MARKET REACTION COMPARISON BETWEEN SOCIALLY RESPONSIBLE AND CONVENTIONAL ETFS DURING THE COVID-19 CRISIS: A CLUSTERING APPROACH

Lappeenranta–Lahti University of Technology LUT

Master’s Programme in Strategic Finance and Analytics

2022

Ella Nurkkanen

Examiners: Post-doctoral researcher Christoph Lohrmann

Associate Professor Sheraz Ahmed

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT School of Business and Management

Business Administration

Ella Nurkkanen

Market reaction comparison between socially responsible and conventional ETFs during the Covid-19 crisis: a clustering approach

Master's thesis

2022

83 pages, 18 figures, 5 tables and 5 appendices

Examiners: Post-doctoral researcher Christoph Lohrmann, Associate Professor Sheraz Ahmed

Keywords: Socially responsible investing, ESG, ETF, crisis, unsupervised machine learning, cluster analysis, k-means

This thesis aims to investigate ESG and conventional investing during the COVID-19 crisis in the U.S. market area and use unsupervised learning, specifically, cluster analysis to examine the behavior of ESG ETFs and conventional ETFs with respect to their price movements. The purpose of clustering analysis is to find similar-behaving ETFs and group them. The clustering methods used are partitional clustering (k-means) and hierarchical clustering (Ward's method), with Euclidean distance as the distance measure. The COVID-19 pandemic will be studied in three different time periods: the early beginning phase, the collapse phase, and the recovery phase. Finally, Morningstar's sustainability rating and financial metrics (e.g., volatilities, min, max, average, and annualized returns) are used to examine the characteristics of the resulting clusters.

Clustering the examined phases of COVID-19 resulted in a different number of clusters as the early beginning phase resulted in the largest number of clusters while the two other phases resulted in fewer clusters. Results indicated the ESG ETFs formed clusters with more similar cluster characteristics compared to conventional ETFs during the examined phases. Overall, the clusters formed purely of ESG ETFs have had generally significantly lower volatilities during the COVID-19 pandemic compared to conventional ETF clusters and ESG ETF clusters have resulted in slightly higher average returns during the early beginning phase as well as the collapse phase. Lastly, the ESG ETF clusters' average Morningstar sustainability rating also varied across the examined phases. A higher rating is not necessarily associated with better performance, as observed during the market downturn (collapse phase). However, when the market stabilized (recovery phase), a higher sustainability rating was associated with lower price volatility and riskiness of the clusters.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT
LUT-kauppakorkeakoulu

Kauppatieteet

Ella Nurkkanen

Markkinareaktioiden vertailu vastuullisten ja tavanomaisten ETF-rahastojen välillä COVID-19-kriisin aikana: klusterointimenetelmä

Kauppatieteiden pro gradu -tutkielma

2022

83 sivua, 18 kuvaa, 5 taulukkoa ja 5 liitettä

Tarkastajat: Tutkijatohtori Christoph Lohrmann ja Apulaisprofessori Sheraz Ahmed

Avainsanat: Vastuullinen sijoittaminen, ESG, ETF, kriisit, valvomaton oppiminen klusterianalyysi, k-means

Tämän tutkielman tavoitteena on tutkia ESG- ja tavanomaista sijoittamista COVID-19-kriisin aikana Yhdysvaltojen markkina-alueella ja käyttää valvomatonta oppimista, erityisesti klusterianalyysia ESG-ETF:ien ja tavanomaisten ETF:ien hintakehityksen tarkastelussa. Klusterointianalyysia käytetään identifioimaan ja ryhmittelemään ETF-rahastoja, joiden hinnat ovat kehittyneet samankaltaisesti. Klusterointimenetelminä käytetään partitionaalista klusterointia (k-means) ja hierarkkista klusterointia (Wardin menetelmä), ja etäisyysmittarina käytetään euklidista etäisyyttä. COVID-19-pandemiaa tutkitaan kolmena eri ajanjaksona, joita ovat: alkuvaihe, romahdusvaihe ja toipumisvaihe. Lopuksi Morningstarin kestävyysluokitusta ja taloudellisia mittareita (esim. volatiliteetteja, min, max, keskimääräisiä tuottoja sekä vuotuisia tuottoja) käytetään löytyneiden klusterien ominaisuuksien tutkimiseen.

COVID-19 kriisin aikana tutkittujen vaiheiden klusterointi johti erilaiseen klusterien määrään, sillä alkuvaiheessa klustereita oli eniten verrattuna kahteen muuhun vaiheeseen. Tulokset osoittivat, että ESG ETF:ien klustereilla oli enemmän samankaltaisia ominaisuuksia kuin tavanomaisilla ETF:illä tutkituissa vaiheissa. Kaiken kaikkiaan pelkästään ESG ETF:istä muodostettujen klusterien volatiliteetit olivat COVID-19-pandemian aikana yleisesti ottaen huomattavasti alhaisemmat kuin tavanomaisten ETF klustereiden, sekä ESG ETF-klusterien keskimääräiset tuotot olivat hieman korkeammat sekä alkuvaiheen että romahdusvaiheen aikana. Tutkituissa vaiheissa ESG ETF klusterien saavuttamat keskiarvot Morningstarin kestävyysluokituksista myös vaihtelivat. Korkeammalla kestävyysluokituksella ei välttämättä ole yhteyttä parempaan suoriutumiseen kuten markkinoiden laskusuhdanteen (romahdusvaihe) aikana havaittiin. Kuitenkin markkinoiden tasaantuessa (elpymisvaihe) korkeampi kestävyysluokitus indikoi pienempää hintavolatiliteettia ja klusterin riskisyyttä.

ACKNOWLEDGEMENTS

The last obstacle has finally been overcome and the great journey in LUT has come to an end. I am truly grateful for the years spent in Lappeenranta, especially the people I have met, and for friendships that will surely last for a lifetime.

Firstly, I would like to thank my supervisor, Christoph Lohrmann for his invaluable feedback throughout this thesis and the great courses during my master's studies. Thank you, Anni, Emilia, Ronja, Satu, and Tia, life in Lappeenranta would not have been the same without you!

I want to thank my family, friends, and Sami for your continuous support during my studies, especially during the thesis writing process. Finally, I have more time to spend with you!

In Helsinki, 7th of December 2022

Ella Nurkkanen

Abbreviations

AUM	Assets Under Management
BIRCH	Balance Iterative Reducing and Clustering using Hierarchies
CURE	Clustering Using Representative
DTW	Dynamic Time Warping
ESG	Environmental, Social, Governance
ETF	Exchange Traded Fund
GSIA	Global Sustainable Alliance
ML	Machine Learning
MPT	Modern Portfolio Theory
MSR	Morningstar Sustainability Rating
PRI	United Nations Principles of Responsible Investing
SC	Silhouette Coefficient
SDG	Sustainable Development Goals
SRI	Socially Responsible Investing
TCFD	Taskforce on Climate-related Financial Disclosures
UNEPFI	United Nations Environmental Program Finance Initiative
US SIF	U.S. Forum for Sustainable and Responsible Investing

Table of contents

Appendices

Figures

Tables

1 Introduction

“*The only thing that is constant is change*”, is an expression of the ancient Greek philosopher Heraclitus (500 BC). Since the Great Depression of the 1930s, Black Monday in 1987, the Asian Crisis of 1997, the bursting of the dot-com bubble in 2001, and the financial crisis of 2008 have illustrated, that the full extent of the crisis’ consequences are impossible to anticipate. (Iraci and Iraci 2020**;** Sornette 2017) The most recent crisis, the COVID-19 pandemic, and the ongoing aftermath have been arduous reminders of the complex nature of crises and what they can cause to modern civilization and global financial markets. (Zhang, Hu and Ji 2020; Claessens, Kose, Laeven & Valencia 2014) At the beginning of the COVID-19 crisis stock markets suffered greatly due to high volatility, with U.S. equity markets losing up to 30% of their value between February and mid-March 2020 (Takahashi and Yamada 2021).

In times of financial and environmental crisis, financial experts and policy makers have identified socially responsible investing as a potential solution to prevent future crises, as transparency, better risk management and good governance are important processes for recovering from crises as well as mitigating the impacts (PRI 2020). Investors’ interest in Environmental, Social and Governance (ESG) investing have been growing rapidly in the last decades. Investors are keen to integrate ESG-investing into their investment portfolios to increase their profitability, along with norm-based reasons and personal values (Dawkins 2018; Nilsson 2009). Considering integrating ESG investing in the portfolio, investors have been able to sustain their portfolio returns in downturn markets. (Pisani and Russo 2021; Bilbao-Terol, Arenas-Parra, Canal-Fernandez & Bilbao-Terol 2016; Ortas, Moneva, Burritt & Tingey-Holyoak 2014; Singh 2020)

A crisis introduces extremely uncertain times to the markets, where the volatility is high, markets are unpredictable and downside risks are great (Stephens, Benik, Gordillo & Pardo-Guerra 2021; Broadstock, Chan, Cheng & Wang 2021). In order to gain information and analyze occurring market trends or find hidden trends in the markets, machine learning tools have been used in many finance-related applications, as asset allocation is one of the key elements in financial markets and portfolio management (Zhang, Zhao, Li, Gao, Kong & Chen 2020).

