assessment methodology, which is based on the identification of potential firms' risks and on the evaluation of the effectiveness of the companies' risk management systems. The main outcome of the study is that the corporations with higher sustainability scores are the ones experiencing a lower frequency of stock-specific risks, avoiding significant drawdowns; furthermore, these companies show, on average, greater profits and higher firm value.

3.2. ESG Ratings Divergence

Having introduced the most relevant players in the ESG rating arena in section [2.3], this part of the literature review focuses on one of the main problems affecting financial investors' decision-making process, namely the divergence between ESG rating agencies' evaluations and the possible consequences that could arise from it. As explained in the previous sections, this heterogeneity in the sustainability assessments is caused by the absence of a clear and univocal methodological evaluation standard. This leads the ESG rating providers to develop and consequently adopt different methodologies, arriving to divergent sustainability grades. With the aim of studying the ESG ratings divergence phenomenon, the first part of this section is focused on the study conducted by Berg et al (2020) [27] and by Capizzi et al (2021) [28].

3.2.1. Divergence Analysis

The study by Berg et al (2020) [27] is based on the data gathered from six rating agencies - KLD (MSCI Stats), Sustainalytics, Vigeo Eiris (Moody's), RobecoSAM (S&P Global), Asset4 (Refinitiv), and MSCI- and it investigates the divergence of ESG ratings, decomposing them into their three main pillars, namely scopes of categories, measurement of categories and weight of categories. The scope represents the set of attributes that together build the overall picture of a company's ESG performance. Concerning the measurement, it is about the indicators that provide a quantitative measure of the attributes. Finally, there is an aggregation rule that combines the indicators into a single comprehensive rating.

According to this paper, ESG ratings divergence could be explained by looking at three principal constituents, which are listed below:

- *Scope divergence*: the ratings are based on different sets of attributes. For example, one rating agency may decide to include lobbying activities, while another might not, causing the divergence of the two ratings.
- *Measurement divergence*: the rating agencies measure the same attribute using different indicators. For example, a firm's attitude towards gender discriminations could be evaluated on the basis of the salaries guaranteed to the employees, or by the number of women present in the board of directors. Both capture aspects of the attribute gender discrimination, but they lead to different assessments.
- *Weights divergence*: the rating agencies give different relative importance to the attributes. For example, the gender discrimination indicator may enter the final rating with greater weight than the lobbying indicator.

In order to go in dept into the problem of divergence, the authors of this paper approached the problem in three main steps:

- The *first phase* consisted in the creation of a taxonomy, built from a starting data set of 709 indicators. In particular, in this taxonomy a category is created every time two indicators from different rating agencies refers to the same attribute. Indicators that are not included into a shared attribute remain unclassified. The categorization of the indicators leads to the creation of 65 categories. This categorization allows to observe the scope of categories covered by each rating as well as to contrast measurements by different rating agencies within the same category. Based on this taxonomy, the rater-specific category scores are computed by averaging indicators that were assigned to the same category.

- In the *second phase*, the original rating is taken as dependent variable (y) and it is regressed using the category scores as dependent variables (x). To do so, it is used a non-negative least squares regression, where coefficients are constrained to be equal to or larger than zero; this regression model allows to compare these fitted ratings to each other in terms of scope, measurement, and aggregation rule.
- In the *third phase*, it is computed the contribution of divergence in scope, measurement, and weights to overall ratings divergence.

3.2.1.1. Finding of the Analysis

Moving to the results of the study, the computed correlations between ESG ratings are on average 0.54, ranging from 0.38 to 0.71. Sustainalytics and Vigeo Eiris experience the highest level of agreement between each other, with a correlation of 0.71, while KLD and MSCI report the lowest correlations with other rating agencies.

Table 3.1: Correlation between different ESG ratings. Source: Berg et al (2020) [27]

	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

Looking at the E-S-G components, the environmental dimension has an average correlation of 0.53, which is slightly lower than the overall correlations. The social dimension is on average correlated at 0.42, while the governance dimension has the

lowest correlation, with an average of 0.30. Table 3.2 shows the firms that are classified in the top and bottom 20% of the common sample across all six rating providers. The first column of Table 3.2 provides an idea of how a sustainable investment portfolio based on a strict consensus of six rating agencies would be.

Table 3.2: Common sets of firms in quantiles. Source: Berg et al (2020) [27]

Common in Top Quantile	**Common in Bottom Quantile**
Akzo Nobel NV	Advance Auto Parts Inc.
Allianz SE	Affiliated Managers Group Inc.
Aviva plc	Amphenol Corporation
AXA Group	Anhui Conch Cement Co. Ltd.
Bayerische Motoren Werke Aktiengesellschaft	Cencosud SA
Dexus Property Group	China Development Financial Holding Corporation
Diageo plc	China Resources Land Ltd.
Industria de Diseno Textil SA	Credit Saison Co. Ltd.
Intel Corporation	Crown Castle International Corp.
Kingfisher plc	DR Horton Inc.
Koninklijke Philips NV	Expedia Inc.
SAP SE	Helmerich & Payne Inc.
Schneider Electric SA	Hengan International Group Company Limited
STMicroelectronics NV	Intuitive Surgical Inc.
Wipro Ltd.	Japan Real Estate Investment Corporation
	MediaTek Inc.
	NEXON Co. Ltd.
	Nippon Building Fund Inc.
	Ralph Lauren Corporation
	Shimano Inc.
	Sumitomo Realty & Development Co. Ltd.
	Sun Pharmaceutical Industries Limited
	Wynn Resorts Ltd.

By looking at the table, it is possible to observe that there are only 15 companies categorized into the top 20% in all ratings, a small number considering that 20% of the sample equates to 184 companies. The second column of the table provides the names of the 23 companies that are included in the bottom 20% in all ratings, and which are expected to be consistently avoided by most of the rational sustainable investors.

To summarize the stylized facts about ESG ratings divergence found by this analysis, the first one is that the divergence exists and it is substantial since correlations between ratings are on average very low (only 0.54). The second

evidence is that the disagreement is stronger for firms that are ranked near the top of the distribution. As a consequence, it is likely that portfolios that are built looking at different ESG ratings have different constituents, and portfolios that select the top performers in all ratings are extremely constrained to very few sets of companies.

After having shown the general results of the conducted analysis, the focus shifts on the divergence contribution of the three pillars of the rating: scope, measurement and weight.

3.2.1.2. Scope Divergence

By looking at the first source of divergence among the ESG ratings, Table 3.3 shows the results of the categorization described in the section [3.2.1]: in particular, for each rating agency, it reports the number of indicators adopted in order to evaluate a specific category. As it is possible to observe, the table highlights the existence of a considerable scope divergence.

Table 3.3: Number of indicators for rater and category. Source: Berg et al (2020) [27]

	Sustainalytics	RobecoSAM	Asset4	Vigeo Eiris	MSCI	KLD
Access to Basic Services	2		1		1	1
Access to Healthcare	6	3	1		1	1
Animal Welfare	2		1			
Anti-competitive Practices			2	1	1	1
Audit	4		5	1		
Biodiversity	1	1	3	1	1	2
Board	6		25	1	1	
Board Diversity	2		1			3
Business Ethics	4	2	1		1	1
Chairperson-CEO Separation	1		1			
Child Labor			1	1		1
Climate Risk Mgmt.		2	1		1	2
Clinical Trials	1		1			
Collective Bargaining	2		1	1		
Community and Society	3	6	10	1		1
Corporate Governance		1			1	
Corruption	2		1	1	1	1
Customer Relationship	1	1	7	1		2
Diversity	2		9	1		3
ESG Incentives	1	1				
Electromagnetic Fields	1	1				
Employee Development	1	2	13	1	1	3
Employee Turnover	1		1			
Energy	3	6	5	1	2	1
Environmental Fines	1		1			1
Environmental Mgmt. System	2		1			1
Environmental Policy	4	2	4	2		
Environmental Reporting	2	1	1			
Financial Inclusion	1				1	1
Forests	1	1				
GHG Emissions	5		5	1		1
GHG Policies	3	2	4			
GMOs	1	1	1			
Global Compact Membership	1		1			
Green Buildings	5	2	1		1	1
Green Products	7	1	20	1	2	1
HIV Programs	1		1			
Hazardous Waste	1	1	1		1	
Health and Safety	7	1	7	1	1	2
Human Rights	2	1	5	1		5
Indigenous Rights	1		1			1
Labor Practices	3	1	16	4	1	3
Lobbying	3	1		1		
Non-GHG Air Emissions	1		2			
Ozone-Depleting Gases	1		1			
Packaging		1			1	1
Philanthropy	3	1	2	1		1
Privacy and IT	1	3			1	2
Product Safety	2	2	13	3	2	6
Public Health	1	3			1	2
Recycling					1	
Remuneration	4	1	15	2	1	4
Reporting Quality	3		5			1
Resource Efficiency	1	3	6			
Responsible Marketing	3	3	1	1		1
Shareholders			16	1		
Site Closure	1	1				
Supply Chain	21	3	4	4	3	6
Sustainable Finance	9	5	3		3	4
Systemic Risk		1			1	1
Taxes	2	1	1			
Toxic Spills	1		2			1
Unions			1			1
Waste	3	2	4	1		3
Water	2	2	3	1	1	2
Unclassified	7	7	42	1	34	2
Sum	163	80	282	38	68	78

More in depth, on the one hand there are categories taken into account by all the rating agencies, representing a sort of common denominator of categories included inside an ESG rating. These categories are, for example, Biodiversity, Employee Development, Energy, Green Products, Health and Safety, Labor Practices, Product Safety, Remuneration, Supply Chain and Water. On the other hand, there are many empty cells, denoting that not all the categories are covered by all the ratings. The logic suggests that the experienced gaps would be detected only in the higher customized categories, but this does not hold in the reality. Indeed, even some

categories such as Taxes, which could be viewed as a fundamental concern in the context of ESG, are not covered by all the ratings. Furthermore, the presence of many unclassified indicators underlines that there are several aspects of ESG that are only measured by one out of six rating agencies. In this regard, Asset4 is the rating provider that adopts the highest number of unclassified indicators (42), which are not considered by any other rating agency.

3.2.1.3. Measurement Divergence

Moving to the second source, in order to compute the measurement divergence, the authors of the paper develop a category score by computing the average of the indicators' values.

The following variables and indexes are used throughout the paper:

Notation	Variable	Index	Range
A	Attributes	i	$(1, n)$
I	Indicators	i	$(1, n)$
C	Categories	j	$(1, m)$
N_{fkj}	Indicators $\in C_{fkj}$	i	$(1, n_{fkj})$
R	Raters	k	$(1, 6)$
F	Firms	f	$(1, 924)$

The category score is computed as

$$C_{fkj} = \frac{1}{n_{fkj}} \sum_{i \in N_{fkj}} I_{fki}$$

for firm f, rating agency k, and category j.

Figure 3.1: Category score computational formula. Source: Berg et al (2020) [27]

The category scores reflect a rating agency's assessment of a specific ESG category. These scores are built on different groups of indicators with the same scope. Therefore, the differences between the scores assigned to each category derive from differences in how the rating agencies decide to measure, rather than what they choose to measure. It follows that the differences between the same categories from different rating providers can be interpretated as measurement divergence.

Table 3.4: Correlation of category scores. Source: Berg et al (2020) [27]

	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
Access to Basic Services	0.08			0.13	0.85			0.49	0.15						0.16	0.31
Access to Healthcare	0.66		0.57	0.49	0.85		0.67	0.56	0.74				0.44	0.71	0.7	0.64
Animal Welfare								0.44								0.44
Anti-competitive Practices		-0.06		0.56	0.76						0	-0.05			0.56	0.30
Audit						0.57		0.66			0.62					0.62
Biodiversity		0.06	-0.08	0.06	0.66					0.61	0.41	0.47	0.47	0.01	0.2	0.29
Board						0.37		0.58			0.51					0.49
Board Diversity								0.8								0.80
Business Ethics	0.04		-0.11	0.4	0.6		0.33	0.03	0.01				-0.1	-0.15	0.38	0.14
Chairperson–CEO Separation								0.59								0.59
Child Labor				0.49												0.49
Climate Risk Mgmt.			0.44	0.42	0.8								0.54	0.54	0.5	0.54
Clinical Trials								0.73								0.73
Collective Bargaining						0.59		-0.04			0					0.18
Community and Society	-0.15	0.25	0.2	0.11		-0.1	-0.19	-0.13		0.51	0.5		0.56			0.16
Corporate Governance														0.08		0.08
Corruption	0.26	0.24		-0.18	0.7	0.54		-0.19	0.37		-0.15	0.33			-0.12	0.18
Customer Relationship	0.38	-0.08	-0.09	0		-0.04	-0.13	-0.05		0.49	0.47		0.41			0.14
Diversity	-0.06	-0.02		0.03		0.61		0.52			0.56					0.27
ESG Incentives																
Electromagnetic Fields							0.68									0.68
Employee Development	0.22	0.29	0.37	0.37	0.73	0.23	0.19	0.36	0.34	0.39	0.29	0.31	0.55	0.45	0.51	0.37
Employee Turnover								0.4								0.40
Energy	0.22	0.13	0.49	0.25	0.8	0.4	0.27	0.27	0.4	0.32	0.41	0.59	0.2	0.4	0.48	0.38
Environmental Fines								0.05								0.05
Env. Mgmt. System	0.65			-0.09				0.46								0.34
Environmental Policy						0.52	0.46	0.46		0.63	0.61		0.62			0.55
Environmental Reporting							0.52	0.25					0.36			0.38
Financial Inclusion	0.29				0.7				0.51							0.50
Forests																
GHG Emissions	0	-0.03		-0.06		0.28		0.31			0.5					0.17
GHG Policies							0.48	0.62					0.41			0.50
GMOs							0.38	0.43					0.25			0.35
Global Compact Member								0.92								0.92
Green Buildings	0.54		0.59	0.21	0.83		0.25	0.26	0.55				-0.02	0.66	0.28	0.42
Green Products	0.23	0.07	0.27	0.34	0.76	0.1	0.37	0.47	0.32	0.31	0.29	-0.05	0.53	0.44	0.53	0.33
HIV Programs																
Hazardous Waste							0.22	0.13	0.34					0.59	0.1	0.28
Health and Safety	0.01	0.27	0.27	0.35	0.73	-0.1	-0.16	-0.16	-0.05	0.63	0.67	0.5	0.57	0.44	0.6	0.30
Human Rights	0	0.19		0.08		-0.01		-0.08			0.42					0.10
Indigenous Rights	0.26			-0.11				-0.46								-0.10
Labor Practices	0.21	-0.04	-0.14	0.07	0.1	0.2	0.14	0.32	0.27	0.54	0.45	0.43	0.35	0.34	0.37	0.24
Lobbying						-0.28										-0.28
Non-GHG Air Emissions								0.28								0.28
Ozone-Depleting Gases								0.44								0.44
Packaging																
Philanthropy						0.42	0.39	0.32		0.48	0.19		0.17			0.33
Privacy and IT	0.48		0.27		0.75		0.17		0.45					0.42		0.42
Product Safety	-0.05	0.06	0.16	0	0.63	-0.14		-0.03	0.07	0.46	0.21	0.11	0.38	-0.03	0.1	0.14
Public Health			0.6		0.74		0.38							0.63		0.59
Recycling																
Remuneration	0.15	0.09	-0.21	0.17		0.71	0.22	0.83		0.25	0.75		0.37			0.33
Reporting Quality								0.48								0.48
Resource Efficiency							0.35	0.42					0.57			0.45
Responsible Marketing	-0.5	-0.06	-0.38	0.24		0.38	0.68	0		0.49	0.05		-0.1			0.08
Shareholders											0.39					0.39
Site Closure																
Supply Chain	0.15	0.17	0.13	0.16	0.62	0.57	0.53	0.56	0.61	0.66	0.62	0.6	0.53	0.34	0.48	0.45
Sustainable Finance	0.58		0.47	0.46	0.52		0.7	0.74	0.7				0.59	0.61	0.59	0.60
Systemic Risk			0.24		0.65									0.24		0.38
Taxes							0.1	0.02					0.01			0.04
Toxic Spills				0.21												0.21
Unions				0.68												0.68
Waste		0.34		0.23							0.33					0.30
Water	0.36	0.36	0.23	0.23	0.67	0.47	0.29	0.31	0.45	0.48	0.32	0.5	-0.02	0.24	0.44	0.36
Average	0.20	0.12	0.20	0.21	0.69	0.29	0.32	0.33	0.37	0.48	0.38	0.34	0.35	0.37	0.38	

Looking at the correlations between the categories, Table 3.4 offers two main insights. On the one hand, the correlation levels are heterogeneous. Indeed, there are categories such as Environmental Policy that show an average level of correlation of 0.55, testifying some level of agreement regarding the existence of the firm's environmental policy. However, there are also categories that measure objective facts that do not experience high level of correlation. In addition, the analysis shows also the presence of some negative correlations, that reflect opposite conclusions given by different agencies. The second interesting insight is that correlations are more likely to increase with granularity. An example of this stylized fact is represented by the categories of Water and Energy, that show an average correlation of 0.36 and 0.38, respectively. These numbers are substantially lower than the correlation of the environmental dimension, with an average of 0.53. This highlights that the divergences compensate each other during the aggregation process.

Another relevant factor that must be taken into account analysing the measurement divergence is the Rater Effect. The Rater Effect represents a kind of bias, where performance in one category influences perceived performance in other categories. In particular, the Rater Effect implies that when the judgement of a company is positive for one particular indicator, it is also likely to be positive for another indicator. It is not clear which are the causes of this phenomenon, but one possible explanation is that ESG rating agencies divide analysts' labour by firm and not by category; consequently, an analyst's overall picture of a company could propagate into the assessments of the different categories.

3.2.1.4. Weight Divergence

Finally, the authors analyse the third source of divergence, namely the weight divergence, by estimating the aggregation rule that transforms the category scores into the final ratings. In order to do so, they perform a non-negative least squares regression, finding the results showed below.

Table 3.5: Non-negative Least Squares Regression results. Source: Berg et al (2020) [27]

	Sustainalytics	**RobecoSAM**	**Asset4**	**Vigeo Eiris**	**MSCI**	**KLD**
Access to Basic Services	0.019	-	0	-	0.138***	0.065***
Access to Healthcare	0.051***	0.004	0	-	0.079***	0.051***
Animal Welfare	0.05***	-	0	-	-	-
Anti - competitive Practices	-	-	0.05***	0.023***	0	0.131***
Audit	0	-	0.026*	0.084***	-	-
Biodiversity	0	0	0	0.028***	0.366***	0.076***
Board	0.072***	-	**0.196*** **	0.113***	0	-
Board Diversity	0.043***	-	0	-	-	0
Business Ethics	0.097***	0.046***	0.008	-	0	0.148***
Chairperson-CEO Separation	0.039***	-	0.016	-	-	-
Child Labor	-	-	0.008	0	-	0.046***
Climate Risk Mgmt.	-	**0.137*** **	0.064***	-	0.069**	**0.234*** **
Clinical Trials	0	-	0	-	-	-
Collective Bargaining	0.051***	-	0.011*	0.072***	-	-
Community and Society	0.079***	0.086***	0.03*	0.001	-	0.14***
Corporate Governance	-	0.048***	-	-	0.198***	-
Corruption	0.049***	-	0.022*	0.072***	**0.388*** **	0.124***
Customer Relationship	0.127***	0.097***	0.086***	0.027***	-	0.104***
Diversity	0.108***	-	0.066***	**0.159*** **	-	0.04***
ESG Incentives	0.006	0	-	-	-	-
Electromagnetic Fields	0.021**	0	-	-	-	-
Employee Development	0.018*	**0.221*** **	0.116***	0.067***	**0.406*** **	0.149***
Employee Turnover	0.024*	-	0	-	-	-
Energy	0.032**	0.016***	0.029**	0.103***	0.194***	0.046***
Environmental Fines	0	-	0	-	-	0
Environmental Mgmt. System	**0.199*** **	-	0.009	-	-	0.205***
Environmental Policy	0.091***	0.098***	0.012	**0.187*** **	-	-
Environmental Reporting	0.043**	0.039***	0.007	-	-	-
Financial Inclusion	0	-	-	-	0.089***	0.061***
Forests	0.008	0.016*	-	-	-	-
GHG Emissions	0.048***	-	0.002	0.033***	-	0.021**
GHG Policies	0.086***	0.008**	0.047**	-	-	-
GMOs	0	0	0	-	-	-
Global Compact Membership	0.029**	-	0	-	-	-
Green Buildings	0.072***	0.071***	0	-	0.304***	0.072***
Green Products	**0.167*** **	0.037***	0.093***	0.024**	0.351***	0.129***
HIV Programs	0	-	0.003	-	-	-
Hazardous Waste	0.021*	0	0	-	0.09***	-
Health and Safety	0.049***	0.042***	0.049***	**0.125*** **	0.148***	0.174***
Human Rights	0.072***	0	0.066***	0	-	0.14***
Indigenous Rights	0.033*	-	0.006	-	-	0.087***
Labor Practices	0.005	0.063***	0.067***	0.153***	0.166***	0.129***
Lobbying	0.091***	0	-	0.013	-	-
Non-GHG Air Emissions	0.014	-	0	-	-	-
Ozone-Depleting Gases	0	-	0	-	-	-
Packaging	-	0	-	-	0.128**	0.033***
Philanthropy	0.028*	0.075***	0.039***	0.073***	-	0
Privacy and IT	0.022*	0.039***	-	-	0.276***	0.124***
Product Safety	0.048***	0.002	0.059***	0.062***	**0.429*** **	**0.216*** **
Public Health	0.022**	0.011*	-	-	0.029	0.074***
Recycling	-	-	-	-	0.119***	-
Remuneration	0	0.054***	**0.117*** **	0.113***	0	**0.223*** **
Reporting Quality	0.123***	-	0.107***	-	-	0
Resource Efficiency	0.014	**0.114*** **	**0.135*** **	-	-	-
Responsible Marketing	0	0.033***	0	0.002	-	0.081***
Shareholders	-	-	0.111***	0.089***	-	-
Site Closure	0.008	0	-	-	-	-
Supply Chain	**0.253*** **	0.061***	0.042**	0.05***	0.188***	0.128***
Sustainable Finance	0.108***	0.079***	0.063***	-	0.275***	0.098***
Systemic Risk	-	0.053***	-	-	0.349***	0.103***
Taxes	0.052***	0.01	0.03**	-	-	-
Toxic Spills	0	-	0.001	-	-	0.113***
Unions	-	-	0.013	-	-	0.158***
Waste	0	0.005	0.035***	0.009	-	0.186***
Water	0.03**	0.016***	0.028**	0	0.035	0.175***
Unclassified Indicators	Yes	Yes	Yes	Yes	Yes	Yes
R2	0.90	0.98	0.92	0.96	0.79	0.99
Firms	924	924	924	924	924	924

An important consideration to be done is that the choice to do a non-negative least squares regression is motivated by the willingness of finding a solution where all the returned coefficients are positive: this is coherent with the fact that, in the aggregation rule of the rating (assumed to be a linear function), the weights associated to the categories are positive.