1.1 Background of the research

The outbreak of the COVID-19 pandemic has been one of a kind, it has not only affected people's health all over the world but has also caused significant restrictions on the socio-economic activities of countries globally (Nayak et al. 2022; Singh 2020). To control the pandemic, many countries have performed strict restriction measures to break the transmission chains. According to Javed, Sarwer, Soto & Mashwani (2020) quarantines, self-isolations and limitations of human-to-human interactions have had a negative effect on individuals' mental health. Grewenig, Lergetporer, Werner, Woessmann, and Zierow (2021) studied the pandemic's impact on education and inequality, particularly how schools' closures affected low- and high-achieving students. They conclude that in both groups time spent in learning was significantly decreased when students were distance-learning. Especially low-achieving students were affected when support from teachers is not available. (Grewenig et al. 2021)

The public health care sector has been also dramatically affected by the pandemic. According to Helsper et al. (2020) since the health care services have been coping with the COVID-19 disease, non-COVID related care has been downscaled and especially cancer-related early diagnoses have been halted. In addition, Almeida et al. (2021) in their study concluded that the crisis has had a severe effect on households' disposable income, especially in the households that belong to the lowest decile of the income distribution. According to Mazur, Dang and Vega (2021) during the times of imposed quarantines, many businesses were unable to operate to their full extent. This made businesses adjust their labor costs which resulted in laying off employees, a sharp reduction in consumption as well as in economic output. (Mazur, Dang and Vega 2021)The outbreak of the COVID-19 pandemic has had a great impact on financial markets. It has been referred to as one of the biggest stock market crashes since the Great Recession in 2008 (Shu, Song & Zhu 2021).

In the U.S. stock market, an emergency measure threshold called the circuit breaker mechanism, which halts the trading activity in case prices decline excessively, was reached four times in a ten day timespan. This mechanism has been previously triggered only once since its inception in 1997 (Shu et al. 2021; Hassan & Riveros Gavilanes 2021; Zhang et al. 2020). In Chinese markets, He, Sun & Zhang (2020) conclude that stock prices on Shanghai Stock Exchange were negatively affected by the pandemic. He et al. (2020) also mention COVID-19's negative effect on traditional industries especially traveling and transportation, whereas some non-traditional industries related to information technology and health, were able to even benefit from the pandemic. The results from the U.S. stock markets are similar to Chinese markets (Mazur, Dang and Vega 2021; Thorbecke 2020). Albuquerque, Koskinen, Yang and Zhang (2020) state that the COVID-19 pandemic introduces an unparalleled shock, where at first abrupt lockdowns were an unexpected shock to global stock markets. Further on, the pandemic has been an exogenous shock that arose out of public health concerns, not from economic conditions (Albuquerque et al. 2020).

1.2 Research questions and the aim of the study

Investment objective selection and construction of an optimal portfolio is investors' purpose to obtain excess returns. The proportions of different assets to make a portfolio that maximizes the expected returns while also minimizing the certain risk level is challenging in even normal market conditions.

According to socially responsible investing (SRI) literature, in times of crisis, one of the reasons ESG investments have been able to sustain downturn markets are loyal shareholders (Albuquerque et al. 2020; Omura, Roca and Nakai 2021). In previous crises, where performance comparisons have been conducted between conventional investments and ESG investments, there is evidence of a trend that ESG investments have been able to hold better financial stability in crises (Nofsinger and Varma 2014) as well as in recent pandemic (Albuquerque et al. 2020; Omura et al. 2021). However, contradicting results have been obtained in different crisis periods (Leite and Cortez 2015; Lean and Pizzutilo 2021). Previous literature mostly considers conventional mutual funds and ESG funds or stocks relationship (Kanuri 2020; Pavlova and de Boyrie 2022). Thus, in this research ETFs

(Exchange Traded Funds) are considered instead due to their growing popularity among investors and to provide insights for them.

This study examines the stock market reaction between conventional ETFs and ESG ETFs generated by the COVID-19 pandemic in the United States market area. Figure 1 below presents the theoretical framework of the study. The impact of a crisis on the stock markets is a highly interesting topic since all crises have their individual features, origins, and impacts on the markets. Comparing two different ETFs provides insights into how these two different types of ETFs have reacted to the crisis. As there is growing interest in ESG investing in times of crisis, this study will give observations on this topic. This will be executed by combining the examination of the market reaction with respect to price movements of conventional ETFs and ESG ETFs by applying unsupervised learning, specifically, using time series clustering analysis on specific time windows before and after the COVID-19 crash.

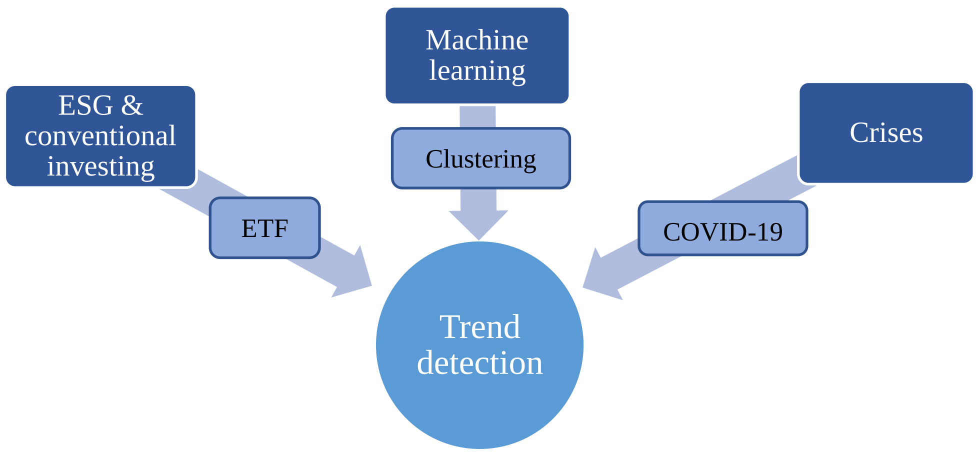


Figure 1. Theoretical framework of the study

Time series clustering analysis has become a popular machine learning approach since cluster analysis is able to capture similarities in data sets and provide information that is very useful for example in constructing portfolios and examining trends (Zhang et al. 2020; Aghabozorgi, Shirkhorshidi and Wah 2015; Liao 2005). Trend detection using clustering analysis utilizing stock market data have been researched for example by Wang, Huang, Zheng, Lee, and Fu (2021), and Dragut (2012) in times of the U.S. financial crisis 2007-2009.

The study will cover the U.S. market area since there are sufficiently ESG ETFs available for analysis. In addition, the value of sustainable financial assets under management in the U.S. market area is the highest worldwide (Statista 2022a). The main questions will investigate the relationship between machine learning and socially responsible investments as well as conventional investments considering ETFs in three different periods. How conventional and ESG investments have reacted to the previous crises in addition to the COVID-19 crisis, will be covered utilizing relevant literature on stock market behavior in uncertain market situations and crises. Furthermore, whether conventional or ESG ETFs have shown better resilience, measured in excess returns, during different phases of a pandemic is addressed from previous studies' point of view.

Combining socially responsible investing with machine learning for time series clustering analysis is a less researched topic, to which this research is contributing. Therefore, the two main questions are formed as:

Q1: "What kind of performance did socially responsible investing instruments show in market crises according to the literature compared to the conventional instruments?"

Q2: "Can clustering analysis identify different kinds of behavior with respect to price movements between ESG and conventional ETFs in different phases of the COVID-19 crisis?"

The research questions are partially motivated by the research of Dragut (2012), Buszko, Orzeszko and Stawarz (2021), and Pástor and Vorsatz (2020). These researchers consider different time windows in market crisis investigation, where clustering analyses were performed.

1.3 Delimitations of the study

The study focuses on the U.S. market area and chosen ESG, and conventional ETFs are following U.S. indices. The analyzed asset class is equity, thus such investment instruments as green bonds, social bonds or green loans are not included in the empirical part of the study, even though they are considered ESG investing instruments. In this research definitions of SRI and ESG investing are used interchangeably since the academic literature is not homogenous considering these expressions. However, their subtle differences are discussed in the literature review.

In the case of data and particularly used ESG ETFs in this research and their ESG ratings needs to be addressed. There are various rating providers which can result the same companies applying ESG rating being able to obtain different ESG scores depending on the rating company used (Billio, Costola, Hristova, Latino, Pelizzon 2021). This research excludes the rating agencies' evaluation differences of ESG scores.

1.4 Structure of the thesis

This thesis is divided into six sections. After the introduction section, the literature review is presented. The literature review will introduce the theoretical framework of the study and the relationship between crisis and socially responsible investing. Also, the stock market's reaction to the COVID-19 crisis is discussed. In addition, the cluster analysis and its usage in trend detection are covered from previous' studies perspectives as well as the concept of socially responsible investing and ETFs are introduced. The section three will introduce the research data followed by the section four concerning the methodology. The section five presents the clustering results, and section six ends this study with conclusions and the implications of the results.

2 Literature review and previous findings

The second chapter presents the literature review and previous findings. The main theory and concepts are introduced in the context of socially responsible investing and conventional investing in uncertain market conditions. Lastly, ETFs as investment objective is presented and the concept of machine learning is introduced followed by a presentation of clustering approaches.

2.1 Socially responsible investing

There has been growing interest in socially responsible investing in the last decades and it has become a quickly growing segment in the U.S. professionally managed assets (Schueth 2003; Bilbao-Terol, Arenas-Parra, Cañal-Fernández and Bilbao-Terol 2015). The origins of socially responsible investing date back to the 1st century when Jewish law provided guidance on how to invest ethically. Socially responsible investing originally has religious roots, and it is assumed that from 1600 to the mid-1700s the Quakers and Methodists immigrants brought the concept of socially responsible investing to the U.S. At that time, the Methodists managed their wealth using so-called social screens, and Quakers practiced investing activities that would not accept investing in slavery or war. (Schueth 2003; Hyrske, Lönnroth, Savilaakso and Sievänen 2012; Puaschunder 2016)

The movement toward the modern state of social investing in the U.S. was formed in the 1960s when strengthening civil rights and women’s rights and objecting to the Vietnamese war expanded the growth of social awareness. Furthermore, the growth of social awareness broadened the perspective to include labor and management issues in the 1970s and 1980s, when social justice themes reached the public’s eye. (Schueth 2003; Camilleri 2021) Lately, global warming, concern regarding nuclear power, and accidents related to oil companies have reached public attention, which has increased the environmental aspect to the investor’s knowledge. Recent concerns have been related to human rights issues and the rapidly increasing consumption habits of the world's growing population, leading to a reduction in scarce resources. (Schueth 2003) Currently, there are acknowledged a consumer-driven phenomenon where investors' value-based choices are influencing the financial sector in terms of increasing the social investments. This has developed a strong demand for SRI. (Camilleri 2021)

Socially responsible investing is originally referring to an approach where investment decisions are based on investors' personal values and ethics, where the goal is "making good" (Revelli 2017). For example, investor may have a personal value related to environment and preserving the ecosystem, thus the investor's primary motivation would be investing in a company, whose mission is in line with the investor's values (Beal and Goyen 1998). Many synonyms are used for referring this type of investment behavior. Sustainable investing, ethical investing, impact investing, green investing, and ESG investing as well as socially responsible investing are used as synonyms for describing such investing styles. However, there are some nuanced differences found when discussing socially responsible investing and ethical investing. Socially responsible investing is based on ethical investing while simultaneously considering the returns aspect (Derwall, Koedijk and Ter Horst 2011). However, ethical investing is purely driven by ethics and morals. Even though there is heterogeneity in terminology, Sandberg, Juravle, Hedesström and Hamilton (2009) found that definitions used in academic literature regarding SRI are consistently denoting that used terminology describes an investment process, which takes into account also non-financial concerns, such as social, environmental and ethical issues in the decision making. (von Wallis and Klein 2015; Hyrske et al. 2012; Schueth 2003)

2.2 ESG framework and socially responsible investing guidelines

The ESG stands for (E)nvironmental, (S)ocial and (G)overnance attributes that are considered in investing process parallel with investment objective's financial data. In ESG analysis, typically environmental aspect is related to energy efficiency, nature's diversity, standards, climate change, and emissions of company's actions and decisions. The social attribute considers human rights, product liability, and labour rights among other aspects. The governance attribute is related to anti-corruption and anti-bribery actions, tax payments, and the actions of the company's CEO and executive management. Also, the composition of board members as well as their independence and the company's reward system are considered in the governance attribute. (Silvola & Landau 2021; Hyrske et al. 2012)

Socially responsible investing is framed by different guidelines and principles, which are formed to standardize SRI processes. One of the largest actors in this field is the Global Sustainable Alliance (GSIA) and the United Nations Principles of Responsible Investing (PRI). GSIA's purpose is to enhance the international collaboration of membership-based sustainable investing organizations and strengthen globally the impact and visibility of sustainable investment organisations and ESG investing. In the U.S. market area, a member of GSIA is the Forum for Sustainable and Responsible Investing (US SIF). The mission of US SIF is to accelerate the shift to sustainable investment practices, focusing on long-term investment that have a positive impact on environmental and social attributes. (GSIR 2020; US SIF 2020a; PRI Association 2022)

Furthermore, the Paris Agreement, the Sustainable Development Goals (SDGs), the Taskforce on climate-related financial disclosures (TCFD), and the United Nations Environmental Program Finance Initiative (UNEPFI) are examples of global developments which have greatly affected the sustainable investing industry by harmonizing processes as well as financial services industry more broadly. (GSIR 2020)

2.3 Socially responsible investing strategies

In socially responsible investing, there are many approaches and so-called strategies that can be utilized. The strategy's purpose is to consider different attributes of responsibility and integrate the different perspectives of SRI into practice. The different strategies can be utilized either one at a time or simultaneously and the investors choose the most suitable approaches for their preferences. There are several strategies, and the most common and frequently used strategies are presented in Figure 2. They are negative and positive screening, thematic investing, and shareholder activism. Other used strategies are for example community investing and impact investing. (Renneboog, Ter Horst and Zhang 2008; Naffa and Fain 2020)

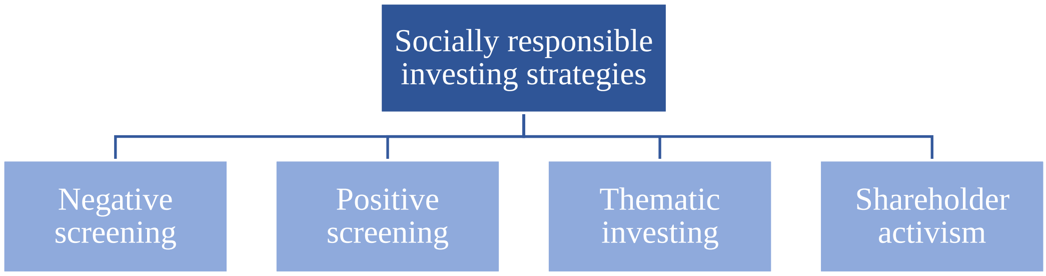


Figure 2. The most common socially responsible investing strategies

2.3.1 Positive and negative screening

The oldest and the most used strategy in socially responsible investing is the exclusion-based strategy called negative screening**.** The focus of this strategy is to evade and exclude businesses and industries that the investor considers controversial, and which do not support investor's social, environmental, and ethical values. Negative screens are also referred to as "sin" screening since specific industries are perceived as controversial by social norms. Such industries are usually related to tobacco, gambling, weapons, alcohol, and animal testing. Renneboog, Ter Horst, and Zhang (2008) find that most SRI mutual funds managers are using multiple screens in the process of determining their portfolios. In the U.S. market area, negative screening is one of the most used strategies. (Leite and Cortez 2015; Schueth 2003)

The relationship between negative screening and financial performance during different economic states was studied by Leite and Cortez (2015) who found that SRI funds perform significantly better in crisis periods than in non-crisis periods compared to conventional funds. However, using purely negative screens as a SRI strategy led to the SRI fund's poor performance in non-crisis periods. (Leite and Cortez 2015)

Screening criteria can be classified into negative and positive screening, and moreover, they can be divided into four wide fields: environmental, social, governance, and controversial business involvement (Capelle-Blancard and Monjon 2014). The second strategy based on screening the investment universe is the aforementioned positive screening. The purpose of this strategy is to select and favor specific investment objectives that follow investor's environmental, social, and ethical values. An essential part of a positive screening strategy is the so-called "best-in-class" approach, where the investor holds only companies who are considered to act most socially and/or environmentally responsible relative to their industry peers. (von Wallis and Klein 2015; Statman and Glushkov 2009; Hyrske et al. 2012)

Funds that only utilize positive screening strategy, have been able to show a good financial performance, particularly in crisis periods compared to conventional funds. Positive screening has been recognized to perform more steadily compared to funds that only use purely negative screening in different market states. (Leite and Cortez 2015) In addition, Statman and Glushkov (2009) concludes that positive and negative screening affected the portfolio's financial performance differently. Negative screening decreased portfolios' financial performance, whereas positive screening did have an increasing effect. Statman and Glushkov (2009) suggest in their study that investors were able to gain the best returns by adopting the "best-in-class" approach in their portfolios.

Moreover, Kempf and Osthoff (2007) investigated the impact of positive, negative, and best-in-class screens on the performance of constructed portfolios, which includes SRI-based companies. Kempf and Osthoff (2007) analyzed if investors can obtain excess returns by utilizing a strategy that is based on SRI ratings. They conclude that investors can gain profit by utilizing a trading strategy, which buys stocks of companies that have high SRI ratings and sell such companies stocks that have low SRI ratings. Thus, their screening activity analysis showed that the best-in-class approach was able to achieve the highest risk-adjusted returns. (Kempf and Osthoff 2007)

2.3.2 Thematic investing

The third strategy presented is thematic investing, which is among the youngest ESG strategies. Thematic investing seeks to identify macro economical, technological, and geopolitical trends that are expected to have a positive impact on the markets and society in long term. The objective of the thematic investing strategy is to recognize possible megatrends and themes that are popular in society considering themes that are related to sustainability. Thematic investing strategy focuses on different challenges and attributes related to ESG. (Naffa and Fain 2020; Ivanisevic Hernaus 2019) Investors can focus on certain sectors and the products and solutions they contain. According to Hyrske et al.

(2012), this could be executed by investing in certain companies' renewable energy technology.

Thematic investing strategy is researched for example by Alvarez and Rodriguez (2015) and Marti-Ballester (2020). Alvarez and Rodriguez (2015) focused on analyzing the financial performance of water sector-based mutual funds and conventional mutual funds. They found no statistically significant differences between the risk-adjusted returns of the funds. A study by Marti-Ballester (2020) centers on analyzing the mutual funds in the healthcare and biotechnology sector related to United Nations sustainable development goal 3 (SDG 3). A comparison of the risk-adjusted returns between the conventional mutual funds and mutual fund focusing on SDG3 resulted during 2007-2018 SDG3 funds outperformed conventional funds. According to Marti-Ballester (2020), the outperformance of SDG3 related funds resulted from fund managers' good ability to select investment objectives in the healthcare and biotechnology sectors, however, there were no significant differences found in performance during crisis and non-crisis periods. In addition, according to Nofsinger and Varma (2014) funds that are especially focused on governance issues were able to show strong risk-adjusted returns in crisis.