In order to assess the robustness of the model, it is computed the coefficient of determination (R2), which provides a measure of how well observed outcomes are replicated by the model, based on the proportion of total variation of outcomes explained by the model. According to the average value of R2 (0.92), the results of the conducted analysis can be considered reliable.

The following formula shows the regression model through which it is possible to compute the weights.

$$R_{fk} = \sum_{j \in (1,m)} C_{fkj} \times w_{kj} + \epsilon_{fk} \qquad (3.1)$$
$$w_{kj} \geq 0.$$

The coefficients (w_{kj}) found through this formula can be interpreted as the category weights, and they reflect the relative importance of a particular category for a specific rating provider. From the coefficients computation results, it emerges that there are substantial differences in the weights for different raters. For example, the three most important categories for KLD are Climate Risk Management, Product Safety, and Remuneration while, for Vigeo Eiris, they are Diversity, Environmental Policy, and Labor Practices. As a consequence, it is clear that there is no agreement

among these two raters on which are the three most important categories to be considered. On the other hand, there are some categories that present zero weight for all rating agencies, such as Clinical Trials and Environmental Fines, GMOs, and Ozone-Depleting Gases. This could mean that these categories are not considered to be statistically relevant for any of the aggregate ratings.

These findings lead to the conclusion that different ESG rating providers have different perspectives about which are the most relevant categories. This testifies the presence of a substantial divergence between the weights adopted by the different raters.

3.2.1.5. Aggregate Analysis

After having demonstrated the existence of scope, measurement and weight divergence, this section is aimed at evaluating the contribution of these three elements to the total amount of ratings divergence. Two methods have been adopted in order to arrive to the same result, but this work focuses only on the first one.

In particular, this method starts with the arithmetical decomposition of the difference between two ratings into the differences due to scope, measurement and weights. This allows to identify the variation caused by each of these three pillars. The variance decomposition is based on the taxonomy, the category scores, and the aggregation weights estimated above. This procedure identifies how scope, measurement, and weights divergence lead to the difference between two ESG ratings assigned to a specific corporation. Scope divergence is computed by considering only the categories that are exclusively contained in one of the two

ratings. Measurement divergence is assessed by calculating both ratings with identical weights, so that differences can only be due to differences in measurement. Finally, weights divergence is represented by the remaining amount of the total difference. Table 3.6 highlights that the measurement divergence is the most relevant driver of ESG ratings divergence, followed by scope divergence and weights divergence. Indeed, measurement divergence causes an average absolute shift of 0.54 standard deviations, scope divergence generates an average absolute shift of 0.48 standard deviations, while weights divergence leads to an average absolute shift of 0.34 standard deviations.

Table 3.6: Rater pairs. Source: Berg et al (2020) [27]

		Scope	Measurement	Weights	Fitted	True
KLD	Sustainalytics	0.27	0.6	0.29	0.73	0.76
KLD	Vigeo Eiris	0.4	0.6	0.27	0.78	0.79
KLD	RobecoSAM	0.28	0.67	0.31	0.8	0.81
KLD	Asset4	0.33	0.6	0.45	0.8	0.86
KLD	MSCI	0.85	0.51	0.51	0.71	0.77
Sustainalytics	Vigeo Eiris	0.39	0.51	0.24	0.54	0.6
Sustainalytics	RobecoSAM	0.32	0.55	0.16	0.58	0.64
Sustainalytics	Asset4	0.19	0.45	0.32	0.53	0.65
Sustainalytics	MSCI	0.86	0.52	0.53	0.76	0.82
Vigeo Eiris	RobecoSAM	0.3	0.39	0.11	0.6	0.61
Vigeo Eiris	Asset4	0.33	0.5	0.19	0.55	0.64
Vigeo Eiris	MSCI	0.78	0.55	0.43	0.81	0.85
RobecoSAM	Asset4	0.26	0.51	0.14	0.62	0.71
RobecoSAM	MSCI	0.86	0.6	0.57	0.83	0.89
Asset4	MSCI	0.85	0.57	0.56	0.78	0.89
Average		0.48	0.54	0.34	0.69	0.75

3.2.2. Italian Market Divergence Analysis

Another relevant paper investigating the sources of ESG ratings divergence is the one by Capizzi et al (2021) [28]. The study is conducted on a sample of Italian listed companies' data coming from six different ESG rating providers: MSCI, Refinitiv, S&P Global, Arabesque S-Ray (hereafter: Arabesque), Truvalue Labs (hereafter: Truvalue), and Inrate. More in depth, the study aims at investigating why ESG ratings diverge. In order to do so, the authors develop a framework able to compare different ESG rating methodologies, splitting the ESG ratings' differences into two main components: ESG values and ESG weights at the pillar and category level. The first important result coming from this study is the confirmation of the low correlation between ESG ratings of different rating providers found by Berg et al (2020) [26]. Furthermore, the second important finding of the paper is that the weight component is the principal driver of ESG ratings divergence with respect to the value component. The weights divergence is particularly concerning, and it indicates a conflict on the relevance of the different ESG categories and pillars. Consequently, even if a corporation receives the same score value for its ESG performance, the final ESG ratings generated by various rating agencies might still differ considerably. Finally, the third main outcome is that the social and the governance indicators are the main responsible for rating divergences. More in depth, for what concerns the social categories, they are the widest in number, with the major part of differences attributable to the weights. On the other hand, for what concerns the governance pillar, not only the weights but also the values are responsible for a significant percentage of divergence among different rating agencies: this evidence derives from a higher subjectivity level in the evaluations of this pillar, with some performance indicators included by one rater, but not by the others.

To conclude, the papers by Berg et al (2020) [27] and by Capizzi et al (2021) [28] demonstrate that ESG rating divergence exists, and that it is not only caused by differences in the analysts' evaluations, but also by disagreement about the underlying building blocks of the ESG rating providers' methodologies.

3.2.3. The Impact of Divergence on Financial Performances

After having analysed the topic of the ESG ratings divergence, this section is dedicated to the implications that it may cause, through the analysis of the study by Billio et al (2020) [29]. Before the draft of this paper, the literature was focused on the analysis either of the disagreement of the rating companies or of the financial performance of the ESG portfolios, treating these two aspects separately. In this respect, this article enters the scene filling this gap, trying to investigate the causal relationship between sustainability rating divergence and ESG portfolios' performance.

3.2.3.1. Methodology and Results of the Study

Starting from the sustainability ratings provided by four rating agencies (namely, STOXX, Dow Jones, Refinitiv and MSCI), the authors of the paper create two different types of portfolios:

- *ESG agreement*: it is the portfolio composed by the stocks of those companies that are considered as ESG leaders by all the rating agencies.
- *NonESG portfolio*: it is the one built through a negative screening approach, including all the stocks of those companies that are excluded by the investment universe.

More in depth, the analysis is divided into two periods: in particular, the first time window ranges from December 1999 to December 2004, while the latter begins in

January 2005 and ends in January 2020. The decoupling point is set in coincidence of 2005 since the term ESG has been coined in that year.

The first step of the analysis is based on the evaluation of the Sharpe and Sortino ratios. They are two indicators that allow to understand the return of an investment compared to its risk. In particular, the Sharpe ratio is computed as the average return earned in excess of the risk-free rate per unit of volatility.

$$Sharpe\ Ratio = \frac{Return\ of\ Portfolio - Risk\ Free\ Rate}{Portfolio\ Excess\ Return's\ Std.\ Dev.} \tag{3.2}$$

On the other hand, the Sortino ratio is a variation of the Sharpe ratio and it utilises the asset standard deviation of negative portfolio returns instead of the total standard deviation of portfolio returns.

Table 3.7: Portfolios' performances. Source: Billio et al (2020) [29]

	Sharpe ratio		*Sortino ratio*	
	ESG	nonESG	ESG	nonESG
2000-2004	0.254	1.224	0.377	1.535
2005-2019	0.870	1.044	1.153	1.287
2000-2019	0.691	1.092	0.945	1.361

From the Table 3.7 it is possible to observe that the values of the two indicators are higher for the nonESG portfolio in the first period, while in the second one the two portfolios present similar values. These results lead to the conclusion that the nonESG portfolio performs better compared to the ESG agreement portfolio from a

financial perspective. However, considering the second portion of the analysis' horizon, the very similar values of the two indicators suggest a possible change of direction in the financial investors' preferences after 2005, year in which there has been the introduction of the ESG corporations' characteristics. As a consequence, the two different types of portfolios experience closer performances in this second time period.

The next phase of the analysis consists in the computation of the Jensen's alpha, which is a risk-adjusted performance measure that represents the average return of a portfolio above the one predicted by a factorial model. In this case, the authors adopted the Carhart four-factor model, which is an extension of the Fama–French three-factor model; it enables to derive the portfolio performance measuring the impact of:

- *The market risk.*
- *The outperformance of small versus big companies.*
- *The outperformance of high book/market versus small book/market companies.*
- *The momentum factor.*

Once these four factors have been calculated and the two alphas for the two portfolios are derived from them, their difference could reveal if the corporate stocks are impacted by Environmental, Social and Governance topics. To evaluate the statistical significance of the difference between the two alphas, the authors of the paper created a long-short portfolio built between the ESG agreement portfolio (long position) and the nonESG one (short position). If the alpha of this portfolio results positive (negative) and significant, this would indicate that the ESG portfolio presents higher performances (is outperformed by) than the nonESG portfolio, while a nonsignificant alpha would mean that there is no statistical difference between the two portfolios. Looking at the alphas on the long-short portfolio, they

are not significant for both the considered periods and, therefore, it appears that there are not relevant differences in the performances of the two types of portfolios.

According to the authors of the study, these results can be explained by making reference to the ESG ratings divergence among the rating agencies, which disperses the effect of the ESG investors' preferences on stocks' prices, arriving to the point that, even when there is agreement, the ESG effect is attenuated and its impact on performances is neutralised.

As a final consideration, the authors conclude that the portfolios' financial performances would be different in the case in which the principal ESG rating providers were able to find an agreement on the methodology adopted in order to assess the companies' grades. If this were the case, the institutional investors would be able to concentrate their financial resources on a common pool of target securities, thus having a significant impact on asset prices.

3.2.4. Investors' Confusion: A Possible Link between Divergence and Performance

This section is dedicated to show how the lack of sustainable portfolios' performance could be caused by the ESG ratings divergence through the confusion created by it in the financial investors. In particular, this section investigates what could be the impact of the confusion caused by a change in the sustainability ratings (namely, an upgrading or a downgrading) on the stock price of a specific corporation.

In this regard, it is worthy to refer to the paper by Rzeznik et al. (2021) [30] This article analyses the specific case of the adoption of a new rating scoring methodology by Sustainalytics in 2019, which led to an inversion of the rating scale: the study is conducted in order to understand the consequences of this disruptive event on the investors' behaviour and, consequently, on the stock prices.

More in depth, the old and the innovative methodologies adopted by Sustainalytics have in common the same scale from 0 to 100, but the main difference is represented by the interpretation of those numbers: indeed, in the old method a higher rating signals lower ESG risk, while in the new method a lower rating indicates lower ESG risk. To make a practical example, a company that is labelled as the best-in-class in terms of ESG risk would have previously received a score closer to 100; with the adoption of the new method, its score will be closer to 0. The inversion of the rating scale has as main consequence the reduction of the ESG scores for a huge number of companies: this decline could be wrongly perceived as a downgrade, so the investors could be incentivised to rebalance their portfolio in line with their expectations, thus impacting the stock prices. In particular, this paper highlights

that one standard deviation reduction in the ESG score leads to a decrease of 1.08% in the monthly four-factor abnormal return.

In order to guarantee that the only impacting factor on the ESG ratings is the inversion of the scale and not the revaluation of the ESG risk, the authors of the paper exploit two distinct methodologies:

- On the one hand, they adopt a variable aimed at measuring the extent to which a company's ESG risk ranking changes relative to its peers before and after the adoption of the new method; furthermore, they introduce a second variable to measure if the change in a corporation's rating is also caused by a reclassification of its ESG risk within Morningstar's ESG Rating Assessment.
- On the other hand, the authors restrict the pool of the companies, considering only the ones that experience a downgrade without reporting a negative change in their ESG exposure. These companies are unambiguously good or unvaried in terms of the new ESG Risk rating, so it is reasonable to expect that they will experience an increase or no change in their stock price. However, this study demonstrates that the investors look at these ESG ratings reductions as a downgrade of the companies, leading to a significant negative implication for the pricing.

The next paragraph focuses on a more comprehensive overview on the analysis conducted by the authors of the article.

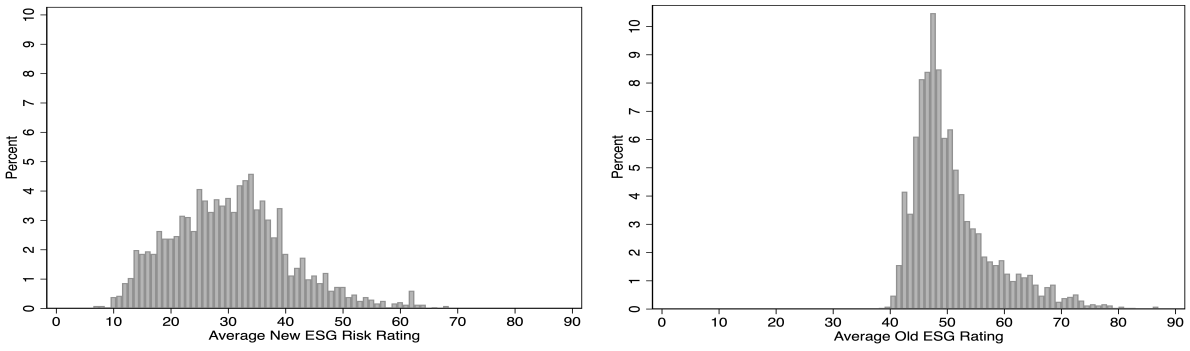
3.2.4.1. Data and Variable Construction

The data on which this paper is built come mainly from two distinct sources:

- The old ESG ratings and the new ESR Risk ratings are taken from Morningstar Direct, checking that they are the same present on Yahoo! Finance.
- The new ESG Risk ratings are collected during the transition period from Sustainalytics.

Starting from this data, the authors develop some relevant variables in order to conduct the analysis:

- $Post_t$: due to the fact that Morningstar publishes the monthly ESG scores at the beginning of the subsequent time bucket, the new September ESG Risk rating is available starting from the beginning of October 2019. Indeed, this variable assumes a value equal to 1 if the month is greater than October 2019 to indicate the adoption of the new ESG Risk rating.
- ΔESG_i: it is the difference in stock i's ESG rating score between October and September 2019.
- $\Delta \overline{ESG_i}$: it is the difference between the mean new ESG rating in the three months period after the adoption of the new methodology, and the average old rating in the three months period before this change.



a) Average new ESG rating distribution. b) Average old ESG rating distribution.

Figure 3.2: Average old and new ESG rating distribution. Source: Rzeznik et al (2021) [30]

Figure 3.2 (a) and (b) show respectively the distribution of the average ESG ratings in the three months before and after the introduction of the new methodology. Figure 3.2 (a) highlights that most of the corporations are attributed an average ESG score between 40 and 80, while in Figure 3.2 (b) most of the firms' average ratings are included in the range from 10 to 60. This stylized fact testifies the inversion in the ratings moving from the old to the new ESG rating methodology.

Other two relevant indicator variables, which isolate the direction of the change in the ratings, are:

- $ESG\ Pseudo{-}Downgrade$: it is equal to 1 if the stock i's change in its average ESG rating is in the lowest quartile of the $\Delta\overline{ESG_i}$ distribution, so it is equal to 1 for the corporations with the largest negative change in their ratings. An example could be the case of Microsoft which, passing from 75 to 13,8 in the ESG rating score, has a $ESG\ Pseudo{-}Downgrad$e equal to 1. The companies with a value of this variable equal to 1 have a very negligible ESG risk both before and after the adoption of the new method; indeed, the decline of their rating is not a signal of a worsened ESG risk exposure, but it is only due to the inversion of the rating scale.
- $ESG\ Pseudo{-}Upgrade$: it is equal to 1 if a company's average ESG rating is greater after the adoption of the new methodology; in this case it is not used the quartile because only the 6% of the firms in the sample under analysis registers a higher ESG rating after the change.

Furthermore, the variations of ESG ratings could be caused also by newly available information on changes on the ESG risk exposure of the companies. In order to monitor and control changes in the ratings due to this fact, the authors of the paper conduct additional robustness analysis. These robustness analyses are based on three main variables:

- $\Delta ESG\ Rank_i$: it is defined as the difference in the stock's relative ranking in the first month after the adoption and in the last month before the adoption of the new method. To compute this variable, all the companies in the sample are ranked assigning the firsts positions to the firms with the lowest ESG risk.
- $Classification\ Upgrade_i$: this variable highlights a potential variation in the company ESG risk exposure according to the Morningstar Classification. It assumes a value equal to 1 if the company experiences an increase in the Morningstar Classification after the implementation of the new methodology.
- $\Delta Sustainalytics\ Rating_i$: it is defined as the difference between the company i's average ESG Risk rating computed with the same new methodology in the three months after and before its adoption. This variable is designed to detect possible exogenous events that could affect the ESG score of a company during the transition between the old and the new methodology.

At this point, in order to assess if the new ESG Risk ratings affect the investors' trading behaviours, the authors compute the abnormal returns based on the following data:

- Daily returns, Prices and Shares outstanding coming from the Center for Research in Security Prices (CRSP).
- Daily and Monthly risk factors for Fama and French three- and Carhart four-factor model from Kenneth French's website. To compute the risk factors and expected returns for each stock, they use daily stock excess returns from July 2018 to June 2019.

The abnormal returns of a company are computed the following month by subtracting the expected return from the actual return.

3.2.4.2. Methodology and Findings of the Study

The analysis of this study is focused on trying to understand the impact that the new ESG Risk rating methodology has on the abnormal returns. In particular, if a considerable portion of financial investors is confused by this change and adjusts its portfolio according to their new expectations, then it is reasonable to expect that abnormal returns will be negative (positive) for those companies that face a "pseudo-reduction" (pseudo-increase) in their ESG score, even if the ESG risk exposure of these corporations is not changed. Furthermore, the authors of the paper predict that what said above is caused by the changes in stock participation by less informed financial actors, namely the retail investors, and not by changes in the behaviour of more sophisticated and institutional investors.

The starting point of the analysis consists in defining the relationship between the stock's abnormal returns in month t ($ARet_{i,t}$) and the stock i's change in its ESG score (ΔESG_i) caused by the adoption of the new Sustainalytics ESG Risk rating method. In particular, $ARet_{i,t}$ is calculated using either a single-, three- or four-factor model. The described equation is shown below.

$$\mathrm{ARET}_{i,t} = \gamma_0 + \gamma_1 \Delta \mathrm{ESG}_i \times \mathrm{POST}_t + \mathrm{D}_i + \mathrm{D}_{s,t} + \varepsilon_{i,t}, \tag{3.3}$$

If the authors' expectations about the effect of financial investor confusion are correct, then it is expected that the coefficient $\gamma 1$ will be positive and significant, indicating that a reduction (increase) in ESG_i pushes investors to sell (buy) the stock, leading to negative (positive) abnormal returns starting from October 2019. Table 3.8 shows the results of the regression.

Table 3.8: Regression results. Source: Rzeznik et al (2021) [30]

Panel A:	Single Factor AbnRet				Three Factor AbnRet				Carhart AbnRet			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\Delta ESG_i \times Post_t$	0.0819***				0.0482***				0.0556***			
	(5.60)				(3.29)				(3.76)			
$\Delta\overline{ESG_i} \times Post_t$		0.0832***				0.0492***				0.0566***		
		(5.58)				(3.28)				(3.76)		
ESG Pseudo Downgrade$_i \times Post_t$			-1.7262***	-1.5670***			-0.8041**	-0.6617*			-1.0765***	-0.9353**
			(-4.35)	(-3.94)			(-2.01)	(-1.65)			(-2.64)	(-2.29)
ESG Pseudo Upgrade$_i \times Post_t$				2.3902				2.1378				2.1196
				(1.58)				(1.41)				(1.40)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768
R^2	0.191	0.191	0.190	0.190	0.179	0.179	0.178	0.178	0.176	0.176	0.175	0.176

Looking at the table, the coefficients of $\Delta ESG_i \times Post_t$ and $\Delta\overline{ESG_i} \times Post_t$ are positive and very significant independently from the methods adopted to compute the abnormal returns. This stylized fact highlights that companies that experience a more significant decline (increase) in their new ESG Risk rating relative to the old one, report more negative (positive) abnormal returns. More in detail, a reduction of one standard deviation in the ESG score of a specific company corresponds to a 1,08% decrease in its monthly abnormal returns.