2.3.3 Shareholder activism

The fourth strategy, shareholder activism also referred to in the literature as shareholder engagement. This SRI strategy is based on shareholders actions, where they actively attempt to influence and support the ethical development of companies and increase their ESG transparency. This is executed through dialogue with management or board of directors and voting at annual general meetings. (Lewis and Mackenzie 2000a; King and Gish 2015; Barko, Cremers and Renneboog 2021) Shareholder activism, therefore, is seeking to make a difference from the inside out. In addition, Oh, Park, and Ghauri (2013) conclude that for example, institutional fund managers have incorporated SRI practices to improve their funds' performance. Therefore, the companies they are holding in portfolios are keen to enhance their practices as well. Consequently, due to the shareholder activism of financial institutions, socially responsible investing practices will be shifting into the companies' CSR practices. Corporate social responsibility (CSR) practices are describing economic, social, and environmental responsibilities and obligations of the company. They are implemented as a part of company’s operations in order to improve stakeholder relationships. (Drucker 1984; Pava 2007) Barko et al. (2021) find that shareholder activism strategy is positively affecting the financial returns of a company when a company is willing to communicate with the activist investor.

2.4 Socially responsible investing and modern portfolio theory

Markets offer countless alternatives for investors to participate in trading. According to Kurtz (2005), the criticism of SRI is often linked to the modern portfolio theory. Socially responsible investing is limiting the investment universe by e.g., using the screening strategy, which ultimately restricts investors' alternatives to invest. One of the key arguments in the modern portfolio theory against SRI investing is limiting the investment universe, which leads to investors having a sub-optimal portfolio. Thus, focusing only on a small and carefully limited group of investment objectives, socially responsible investing might negatively affect the investor's portfolio. (Kurtz 2005; Barnett and Salomon 2006) The relationship between risk and return profile of investment is considered next.

The first theory introduced is the modern portfolio theory (MPT) by Markowitz (1952), which defines the relationship risk-return profile of the investment. Markowitz’s (1952) study on the selection of an optimal portfolio is based on two key assumptions. The first assumption is that the investor maximizes profits while the riskiness of the security increases. The second assumption is based on the fact that, the investor will only accept a higher risk if it is compensated by a higher expected return. Thus, investors are assumed to be rational and maximize returns while simultaneously minimizing the level of risk. Therefore, a rationally behaving investor will choose an investment objective that gives the highest return at a certain level of risk. Alternatively, the investor chooses the lowest possible level of risk and the level of expected return. According to Markowitz (1952), the optimally diversified portfolio holds securities whose returns would correlate as little as possible.

Modern portfolio theory describes two types of risk that an investment portfolio bears: unsystematic and systematic risk (Markowitz 1952; Sharpe 1964; Fama 1971). The unsystematic risk or so-called specific risk which can be associated with the volatility of a single individual stock in the portfolio can be offset by using diversification. Systematic risk is the risk associated with the volatility of the entire capital markets however cannot be offset by diversification. Thus, according to MPT, diversification means that investors can construct such portfolios where the specific risk of an individual stock is compensated by the other individual stock's specific risk. In efficient markets, investors are rewarded for tolerating systematic risk, however, since diversification is possible, bearing specific risk does not offer capitalization opportunities for investors. In the case where a mutual fund holds systematic risk, the risk and return trade-off is not maximized since the fund does not optimize its position in the efficient frontier. (Barnett and Salomon 2006) The efficient frontier describes the relationship between risk and return. The optimal ratio of expected return and volatility boundary is formed by a set of feasible portfolios, and they illustrate the efficient frontier where the risk-adjusted return is maximized. This denotes that the portfolios are located in a such way that the maximum return for a given level of risk is shown. (Markowitz 1952)

Previously introduced socially responsible investing strategies are playing an essential role, since utilizing them constrains the potential investment universe, which is illustrated in Figure 3.

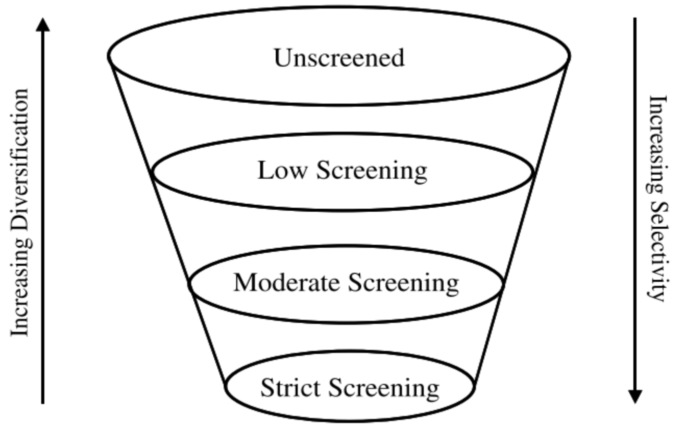


Figure 3. The effects of social screening on the universe of stock choices (Adapted from Barnett & Salomon 2006)

Since certain firms or whole industries are excluded from the potential universe of stocks, a question arises whether the investor's portfolio's specific risk increases. According to

Barnett and Salomon (2006) fund's greater screening intensity decreases the universe of potential investment objects.

In the case of mutual funds, this has been researched by Humphrey and Tan (2014); Lee, Humphrey, Benson & Ahn (2010); Bello (2005), and Kurtz (1997). Studies conclude that ESG portfolios overall did not suffer from weaker levels of diversification compared to conventional mutual funds. Lee et al. (2010) investigated whether there is a relationship between screening intensity and idiosyncratic risk. The used method to inspect the relationship was CAPM and Carhart's (1997) model residuals as a proxy for idiosyncratic risk. They conclude that analyzed SRI funds' unadjusted returns were not impacted by screening intensity. However, they found that the relationship between screening intensity and risk could be curvilinear, furthermore, screening intensity has a negative relation with systemic risk. (Lee et al. 2010)

Considering ETFs, Rodriguez and Romero (2019) studied the international diversification value of SRI ETFs compared to conventional ETFs from U.S. investors' point of view. The used methodology was based on orthogonal returns in a two-factor model. Rodriguez and Romero (2019) argue that SRI ETFs can offer more international diversification value. Mixed results were found by Verheyden, Eccles & Feiner (2016), who compared the unscreened and ESG-screened investment universes diversification ratios. They concluded that three out of four ESG-screened universes were able to challenge the key argument of ESG screening sacrifices diversification of portfolio. However, one ESG-screened universe still resulted in a slightly lower diversification value than the unscreened universe. Nevertheless, Verheyden, Eccles & Feiner (2016) found that utilizing ESG screening does improve the risk-adjusted returns.

Finally, Hoepner (2010) divided portfolio diversification into three sections: the number of selected stocks, the correlation between selected stocks, and selected stocks' average specific risk. Hoepner (2010) discusses that ESG screening intrinsically decreases the number of available stocks which reduces diversification. In addition, fewer stocks that have higher ESG value tend to be more heavily correlated versus unscreened stocks and therefore this decreases diversification. Nevertheless, Hoepner (2010) found that those stocks which obtained higher ESG scores tend to have lower specific risk thus in the end, the overall effect of ESG portfolio screening may not be negative.

2.5 Value creation through sustainability and responsibility

There has been a debate on the question, of whether it is possible to gain excess returns by utilizing SRI. Not to mention, how the value is created eventually along with investors' behavior and mindset towards socially responsible investing. This is discussed in this subsection. The second theory introduced in this context is the stakeholder theory proposed by Freeman (1984). Essentially, stakeholder theory suggests that maximizing stakeholders' interests may result in higher firm productivity and overall value (Freeman 1984; Wicks, Berman & Jones 1999). Companies' role in society is great, therefore companies should also assimilate their own societal commitment and responsibility. Theory implicates that not only shareholder value should be optimized. Companies should carefully consider also other stakeholders, especially customers and employees. Freeman (1984) and Lee et al. (2010) argue that companies that consider different aspects of responsibility tend to manage their resources more efficiently. Companies that invest in stakeholder relations have increased their return on equity and return on assets (Graves & Waddock 2000). This leads to improved financial performance and lowers governance costs.

Moreover, a model created by Albuquerque, Koskinen and Zhang (2019) considers firms that invest in ESG policies and utilize this approach as a differentiation strategy for their products will benefit from it. This strategy benefits ESG firms since it creates a more loyal customer base. Therefore, a demand which has less price-elasticity offers the possibility to the ESG firms to be more flexible with pricing, thus profit margins can be higher. Albuquerque et al. (2020) argue that if the COVID-19 crisis influences consumer demand, ESG firms' customer loyalty is the factor that will benefit ESG companies' stock performance in the crisis.

In literature, socially responsible investors have been approached as a quite homogenous group. However, there are different investor types within SRI. Derwall, Koedijk and Ter

Horst (2011) acknowledge that investor behavior in SRI markets can be approximately divided into two segments. There are a values-driven segment and a profit-seeking segment. The values-driven investor is not motivated by the profit and the investor chooses to hold assets for reasons that are unrelated to profit aspect. The profit-seeking investor specifically chooses socially responsible investments to be able to outperform conventional investments. (Derwall et al. 2011)

Similar divisions based on investor types were made by Nilsson (2009) who furthermore divided socially responsible investors into three segments. The first segment is “socially responsible and returns-driven”. This investor type is characterized by being interested in both, a high level of the socially responsible profile of investment as well as high financial return. The second segment of investors is “primarily concerned about the profit”. In this segment, the investor is not interested in the socially responsible related matters as much as the other two investor types presented and holds the least amount of SRI assets in their portfolios compared to the first and the next presented, third segment. The third segment is “primarily concerned about social responsibility” investors. They consider the profit factor of SRI less important than social responsibility and therefore they hold more socially responsible investments in their portfolios. (Nilsson 2009) In addition, Kinder (2005) also offers a taxonomy of social investors, where there are also three different segments of investor behavior presented. The value-based investors consider their moral standards when investing, whereas value-seeking investors would invest purely to increase their portfolio performance. Lastly, the value-enhancing investors would particularly use the shareholder activism strategy to improve investment value. (Kinder 2005)

Previous studies of socially responsible investors’ behavior conclude that there are similarities in the behavior of ESG investors, as several similar segments of investment behavior have been identified. Investors are investing to enhance the performance of their portfolios while putting their personal values into practice as SRI is part of their lifestyle. Some investors solely concentrate on achieving the highest returns and some investors consider returns a secondary concern. (Derwall et al. 2011; Nilsson 2009; Kinder 2005; Lewis and Mackenzie 2000b)

2.6 Socially responsible investing and crises

Mierau and Mink (2013) define financial crises as: "*Financial crises are characterized by the sudden and simultaneous materialization of risks that in tranquil times were believed to be independent.*" Next, the most recent crises and their impact on SRI instruments are presented briefly in chronological order.

The dot-com crisis or technology bubble burst in 2000, was denoted by a rapid price rise of technology sectors and Internet-related business stocks in U.S. According to Ofek and Richardson (2003), the overly optimistic behavior of investors towards funding the Internet start-ups boosted the stock prices heavily despite the fact companies were showing little to no evidence of being profitable. Eventually, the firms' valuation's real state were comprehended by the investors, which led to the panic selling of the stocks. Becchetti, Ciciretti, Dalo and Herzel (2015) discovered that in the dot-com crisis, socially responsible funds' performance measured in Jensen's alpha, were lacking compared to conventional funds. They found that this was caused by socially responsible funds relatively high exposure to high-tech stocks. However, Nofsinger and Varma (2014) analyzed SRI mutual funds' performance in crisis and non-crisis periods during 2000-2011. According to them, in the dot-com crisis, funds that focused on ESG issues were able to outperform conventional funds measured in annualized returns.

Excessive liquidity and easy monetary policy of the Federal Reserve and low interest-rates were identified as the initial sources of the 2007-2010 US financial crisis (Choi 2013). The active asset securitization, which means that assets such as mortgages can be repackaged into interest-bearing securities (Jobst 2008), combined with low-interest rates allowed excessive liquidity to flow into the subprime mortgage market. Consequently, an economic downturn followed shortly after the rise in interest rates. Afterward, the severe liquidity freeze caused by the mortgage market breakdown aggravated the poor economic state. (Choi 2013)

The performance of SRI funds in the global financial crisis was researched by Syed (2017). This study considers the UK and France area, and the timeframe is divided into two; the pre-crisis period from 2004 to 2007 and the crisis period from 2007-2009. The performance is compared to the market benchmark, which is the Dow Jones Euro Stoxx index in the case of French SRI funds and the FTSE All-Share index in the case of UK SRI funds. The study concludes that in the pre-crisis period UK SRI funds were able to outperform the benchmark, however, the results of outperformance are not statistically significant, and the same conclusion is made for French SRI funds. In addition, in the crisis period, SRI funds were able to outperform their benchmarks measured by the modified Sharpe ratio, however, Jensen alpha and Treynor ratio showed different results, overall being statistically insignificant. Therefore, according to Syed (2017), no significant performance difference were found between pre-crisis and crisis periods in contrast to the used benchmark indices. However, Syed (2017), argues that there is evidence that the SRI funds are less risky regarding their beta values compared to both market indices used.

Moreover, similar conclusions were made by Becchetti et al. (2015). They examined the U.S. market area in the global financial crisis, where SRI funds and conventional funds' performance were compared by their risk-adjusted returns. The funds collected for the study were from the Morningstar database and examined across different market segments (global, Europe, Asia, and North America). SRI funds outperformed conventional funds in all market segments, as measured by Jensen's alpha, during the global financial crisis. (Becchetti et al. 2015) In addition, Nofsinger and Varma (2014) found that during the global financial crisis, the risk-adjusted returns of SRI funds were higher than conventional funds, however, they stated that in the non-crisis period, the return of SRI funds fell slightly compared to conventional funds. Das, Chatterje, Ruf and Sunder (2018) express similar results with respect to mutual funds and ESG ratings. They compare higher ESG ratings to lower and medium ESG ratings and argue that fund's higher ESG ratings were able to mitigate fund's financial losses in the crisis period. (Das et al. 2018)

The pandemic is still a quite new phenomenon, thus more and more studies are expanding the academic literature. Given the purpose of this research, the COVID-19 pandemic and its effects on socially responsible instruments are presented next in greater detail.

At the end of December 2019, the first case of an unknown virus was reported in China in the city of Wuhan. The World Health Organization (WHO) declared the coronavirus (COVID-19) a global pandemic on the 11th of March 2020 due to the rapid spread of the virus. The number of confirmed cases has exceeded over 514 million, at the time of writing (5/2022), as well as new cases have not stopped increasing and new variants are arising (WHO 2020; WHO 2022). All-encompassing consequences of the novel health crisis, the COVID-19 pandemic, are not entirely visible yet. However, the financial markets have been reacting strongly and investors have been suffering great losses in a short period. (Zhang et al. 2020; Pavlova and de Boyrie 2022)

Shortly after the outbreak of the COVID-19 virus, Albuquerque et al. (2020) studied the stocks that contained higher ES (environmental and social) ratings in the COVID-19 market crash. They found that in the first quarter of the crisis (Q1 2020), higher ES scores contributed to stocks' higher returns as well as higher operating profit margins and lower return volatility (Albuquerque et al. 2020).

U.S, Japan, and Europe market areas are considered in the research by Omura et al. (2021). They applied asset-pricing models, the Fama-French five-factor model, and the Sharpe ratio, in order to analyze U.S. ESG ETFs and SRI indices' risk-adjusted performance before as well during the pandemic in contrast to conventional indices. Omura et al. (2021) remark that before and during the COVID-19 crisis, SRI indices were able to outperform the conventional indices. However, the analyzed ESG ETFs were not able to obtain similar results compared to the used benchmark indices in examined periods. Omura et al. (2021) contemplate whether the results of ETFs performance have been impacted by the socially responsible strategies used within the funds. Takahashi and Yamada (2021) also investigated the Japanese stock market during the COVID-19 crisis. They conclude that they did not find evidence that high scores of ESG stocks resulted in high abnormal returns over the pandemic. Nevertheless, they did associate that on average, firms that incorporate ESG into their operations were able to generate higher returns, compared to conventional firms that do not integrate ESG into their operations (Takahashi and Yamada 2021).

Pastor and Vorsatz (2020) state that in the COVID-19 crisis period defined as 10 weeks (20.2.-30.4.2020), U.S. equity mutual funds that have obtained high Morningstar's sustainability ratings were able to gain higher benchmark-adjusted returns. The benchmark used were Russell/FTSE indices. Pastor and Vorsatz (2020) argue that the environmental aspect is largely behind the outperformance result.

Pavlova and de Boyrie (2022) also investigated the performance of sustainable ETFs' risk-adjusted returns in the pre-COVID and after-COVID crash timeframe and compared them to conventional ETFs' risk-adjusted returns. Pavlova and de Boyrie's (2022) approach focused on dividing ESG ETFs into two groups based on their sustainability rating, lower rated ESG and higher rated ESG ETFs. They found that risk-adjusted returns of all ESG ETFs did not show differences with respect to the two rating groups performance during the COVID-19 crash and in addition, used ESG ETFs did not show evidence of underperforming the market. Nevertheless, before the crash, lower rated ESG ETFs outperformed the higher rated ESG ETFs. In the end, Pavlova and de Boyrie (2022) argue that the comparison between conventional ETFs and ESG ETFs did not show superiority in favor of ESG ETFs. Therefore, higher sustainability ratings did not protect the ESG ETFs' returns in the market downturn, however, ESG ETFs did not result in worse performance than the markets generally. (Pavlova and de Boyrie 2022)

Considering the Chinese market, Broadstock et al. (2021) analyzed ESG portfolios constructed from stocks that are contained in the CSI300 index. During the time of the COVID-19 crash, they found that the reaction concerning the stock prices measured using a multi-factor asset pricing model and event study, was more resilient as the stock prices resulted in smaller declines in the ESG companies during the COVID-19 market crash. They conclude that ESG stocks performance is positively correlated with short-term cumulative returns during the COVID-19 crash. Thus ESG-related performance can be a signal of possible risk mitigation in times of crisis. (Broadstock et al. 2021)

Singh (2020) examined the spillover effect across different investment styles that are considered to be safer during the time of crisis, such as the COVID-19 pandemic. Analyzed investment styles were defensive-cyclical sectors, EAFE (stocks from Europe, Australasia, and the Far East) and ESG based portfolios. In economic downturns, according to Singh (2020) investors became more observant of corporate fundamentals. The outperformance of the ESG portfolio was the result of an increase in the capital flows away from the other investment styles that are seen as safe alternatives, to the ESG portfolios in crisis periods. Investors find security in the ESG investing approach because it focuses on the long-term sustainability of businesses. (Singh 2020)

There are different results achieved in the performance of socially responsible investment instruments during the COVID-19 crisis. The exchange-traded fund's financial performance and ESG rating relationship during the COVID-19 crisis were examined by Folger-Laronde, Pashang, Feor, and ElAlfy (2020). Research is comparing sustainable ETFs' weekly logarithmic returns before and during the pandemic's market crash. Used methodologies were multivariate linear regression and ANOVA. They conclude that the ETFs that have higher Eco-fund ratings during the COVID-19 market crash, did not benefit from the sustainability aspect since the returns were negatively related to the Eco-Fund rating (Folger-Laronde et al. 2020).