A further step to analyse the relationship between the variation in the ESG score and its impact on the abnormal returns consists in investigating whether abnormal returns are differently impacted for companies that experience huge ESG rating declines, and for firms that report an increase in their rating. The coefficient on the interaction term $ESG\ Pseudo{-}Downgrade_i \times Post_t$ is negative and significant,

meaning that abnormal returns are smaller for corporations that experienced a very significant reduction in their ESG score compared to their industry peers. At this point, it is important to highlight that a huge portion of the companies that have a value of 1 for $ESG\ Pseudo{-}Downgrade_i$ also have absolutely no change in their Morningstar Classification: this fact testifies that the reduction in the abnormal returns is very likely to be caused by a wrong interpretation in the meaning of the change in the rating scale. Differently from what has been noted for the $ESG\ Pseudo{-}Downgrade_i$, the coefficient on the interaction term $ESG\ Pseudo{-}Upgrade_i \times Postt$ is positive but insignificant; this finding could be explained by the small number of companies that experienced an increase in their ESG rating.

To better investigate the effect of changes in the ESG rating on the abnormal returns, the authors of the paper conduct a semi-parametric regression where:

- The *dependent variable* is represented by the company's change in its abnormal return.
- The *independent variable* is represented by the variation in the company's average ESG rating after the adoption of the new methodology.

The image below shows the results of this analysis.

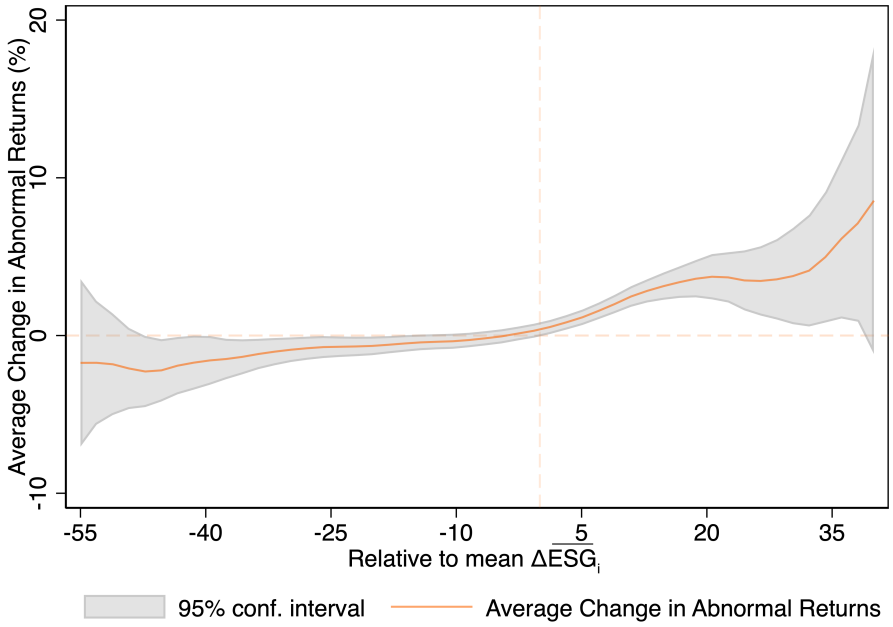


Figure 3.3: Semi-parametric regression results. Source: Rzeznik et al (2021) [30]

It is possible to see that the change in the ESG score of a company relative to the mean leads to a corresponding variation in the firm's abnormal returns. More in detail, the change in the abnormal returns declines (increases) when the variation in the ESG score becomes more negative (positive). This observation testifies that financial investors that care about ESG ratings have a significant impact on the stock returns, even in the case in which they do not give a proper interpretation to what caused the change in the company's ESG score.

The next step of the study consists in trying to reinforce the results showed above by analysing the impact of new information on returns and also by restricting the pool of companies, considering only the ones for which the interpretation of the variation in their ESG rating is unambiguous. In order to separate the effect of the

scale inversion from the possible effect of new information about the ESG risk exposure of a company on its ESG rating, the authors checked two different types of information:

- The first one is represented by the possible occurrence of an ESG event during the transition period from the old rating methodology to the new one. Here it is used the same regression model that has been already explained above, adding as independent variable $\Delta Sustainalytics\ Rating_i$.
- The second one is represented by the information that is generated by the innovative Sustainalytics rating methodology on the variation of a company's ESG risk exposure; this kind of information is analysed introducing two variables in the analytical model: $\Delta ESG\ Rank_i$ and $Classification\ Upgrade_i$.

Table 3.9: Regression results (2). Source: Rzeznik et al (2021) [30]

Panel B:	Single Factor AbnRet				Three Factor AbnRet				Carhart AbnRet			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\Delta ESG_i \times Post_t$	0.0864***				0.0501***				0.0579***			
	(5.88)				(3.39)				(3.90)			
$\Delta \overline{ESG}_i \times Post_t$		0.0869***				0.0504***				0.0584***		
		(5.83)				(3.36)				(3.87)		
ESG Pseudo Downgrade$_i \times Post_t$			-1.7660***	-1.6019***			-0.8155**	-0.6737*			-1.0939***	-0.9519**
			(-4.44)	(-4.02)			(-2.03)	(-1.68)			(-2.68)	(-2.32)
ESG Pseudo Upgrade$_i \times Post_t$				2.5181*				2.1756				2.1794
				(1.67)				(1.44)				(1.45)
ΔESG Rank$_i \times Post_t$	-1.6324**	-1.6349**	-1.3451*	-1.4726**	-0.5588	-0.5604	-0.3805	-0.4906	-0.7856	-0.7880	-0.5880	-0.6984
	(-2.38)	(-2.38)	(-1.95)	(-2.15)	(-0.82)	(-0.82)	(-0.55)	(-0.72)	(-1.14)	(-1.14)	(-0.85)	(-1.01)
ΔSustainalytics Rating$_i \times Post_t$	0.1402	0.0775	0.0801	0.0677	0.0878	0.0515	0.0528	0.0421	0.0866	0.0446	0.0463	0.0355
	(0.98)	(0.54)	(0.54)	(0.47)	(0.63)	(0.37)	(0.37)	(0.30)	(0.62)	(0.32)	(0.33)	(0.26)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768
R^2	0.192	0.192	0.190	0.191	0.179	0.179	0.178	0.178	0.176	0.176	0.176	0.176

The output table above shows that the coefficient of Sustainalytics Rating$_i$ is not significant, meaning that the abnormal returns are not impacted by a change in the mean Sustainalytics ESG Risk score. This means that the majority of companies does

not report any ESG event able to potentially affect investors' behaviour and, consequently, to have an impact on returns. For what concerns the coefficient of the variation in the relative ranking of the firm - ESG Rank$_i$ - it is significant in the case in which the dependent variable is the single-factor abnormal return. This result shows that abnormal returns are higher when the company experiences an improvement in the rank.

Table 3.10: Regression results (3). Source: Rzeznik et al (2021) [30]

Panel C:	Single Factor AbnRet				Three Factor AbnRet				Carhart AbnRet			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\Delta ESG_i \times Post_t$	0.0810***				0.0439***				0.0514***			
	(5.59)				(3.03)				(3.54)			
$\Delta \overline{ESG}_i \times Post_t$		0.0814***				0.0441***				0.0518***		
		(5.53)				(3.00)				(3.51)		
ESG Pseudo Downgrade$_i \times Post_t$			-1.5834***	-1.4692***			-0.5845	-0.4869			-0.8584**	-0.7616*
			(-3.94)	(-3.62)			(-1.45)	(-1.19)			(-2.09)	(-1.83)
ESG Pseudo Upgrade$_i \times Post_t$				2.2756				1.9460				1.9301
				(1.51)				(1.29)				(1.29)
ΔSustainalytics Rating$_i \times Post_t$	0.1225	0.0637	0.0733	0.0587	0.0871	0.0552	0.0613	0.0488	0.0833	0.0459	0.0525	0.0401
	(0.86)	(0.44)	(0.50)	(0.41)	(0.63)	(0.40)	(0.44)	(0.36)	(0.60)	(0.33)	(0.37)	(0.29)
Classification Upgrade$_i \times Post_t$	-0.3289	-0.3303	-0.3041	-0.2310	-0.5636	-0.5643	-0.5782	-0.5156	-0.6384	-0.6391	-0.6378	-0.5757
	(-0.80)	(-0.80)	(-0.73)	(-0.56)	(-1.37)	(-1.37)	(-1.40)	(-1.24)	(-1.52)	(-1.52)	(-1.51)	(-1.36)
Classification Donwgrade$_i \times Post_t$	0.0232	0.0162	0.3911	0.2831	0.2585	0.2546	0.5078	0.4154	0.1910	0.1853	0.4514	0.3598
	(0.04)	(0.03)	(0.68)	(0.50)	(0.46)	(0.46)	(0.90)	(0.75)	(0.34)	(0.33)	(0.79)	(0.65)
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768	12768
R^2	0.191	0.191	0.190	0.191	0.179	0.179	0.178	0.179	0.176	0.176	0.176	0.176

Similar findings are shown in the Table 3.10, which contains information about the indicator variables signaling if the company reports a downgrade or an upgrade in its Morningstar Classification. In this case, the coefficients on the Morningstar Classification upgrade (downgrade) variable are always (generally) insignificant. Looking at the R2 of Tables 3.9 and 3.10, they result equal to the one shown in Table 3.8, meaning that abnormal returns are impacted by the wrong interpretation of the rating by the financial investors and not by the availability of new information on the market.

As a final test, the authors of the paper reduce the pool of companies to be analysed, considering only the ones whose decline in the ESG rating has not a corresponding negative variation in the sustainability ranking. Analysing only this set of companies, it is possible to focus the analysis uniquely on the cases in which the investment portfolio rebalances are caused by the misunderstanding of the new rating methodology. If investors wrongly interpret the rating scale, thinking that a company's ESG risk exposure has increased when its rating decreases, then the result of this analysis should remain unchanged also in the case with the restricted pool of firms. The table below shows the results of the regression run on this sample of companies. In line with the expectations of the authors, the coefficients of ESG_i and the corresponding R2 remain similar to the ones of Table 3.8 reported above. At confirmation of the fact that investors are misled by the new rating methodology, the best-in-class companies in terms of ESG risk exposure keep on experiencing the most negative abnormal returns.

The results obtained above show that stock prices can be impacted by the financial investors' confusion about how to interpret the ESG ratings obtained with the new methodology. However, this impact should be assumed as temporary, and it could be expected to be reversed once the investors realize their wrong interpretation. The results of this study testify the magnitude of the ESG ratings divergence problem, showing how the confusion generated by misaligned sustainability evaluations could lead to a lack of ESG portfolios' performances.

4 Research Objectives and Methodology

This chapter illustrates the research questions of the present work and the methodology adopted in order to investigate them. After having introduced the objectives and the theoretical methodology in the first two sections, the last ones are dedicated, on the one hand, to the description of the data sample used to feed the model and, on the other, to the implementation of the analysis.

4.1. Objectives of the Analysis

The chapter [3] has been dedicated to the exploration of the literature concerning the implications between ESG and financial and economic performances. Starting from the definition of the ESG paradigm, and investigating the papers treating the topic about its relationship with corporate performance (e.g., Friede et al (2015) [23]), it has emerged that in most of the cases ESG investing enhances the returns, and so the value, of both corporations and sustainable-oriented portfolios. This positive correlation is leading financial investors and firms to pay more attention to the sustainability themes, generating the need for reliable and synthetic metrics able to measure in a comprehensive way the environmental, social and governance performances. Hence, the section [2.3] describes the assessment methodologies

adopted by the main rating providers; in this regard, the lack of a common measurement standard brings to different (and sometimes opposite) ESG grades for the same target companies, as shown by Berg et al (2020) [27] and Capizzi et al (2021) [28]. This rating divergence phenomenon creates confusion in the investment processes, leading to a dispersion of the financial investors' preferences. Indeed, as demonstrated by Billio et al (2020) [29], this stylized fact jeopardises the potential benefit of the ESG portfolios, making their performance being equal to the traditional ones. This finding undermines the positive relationship between ESG and corporate performances which has been demonstrated by most of the previous literature.

Having explained the state of the art of the existing literature, the present work started from this context with the aim of investigating the reaction of financial markets to rating agencies' ESG grade updates, understanding whether possible variations in the sustainability ratings of the corporations have a direct impact on their value. In doing so, the purpose of this work was to try to answer to the following research questions:

I. If the changes in the ESG ratings of the corporations exert a direct impact on their market value.

II. If the reaction of the financial market to the sustainability ratings updates has become stronger in the last two years, due to a higher consciousness of the investors about the ESG-related topics. This research question starts from the evidence that sustainable investment is gaining more and more relevance in the preferences of retail and institutional investors, as reported by the Global Sustainable Investment Alliance (2021) [6].

Furthermore, the present work addresses the topic of the sustainability rating divergence, showing how rating agencies diverge in their evaluations and investigating whether financial markets exhibit different reactions to their ratings updates. In particular, among the five rating agencies presented in the section [2.3], only the cases of Refinitiv and MSCI are debated in the analysis.

In order to deepen the above-mentioned research questions, the methodology adopted is the Event Study, which is presented in detail in the following section, where the theoretical methodology and its historically demonstrated robustness are illustrated (sections [4.2] and [4.2.3], respectively). In this context, the innovation brought by the analysis of the present work is to apply this consolidated methodology in the ESG ratings universe, in order to try to respond to the questions reported above.

4.2. Event Study Methodology

The Event Study methodology is used in order to assess the impact of an economic event on the valuation of a corporation. This quantitative approach is built on one fundamental pillar: the rationality of financial markets, meaning that every single economic event is instantaneously incorporated in the price of financial assets. As a consequence, the effects of an economic event can be assessed by observing the company's price in a specific short time window. In the academic literature, the Event Study methodology has been adopted for a wide range of different firm-specific phenomena, for example merger and acquisitions and earning announcements. Furthermore, this approach can be used not only to investigate strictly economic events, but it could also be applied to a heterogeneous variety of events, such as changes in regulations or legal-liability cases. Some relevant applications of this methodology will be discussed later in the dedicated section [4.2.3]. In the majority of cases, the methodology is aimed at evaluating the consequence of an economic impact on the common equity of the firms, but it could be potentially applied to debt securities. As explained by Campbell et al (1997) [31], the typical structure of an Event Study is made up of 7 steps:

- *Event definition*: it consists in defining the event of interest and the event window, which is the time horizon in which the analysis of the involved company's stock prices is conducted. The choice of an appropriate event window is crucial in order to capture the effect of the economic event on the price of the stocks. Furthermore, the periods prior to and/or after the event could also be of interest, and thus included in the analysis: for example, in some cases it is possible that the market acquires relevant pieces of information before the actual occurrence of the event.

- *Selection criteria*: in this step there is the definition of the selection criteria used in order to determine which companies have to be included in the sample of interest. One example of these criteria could be represented by the membership of a firm in a specific sector. Moreover, another important task to be conducted at this stage it is the listing of the potential biases which could be implicitly introduced through the selection of the pool of companies.

- *Normal and abnormal return*: in order to evaluate the effect of an event, this methodology recurs to the measurement of the abnormal returns. The abnormal return is defined as the actual ex-post return experienced by the company over the event window minus the normal return of the security over the same time period. The normal return, instead, is the return expected in a normal condition, without the occurrence of the specific economic event. More in detail, there are two different ways in order to model the normal return. The first one is the Constant-Mean-Return Model, in which the normal return of a specific security is constant over time. The second option is represented by the Market Model, which assumes a stable relation between the market return and the normal return. In particular, after having computed the normal return, the abnormal return could be found through the following equation:

$$\epsilon_{it}^{*} = R_{it} - E[R_{it}|X_t] \tag{4.1}$$

where ϵ_{it}^{*}, R_{it} and $E(R_{it})$ are the abnormal, actual, and normal returns, respectively, for the time period t. X_t is the conditioning information for the normal performance model.

- *Estimation procedure*: after the selection of the normal performance model, the following step consists in the estimation of the model's parameters; in order to do so, it is used a sample of data defined as the estimation window. Usually, the estimation window does not include the focal event to avoid that it could alter the estimation of the normal performance model's parameters.

- *Testing procedure*: having estimated the set of all the needed parameters, the abnormal returns can be computed. Furthermore, in this phase there is the definition of the testing framework, which is important to understand and assess the relevance of the event on the stock prices fluctuations. Another crucial point of this phase is the decision about the null hypothesis and the aggregation rule for the abnormal returns.

- *Empirical results*: this step aims at presenting the empirical results obtained after the setting and the running of the model, in order to extract insights from the analysis.

- *Interpretation and conclusions*: this is the final phase, in which the interested stakeholders have the opportunity to understand possible causal effects between the event and the prices fluctuations.

4.2.1. Models for Measuring Normal Returns

In the literature there are several models available in order to compute the normal returns of a security. These approaches can be clustered into two main categories: statistical and economic.

4.2.1.1. Statistical Models

These models are based on statistical hypotheses about the behaviour of the stock returns, and they are independent from any economic considerations. For what concern these models, they are usually founded on the assumption that stock returns are jointly multivariate normal and independently and identically distributed over time. This hypothesis about returns distribution is sufficient to completely define two statistical models that are presented in the following paragraphs, namely the Constant-Mean-Return Model and the Market Model. Moreover, this assumption allows to extract exhaustive results from the analysis, which are also robust to deviations from the hypothesis. The next three paragraphs are dedicated to the presentation of the three main statistical models.

Constant-Mean-Return Model

The Constant-Mean-Return model is considered one of the simplest models but, in the literature, the study conducted by Brown and Warner (1980, 1985) [32,33] has shown that the results obtained through this methodology are comparable to the

ones returned by more sophisticated models. This evidence is due to the fact that more sophisticated models often do not reduce the variance of the abnormal return with respect to this simpler version.

Going more in dept with the technical side of the model, it is described by the formulation reported below:

$$R_{it} = \mu_i + \xi_{it} \tag{4.2}$$

$$E[\xi_{it}] = 0 \tag{4.3}$$

$$Var[\xi_{it}] = \sigma_{\xi_i}^2 \tag{4.4}$$

where μ_i is the mean return for the *i*-th asset, R_{it} is the return of the security in the time bucket *t*, the ξ_{it} is the disturbance term.

Market Model

The Market Model is a statistical model that derives the return of a given security from the return of the market. It is a further improvement of the Constant-Mean-Return Model: it allows to reduce the variance of the abnormal returns, isolating them from the market's returns. As a consequence, it enhances the capability to observe relevant event's effects. The model assumes a linear relation between the two returns mentioned before. In particular, it is described by the following formulation:

$$R_{it} = \alpha_i + \beta_i R_m + \epsilon_{it} \tag{4.5}$$

$$E[\epsilon_{it}] = 0 \tag{4.6}$$

$$Var[\epsilon_{it}] = \sigma^2_{\epsilon_i} \tag{4.7}$$

where R_{it} represents the return in the period t of the i-th stock, while R_m represents the return of the market in the same time period. ϵ_{it} is the noise of the model, which is characterized by a mean equal to zero. α_i, β_i e $\sigma^2_{\epsilon_i}$ are the model's parameters, which have to be calculated in the estimation window. The reliability of the model can be captured by the parameter R^2: the higher its value, the smaller will be the variance of the abnormal return.

Other Statistical Models

In the literature, different statistical models have been applied in order to assess normal returns. The most common model are the Factor models, a family of approaches which describe the stock return through a multi-factor linear regression. The Market Model explained before belongs to this family, since it uses a single factor in order to evaluate the stocks' returns. Two examples of the Factor models are the following one:

- Multifactorial models, that take into account not only the market, but also industry-related indexes, as the ones proposed by Sharpe (1970) [34] and Sharpe et al (1995) [35].

- The second example is represented by a variant of the Factor model, in which the abnormal return is defined as the difference between the actual return and a portfolio of firms of similar size, where the size is given by the market capitalization. This model is built on the assumption that normal returns are proportionally related to the size of the company.

The benefits of adopting these more sophisticated models are not so greater with respect to the ones obtainable through a simpler single-factor Market Model. Indeed, the reduction of the variance of the abnormal return is not so relevant if compared with the one returned by considering only the market factor.

4.2.1.2. Economic Models

This family of models aims at reducing the number of parameters with respect to the ones employed in the statistical models, in order to return more simplified and constrained versions of the normal return models. The most common economic models available in literature are the following ones:

- *Capital Asset Pricing Model (CAPM)*: it has been proposed by Sharpe (1964) [36] and Lintner (1965) [37], and it is defined as an equilibrium theory where the expected return of a specific security is linked with the covariance with the market portfolio's return through a linear function.
- *Arbitrage Pricing Theory (APT)*: this model has been introduced by Ross (1976) [38], and it is an asset pricing theory where the expected return of a specific stock is assessed by its covariance through the estimation of multiple factors in absence of arbitrage.

4.2.2. Measuring and Analysing Abnormal Returns

This section is dedicated to the measurement of the abnormal returns. Among all the models that have been presented before, the selected one is the Market Model, but all the reasonings are identical for the others statistical models.

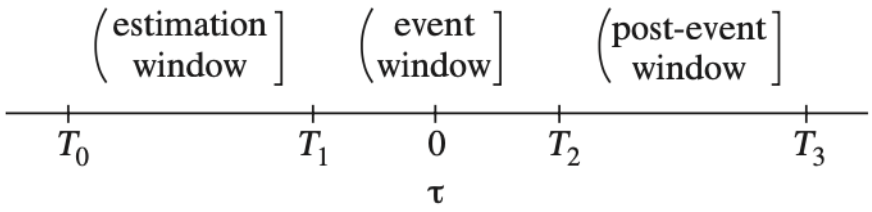


Figure 4.1: Event Study Timeline. Source: Campbell et al (1997) [30]

The first step consists in presenting the legend of some key terms and notations of the analysis. All the terms that are defined below refer to the timeline showed in the figure 4.1.

- τ is the variable describing the timing of the event. $\tau = 0$ represents the event date. The time interval going from $\tau = T_1 + 1$ to $\tau = T_2$ represents the event window, while the period between $\tau = T_0 + 1$ to $\tau = T_1$ is the estimation window.
- $L_1 = T_1 - T_0$ and $L_2 = T_2 - T_1$ are the length of the estimation window and the length of the event window.

The post-event window is the time period between $\tau = T_2 + 1$ and $\tau = T_3$, while $L_3 = T_3 - T_2$ represents its length. To make a practical example, if the event to be studied is an announcement on a specific date, T_2 will be $T_1 + 1$, and L_2 will be equal to 1. In this methodology, the abnormal return over the event window is considered as a proxy of the impact of the event on the company's value. As a consequence, the

focal event is interpreted to be exogenous with respect to fluctuations of the stock prices attributable to the market. While designing this model, one typical choice is the one not to make the estimation window and the event window to overlap. Structuring the model in this way, it is possible to estimate the needed parameters of the normal return model avoiding any influences by the event-related returns. Indeed, including the event window in the estimation window, there is the possibility that the event return carries out a significant influence on the normal return measure. Therefore, in this case the impact of the event is captured not only by the abnormal returns, but also by the normal ones. This is not consistent with the structural hypothesis of the methodology, which is founded on the assumption that the event impact is reflected by the abnormal returns.

4.2.2.1. Estimation of the Market Model

The formulation of the Market Model for the i-th security and for the observation τ is:

$$R_{i\,\tau} = \alpha_i + \beta_i R_{m\,\tau} + \epsilon_{i\,\tau} \tag{4.8}$$

The estimation window observations can be defined though a regression system, as shown by the formula below:

$$R_i = X_i \theta_i + \epsilon_i \tag{4.9}$$

where:

- R_i is the vector containing the returns of the stock in the estimation window.
- X_i is a matrix made up of a vector of ones in the first column, and the vector of market observations in the second column (R_m).
- θ_i is the parameter vector, which contains the model's parameters to be estimated (α_i, β_i).

Usually, the Ordinary Least Squares (OLS) procedure is the most efficient and consistent methodology in order to estimate the Market Model parameters. The estimators of the Market Model parameters returned by the application of the OLS are:

$$\hat{\theta}_i = (X_i' X_i)^{-1} X_i' R_i \tag{4.10}$$

$$\hat{\sigma}_{\epsilon i}^2 = \frac{1}{L_1 - 2} \hat{\epsilon}_i' \hat{\epsilon}_i \tag{4.11}$$

$$\hat{\epsilon}_i = R_i + X_i \hat{\theta}_i \tag{4.12}$$

$$Var[\hat{\theta}_i] = (X_i' X_i)^{-1} \sigma_{\epsilon_i}^2 \tag{4.13}$$

4.2.2.2. Statistical Properties of Abnormal Returns

Once the Market Model parameters estimates have been calculated through the OLS procedure, it is possible to compute and then analyse the abnormal returns. Defining $\hat{\epsilon}_i^*$ as the vector containing the company *i*'s abnormal returns during the considered event window, it is possible to compute it by subtracting the estimated normal return from the actual return:

$$\hat{\epsilon_i}^* = R_i^* - X_i^* \hat{\theta}_i \tag{4.14}$$

The computed abnormal returns will be normally distributed with a zero conditional mean and a conditional covariance matrix V_i, as shown by the formulations below:

$$E[\hat{\epsilon}_i^* | X_i^*] = 0 \tag{4.15}$$

$$V_i = I\, \sigma_{\epsilon_i}^2 + X_i^* (X_i' X_i)^{-1} X_i^{*\prime} \sigma_{\epsilon_i}^2 \tag{4.16}$$

From the first formula, it is possible to observe how the abnormal return vector is unbiased.

The second formula, instead, shows that the covariance matrix of the abnormal return vector it is made up by two parts:

- The first term of the sum represents the variance attributable to future disturbances.
- The second term of the sum reflects the additional variance given by the sampling error in $\hat{\theta}_i$. In particular, this sampling error will tend to zero as much as the length of the estimation window becomes wider; as a consequence, the abnormal returns across the time period will become independent asymptotically.

If the null hypothesis is true, thus the focal event is assumed not to have any influence on the mean and on the variance of the returns, then for the abnormal return vector of event window the following relation holds:

$$\hat{\epsilon}_i^* \sim N(0, V_i) \tag{4.17}$$

This equation returns the distribution of any abnormal return observations, and it represents the starting point for the subsequent abnormal returns' aggregation process.

4.2.2.3. Aggregation of Abnormal Returns

As explained Campbell et al (1997) [31], in order to detect if the event of interest has an impact on the security's price, it is necessary to perform the aggregation of the abnormal return observations. This aggregation can be conducted along two dimensions, namely through time and across securities. The first step consists in aggregating through time for a specific security; after this process is completed, it is then possible to proceed with the aggregation both across securities and through time.

In order to do so, it is necessary to define the matrix of Cumulative Abnormal Returns (CAR) for the security i, named as CAR_i. Furthermore, γ is a vector constituted by ones in the positions from $\tau_1 - T_1$ to $\tau_2 - T_1$, which is crucial in order to compute the aggregation of the abnormal returns.

$$\widehat{CAR}_i(\tau_1, \tau_2) = \gamma' \hat{\epsilon}_i^{\,*} \tag{4.18}$$

$$Var\left[\widehat{CAR}_i(\tau_1, \tau_2)\right] = \sigma_i^2(\tau_1, \tau_2) = \gamma' V_i \gamma \tag{4.19}$$

Applying the null hypothesis to the CAR, it becomes:

$$\widehat{CAR}_i(\tau_1, \tau_2) \sim N(0, \sigma_i^2(\tau_1, \tau_2)) \tag{4.20}$$

In order to test the null hypothesis H_0, it can be computed the Standardised Cumulative Abnormal Return, which is defined by the formula below:

$$\widehat{SCAR}_i(\tau_1, \tau_2) = \frac{\widehat{CAR}_i(\tau_1, \tau_2)}{\sigma_i^2(\tau_1, \tau_2)} \tag{4.21}$$

where $\sigma_i^2(\tau_1, \tau_2)$ is computed using $\sigma_{\epsilon_i}^2$ in the equation:

$$\hat{\sigma}_{\epsilon i}^2 = \frac{1}{L_1 - 2} \hat{\epsilon}_i{}' \hat{\epsilon}_i \tag{4.22}$$

Due to the fact that the null hypothesis is assumed to be true, the distribution of the $\widehat{SCAR}_i(\tau_1, \tau_2)$ is the Student *t*, with $L_1 - 2$ freedom degrees. According to the characteristics of the Student *t* distribution, it derives that:

- The expected value of $\widehat{SCAR}_i(\tau_1, \tau_2)$ is zero.
- The variance is equal to $\frac{L_1-2}{L_1-4}$.

Furthermore, if the estimation window is sufficiently wide (L_1>35), then the distribution of $\widehat{SCAR}_i(\tau_1, \tau_2)$ can be approximated with the standard normal.

What have been discussed so far refers to the aggregation of the abnormal returns for a single security. The next step consists in performing the aggregation across securities and through time. In order to do so, it is necessary to assume that the abnormal returns of different companies are not correlated; this is generally true if there is not the presence of any clustering, meaning that there is not any overlap between the focal securities' event windows. If the absence of any overlap is demonstrated, this is sufficient to affirm that the abnormal returns and the cumulative abnormal returns will be independent across the securities.

The first step is to define the average abnormal return vector:

$$\bar{\epsilon}^* = \frac{1}{N} \sum_{i=1}^{N} \hat{\epsilon}_i^* \tag{4.23}$$

$$Var[\bar{\epsilon}^*] = V = \frac{1}{N^2} \sum_{i=1}^{N} V_i \tag{4.24}$$

where N represents the number of securities.

After that this process is completed, the elements of the average abnormal return vector are aggregated though time using the same approach explained before, defining CAAR as the Cumulative Average Abnormal Return:

$$CAAR(\tau_1, \tau_2) = \frac{1}{N}\sum_{i=1}^{N} \widehat{CAR}_i(\tau_1, \tau_2) \tag{4.25}$$

$$Var[CAAR(\tau_1, \tau_2)] = \bar{\sigma}_i^2(\tau_1, \tau_2) = \frac{1}{N^2}\sum_{i=1}^{N} \sigma_i^2(\tau_1, \tau_2) \tag{4.26}$$

Under the assumption that the event windows of the N securities do not overlap, and assuming as true the null hypothesis H_0 , the CAAR can be described by the following formulation:

$$CAAR(\tau_1, \tau_2) \sim N(0, \bar{\sigma}_i^2(\tau_1, \tau_2)) \tag{4.27}$$

After that, it is possible to proceed with the test of the null hypothesis, by using the following equation:

$$J_1 = \frac{CAAR(\tau_1, \tau_2)}{[\bar{\sigma}^2(\tau_1, \tau_2)]^{1/2}} \sim N(0,1) \tag{4.28}$$

In the econometrics literature there is also a second method available in order to conduct the test of the H_0 hypothesis. This second method of aggregation assigns the same weighting to all the SCAR, and then defines the average SCAAR over N securities as follows:

$$SCAAR(\tau_1, \tau_2) = \frac{1}{N}\sum_{i=1}^{N} \widehat{SCAR}_i(\tau_1, \tau_2) \tag{4.29}$$

Under the null hypothesis H_0, the SCAAR will be normally distributed, with a mean equal to zero and a variance equal to $\frac{L_1-2}{N(L_1-4)}$. Finally, the null hypothesis can be tested using:

$$J_2 = \left(\frac{N(L_1-4)}{L_1-2}\right)^{1/2} SCAAR(\tau_1, \tau_2) \sim N(0,1) \tag{4.30}$$

4.2.3. Event Study Applications

The Event Study methodology has been applied several times in the econometrics literature. This section is dedicated to the presentation of some relevant examples of studies in which this methodology has been adopted to assess the impact of events of heterogeneous nature on the financial performances of corporations.

The first historical application of this method can be traced back to 1933, when it was adopted by Dolley (1933) [39] in order to assess the impact of stock splits on nominal price changes. More in detail, the author selected a pool of 95 splits in a 10 years time-horizon, finding that the price increased in 57 cases, declined in 26 cases and remained unchanged in the other 12. During the period between the 1930s and 1960s, the Event Study methodology experienced a boost in the level of its sophistication, due to the removal of confusing events. In this regard, the authors that contributed to this improvement with their papers are Myers and Bakay (1948) [40], Barker and Austin (1956) [41], and Ashley (1962) [42]. In the following years, an important and remarkable example is represented by the innovation introduced by the authors Ball and Brown (1968) [43] and later by Fama et al (1969) [44], who examined the impacts of stock splits after having isolated the effects of simultaneous dividend increases; these works are crucial for the development of the Event Study methodology, because they paved the way for the refinement of the methodology that is essentially still used today. After the publication of these forerunner studies, other authors proposed various improvements to the basic methodology. These modifications give the possibility to manage the criticalities deriving from violations of the statistical assumptions used in the early work. In this regard, Brown and Warner (1980, 1985) [32,33] are interesting papers which show the huge importance of the application of these implementations. More in dept, the first

research considers the application criticalities regarding the data sampled at a monthly interval, while the second study focuses on the same thematic but with a daily interval.

In 1995, in the study by Ikenberry et al (1995) [45], the authors tried to investigate the impact of open market share repurchase announcements on the financial performances of the companies. More in dept, the sample of analysis has been built by identifying all the announcements listed in the Wall Street Journal in the time-period going from January 1980 to December 1990, in which it is stated that a specific firm expressed the intention of repurchasing its own common stocks through open market transactions. The authors of the paper analysed all the repurchase announcements, without considering whether they were actually completed or not. For what concerns the methodology adopted in this study, both short-term returns surrounding the announcement and long-term performance following the announcement have been assessed. In particular, the short-term performance has been computed considering an overall time-window going from 20 days before to 10 days following the announcement event. Moving to the computation of the long-run performance, it is important to take into account that the results can be sensitive to the procedures adopted. In this paper, two different approaches have been used but, considering the aim of the present work, only the technique based on cumulative abnormal returns (CARS) is reported. More in dept, under the CAR approach, abnormal returns are computed with a monthly frequency relative to a benchmark, and then they are aggregated over time. At the end, the abnormal performance is calculated using the returns to all companies existing in a given event month.

Moving to the analysis of the results of this study, the paper focuses firstly on the short-term abnormal returns surrounding the repurchase announcements. In

particular, it has been observed that there are negative abnormal returns prior to the announcement, measured from days - 20 to - 3, for a total of - 3.07%. For what concern the average market reaction measured in the window going from two days before to two days following the announcement in the Wall Street Journal, it is 3.54%. After the announcement event, the returns, on average, seem to be very similar to those of the market. These results are extremely comparable to the findings obtained by other researchers examining repurchase announcements.

The paper moves then to the analysis of the long-term abnormal returns. The figure below shows the CARs up to 48 months following the repurchase announcement, using four different benchmarks. It is important to specify that these CARs are computed beginning in month 1, and thus excluding the initial market reaction to the repurchase announcement.

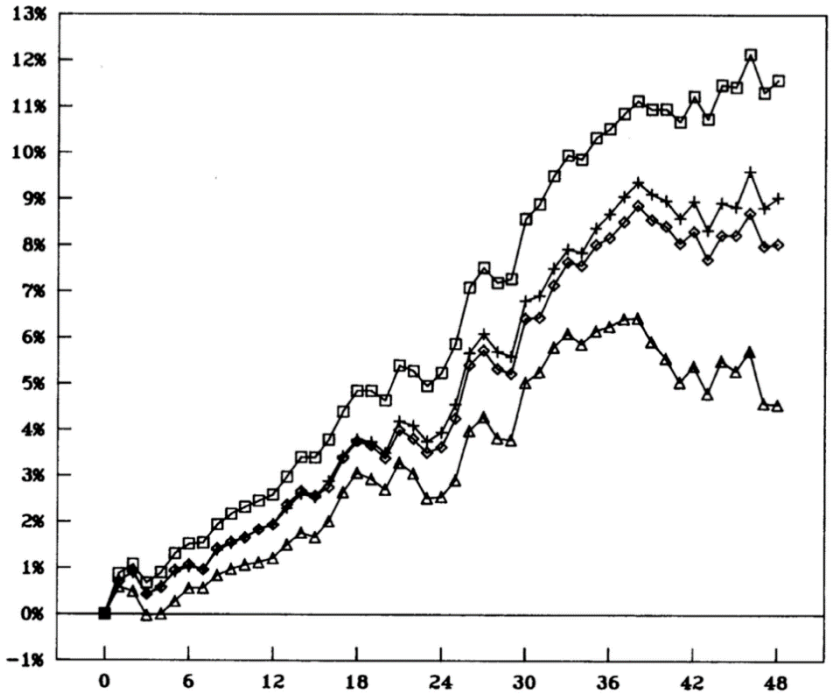


Figure 4.2: CAR after the repurchase announcement. Source: Ikenberry et al (1995) [45]

What emerges is that the corporations which announce an open market stock repurchase tend to perform abnormally well in the long run.

To sum up, this study focuses on the analysis of the consequences of open market stocks repurchase announcements by the companies on their abnormal returns. Although in the literature these announcements are interpretated as a signal for an undervaluation of the stock's price by the market, and so this should be a positive signal encouraging the market to buy the company's shares, what actually happens is that the average market response to the news of an open market share repurchase is only 3.5%. Hence, the conclusion of this paper is that the market substantially ignores this undervaluation signal, and it underreacts in the short-term to open market share repurchase announcements. On the contrary, the study shows that the impact of these announcements does not end in the short-term period, but it continues also in the long run, making the interested companies experience a substantial growth in their abnormal returns.

In 2015, the study by Miyamoto (2015) [46] investigates which is the impact of the credit rating changes by the Japanese rating agency on the Japanese stock market. More in dept, the definition of credit rating the author refers to is the following one: "credit ratings are the overall debt capacity and creditworthiness to pay the debt". Similarly to what happens with ESG ratings, also credit ratings are usually expressed in the form of alphabetical codes, ranging from D (the worst evaluation) to AAA (the best possible grade). In order to conduct the analysis, the author uses the Event Study method with the aim of isolating the events and seeing what the consequences of their occurrence on the abnormal returns are. In the literature there were already other studies trying to investigate the same topic under different conditions. For example, the authors Holthausen and Leftwich (1986) [47] studied 1,014 rating changes provided by S&P and Moody's, analysing their effects on the

corporations in the US, covering the period from 1977 to 82. What they found is that negative credit rating changes led negative abnormal returns while, on the other hand, positive credit rating variations showed were not associated with positive abnormal returns. Following the example of this research, other authors applied the same methodology in order to study the impact of credit rating changes on other national financial markets, as the UK, the Swedish and the Portuguese markets.

Going back to the study of Miyamoto, it starts from the formulation of the following two hypotheses:

- *hypothesis 1*: credit assignments for Japanese firms are expected to be followed by a significant positive or negative market reaction.
- *hypothesis 2*: a positive (negative) variation in the credit rating is expected to be associated with a positive (negative) share market reaction.

To assess validity of the above-mentioned hypothesis, the author sampled the rating assignment announcement by Rating and Investment Information, for the period of going from November 2000 to October 2007. For the scope of the analysis, only those companies having a credit rating for each of the seven-year period are considered. The sample resulting from this data gathering process is composed by 221 corporations belonging to different industrial sectors, and experiencing an overall number of credit rating changes equal to 383. The Table 4.1 shows the numbers of credit rating changes by industries.

Table 4.1: Number of events. Source: Miyamoto (2015) [46]

	Number of Corporations	Number of Rating Changes
Total	221	383
Down Grade	-----	187
Up Grade	-----	196
Manufacturings	136	236
Down Grade	-----	119
Up Grade	-----	117
Non-manufacturing	85	147
Down Grade	-----	68
Up Grade	-----	79
Financial services	25	49
Down Grade	-----	24
Up Grade	-----	25
Non-financial services	196	334
Down Grade	-----	163
Up Grade	-----	171

From a methodological perspective, in order to compute the expected market returns (normal returns) the author adopts the Market Model on an estimation window that goes from 139 to 10 days before the occurrence of the event. After having estimated the model parameters, the significance of the event is evaluated using the Cumulated Average Abnormal Return methodology (CAAR), on a 21-day event window, arriving to 10 market days after the event.

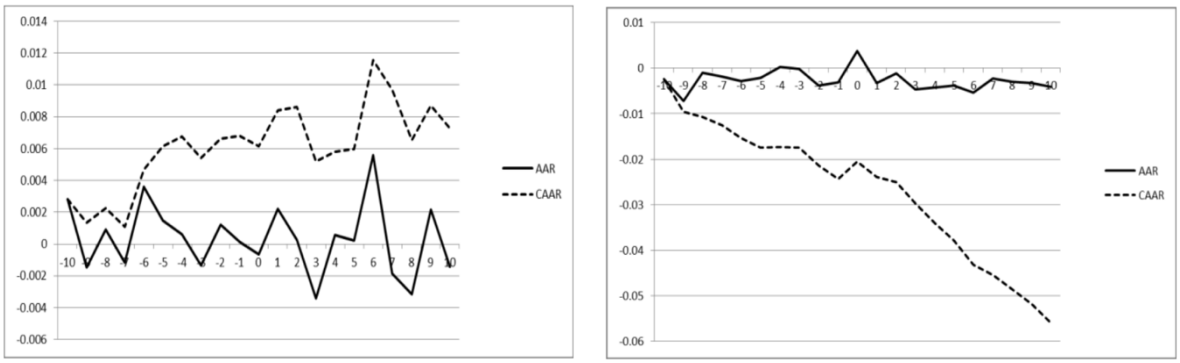


Figure 4.3 Aggregated AAR and CAAR for downgrade and upgrade.

Source: Miyamoto (2015) [46]

The analysis has been conducted considering credit rating upgrades and downgrades separately. The Figure 4.3 shows the results for aggregated AAR and CAAR for downgrades and upgrades respectively. Regarding the downgrades, it is possible to see that the reaction of CAAR is unstable, going up and down for declass announcement. On the other hand, CAAR experience negative reactions for the upgrade announcement. To sum up the results of the paper, the Japanese market reacted positively with negative announcement; this can be explained by the fact that stock prices seem to react before the information of rating changes is disclosed. On the contrary, the reaction to upgrades seems to be negative since the investors are able to anticipate the positive credit rating changes, buying the stocks before and selling them after the occurrence of the event. In the end, this study led to the

validation of the hypothesis number 2, which stated that a positive (negative) outlook is expected to be associated with a positive (negative) share market reaction. This means that, before the information about rating change is announced, the stock prices react. Therefore, it is possible to state that market participants seem to act on rumours of rating changes.