Similar conclusions to the abovementioned research were made by Demers, Hendrikse, Joos & Lev (2020) considering mutual funds. Demers et al. (2020) analyzed ESG scores of U.S. equity share prices in the first COVID-19 crisis quarter (Q1 2020) and second quarter of 2020 (Q2 2020) using regression analyses. They argue that ESG scores do not have explanatory power on companies' returns during the COVID-19 market crash. They find that during the crisis period (Q1 2020) ESG variable is insignificant. Researchers also found that during the recovery period (Q2 2020), there is a negative relationship between ESG and returns. Therefore, they concluded that ESG does not offer enhancement or safe havens for returns during crisis periods. (Demers et al. 2020)

2.7 Exchange-Traded Funds

Exchange-traded funds are funds that follows a specific asset class, sector, or index, such as the S&P 500 or the EURO STOXX50. If a specific ETF tracks the S&P500 index, it consist of the basket of assets that the underlying index holds in one investment vehicle, which can be traded like a stock throughout the trading day on the stock exchange. An ETF eventually combines the benefits and flexibility of a mutual fund and a share in one investment instrument. By investing in ETFs, the advantages of diversification can be utilized without having to directly invest in several stocks individually. (justETF 2022; Fidelity 2022)

ETFs in the markets can be active or passive managed ETFs. The basic difference between them is that passive ETFs ETF follows solely a predetermined index, while active ETFs aim to beat the market and are actively managed by portfolio managers. An active ETF is typically assigned to follow a certain benchmark index, while simultaneously there are added for example ETF manager's top picks or ETF is assigned to partly mirror an existing mutual fund. Overall active ETF's target is to offer above average returns to the investor. ETFs also offer lower costs and tax efficiency compared to mutual funds (Pavlova and de Boyrie 2022; Zhang et al. 2020; Rompotis 2013).

ETFs were originally launched in the U.S. market area in 1993, where the first ETF, SPDR 500, tracks the S&P 500 index. The ETF industry consists of different asset categories: equity ETFs, fixed-income ETFs, commodities and real estate ETFs and currency and multi-assets ETFs and assets of ETFs by geographical region. (justETF 2022; Meziani 2016) The ESG ETFs popularity among private investors has grown greatly and worldwide expansion is shown in Figure 4. North America has the largest assets under management by ETFs focusing on the ESG approach. (Statista 2020b)

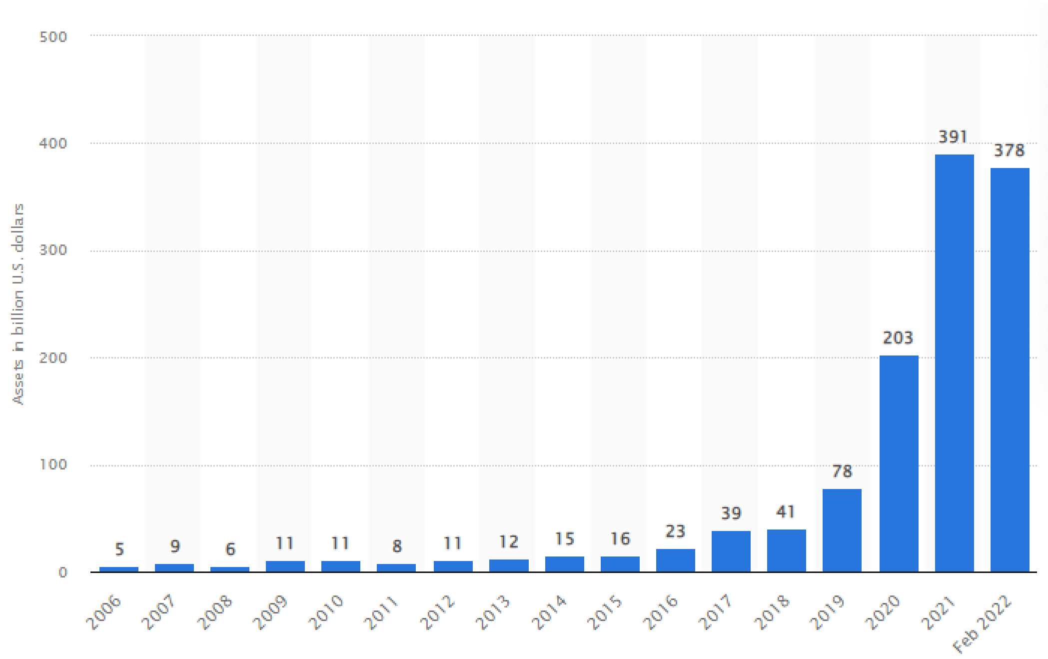


Figure 4. Global ESG ETF assets from 2006-2022. (Statista 2022b)

2.8 Machine learning

One of the earliest definitions of machine learning (ML) was introduced by Arthur Samuel (1959), where ML is defined as ”*the field of study that gives computers the ability to learn without being explicitly programmed*.” Mahesh (2020) explains that to solve certain data problems, ML relies on different algorithms. Machine learning is a subset of artificial intelligence, and it is used in various applications in order to identify complex patterns. By default, there are no ready answers to the question of which algorithm is generally the best to solve a problem. It all depends on the problem one wishes to solve and examine. In addition, data availability and a number of variables have an influence on the problem setting. (Mahesh 2020; Henrique, Sobreiro and Kimura 2019; Alpaydin 2014) Machine learning can be divided into three basic paradigms: supervised learning, unsupervised learning and reinforcement learning (Jung 2022; Sutton & Barto 2018), as illustrated in Figure 5 below, where the white color highlights what aspect of the machine learning part is used for this thesis. Next, they are explained, and after that unsupervised learning and furthermore, clustering and their approaches are presented.

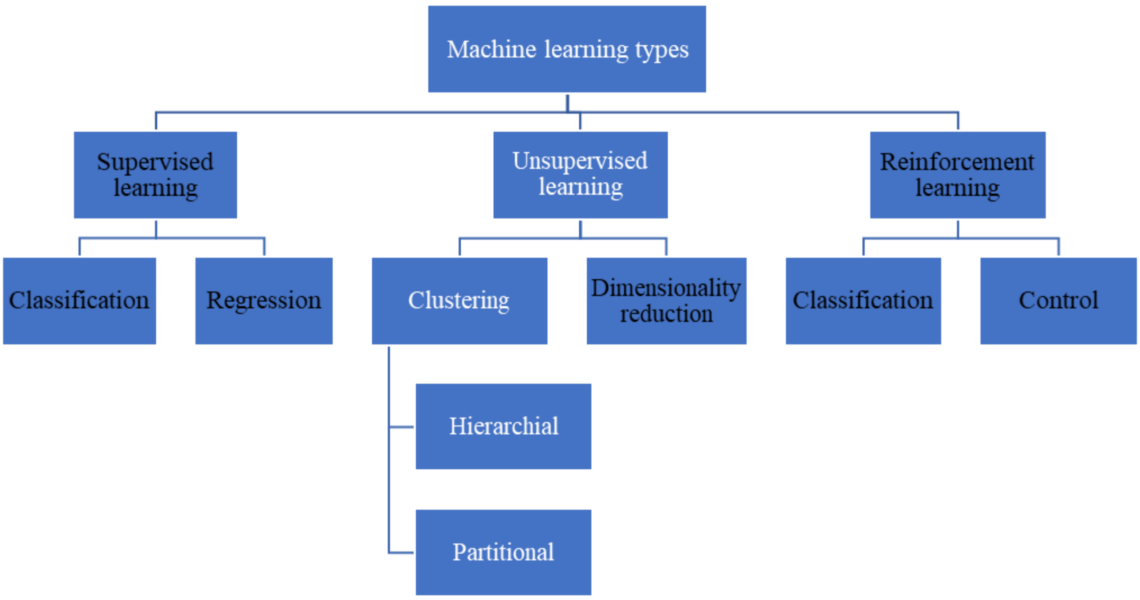


Figure 5. Machine learning roadmap for selected clustering methods (Jung 2022; Mahesh 2020; Sutton & Barto 2018)

Supervised learning requires known data input and output, to predict future outputs. Supervised learning's goal is to find and learn from a relationship between input X and output Y and is used for example for classification and regression problems. The unsupervised learning technique finds hidden patterns or structures in the input data. Unsupervised learning problems usually focus on feature reduction and clustering (Mahesh 2020). Clustering analysis, presents algorithms like k-means clustering, hierarchical clustering, self-organizing maps (SOM), hidden Markov models and Gaussian mixture. (Dixon, Halperin and Bilokon 2020) Reinforcement learning is based on action and reward, where the software agent instead of being told what the correct action is, will discover which actions generate the greatest reward (Mahesh 2020; Sutton & Barto 2018). Reinforcement learning problems are often related to goal-orientated learning, control and decision-making and classification (Sutton & Barto 2018).