Another interesting study in which the Event Study methodology has been applied to a real context is represented by the paper by Cummings et al (2018) [48]. In particular, the authors focus on the possible impact that the adoption of a public social media could have on the market value of the corporations, and whether these effects are influenced also by the timing of the adoption. More in dept, the study exploits the Event Study methodology in order to analyse variations in the value of companies after the firm's adoption of two public social networking sites: Facebook and Twitter. This research differs from similar studies conducted in the literature, which were used to focus on technologies that provide hard-to-copy competitive advantages. On the contrary, this paper is oriented toward the analysis of public technologies operated by third party vendors (e.g., social media), which could be easily accessed also by all the competitors, hence they may represent a limited source of competitive advantage. As a consequence, the aim of this paper is the one of understanding whether these technologies could bring a value to the firms and, if this is the case, when to adopt them. Indeed, these technologies are characterised by the presence of the network externalities, enhancing their value proportionally to the number of the users; so being a first mover may not represent the optimal strategy. In order to evaluate the impact of adoption of SNSs on companies' value, the Event Study methodology has been adopted in this paper. The sample for this study is composed exclusively by companies belonging to consumer-based industries, given the fact that SNSs are consumer-oriented technologies. This population comprehends firms which have adopted either Facebook or Twitter. In

order to choose the companies forming the sample of the study, two criteria were adopted. Firstly, given the fact that the dependent variable of the study is stock returns, only U.S. publicly traded firms were considered. Secondly, the authors selected only companies operating within business-to-consumer (B2C) industries, since these corporations would adopt public SNSs primarily in order to engage existing customers and to reach out to potential ones. Starting from an initial sample of over 6000 publicly traded companies, 840 firms operated within a consumer-based industry, and therefore they were selected for the study. The final sample for Facebook contained 243 firms and the final sample for Twitter contained 303 firms. For what concerns the timing of the technology adoption, the study refers to the categorisation of technology adopters made by Rogers (1995) [49]. In particular, the paper considers only 3 categories out of the 5 proposed by Rogers:

- *Innovators*: they are defined by Rogers as the first 2,5% of adopters of a technology within a population.
- *Early adopters:* they represent the next 13,5% of adopters of a technology.
- *Early majority*: according to Rogers, they are the next 34% of adopters of a technology.

Applying these percentages to the sample of interest for this study, it results that innovators and early adopters are 21 and 113 respectively, while the remaining part of the firms has been considered as constituents of the early majority. Moving to the description of the Event Study methodology adopted, this paper uses a 220-day estimation window (between days 250 and 30 before the event) in order to estimate the normal or expected returns. The decision of stopping the estimation window 30 days prior to the event has been taken in order to avoid the possibility of having any event-related information incorporated into the estimation of the normal returns. For what concerns the event window, it is composed by the day in which a

specific firm adopts a SNS and the day afterward, or day t = 0 and t = 1. In order to evaluate the market's overall reaction to companies which adopt SNSs, the mean cumulative abnormal returns (MCAR) are analysed. To evaluate if the adoption of a new technology leads to increased firm performance, MCARs are analysed. In particular, separate MCARs are estimated for the two samples (one for Facebook and one for Twitter). The results show that on average corporations significantly benefit from adopting SNSs. Indeed, companies which adopted Facebook experienced an abnormal increase of 1,20 percent in their stock prices during the event-window. On the other hand, companies that adopted Twitter experienced a 0,67 percent abnormal increase in their stock prices during the same two-day window. After having obtained these results, the paper goes on with the analysis regarding the timing of the technology adoption. More in dept, the authors exploit a multifactorial model in which the 3 categories of adopters (namely innovators, early adopters and early majority) are included as dummies variables; this regression is aimed at describing the abnormal returns, which represent the dependent variable of the model. Estimating the parameters of the multifactorial model through a cross-sectional generalised least squares regression, it has been possible to determine what is the impact of belonging to a specific adoption category on the abnormal returns. From this analysis, it emerges that only the early adopters experience a significant parameter, meaning that they are the ones that benefit the most from the adoption of the new social media technologies. This stylized fact led the authors to conclude that, due to the presence of network externalities, these technologies are characterized by a second-mover advantage, and not by the typical first mover advantage as the other traditional technologies.

Having seen the elevated number of different applications in which the Event Study methodology can be worthy, this proves the robustness and the reliability of the method itself. In particular, it is interesting to notice that this methodology enhances its value when the scope of the analysis is to understand the magnitude of any event on the short-term stocks' performances of the interested corporations. More in depth, as it is possible to deduct from the first two studies cited above, this model is particularly fitting when dealing with financial events having a direct impact on the perceived value of the company, such as corporate announcements and rating updates. In this regard, the above-mentioned paper by Miyamoto (2015) [46] is particularly relevant since it applies this specific methodology to the ratings published by the credit rating agencies, and it testifies its robustness in this specific application field. Indeed, as discussed in the introduction of the present work, more and more in the last years the ESG ratings have assumed a relevant role in the decision-making process of the financial investors, arriving to have an impact comparable to the ones of the credit ratings. For this reason, the application of the Event Study methodology is especially suitable for the analysis of the focal event of this work, namely the ESG ratings' updates announcements.

4.3. Input Data

After having introduced the theoretical methodology of the Event Study, the next sections are dedicated to the description of the data set used in order to feed the model and run the analysis. In particular, the pieces of information that made up the model's data set can be classified into three clusters:

- *The pool of companies analysed.*
- *The rating providers and their evaluations.*
- *The stocks' prices of the selected firms.*

Before going in dept with the description of the clusters of data, it is important to specify that during the collection of the data all the relevant choices have been made according to two main concerns: on the one hand the availability of the data itself; on the other hand, some assumptions were needed in order to make the analysis feasible in terms of timing and volume, without losing the statistical relevance of the study.

4.3.1. Set of Companies Analysed

The first step of the analysis consists in the choice of the pool of companies that are at the centre of the study. More in depth, the choices made in this regard have been taken according to three main drivers, namely the pool dimension, the reference financial markets and, eventually, the market of sectors the selected companies belong to. For what concerns the dimension of the set of companies, the main constraint for this important decision is represented by the trade-off between the statistical significance of the analysis and the amount of data to be processed in the

model. Indeed, the higher the number of companies forming the testing cluster, the higher are the robustness and the reliability of the analysis' results but, on the other hand, also the volume of the data to be collected and managed increases exponentially, requiring a higher processing effort. For these reasons, the choice has been the one of selecting a set of 75 firms in order to have a good balance between the two sides of the trade-off presented before. The next table shows the names of the corporations that have been selected for the study.

Table 4.2: Pool of companies analysed

Enel SpA	Tesla Inc	Royal Dutch Shell PLC
Stellantis NV	Berkshire Hathway Inc	Unilever PLC
Intesa Sanpaolo SpA	Starbucks Corp	HSBC Holdings PLC
Eni SpA	JPMorgan Chase & Co	AstraZeneca PLC
Ferrari NV	Visa Inc	BP PLC
Amplifon SpA	PepsiCo Inc	GlaxoSmithKline PLC
Assicurazioni Generali SpA	Taiwan Semiconductor CO LTD	British American Tobacco PLC
UniCredit SpA	UnitedHealth Group Inc	Rio Tinto PLC
FinecoBank Banca Fineco SpA	Johnson & Johnson	Barclays PLC
Exor NV	Home Depot Inc	Rolls-Royce Holdings PLC
Nexi SpA	Exxon Mobil Corp	Nestle SA
Moncler SpA	Coca-Cola Co	Novartis AG
Snam SpA	Chevron Corp	Roche Holding AG
Poste Italiane SpA	Abbott Laboratories	Zurich Insurance Group AG
Davide Campari Milano SpA	AT&T Inc	Credit Suisse Group AG
Atlantia SpA	Keyence Corp	Gazprom PAO
Terna Rete Elettrica Nazionale SpA	Toyota Motor Corp	NK Rosneft' PAO
Tenaris SA	SoftBank Corp	NK Lukoil PAO
Recordati Industria Chimica e Farmaceutica SpA	Sony Group Corp	GMK Noril'skiy Nikel' PAO
DiaSorin SpA	Nippon Telegraph and Telephone Corp	Novatek PAO
Microsoft Corp	Nintendo Co Ltd	Industrial and Commercial Bank of China Ltd (ICBC)
Apple Inc	Recruit Holdings CO LTD	China Construction Bank Corp
Amazon.com Inc	Nidec Corp	Agricultural Bank of China Ltd
Alphabet Inc	Shin-Etsu Chemical CO LTD	Ping An Insurance Group Co of China Ltd
Meta Platforms Inc (Facebook)	Denso Corp	Bank of China Ltd

From the list of companies shown in Table 4.2, it is possible to figure out the other two drivers that guided the pool selection, namely the financial markets and the market sectors. Regarding the selection of the financial markets, the firms selected have been chosen with the aim of giving a higher priority to the Italian region and, enlarging the scope, to the European one; indeed, even though the number of Italian

companies is equal to the American companies' one, the USA market should have been given a significant higher weight due to its dimension. Furthermore, the selection of the companies has been done trying to preserve as much as possible the heterogeneity at the world level, including companies belonging to the main worldwide financial markets. After the selection of the companies with their belonging capital markets, the next step consisted in including in the analysis the market indexes, which reflect the normal returns of the companies according to the Market Model. The table below shows the selected stock exchanges with their market indexes.

Table 4.3: Number of companies for financial markets

Financial Market	Market Index	Number of Companies
Milano Stock Exchange	FTSE MIB	20
NASDAQ/NYSE	S&P 500	20
London Stock Exchange	FTSE 100	10
Tokyo Stock Exchange	TOPIX	10
SIX Swiss Exchange	SMI	5
Moscow Interbank Currency Exchange	MOEX	5
Shanghai Stock Exchange	CSI300	5

In particular, the usage of different indexes for the different capital markets is coherent and consistent with the implementation of the Event Study methodology, which consists in extracting the abnormal returns from the total returns of the corporations, by subtracting the market returns. For this reason, the best accurate way to measure the market returns is to choose the specific market index for each company in the analysis's pool.

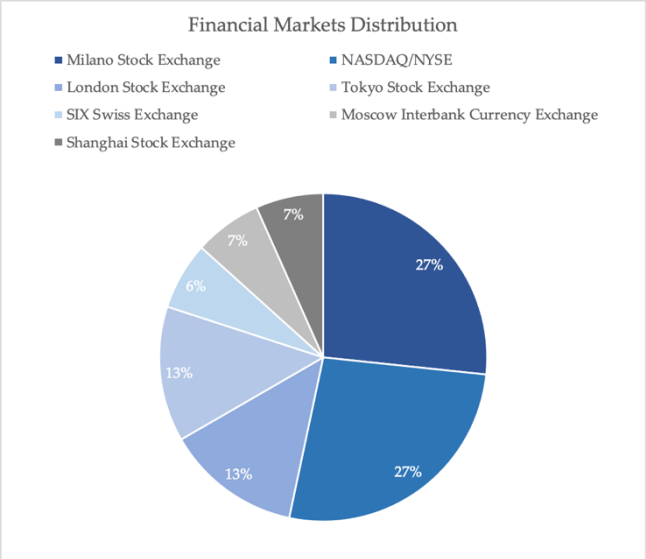


Figure 4.4: Financial market distribution

It is evident that the financial markets have not been given the same weights: the higher importance recognised to the Italian country is testified by the number of companies belonging to the Milan Stock Exchange, namely 20. The USA market is represented by a pool of 20 corporations too but, looking at the ratio between the weight given in the analysis and the capitalisation of the corresponding market, it is clear that a higher importance is attributed to the Italian financial market. On the other hand, a set of 10 companies have been selected for the English and the Japanese financial markets. Finally, 5 corporations have been selected for the Swiss, the Russian and the Chinese markets. At this stage, in order to choose the list of the firms, the criteria used for the selection has been the one of picking the companies reporting the highest market capitalization, trying to keep a good level of heterogeneity in terms of the belonging industry. In the Table 4.4 there is the classification of the companies according to the industry sector in which they operate and the financial markets they belong to. The exact denominations of the belonging sectors presented in the table have been taken from Refinitiv Eikon; this platform is better explained in the following section. Looking at the different industries reported in the table above, it is possible to see that they are huge in number, showing diversified degrees of granularities. Indeed, a considerable set of sectors has only a single firm as representative, while other sectors such as “Banks” and “Auto & Truck Manufacturers” contain a higher number of corporations inside them. An annotation is worthy to be done at this

stage: some sectors are denominated by the platform Refinitiv Eikon with slightly different denominations, even though they could be seen as a unique and comprehensive cluster. This is the case, for example, of all the "Oil & Gas" related sectors.

Table 4.4: Companies by industrial sector

Industry Sector	Number of Companies
Aerospace & Defense	1
Apparel and Accessories	1
Auto & Truck Manufacturers	5
Auto, Truck & Motorcycle Parts	1
Banks	11
Business Support Services	1
Consumer Goods Conglomerates	1
Courier, Postal, Air Freight & Land-based Logistics	1
Department Stores	1
Distillers & Wineries	1
Diversified Chemicals	1
Diversified Mining	1
Electric Utilities	2
Electrical Components & Equipment	2
Employment Services	1
Food Processing	1
Highway & Rail Tracks	1
Home Improvement Products & Services Retailers	1
Household Electronics	1
Integrated Oil & Gas	5
Integrated Telecommunications Services	1
Life & Health Insurance	2
Managed Healthcare	1
Medical Equipment Supplies & Distribution	3
Multiline Insurance & Brokers	1
Non-Alcoholic Beverages	2
Oil & Gas Exploration and Production	1
Oil & Gas Refining and Marketing	3
Oil & Gas Transportation Services	1
Oil Related Services and Equipment	1
Online Services	3
Personal Products	1
Pharmaceuticals	6
Phones & Handheld Devices	1
Restaurants & Bars	1
Semiconductors	1
Software	1
Specialty Mining & Metals	1
Tobacco	1
Toys & Children's Products	1
Wireless Telecommunications Services	2

It is clear that, by considering the companies part of these sectors as belonging to a common cluster, the granularity of the resulting industry decreases, but the homogeneity between the sectors increases. Furthermore, looking at the distribution of the sectors in the different stock exchange markets, it is evident that there are some financial markets that are characterized by a higher degree of heterogeneity. This is the case of the Milano Stock Exchange and of the London Stock Exchange, for which the selected companies belong to very different market sectors. On the other hand, there are financial markets such as the Moscow Interbank Currency Exchange and the Shanghai Stock Exchange that present a high level of homogeneity from the industry perspective. This lack of heterogeneity is due to the market capitalization selection criterion and, on the other hand, by the inner nature of these financial markets. Indeed, on the one hand, the Russian market is historically dominated by firms operating in the business of the natural resources production; on the other hand, instead, the Chinese arena is known for being a strongly government-driven market and, as a consequence, it is dominated by governmental institutions like banks.

4.3.2. Rating Providers

After having explained the choice of the selected companies, this chapter is dedicated to the explanation of the criteria used to decide the rating agencies to use in the analysis. In particular, the first concern has been the one of having more than one provider: this criterion derives from the aim of the analysis itself, which is the one of understanding not only if the changes in the sustainability ratings could lead to changes in the stocks' prices, but also, given the existence of the divergence between the different ratings, if one rating agency has more influence on the

companies' stock prices with respect to the others. The second constraint has been represented by the availability of data: usually, indeed, in order to have access to the ESG ratings, it is necessary either to pay a subscription fee to the rating provider or to be licensed to enter in possession of this confidential data. For these reasons, among the several ESG rating providers presented at the beginning of the present work, the choice has fallen on the selection of Refinitiv and MSCI.

- *Refinitiv*: the data about the ESG ratings provided by Refinitiv has been taken from the proprietary platform of the rating provider itself. This technological solution is called Refinitiv Eikon, and it is an open platform which provides access to industry data, insights and exclusive news. This technological application enables financial market professionals to exploit analytical tools in order to extract valuable pieces of information from the available market materials and the Refinitiv's proprietary researches. This platform provides for each company a dashboard with financial and non-financial data, as general information about the companies (i.e., the reference operating industry and the stock exchange in which they are listed), the stock prices fluctuations, and specific sectorial researches. In particular, the pieces of information relevant for this study are taken from the Refinitiv Eikon's section dedicated to the ESG evaluations. In this regard, the image above shows this sustainability section, in which it is possible to see the annual ESG ratings assigned to a specific company and divided by the three sustainability pillars, namely Environmental, Social and Governance.

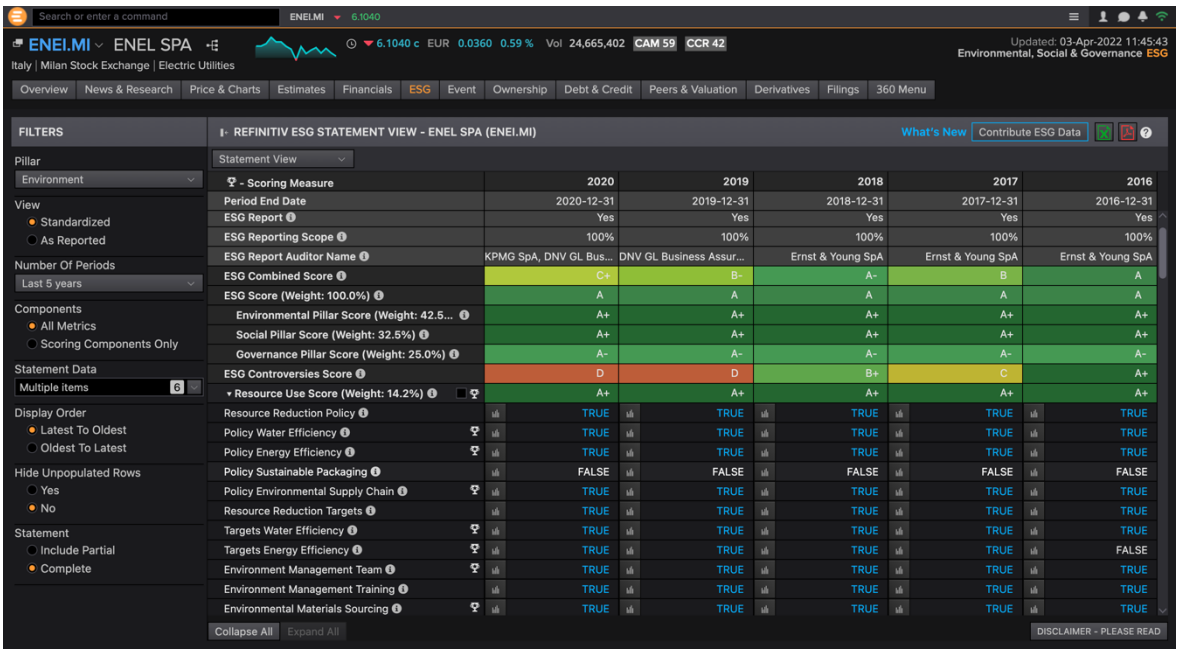


Figure 4.5: Refinitiv Eikon ESG Statement view

As it is shown by the Figure 4.5, these ratings are published annually by Refinitiv: in particular, for most of the companies analysed, this data are updated annually on the 31 of December. For what concerns the Japanese firms, instead, the ESG ratings in this case are provided annually on the 31 of March.

- *MSCI*: for what concern the ESG ratings provided by Morgan Stanley Capital International, they have been taken by the proprietary database of the MIP Graduate School of Business. In particular, this database gives access to all the companies' yearly ratings provided by MSCI, in the form of an excel sheet. Even in this case, the frequency with which these ratings are published is annual; but, contrarily to what happens in the case of Refinitiv, whose grades are announced at the end of each year, the sustainability evaluations provided by MSCI follow the reporting period of the companies analysed.

At this stage, it is necessary to explain the decision about the time-horizon selected for the analysis. Considering that the attention to the ESG themes has increased exponentially in the past years and keeping in mind the existing trade-off between the amount of data and the effort required to manage it, the decision has been to consider a time period going from the 2016 to the 2020. As a consequence, given the fact that the updating frequency of the ESG ratings is annual, for each company there is a set of 5 events available; but there are some exceptions, represented by companies for which there is a number of events higher or lower than 5.

Looking at the collected data, it is possible to observe practically the existence of the ESG rating divergence that has been explained theoretically in the third chapter of the present work. Some explanatory cases are reported in the table below.

Table 4.5: Examples of rating divergence

	2020	**2019**	**2018**	**2017**	**2016**
Enel SpA					
MSCI	AAA	AAA	AA	AA	AA
REFINITIV	B-	C+	A-	B	A
Intesa Sanpaolo SpA					
MSCI	AAA	AAA	AAA	AAA	AAA
REFINITIV	B	C+	A-	B+	A-
Tesla Inc					
MSCI	A	A	AA	AAA	AAA
REFINITIV	C-	C-	C-	C-	C-
Rolls-Royce Holdings PLC					
MSCI	BBB	BBB	BBB	BBB	BBB
REFINITIV	A-	A-	B+	C+	B-
Nintendo Co Ltd					
MSCI	A	AA	AA	A	BBB
REFINITIV	B-	B+	B+	B	

Before starting with the observations about the table above, it can be noticed that the two selected ratings adopt the same grading scale (i.e., letters from D to AAA): as a consequence, the ESG raters' evaluations are directly comparable. The first important consideration that can be done is that the companies selected as examples belong to various industries and also to different countries, meaning that the divergence phenomenon affects all the companies indiscriminately, without being linked to particular firm-specific characteristics. Furthermore, comparing the sustainability evaluations provided by the two ESG rating agencies, it can be seen that MSCI usually gives higher scores to the firms with respect to its competitor Refinitiv. However, there are also some cases in which this tendency is not confirmed, like for example the case of Rolls-Royce Holdings PLC, which has received a higher ESG rating by Refinitiv. Another stylized fact resulting from the comparison is that the evaluations given to the companies change their value over time with different patterns. In particular, in the majority of cases the evaluations provided by MSCI remain more stable along time, while the ones given by Refinitiv tend to fluctuate. This can be seen looking at the cases of Intesa Sanpaolo Spa and Rolls-Royce Holdings PLC.