2.9 Clustering

A specific approach in unsupervised learning is determined as clustering, which groups objects together without prior knowledge of datasets. According to Rokach and Maimon (2005) clustering process ultimately divides data patterns according to their similarity, thus similar patterns are identified and clustered together. The purpose of clustering algorithms is to find and identify high-quality groups of objects that are similar and identify hidden patterns from the dataset. Clustering problems consider dividing the given dataset into clusters or groups where the determined points within the specific cluster are like each other (cohesion) and dissimilar to points in different clusters (separation). Similarities and dissimilarities of clusters are analyzed utilizing distance measures. (Rodriguez et al. 2019; Buszko, Orzeszko and Stawarz 2021; Han, Pei & Kamber 2012; Bandyopadhyay and Saha 2013)

Clustering techniques have developed over time and are constantly evolving (Murty & Devi, 2015). According to Han, Pei and Kamber (2012) clustering methods can be divided into four major categories such as: partitioning methods, density-based methods, hierarchical methods, and grid-based methods. Madhulatha (2012) also adds model-based clustering into the abovementioned clustering method list by Han, Pei and Kamber (2012).

2.9.1 Hierarchical and partitional clustering methods

The two types of clustering methods discussed in this thesis are hierarchical and partitional clustering methods. The methods consider different strategies when focusing on minimizing the inter-group similarity as well as maximizing the intra-group similarity. Considered one of the most popular clustering methods, partitional clustering algorithms minimize the distance between cluster centers, called centroids, of each cluster and data points included in that specific cluster. The defined centroid does not have to be a data point, depending on the used algorithm. However, the number of clusters needs to be pre-defined. Self-organizing map (SOM), k-means, k-medoids, and Fuzzy C-means are examples of partitional clustering algorithms. (Javed, Lee and Rizzo 2020; Irani, Phatak and Pise 2016)

Hierarchical clustering is widely used due to its ability of its visualization power as well as its ability to execute clustering without the need of providing the number of clusters by default (Keogh and Lin 2005). Hierarchical clustering algorithms can be subdivided into two groups called divisive and agglomerative clustering. Initially, the agglomerative (bottom-up) hierarchical clustering algorithm presumes that each object forms its own cluster, whereas divisive (top-down) hierarchical clustering presumes that at first, all objects belong jointly to one single cluster. Thus, in the case of a divisive clustering algorithm, it eventually splits bigger clusters into smaller ones, whereas agglomerative clustering merges smaller clusters to create bigger clusters (Murty and Devi 2015). Irani, Phatak and Pise (2016) mention examples of hierarchical clustering algorithms such as BIRCH (Balance Iterative Reducing and Clustering using Hierarchies) and CURE (Clustering Using Representative). (Rodriguez et al. 2019; Koutroumbas and Theodoridis 2009; Javed, Lee and Rizzo 2020)

2.9.2 Distance measure & performance evaluation criterion

In clustering, there are considered two significant design criteria: similarity/distance measure and clustering method. The clustering methods were described above at a general level in previous chapter 2.9.1. Similarities and differences between data points in clusters are analyzed using a distance measure to determine the similarity between two objects. Therefore, a similarity measure defines the distance between several data points, called members, within a cluster. Hence, clustering performance is directly impacted by the distance measure choice since many alternative similarity measures exist. (Javed, Lee and Rizzo 2020; Bandyopadhyay and Saha 2013; Irani, Phatak and Pise 2016)

There are multiple different distance measures based on the strategy of the clustering method applied. Considering time series clustering, Euclidean distance as well as, shape-based distance, for example Dynamic Time Warping (DTW), are commonly used distance measures for the partitional clustering method (Izakian, Pedrycz and Jamal, 2015). In the case of the hierarchical clustering method, there are different distance measure categories based on the calculation of the distance between data points. They are complete-linkage clustering, single-linkage clustering, and average-linkage clustering. Other popular distance measures used for agglomerative clustering are the Manhattan distance function, Cosine distance as well as Jaccard distance. (Liao 2005; Madhulatha 2012; Irani, Phatak and Pise 2016; Saxena et al. 2017)

The clustering results should be evaluated against some criterion. Most of the clustering algorithms are based on specific criteria that determine the clusters in which dataset can be partitioned. The objective of clustering evaluation is to find which partition method fits the underlying data the best and give an indication of the quality of the partition. (Halkidi, Batistakis and Vazirgiannis 2001) The clustering algorithm results, formed clusters, are validated according to their compactness. The evaluation measure criteria measures the intra-cluster homogeneity and inter-cluster separability or their combination. According to Milligan and Cooper (1985) performance evaluation approaches like cluster validity indices are used to help select the suitable and appropriate number of clusters. Sum of Squared Error (SSE), Silhouette index/silhouette coefficient, Calinski-Harabasz and C-index are examples of evaluation criteria. (Aghabozorgi, Shirkhorshidi and Teh 2015; Liao 2005; Gagolewski, Bartoszuk and Cena 2021; Saxena et al. 2017; Halkidi, Batistakis and Vazirgiannis 2001) In this thesis the silhouette coefficient is considered for the clustering results evaluation and criteria of maximizing the silhouette coefficient is applied, which is discussed in the methodology chapter.

2.9.3 Time series clustering & taxonomy

Javed, Lee and Rizzo (2020) define a time series as a sequence of values of variables ordered by time. Time series data is one of the most popular data types used in clustering problems, especially in finance-related issues like anomaly detection, portfolio optimization and examining similar behaving stocks, where for example stock prices and exchange rates data are used as they are representations of time series. Thus, time series clustering is a special type of clustering and is commonly used to examine and discover patterns on datasets that are stored in time series type. (Aghabozorgi, Shirkhorshidi and Wah 2015; Javed, Lee and Rizzo 2020) The taxonomy of time series is illustrated in Figure 6 below.

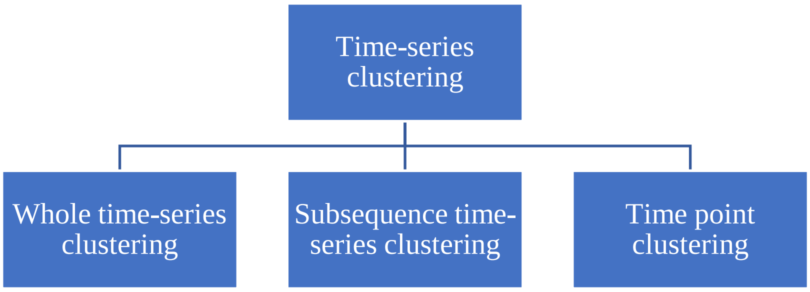


Figure 6. Taxonomy of time series clustering (Adapted from Zolhavarieh, Aghabozorgi and Teh 2014)

Time series clustering can generally be classified into three types: whole time-series clustering, subsequence time-series clustering and time point clustering. Whole time series clustering is considered when a set of individual time-series are clustered with respect to their similarity, thus multiple time series are considered in this clustering process. Subsequence time series clustering means that a set of subsequences of a single time series are extracted utilizing a sliding window approach. Then, clustering is executed on the extracted proportions of an individual long time series. Thus, subsequence time series clustering considers one time series at a time, when different timepoints within the time series are clustered. Lastly, time point clustering also focuses on one time series at a time, where time points are clustered regarding their combination of temporal proximity as well as the similarity of the corresponding values. (Zolhavarieh, Aghabozorgi and Teh 2014; Aghabozorgi and Teh 2014; Aghabozorgi, Shirkhorshidi and Teh 2015; Durán-Rosal et al. 2017; Keogh and Lin 2005)

This thesis aims to find fluctuations concerning the price movement in the time series of chosen ETFs in different pandemic phases and group them with similar trending time series. Therefore, for this thesis' clustering problem, the whole time-series clustering is identified. The Euclidean distance measure is a very commonly used distance measure in time series clustering applications. Wang et al. (2013) compared multiple distance measures' performance and they conclude the Euclidean distance has performed with reasonably high accuracy. In addition, the literature shows it has been applied efficiently in time series clustering applications and the most popular similarity measures are DTW and Euclidean distance (Aghabozorgi, Shirkhorshidi and Teh 2015).