4.3.3. Stocks' Returns

The last phase of the data collection process is represented by the stocks' prices gathering. In this regard, the most relevant annotations to be done are the following ones:

- *Estimation and event window*: when talking about the stocks' prices collection, the most important decision to be taken is the one concerning the time periods in which the Event Study analysis is structured. In particular, given

the fact that the selected methodology in order to compute the normal returns is the Market Model, the two time periods to be estimated are the estimation window and the event window. At this stage, it is worthy to mention the main trade-offs related to the choice about the length of this two time-horizons. The trade-off related to the estimation window's lengths is the following: the wider is the interval considered, the more statistically relevant is the resulting estimated model but, on the contrary, the higher is the computing effort required to manage the data and derive the model. As a consequence, the choice about the estimation window has been to consider a length of 30 market days, arriving to 2 market days before the arising of the focal event. For what concerns the event window, instead, the most relevant trade-off is the following: considering a shorter time interval, it is reasonable to assume that the reaction in the market prices after the event is due only to the event itself but, on the contrary, the model experiences a lack of market visibility in the long-term. For these reasons, the decision about the length of the event window has been to consider a relatively short period of 12 market days, starting from 2 days before the arising of the event. This choice has been driven mainly by two reasons; the first one is the conviction that the main effect of eventual upgrading or downgrading will show immediately in the days after the ESG rating publications, given the high efficiency of the markets. The second reason is given by the willingness to purify the reaction of the markets from other possible concurrent events.

- *Returns*: after having taken the decision about the two temporal windows, the next step consists in determining the exact moment of the day in which the returns are collected. In order to have data as much as possible standardised among the pool of companies, the returns considered are the closing ones.

Last but not least, the data about the stocks' returns have been gathered exploiting the Excel Bloomberg API, which allows to extract the data directly from the proprietary Bloomberg's database, by simply using Excel formulas. More in detail, the specific formula utilised for this study is the following one:

$$bdh(\text{Bloomberg Ticker, pct\_chg\_1d}, starting\ date, ending\ date) \tag{4.31}$$

Thanks to this formula, it is possible to download the data about the stocks' returns, and automatically upload it in an Excel sheet.

4.4. Implementation

After having introduced the main data gathered in other to do the analysis, this section is dedicated to the presentation of the Event Study conducted. The starting point is to define the analysis, which is divided into two different Event Studies, each one corresponding to the two ESG rating providers selected. In order to implement the analysis, the paper refers to the typical Event Studies' seven steps presented in the methodology (section [4.2]). In this regard, the first phase consists in the definition of the event of interest and of the event-window. In particular, the focal event of the present work is represented by the sustainability ratings publications by the selected ESG rating providers. However, this event is made up of three sub-events, which are upgrading, downgrading or confirmation of the evaluations; these sub-categories are investigated through three distinct and independent Event Studies. As a consequence, the total number of studied events is equal to 6. In order to study these events, the event-window is represented by a 12 market days time-horizon, and it starts two days before the arising of the focal event. This choice about the event-window's length, as explained in detail in the chapter before, is due to two main reasons: on the one hand, it is motivated by the willingness of eliminating disturbing noises caused by other possible concurrent events. On the other hand, it is also driven by the belief that the capital markets are efficient, so they immediately process the new pieces of information available after the event, adjusting the prices of the stocks. The second phase of the methodology refers to the choice of the selection criteria utilised in order to define the sample of companies to be used for the study. How it is explained in the previous chapter, the main drivers for the companies' selection are the following three, and they are oriented to the achievement of a diversified pool of companies, overweighted toward the Europe. More in depth, the first screening is done according to the

reference stock exchange markets. After this choice, the companies experiencing the highest market capitalizations are selected, not forgetting to take into account as third driver also the industry sectors the firms belong to. The third step is represented by the computation of the abnormal returns, which are utilised in order to estimate the effect of the focal event on the prices of the stocks. As explained in the section [4.2.2], in order to calculate the abnormal returns, it is necessary to evaluate the normal returns, exploiting a statistical model called Market Model. In particular, this model derives the returns of a specific stock from the return of the market through the formulation presented below:

$$R_{it} = \alpha_i + \beta_i R_m + \epsilon_{it} \tag{4.32}$$

Starting from this formula, after having estimated the parameters of the Market Model, it is possible to compute the abnormal returns by subtracting the expected returns (that are the ones of the market) from the actual returns of the stock.

$$\epsilon_{it} = R_{it} - (\alpha_i + \beta_i R_m) \tag{4.33}$$

After having chosen the methodology that is used to compute both the normal and the abnormal returns, the following phase consists in the estimation of the Market Model's parameters. In order to implement this step, it is necessary to define the estimation window, which is the time period prior to the arising of the event and used to assess the value of the parameters characterizing the Market Model's

equation. Even for this decision, as in the case of the event-window, there is a trade-off; in particular, as the length of the estimation window increases, so does the amount of the effort needed to manage the available data, but on the other hand the robustness of the estimated model increases too. As presented in the chapter before, all the mentioned elements drove to the decision of setting the length of the estimation window equal to 30 market days, until two market days before the focal event. From a practical point of view, in order to estimate the model's parameters necessary to compute the normal returns, it has been decided to adopt the software Gretl, which is presented in all its details in the following section.

4.4.1. Parameters' Estimation: Gretl

As presented above, the normal returns have been computed through the implementation of Gretl, which is an open-source cross-platform software package. In particular, this software is written in the C programming language, and it is aimed at supporting in the realisation of econometric analysis. It is characterized by the presence of easy intuitive interfaces, which are available in several languages apart from English. More in depth, a wide portfolio of estimators is included inside this software package, for example, the most relevant ones are the least squares, the maximum likelihood and the regularized least squares (LASSO, Ridge, elastic net).

In order to conduct the econometric analysis, Gretl provides a huge variety of time series models, among which the most relevant ones are the ARIMA, the univariate GARCH-type models and, eventually, the VAR models. Another strength of this is represented by the availability of a wide set of programming tools and matrix operations in order to handle the data and extract valuable insights from it.

Furthermore, the open-source software offers some models in order to visualize the analysis' output in LaTeX files, tabular and/or equation format and, on the other hand, it enables an easy exchange of the results thanks to the exploitation of some facilities. Before starting with the implementation of the script of the Gretl code, it is important to define all the relevant variables of the model. In particular, these Event Studies are characterised by a number of events equal to the ones shown by the table below.

Table 4.6: Number of events

Rating agencies	**Announcement**	**Number of events**
MSCI	Upgrade	80
	Confirmation	304
	Downgrade	27
Refinitiv	Upgrade	94
	Confirmation	108
	Downgrade	89

As a consequence, for each event it is reasonable to expect a number of time series equal to the number of events multiplied by two: indeed, in order to run the analysis, it is necessary to have not only the returns of the company, but also the market returns in the same days. As explained in the chapter before, the two fundamental time-periods for this study are represented by the estimation window and the event-window, which have a length of 30 and 12 market days respectively; hence, the total number of days is equal to 42. The timeline used to analyse the event starts from the first observation day, for which the descriptive time-variable τ is equal to 1. For this reason, the focal event happens at a τ equal to 32. At this stage,

it is fundamental to introduce the variable $TIT_{i\tau}$, which corresponds to the returns of the event i at day τ. In particular, the values of i range from 1 to the corresponding number of events, while the values of τ belong to the interval that goes from 1 to 42. Furthermore, another relevant variable of the model is $MKT_{i\tau}$, which represents the market return for the *i-th* stock in day τ. For what concerns both the returns, they are computed as the percentual change in the prices of the stocks with respect to their values in the previous market day. Hence, for each event, a number of returns equal to 84 is collected. At this point, it is possible to introduce the script of the GRETL code applied in order to conduct the study, which is shown by the image below.

```
open ref_eq.gdt
pre = 30
eve = 12
num_events = 109
smpl 1 pre
loop i=1..num_events
  ols TIT$i const MKT$i
  coeff$i = $coeff
  sigma$i = $sigma
  xx$i = $vcv
endloop
begs = pre+1
ends = pre+eve
smpl begs ends
loop i=1..num_events  --quiet
  matrix ystar$i =  { TIT$i }
  matrix xstar$i =  { const , MKT$i }
  AR$i = ystar$i-xstar$i*coeff$i
  v$i = (sigma$i^2)*I(eve)+xstar$i*xx$i*xstar$i'
endloop
matrix ccc=I(eve)
loop for i=2..eve --quiet
    loop for j=1..i --quiet
        if i>j
            ccc[i,j]=1
        endif
    endloop
endloop
print ccc
loop i=1..num_events  --quiet
  CUM_AR$i = ccc*AR$i
  vvv$i=ccc*v$i*ccc'
  series CAR$i = CUM_AR$i
endloop
series CAR = (1/num_events)*CAR1
vvv = (1/num_events^2)*vvv1
loop i=2..num_events  --quiet
  series CAR = CAR+(1/num_events)*CAR$i
  vvv = vvv+(1/num_events^2)*vvv$i
endloop
series sigma2_CAR = diag(vvv)
series LB = -1.96*sqrt(sigma2_CAR)
series UB =  1.96*sqrt(sigma2_CAR)
gnuplot CAR LB UB { set title "CAAR"; set yrange [-5:5]; set y2range [-5:5]; }
```

Figure 4.6: Gretl code

The first line of the code has the aim to open the file in which the returns are stored.

```
open ref_eq.gdt
pre = 30
eve = 12
num_events = 109
```

Figure 4.7: Gretl code (2)

The first part of the code is dedicated to the declaration of the variables that are employed later. More in dept, three are the variables defined at this stage:

- *pre*: it represents the length of the estimation window, and so it assumes a value of 30.
- *eve*: it is the length of the event-window and it is assigned a value of 12.
- *num_events*: it is the variable used to declare the number of the events considered in the study. It assumes a value that is different for each different type of event (namely upgrading, downgrading and confirmation of the grade) and for each ESG rating providers (namely MSCI and Refinitiv).

After having declared the different variables, the next step consists in the estimation of the Market Model for a number of times equal to the number of events present in the study (*num_events*). Considered the high number of iterations required in this phase, it is used the loop shown in the script below.

```
smpl 1 pre
loop i=1..num_events
  ols TIT$i const MKT$i
  coeff$i = $coeff
  sigma$i = $sigma
  xx$i = $vcv
endloop
```

Figure 4.8: Gretl code (3)

In order to estimate the model's parameters, it is necessary to use only the data belonging to the first 30 market days time-window; for this reason, the function "*smpl*" allows to consider only the relevant market days. Having selected the useful

set of data, it is possible to enter the estimation loop, which is enclosed between the key words "*loop*" and "*end loop*". The variable *i* represents the counter of the loop, which assumes the values ranging from 1 to the total number of events (num_events), and it increments its value by 1 at the end of every cycle. The first line of this portion of the code exploits the function *"ols"* in order to run the regression, taking as dependent variable the market returns of the company (TIT$i) and as independent variable the returns of the market (MKT$i). The parameters estimated through the Ordinary Least Square method are stored in the variables "*coeff$i*" and "*sigma$i*". These two variables are created in a number equal to the number of events (num_events); more in dept, "*coeff$i*" is a vector containing the coefficients of the Market Model, while "*sigma$i*" is a scalar containing the variances. The last line of the loop is aimed at storing the variance-covariance matrixes of the coefficients *(xx$i*), which are one of the outputs of the OLS methodology. The next step consists in the computation of the abnormal returns in the event-window, and of the corresponding variance-covariance matrix. In order to do so, the first thing to do is to define the time-period of interests, namely the event-window; this makes necessary the definition of two variables, which contains the boundaries of the interval.

In particular, the first variable is "*begs*", which corresponds to the first observation of the event-window; on the other hand, the second variable is "*ends*", which is the date (τ =42) of the last observation. The function "*smpl*" is used again in order to consider only the data belonging to the interval going from "*begs*" to "*ends*".

```
begs = pre+1
ends = pre+eve
smpl begs ends
loop i=1..num_events  --quiet
  matrix ystar$i =  { TIT$i }
  matrix xstar$i =  { const , MKT$i }
  AR$i = ystar$i-xstar$i*coeff$i
  v$i = (sigma$i^2)*I(eve)+xstar$i*xx$i*xstar$i'
endloop
```

Figure 4.9: Gretl code (4)

At this stage, the abnormal returns and the variance-covariance matrixes are computed though the execution of the loop shown above. Even in this case there is a counter (i), which increments itself and goes from 1 to *num_events*. Inside this loop, two matrixes are created:

- *ystar$i*, which contains the actual returns of the company.
- *xstar$i*, which contains the returns of the market.

The reason why these two matrixes are defined is to exploit the possibility to make mathematical operations between them. Once these variables are defined, the abnormal returns and the variance-covariance matrixes are computed and the results coming from these operations are stored in AR$i and in V$i.

```
matrix ccc=I(eve)
loop for i=2..eve --quiet
    loop for j=1..i --quiet
        if i>j
            ccc[i,j]=1
        endif
    endloop
endloop

loop i=1..num_events  --quiet
  CUM_AR$i = ccc*AR$i
  vvv$i=ccc*v$i*ccc'
  series CAR$i = CUM_AR$i
endloop
```

Figure 4.10: Gretl code (5)

In the lines of code shown in the image above, the computation of the cumulated abnormal returns is presented. It is crucial to notice that the cumulation of the abnormal returns is done along two dimensions, which are time and events. The cumulation through time is done by creating the identity matrix "ccc", whose dimensions are given by the length of the event-window. This matrix is then fulfilled though the execution of the double loop shown in the script, which is aimed at putting a series on ones in the portion of the matrix situated below the diagonal. The resulting matrix is then multiplied by the vector of the abnormal returns in order to compute the cumulation through time. After that, the variance-covariance matrix associated to the cumulated abnormal returns is calculated. At this point, the variable CAR$i is created in order to store the cumulated abnormal returns as a series. After that, the cumulation is done also along the second dimension, namely through events. In order to do so, the cumulated abnormal return among all the events is computed. As shown by the portion of the code in the imagine below, the average is performed dividing the cumulated abnormal returns of each event by the total number of events, and then summing them through the execution of a loop. The same procedure holds true also for the variance-covariance matrix associated to the cumulated abnormal returns.

```
series CAR = (1/num_events)*CAR1
vvv = (1/num_events^2)*vvv1
loop i=2..num_events  --quiet
  series CAR = CAR+(1/num_events)*CAR$i
  vvv = vvv+(1/num_events^2)*vvv$i
endloop
series sigma2_CAR = diag(vvv)
series LB = -1.96*sqrt(sigma2_CAR)
series UB =  1.96*sqrt(sigma2_CAR)
gnuplot CAR LB UB { set title "CAAR"; set yrange [-5:5]; set y2range [-5:5]; }
```

Figure 4.11: Gretl code (6)

The last lines of the code are the ones allowing the creation of the plot necessary to assess the eventual impact of the focal event. In particular, it is important to notice that, in order to create a plot in GRETL, all the input variables must be transformed in the form of a series. After that, the next passage consists in the definition of the two confidence boundaries, which are depicted by the variables "LB" and "UB", and whose formulations are shown in the image. Eventually, the function "gnuplot" draws the final plot.

5 Analysis of the Results

This chapter is dedicated to the presentation of the results arising from the running of the Event Study methodology. In particular, in a first instance the analysis is conducted at a macro level, considering the entire pool of the companies, regardless of their belonging sector and referring financial market. After that, in order to go more in dept with the investigation, the specific case of the Italian market is analysed: only this geographical area is considered due to its statistical relevance for the analysis with respect to the other financial markets. These two analyses described above have the objective to respond to the first research question, which is reported in the section [4.1]. After this test, the second research question is investigated, splitting the events according to their year of occurrence. In particular, the two considered time-periods are 2016-2018 and 2019-2020 respectively.

Before moving to the explanation of the results, it is fundamental to explain the criterion applied in order to interpret the output of the Gretl's algorithm. The particularity of the applied model is given by the fact that the interpretation of its results can be visualised though a graphical representation.

Indeed, the outcome of the model is a graph composed by two fundamental elements:

- *Two boundaries*, which delimit the confidence region in which the event can be considered as not impacting.

- The *CAAR* which, as previously explained in the methodology section [4.2], represents the cumulated value of the abnormal returns through time and through securities.

In order to understand if an event can be considered as relevant or not by a statistical viewpoint, the following rule is applied:

- An event is defined as *impactful* for the value of the companies if the CAAR exceeds one of the two confidence boundaries.
- An event is considered as *not relevant* for the value of the companies in the case in which the CAAR remains between the two confidence boundaries.

5.1. General Analysis

Moving to the analysis of the obtained results, the first outcomes returned by the algorithm are the estimated parameters computed through the running of the OLS for each one of the Market Model regressions. Focusing on the interpretation of the core results of the Event Study, it can be seen from the Figures 5.1, 5.2 and 5.4 that the CAARs remain inside the confident region in all the 6 cases, meaning that the ESG rating updates do not have any relevant impact on the value of the companies involved in the study. More in depth, the following sections is dedicated to the description of the results, divided according to the typology of the rating update (namely upgrade, downgrade and confirmation of the evaluation).

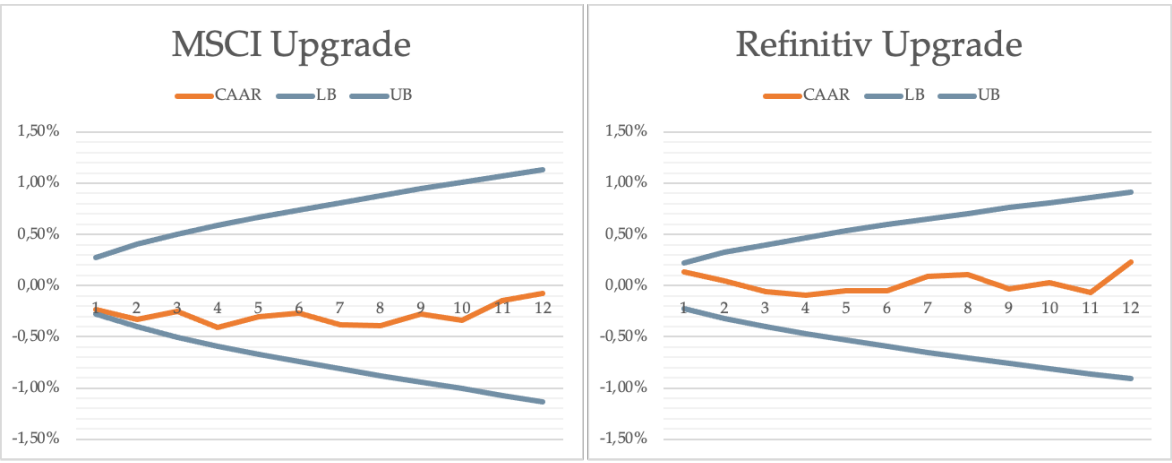


Figure 5.1: CAAR of MSCI and Refinitv upgrades

Starting from the results for the upgrade events shown in the Figure 5.1, the trend of the CAAR is mostly flat. The only difference that can be noticed is that, on the one hand, the CAAR of Refinitiv tend to remain stable over time while, on the other hand, the response to the MSCI's upgrade seems to be negative in the first observation days, but without exceeding the lower confidence boundary. This stylized fact is in line with what has been observed in the above cited study by Miyamoto (2015) [46]: in this paper the author found that the Japanese stock market reacted negatively to the upgrade of the credit rating. In that case, the author explained this dynamic with the capacity of financial investors to anticipate the positive announcement of the change in the rating, buying the stocks before and selling them after the occurrence of the event, leading to a depreciation of the stocks. The same dynamic could hold true in minimal terms also in the case of positive changes in the ESG ratings, explaining the slightly downward fluctuation of the CAAR in the first time buckets of the observation period.

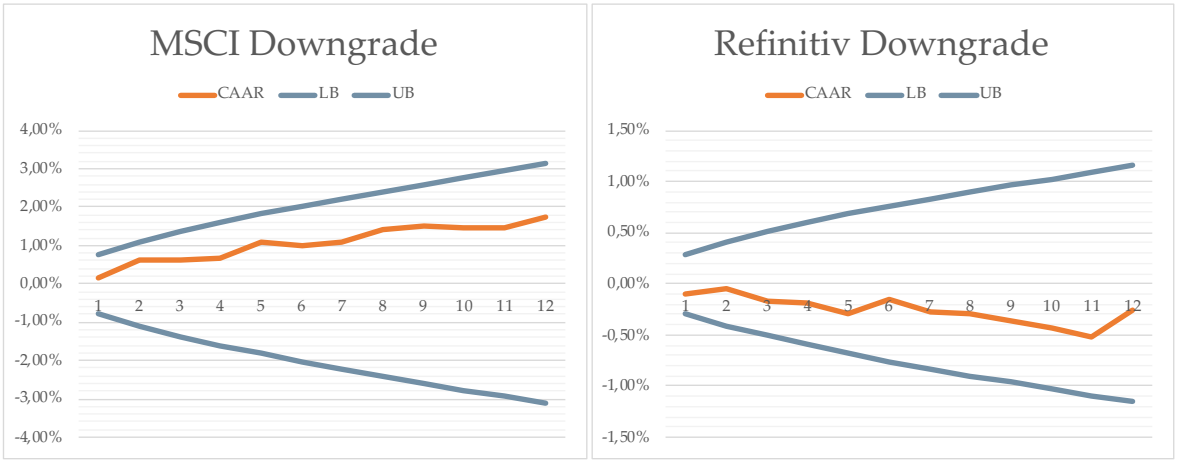


Figure 5.2: CAAR of MSCI and Refinitiv downgrades

Moving to the downgrade event, even in this case the CAAR's tendency is to remain stable over time for both the rating providers. In particular, the reaction of the market to the MSCI announcement has an opposite trend with respect to the one observed in the upgrade case: as shown in the graph, it can be observed that this reaction is slightly positive, even if the downgrades keep on being not statistically relevant. Also in this case, this dynamic is coherent with the results of the above-cited study by Miyamoto (2015) [46], but with a specular explanation with respect to the case of the upgrades. According to this paper, one possible explanation for this stylized fact could be that the financial investors, knowing in advance the negative change in the sustainability ratings, could decide to short-sell the corporations' stocks, causing an increase in the securities' prices after the occurrence of the event. For what concern Refinitiv, as in the upgrade case the CAAR maintains more stability over time, without presenting any particular trend. Before moving to the comment of the results obtained in the case of grade confirmations, the next step consists in conducting a dedicated analysis in order to test the validity of Miyamoto's theory. In particular, if the Japanese author was right, the expectation should be the one of observing a positive (negative) trend of the CAAR in the period

before the occurrence of the event in the case of upgrade (downgrade). To verify this hypothesis, the testing analysis is conducted with the same methodology, but anticipating the event-window by 5 market days. As it is possible to see from the Figure 5.3, the CAAR does not show any trend in the days before the event occurrence, both for the upgrade and the downgrade cases. This leads to the conclusion that the theory of Miyamoto cannot be used in order to justify the results obtained from the previous analysis.

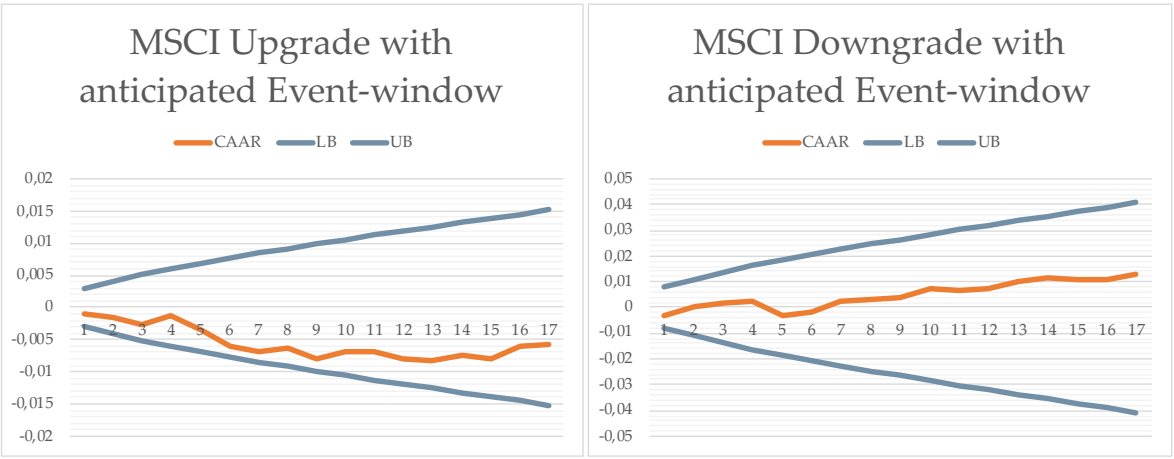


Figure 5.3: CAAR of MSCI upgrade and downgrade with anticipated event-window

Finally, looking at the case of the ESG grade confirmation, it is possible to observe that there are not specific patterns in the fluctuation of the CAAR. In particular, for both the rating providers, the CAAR remains permanently inside the confident region, without never approaching the two external boundaries. Given the fact that both upgrades and downgrades did not result impactful for the market companies, what is found in the case of grade confirmation is reasonable and coherent.

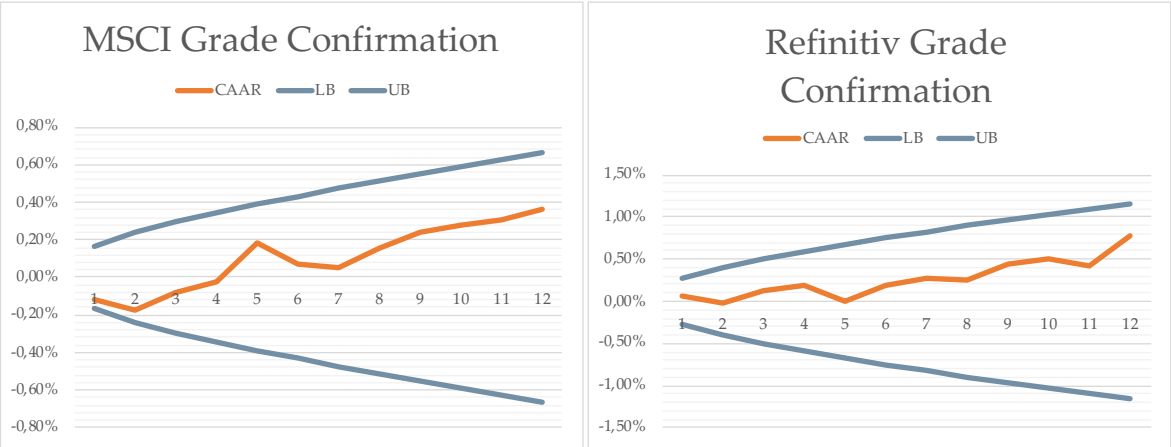


Figure 5.4: CAAR of MSCI and Refinitiv grade confirmations

5.2. Italian Market Analysis

After having described the results found for the overall set of companies, this paragraph is focused on the analysis of the Italian stock exchange market. In this regard, the first important premise to be done is that in some cases, as the one of the MSCI's downgrades, the results of the analysis cannot be considered as always relevant under a statistical viewpoint due to the scarcity of the number of events. Moving to the comment of what is returned by the algorithm for the Italian companies, it is possible to observe from the graphs that the overall tendency of the market seems to be aligned with the one found in the analysis conducted on the complete pool of companies. Indeed, except for the case of MSCI's upgrades, in the other five cases the reaction of the financial investors to the ESG rating agencies' announcements appears to be as not impactful for the value of the corporations. Indeed, the CAAR keeps on being inside the confidence region but, differently from what happened in the previous general analysis, it shows an even flatter trend over time.

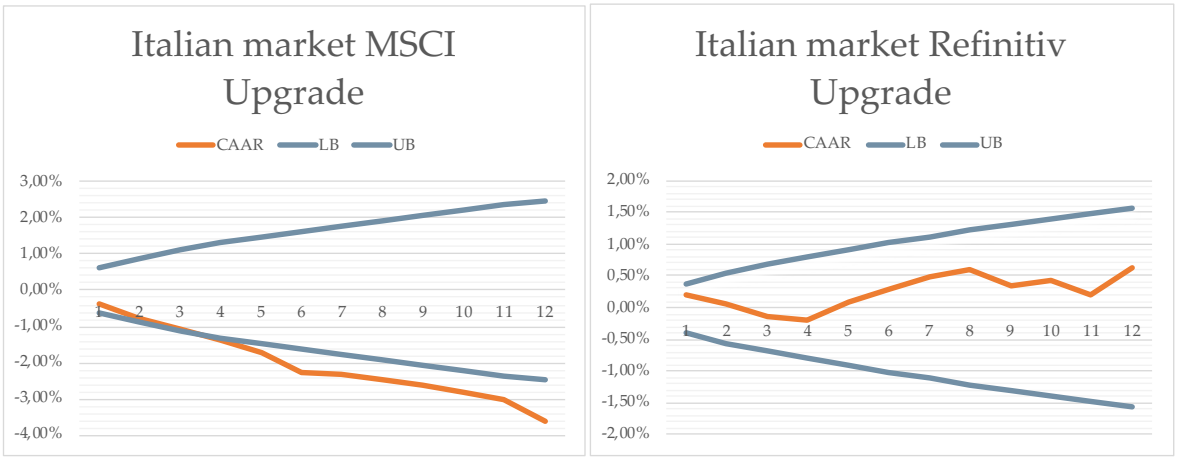


Figure 5.5: CAAR of MSCI and Refinitiv upgrades in the Italian market

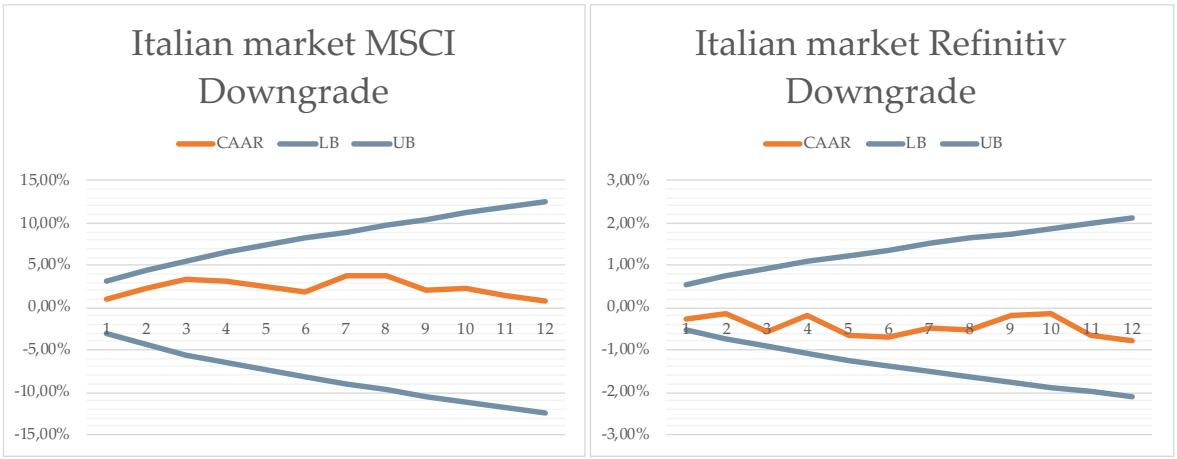


Figure 5.6: CAAR of MSCI and Refinitiv downgrades in the Italian market

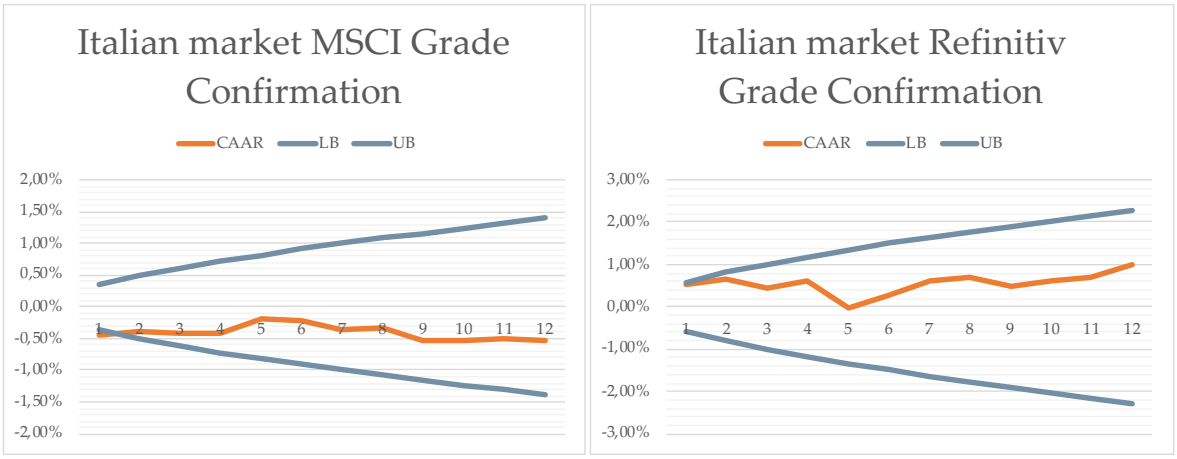


Figure 5.7: CAAR of MSCI and Refinitiv grade confirmations in the Italian market

For what concern, instead, the case of MSCI's upgrades, how it is possible to see from the Figures 5.5, 5.6 and 5.7, this is the only case in which the event appears to be relevant, since the CAAR exceeds the lower bound. Even in this case, one possible explanation for this negative reaction of the market could be found in the contribution provided by Miyamoto (2015) [46]. To test this hypothesis, it is decided to apply the Event Study methodology anticipating the starting point of the event-window by 5 working days. This analysis allows to understand if this negative reaction of the market is determined by the fact that the financial investors are able to acquire in advance the pieces of information about the increase in the ESG rating. More in dept, if the hypothesis proposed by Miyamoto holds true, the expectation is to observe an increase of the abnormal returns and, consequently, a positive trend of the CAAR in the period before the event. The Figure 5.8 shows the result of the conducted analysis. Looking at the chart, it is possible to see that there is not any positive trend in the return in the days before the occurrence of the event, meaning that the hypothesis mentioned above has to be rejected. This result is coherent with the test conducted on the overall set of companies.

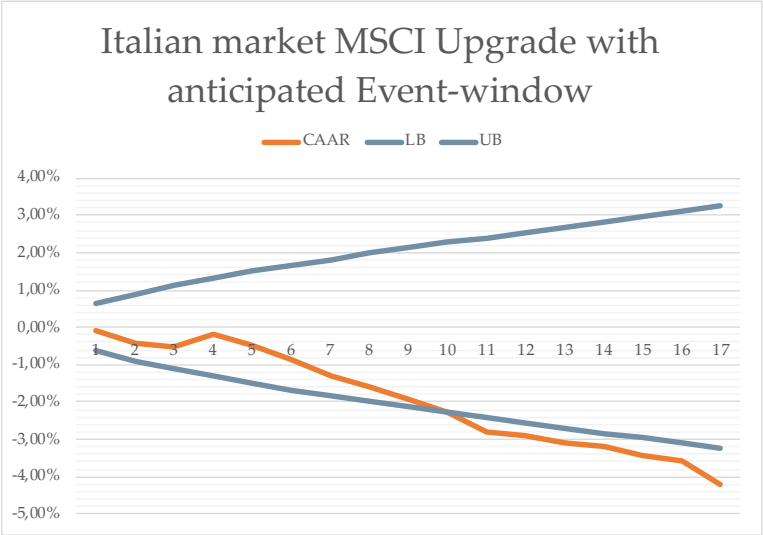


Figure 5.8: CAAR of MSCI upgrade with anticipated event-window in the Italian market

Excluding this hypothesis, it is natural to conclude that nowadays ESG ratings updates are not associated in most of the cases to creation of value both for

companies and financial investors. The result obtained leads to the conclusion that the market value of the corporations is not impacted by the ESG ratings updates. One possible explanation for the finding concerning the first research question could be found by referring to the study conducted by Billio et al (2020) [29]: the existing divergence between the sustainability grades disperses the preferences of the financial investors, which are then disincentivised by integrating the ESG ratings in their investment choices. Hence, the trend of the stock prices does not follow the one of the sustainability grades updates.

5.3. Temporal Analysis: A Comparison between 2016-2018 and 2019-2020

The next section is dedicated to the investigation of the second research question of the present work, with the aim of understanding if the ESG ratings have assumed a higher value in the eyes of the investors in the last years. In order to do so, the same analysis is conducted dividing the initial 5-year time horizon into two temporal subsets of 3 and 2 years respectively. In particular, the first time period goes from 2016 to 2018, while the second one ranges from 2019 to 2020. The last two years are considered coupled and not separated in order to have a wider pool of data and, consequently, a higher statistical relevance of the analysis' results; in this regard, it is important to note that, as in the case of the Italian market's analysis, the MSCI downgrade case cannot be considered as highly statistically relevant due to the reduced number of events. In order to make the comment of this analysis' results more understandable, this section is divided according to the two rating agencies.

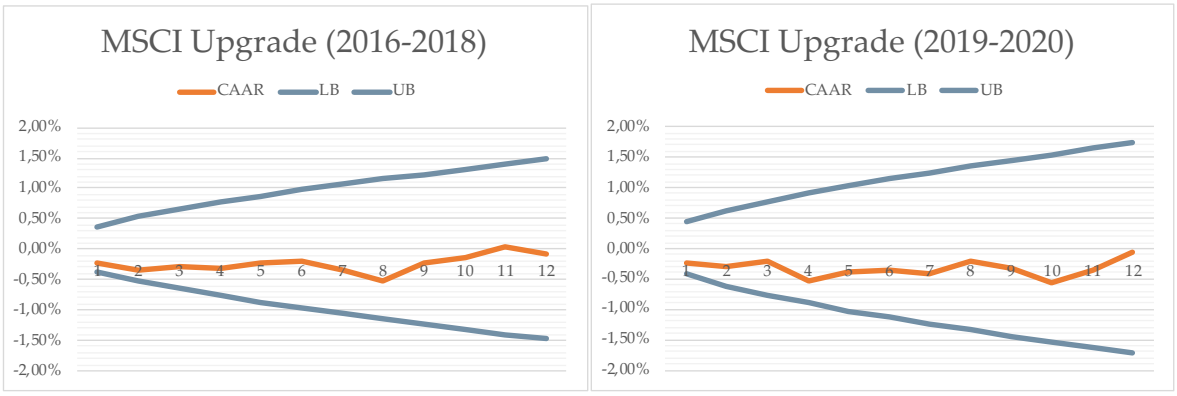


Figure 5.9: CAAR of MSCI upgrades, 2016-2018 vs 2019-2020

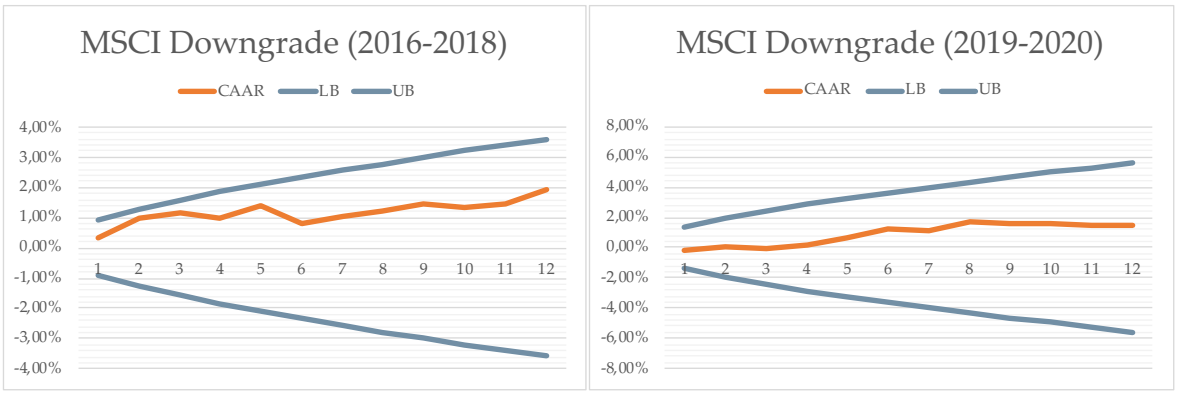


Figure 5.10: CAAR of MSCI downgrades, 2016-2018 vs 2019-2020

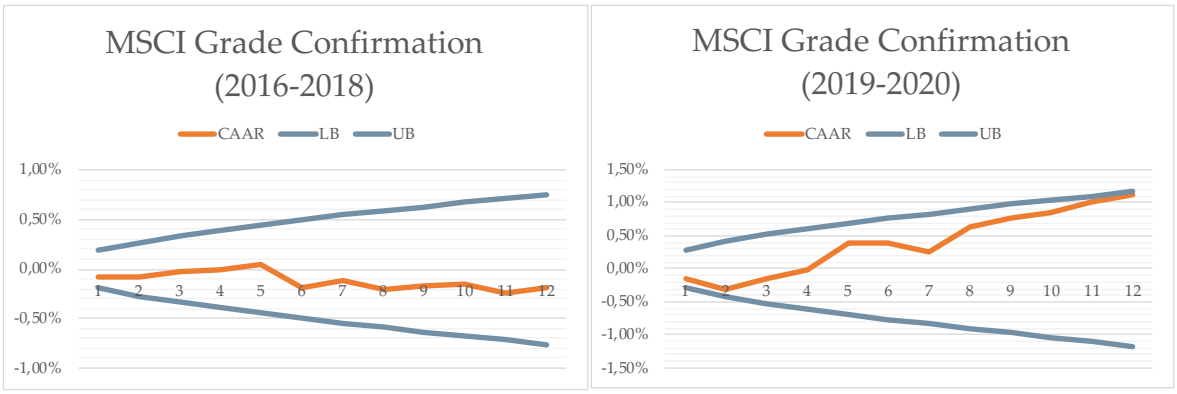


Figure 5.11: CAAR of MSCI grade confirmations, 2016-2018 vs 2019-2020

Starting from MSCI, the first general consideration to be done is that in all the 6 cases the CAAR does not cross neither the upper boundary nor the lower one, remarking the fact that the impact of the ESG rating updates is always not relevant. Going more in dept with the analysis of the different events, the upgrade case's CAAR appears to show the same trend in both the time periods. On the other hand, in the interval 2019-20 the CAAR of the downgrade case seems to be flatter and closer to the lower boundary with respect to the previous period, in which instead it shows an unexpected increase of the abnormal returns in the first days of observation; this change in the trend of the cumulated average abnormal returns is more coherent with the intuitive expectations. Finally, for what concern the case of the grade confirmation, it returns the most particular results among the three event typologies. Indeed, in the first period of analysis, the CAAR remains constantly near to the zero value. Instead, after 2019, it assumes a counterintuitive tendency, approaching the lower boundary in the first market days and the upper boundary in the last ones. The reason why this trend is not coherent with the expectation is that, being this the most neutral typology of event, a rational reasoning would expect at least a confirmation of the flat tendency.

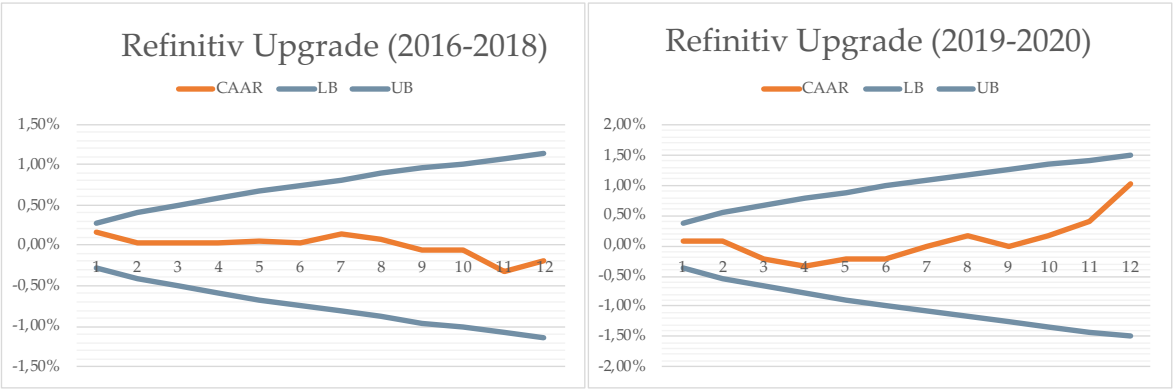


Figure 5.12: CAAR of Refinitiv upgrades, 2016-2018 vs 2019-2020

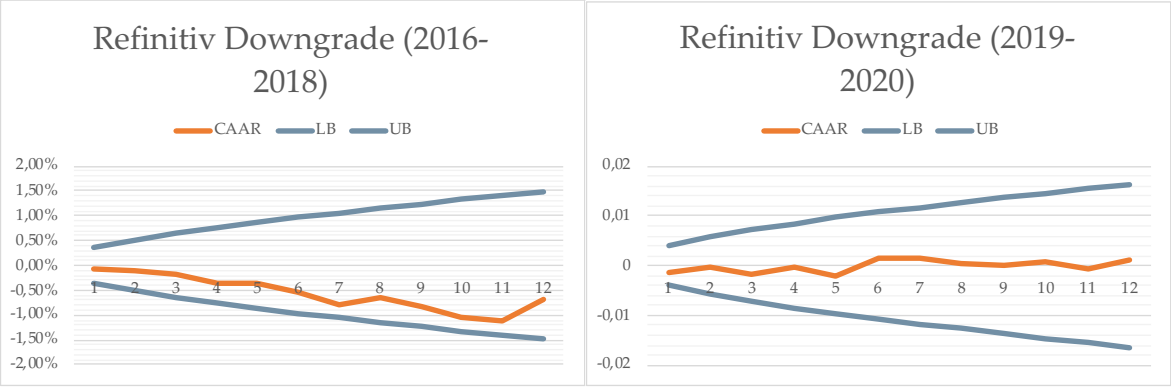


Figure 5.13: CAAR of Refinitiv downgrades, 2016-2018 vs 2019-2020

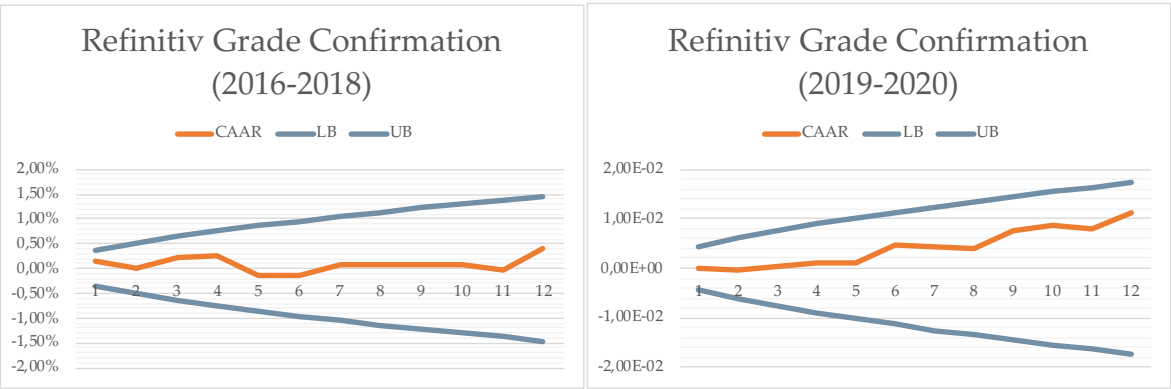


Figure 5.14: CAAR of Refinitiv grade confirmations, 2016-2018 vs 2019-2020

Moving to Refinitiv, even in this case there is not any significant event, regardless of the typology. For what concern the upgrade case, it seems to be the only one showing a slightly positive change in the reaction of the markets to the ESG grade updates in the last two years; indeed, as it is possible to observe from the graphs above, the CAAR remains constantly stable around the zero in the first time-period, while it assumes a positive tendency in the last considered interval. Differently from the upgrade case, the downgrade one presents a less coherent result: on the one hand, in the 2016-2018 period the CAAR is more aligned with the rational expectations, being near to the lower bound and so being close to have a negative impact. On the other hand, in the last two years it loses its negative trend, and it remains permanently in the central part of the confidence region. Finally, the case

of the grade confirmation returns a result that is very comparable to the one obtained in the corresponding event of the MSCI's analysis. As a matter of fact, the CAAR exhibits a flat trend before 2019 while, after this threshold, it assumes a counterintuitive positive trend, moving parallelly to the upper bound. As in the MSCI case, even in this occurrence the empirical finding is not so coherent with the neutral nature of this event typology.

After the description of the results obtained from the temporal analysis conducted on the two rating agencies, it is possible to conclude that the reaction of the financial markets to the sustainability ratings updates has not become stronger in the last two years, thus responding to the second main question of the present work. This finding is motivated by the fact that the results of the last two years do not show a clear change in the market reaction to the ESG rating updates, with the CAAR never crossing the upper or the lower confidence boundary regardless of the event typology and the rating agency considered. A possible explanation for this finding could be the following one: even if the growth of the sustainable assets under management testifies an increasing attention of investors toward sustainability, the ESG finance is far from being mature, with only a narrow fraction of the financial actors integrating ESG ratings in their investment processes; as a result, the ESG performances of corporations have not a direct impact on their market capitalizations.

5.4. The impact of MSCI and Refinitiv grades' announcements on the market

Having investigated the two research questions presented in the section number [4.1], the last point to be discussed regards an overview on the market reaction to the two different rating agencies' grades publications. As extensively explained in the present work, the first consideration to be done is that the two raters differ for the methodology applied in order to assess the sustainability evaluations of the corporations, as shown in the sections number [2.3.1] and [2.3.2]. The consequence of this methodical diversity leads to a grade divergence, not only in absolute terms but also in the evolution of the ESG ratings over time. Indeed, as shown in the section number [4.3.2], MSCI is the rating agency that tends to remain more stable in its evaluations along the years while, on the contrary, Refinitiv changes its grades with a higher variability over time. It is important to remark that the comments that will be done in this paragraph refer to the results reported in the section number [5.1], given the fact that the temporal analysis done on the last two years compared with the previous three ones showed that the significance attributed to ESG ratings' updates has not changed over time. Looking at the six charts, it is possible to observe that neither MSCI nor Refinitiv is never impacting for the stocks' valuations for the reasons explained above. For this reason, the comparison between the two rating agencies is done according to the fluctuations of the CAAR inside the confident region. Focusing on the different event typologies, and starting from the upgrade case, the response of financial investors to the rating agencies' upgrades is very similar, with the CAAR presenting a flat trend over time. However, the market reaction to Refinitiv appears to be more coherent with the nature of the event; indeed, in this case the CAAR assumes positive values in the event window, while for MSCI it remains constantly below the zero-value threshold. This difference is

exacerbated in the downgrade case. For this event, the response of investors to MSCI is counterintuitive, showing a positive trend. On the other hand, Refinitiv, coherently with the rational expectations, affects the market evaluations negatively, with the CAAR always remaining below the zero with a negative slope. A remark has to be done for the MSCI case; indeed, for this rating agency the number of events at disposal of the analysis is considerably reduced if compared to the Refinitiv case, decreasing the statistical relevance of the findings. Differently from the other two cases, in the grade confirmation one the CAAR of MSCI and Refinitiv presents the same tendency, moving in parallel to the upper bound with a positive slope. Hence, for this event typology the two rating agencies seem to have a similar influence on the stocks' returns. To sum up, it is possible to conclude that both the rating agencies have not a direct impact on the stocks' prices in compliance with the analysis previously conducted. However, among the two, Refinitiv is the one that seems to influence the financial market's actors more coherently with respect to the nature of the event typology.

6 Conclusion

The present work comes to life in a context where sustainability-related topics are assuming in the last years more and more relevance in the financial world, through the ESG paradigm and its integration in the whole economic system.

In this scenario, the analysis is inspired by two stylised facts: on the one hand, the positive correlation between the sustainability integration and the corporate financial performance demonstrated by most of literature and described in the section number [3.1]. On the other hand, the existence of divergence in the sustainability evaluations provided by the rating agencies, which could lead to a dispersion of the value associated to the ESG topics, as shown in section [3.2].

Starting from these two drivers, this work aims at investigating the reaction of financial markets to rating agencies' ESG grade updates, understanding whether possible variations in the sustainability ratings of the corporations have a direct impact on their value.

In doing so, the purpose of this study is to try to answer to the following research questions:

I. If the changes in the ESG ratings of the corporations exert a direct impact on their market value.

II. If the reaction of the financial market to the sustainability ratings updates has become stronger in the last two years, due to a higher consciousness of the investors about the ESG-related topics.

In order to investigate the two previously mentioned research questions, the methodology adopted is the Event Study, which is a model used in order to evaluate the impact of an economic event on the valuation of the corporations through the assessment of their Cumulated Average Abnormal Returns (CAAR). This quantitative approach is built on one fundamental pillar: the efficiency of financial markets, meaning that every single economic event is instantaneously incorporated in the price of financial assets. Hence, the effects of an economic event can be assessed by observing the company's price in a specific short time window. In the present work, the economic events are represented by the ESG evaluations updates provided in the period ranging from 2016 to 2020 by two rating providers, namely MSCI and Refinitiv. These events are tested on a pool of 75 companies belonging to the main stock exchange markets, and heterogeneous from the belonging industry point of view. The analysis is conducted on two dimensions: on the one hand, the typology of the grade update (namely upgrade, downgrade or confirmation of the score) and, on the other hand, the rating providers (MSCI and Refinitiv). As a consequence, this quantitative work presents six Event Studies, which are conducted on more than 700 ratings updates (events).

In addition, this study deepens the sustainability ratings divergence topic, showing how the two rating agencies provide different, and sometimes even opposite, evaluations and investigating whether financial investors show different reactions to ESG ratings updates.

The first research question is tested in a first instance on all the events regarding the entire pool of companies. The outcome of this first step reveals that ESG ratings updates are not impactful for the companies' value, regardless of the sustainability grade update typology and of the rating agency making the announcement. Furthermore, the same analysis is conducted also on the events belonging exclusively to the Italian stock exchange, returning the same results of the overall previous analysis. The fact that the CAAR results not significant for any considered case leads to the conclusion that nowadays ESG ratings are not associated to a change of value of the corporations in the eyes of the financial investors. Two possible explanations could motivate this finding: on the one hand, as demonstrated in the study by Billio et al (2020) [29], the divergence between rating providers' methodologies and, consequently, between their ESG assessments, could disperse the preferences of financial investors, who are then disincentivised by looking at sustainability grades during their asset allocation processes. On the other hand, the second explanation could be that most of the financial actors are not interested in integrating sustainability, and so ESG ratings, into their investment decisions. As a result, the trend of the ESG ratings updates is not reflected by the market evaluation of the subject companies.

The second research question is investigated by comparing the market reactions to the events belonging to the period 2016-2018 versus the ones belonging to the period 2019-2020. Even in this case, the events are not influencing the companies' stock prices and, moreover, the CAAR does not show a significant change in its trend passing from the first time interval to the second one. The results obtained from this temporal analysis make possible to conclude that the reaction of the financial market to ESG ratings updates has not changed in the last two years. A possible explanation for this evidence could be the following one: even if

sustainability-related topics are gaining increasing importance in the financial world, as shown in the section number [2.2], the ESG finance is not mature yet, still being in a transition phase. For sure, an inversion of the trend toward sustainability has been observed in the last years, helped by the occurrence of extreme weather events, social scandals and, last but least, by the Covid pandemic. However, this process is not discrete, but it is characterised by a gradual and continuous nature, so more time is needed in order to observe a concrete switch of the financial investors' preferences toward sustainability.

For what concern the third objective of the analysis, namely the comparison between MSCI and Refinitiv ratings updates, the first consideration is related to the different evaluations provided by the two raters: by adopting different methodologies (as shown in the section number [2.3]), the evaluations of MSCI and Refinitiv differ not only in absolute terms, but also in their evolution over time, with the ones of MSCI being more stable along the years. Moving to the market reaction to the two rating agencies' updates, both MSCI and Refinitiv announcements seem to be not impactful for stocks' prices; however, the comparative analysis shows how, in the case of Refinitiv, the direction of the financial investors' reaction is more coherent with the nature of the ratings updates events.

After having presented the main results of the conducted study, the last section of the present work is dedicated to the discussion of the limits of the analysis, its value added to the literature and, finally, the possible further improvements.

The limitations of this work are mainly related to the data gathering process. Indeed, the study is done on 2 rating providers, 75 companies and considering a 5-years time period, for a total of 702 ESG ratings updates events. In order to make

the analysis more robust, it should be considered a higher number of rating providers and companies, enlarging the temporal horizon and, consequently, increasing the number of events feeding the Evet-Study methodology. By the way, it is important to remark the underlying cause of this limitation, which is represented by the existing trade-off between the robustness of the analysis and the effort needed to conduct it. Indeed, when the amount of data increases, the statistical relevance of the analysis does the same but, on the other hand, the data management effort needed in order to process all the pieces of information increases exponentially. Consequently, the choices regarding the data at the basis of this analysis have been made to have the best balance of the above-mentioned trade-off.

The value provided by the present work can be found in the investigation of how the financial markets price the ESG ratings announcements. The results of the present work integrate the exiting literature, showing that nowadays the ESG finance is far from being mature, with most of the financial investors not looking at ESG ratings in their investment processes. As a result, the companies' market capitalizations are not impacted by changes in their sustainability performances.

The last paragraph is dedicated to the possible future developments that could be undertaken in order to complement the contribution provided by this study. In particular, a possible variation may regard the adopted approach: the temporal horizon of the analysis could be changed, looking at the medium-long term companies' performances instead of focusing only on the short-term. Furthermore, recalling the results found through the investigation of the second research question, it emerges that the transition towards a complete ESG integration into financial investments is still on going. Hence, the analysis conducted in the present

work could be replicated in future, with the aim of understanding if the sustainability trend will have reached the mature phase or, on the other hand, if it will present the same profile shown by the present analysis.

Bibliography

[1] World Commission on Environment and Development, "Our Common Future", 1987.

[2] M. Friedman, "The Social Responsibility of Business is to Increase its Profits", The New York Times Magazine, 1970.

[3] Morgan Stanley, "Sustainable Value: How Emerging Public Companies Can Deliver on ESG Expectations", 2020.

[4] UN Global Compact, "*Who Cares Wins*", 2004.

[5] European Banking Authority (EBA), "ON MANAGEMENT AND SUPERVISION OF ESG RISKS FOR CREDIT INSTITUTIONS AND INVESTMENT FIRMS", 2020.

[6] Global Sustainable Investment Alliance, "Global Sustainable Investment Review 2020", 2021.

[7] MSCI, "MSCI ESG Ratings Methodology", 2020.

[8] Refinitiv, "Environmental, Social and Governance (ESG) Scores from Refinitiv", 2021.

[9] Sustainalytics, "ESG Risk Ratings – Methodology Abstract",2021.

[10] Vigeo Eiris, "ESG Assessment Methodology", 2020.

[11] S&P Global, “Measuring Intangibles – The S&P Global Corporate Sustainability Assessment Methodology”, 2021.

[12] J. Derwall, N.Guenster, R. Bauer and K. Koedijk, “The Eco-Efficiency Premium Puzzle”, Financial Analysts Journal, 2004.

[13] S. Manrique and C.P. Marti-Ballester, “Analyzing the Effect of Corporate Environmental Performance on Corporate Financial Performance in Developed and Developing Countries”, Sustainability, 2017.

[14] A. Edmans, “Does the stock market fully value intangibles? Employee satisfaction and equity prices”, Journal of Financial Economics, 2011.

[15] P.A. Gompers, J.L. Ishii and A. Metrick, “CORPORATE GOVERNANCE IN EQUITY PRICES”, Quarterly Journal of Economics, 2003.

[16] P. Velte, “Does ESG performance have an impact on financial performance? Evidence from Germany”, Journal of Global Responsibility, 2017.

[17] W.C. Wong, J.A. Batten, A.H. Ahmad, S.B. Mohamed-Arshad, S. Nordin and A.A. Adzis, “Does ESG certification add firm value?”, Finance Research Letters , 2020.

[18] A. Fatemi, M. Glaum and S. Kaiser, “ESG performance and firm value: The moderating role of disclosure”, Global Financial Journal, 2017.

[19] L. Brogi and V. Lagasio, “Environmental, social and governance and company profitability: Are financial intermediaries different?”, Corporate Social Responsibility and Environmental Management, 2018.

[20] I. Sila and K. Cek, “The Impact of Environmental, Social and Governance Dimensions of Corporate Social Responsibility on Economic Performance: Australian Evidence”, Elsevir B.V., 2018.

[21] A. Aouadi and S. Marsat, “Do ESG Controversies Matter for Firm Value? Evidence from International Data”, Journal of Business Ethics, 2016.

[22] G. Giudici, A. Annese and M. Corradini, “The dynamics of ESG ratings and the return on bonds listed in Europe”, Banor SIM-Politecnico di Milano, 2020.

[23] G. Friede, T. Busch and A. Bassen, “ESG and financial performance: aggregated evidence from more than 2000 empirical studies”, Journal of Sustainable Finance & Investment, 2015.

[24] J. Nofsinger and A. Varma, “Socially Responsible Funds and Market Crises”, Journal of Banking & Finance, 2014.

[25] H.M. Henke, “The effect of social screening on bond mutual fund performance”, Journal of Banking & Finance, 2016.

[26] G. Giese and L.E. Lee, “WEIGHING THE EVIDENCE: ESG AND EQUITY RETURNS”, MSCI ESG Research LLC, 2019.

[27] F. Berg, J.F. Koelbel, R. Rigobon, "Aggregate Confusion: The Divergence of ESG Ratings", MIT Sloan, University of Zurich, 2020.

[28] V. Capizzi, E. Gioia, G. Giudici and F. Tenca, "THE DIVERGENCE OF ESG RATINGS: AN ANALYSIS OF ITALIAN LISTED COMPANIES", Politecnico di Milano, 2021.

[29] M. Billio, M. Costola, I. Hristova, C. Latino, L. Pellizzon,"Inside the ESG ratings: (Dis)agreement and performance", Ca Foscari, Libniz Institute for Financial Research SAFE, Goethe University, 2020.

[30] A. Rzeznik, K.W. Hanley, L. Pelizzon, "The Salience of ESG Ratings for Stock Pricing: Evidence From (Potentially) Confused Investors", Libniz Institute for Financial Research SAFE, 2021.

[31] J.Y. Campbell, A.W. Lo and A.C MacKinlay, "The Econometrics of Financial Markets", Princeton University Press, 1997.

[32] J.S. Brown and Warner, "Measuring Security Price Performance", Journal of Financial Economics, 1980.

[33] J.S. Brown and Warner, "Derived Factors in Event Studies", Journal of Financial Economics, 1985.

[34] W.F. Sharpe, "Portfolio theory and capital markets", New York: McGraw-Hill,1970.

[35] W.F. Sharpe, G.J. Alexander and J.V. Bailey, "Investments", Fifth Ed. Englewood Cliffs, New York: Prentice-Hall, 1995.

[36] W.F. Sharpe, “Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk”, Journal of Finance, 1964.

[37] J. Lintner, “The Valuation of Risky Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets”, The Review of Economics and Statistics, 1965.

[38] S.A. Ross, “The Arbitrage Theory of Capital Asset Pricing”, The Journal of Economic Theory, 1976.

[39] J.C. Dolley, “Characteristics and Procedure of Common Stock Split-Ups”, Harvard Business Review, 1933.

[40] H.J. Myers and A.J. Bakay, “Influence of Stock Split-Ups on Market Price”, Harvard Business Review, 1948.

[41] Barker and C. Austin, “Effective Stock Splits”, Harvard Business Review, 1956.

[42] J.W Ashley, “Stock Prices and Changes in Earnings and Dividends: Some Empirical Results”, Journal of Political Economics, 1962.

[43] R. Ball and Brown, “An Empirical Evaluation of Accounting Income Numbers”, Journal of Accounting Research, 1968.

[44] F.E. Fama et al., “The Adjustment of Stock Prices to New Information”, Int. Econ. Rev., 1969.

[45] D. Ikenberry, J. Lakonishok and T. Vermaelen, “Market underreaction to open market share repurchases”, Journal of Financial Economics, 1995.

[46] M. Miyamoto,“Event Study of Credit Rating Announcement in the Tokyo Stock Market”, Journal of Economics, Business and Management, 2015.

[47] W.R. Holthausen and R.W. Leftwich, “The effect of bond rating changes on common stock prices”, Journal of Financial Economics, 1986.

[48] J. Cummings, A. Dennis, S. O’Donoghue and K. Cooney, “The Benefit of Being Second: An Event Study of Social Media Adoption”, ,2018.

[49] E.M. Rogers, “Diffusion of Innovations”, New York: Free Press, 1995.

[50] Sustainability Accounting Standards Board, “*SASB Conceptual Framework*”, 2017.

[51] A. Chiu, L. Chen and J. Hu, “A Study of the Relashionship between Corporate Social Responsibility Report and the Stock Market”, Sustainability, 2020.

Index of Figures

Index of Tables