2.9.4 Time series clustering in the literature

Identifying trends in the stock market using time-series clustering and segmentation were examined by Durán-Rosal et al. (2017). Their objective was to identify market behavior in different socioeconomic cycles from 1992 to 2014 in the European stock market using multiple indexes' closing prices as the data. Durán-Rosal et al. (2017) combined a time-series segmentation as well as a characterization method in a hybrid genetic algorithm. Then, clustering was used to group common trends and patterns from the dataset. The clustering method used was k-means, where the determined number of clusters was five. The Euclidean distance was applied as a distance measure. The quality of the clustering process was measured via an internal criterion, the COP index. Durán-Rosal et al. (2017) conclude that obtained clusters were able to illustrate patterns that acted as early signals for a trend change. For example, the financial crisis of 2008 and the dot.com crisis, which impacted the stock markets were identified. (Durán-Rosal et al. 2017)

The relationship between the COVID-19 pandemic and the stability of the Warsaw Stock Exchange sector indices was researched by Buszko, Orzeszko and Stawarz (2021). They investigated how different industries represented by stock market indices were affected by the COVID-19 stock market shock. The similarity of stock market movements with respect to the indices' behaviour was investigated using clustering in different time frames. The compared time frames during the pandemic consisted of the first six months of the pandemic 7.1.-6.7.2020. Due to the rapid changes in the markets during March 2020, the initial pandemic time frame was divided into a short term (2 weeks and 1 month) and a medium term (3 and 6 months) time frame to examine the differences in sector indices. Clustering methods used were k-means and Ward's method and silhouette coefficient and SSE were utilized as the clustering performance validation measures. Buszko, Orzeszko and Stawarz (2021) conclude that in short term they were able to observe five clusters and in the medium term, four clusters were identified. The most responsive, therefore viewed as the most unstable sector considering the volatility in this study was the pharmaceutical sector. (Buszko, Orzeszko and Stawarz 2021)

Nanda, Mahanty and Tiwari (2010) used a clustering approach for stock market data (Bombay Stock Exchange) in portfolio building beginning of the financial crisis in 2007. Stocks were clustered based on their returns and valuation ratios such as P/E (price to earnings ratio) and P/BV (price to book value). The best performing stocks were selected from obtained clusters and were used to construct an optimal portfolio. The used clustering methods were k-means, SOM and Fuzzy C-means using the Euclidean distance as the distance measure. Nanda, Mahanty and Tiwari (2010) determined the tested range of k from 2 to 12. They find that k-means was able to build the most compact clusters and the optimal number of clusters was five or six with respect to the given data. This is measured by using Intraclass inertia F(K), which considers the relationship between cluster compactness and a fixed number of clusters. Other validity indices used were for example Silhouette index, Dunn's index and Calinski-Harabasz. (Nanda, Mahanty and Tiwari 2010)

Dragut (2012) focuses on detecting trends in the stock market, where the open and closing prices of the S&P 500 were utilised during the financial crisis of 2007-2009 in the U.S. The raw data clustering algorithm, Subsequence Time-Series (STS) clustering, was deployed and clustering of data streams where used in order to extract trends and identify market situation "warp" points. The used similarity measure was the cross-correlation similarity measure. Dragut (2012) concludes that the cluster's intra-class similarity became smaller and inter-class similarity grew larger when the market crisis deepened. When the crisis times were eased, the intra-class similarity increased and simultaneously the inter-class similarity decreased. (Dragut 2012)

In the literature introduced in this chapter, multiple applications of time series clustering analysis have been used in finance to extract information from the dataset. Portfolio formation (Nanda, Mahanty and Tiwari 2010; Gubu, Rosadi and Abdurakhman 2019), stock price prediction (Nakagawa, Imamura and Yoshida 2019), and stock behaviour in crisis analysis using clustering methods, mostly k-means and the Euclidean distance measure in previous studies. (e.g. Durán-Rosal et al. 2017; Buszko, Orzeszko and Stawarz 2021; James and Menzies 2021) In the case of socially responsible investing, Ivanisevic Hernaus (2019) clustered SRI mutual funds according to their SRI strategies using k-means and Ward's method. Based on the previous applications of the k-means method in parallel with Ward's method is applied in this thesis as well.

3 Data

The used data and its collection process is introduced in this section. For further analysis, the data pre-processing, and the split into different timeframes during COVID-19 are also explained. The used ETFs in this study are considering conventional ETFs and socially responsible investing style-related ETFs, focusing more closely on ESG ETFs.

3.1 Data collection process

This research focuses on the investigation of ESG ETFs' and conventional ETFs' trends during the COVID-19 crisis. The first phase of the data collection process is the identification of suitable ETFs for the purpose of the study. Information on ETFs is collected from Morningstar (2022), ETFdb (2022) as well as ETF.com (ETF.com 2022). In the case of ESG ETFs, ETFdb's ESG score screener was used to assist in filtering the ESG ETFs from the database to identify lower and higher ESG scores containing ETFs for analysis. In addition, Morningstar's sustainability rating (Morningstar 2021) was considered in this process, which is explained in sub-chapter 3.2.

Equity is chosen as the asset class since the thesis focuses on stock markets. The geographical area of focus is the U.S. market area, and ETFs' currency is dollars. Finally, ETFs were filtered based on the assets under management (AUM) where the smallest ETFs were excluded using 100 M$ AUM as the threshold. However, for ESG ETFs, a smaller threshold was accepted to include more ESG ETFs in the analysis, and 7 out of 31 ESG ETFs have AUM under 100 M$. Lastly, actively managed ETFs were excluded.

3.2 Pre-processing and description of the research data

After the collection process, there is a total of 68 selected ETFs, where 31 are ESG ETFs and 37 are conventional ETFs. They are presented in Appendix 1 and Appendix 2, where also ETFs’ AUM and inception dates are shown. The largest conventional ETF is Vanguard Dividend Appreciation ETF (63,38 Bn$) and the largest ESG ETF is iShares ESG Aware MSCI USA ETF (22,95 Bn$).

After obtaining the suitable ETFs, their adjusted closing prices are collected from Refinitiv Eikon Datastream (2022) and they are used in this research. Corporate actions, such as dividends and stock splits, are corrected before the next opening day, which the adjusted closing prices are reflecting (D’Urso, De Giovanni and Massari 2021).

Following the approaches by Pavlova and de Boyrie (2022), Pisani and Russo (2021) and Pástor and Vorsatz (2020), in this study every ESG ETF’s sustainability measure is adopted from Morningstar. Morningstar's ESG rating is derived from Morningstar's own ESG valuation methodology as well as Sustainalytics’ ESG risk rating. Morningstar expresses the ESG rating as a category of one to five “globes”. A higher number of globes reflects a lower ESG risk relative to an ETF’s peer group. (Morningstar 2021) The ESG ETFs’ sustainability ratings are provided in Table 1. As shown in Table 1, most of the selected ESG ETFs have obtained high sustainability ratings, however lower sustainability ratings obtained ESG ETFs were included to reflect the diversity of existing ETFs in the markets.

Table 1. The selected ESG ETFs’ sustainability ratings categories

Number of ESG ETF	Morningstar Sustainability Rating
8	
11	
10	
1	
1	

The dataset is composed of the daily adjusted closing prices of every obtained ETF for a timeframe of half a year. Thus, each ETF has 126 observations as this represents half a year of trading days. The dataset did not contain any missing values. The raw data is normalized to the interval [0,1] using the z-normalization (Equation 1), which is one of the most used data normalization methods for time series (Łuczak 2018). Normalization allows similar patterns to be identified in time series with different fluctuation rates (Keogh and Lin 2005). According to Łuczak (2018) time series z-normalization is a transformation that results in a new time series of the same length, where the arithmetic mean of the time series is zero and the standard deviation is one, which is shown in Equation 1 below.

$$Z(x_i) = \frac{x_i - \mu(x_i)}{\sigma(x_i)}, \tag{1}$$

where $\mu(x_i)$ states the mean of the univariate time series of x_i and the standard deviation of x_i is $\sigma(x_i)$. The Z-score normalized research data is shown in Figure 7 below, where the ESG and conventional ETFs are presented separately.

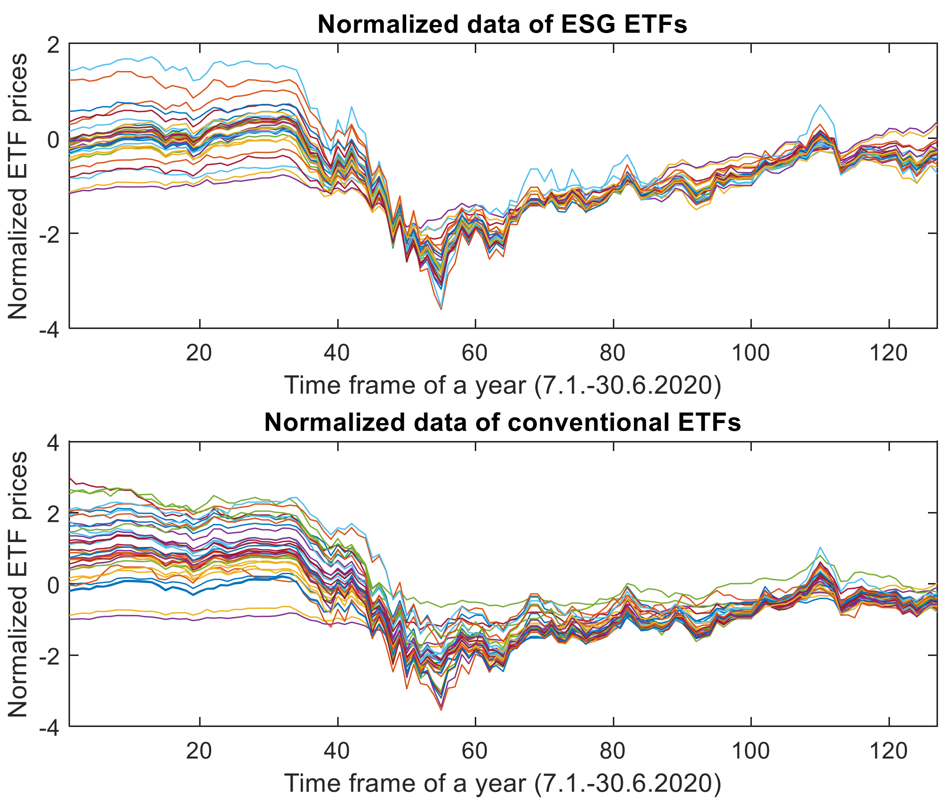


Figure 7. Normalized research data

3.3 Data split into different COVID-19 phases

The collected data timeframe is from 7.1.2020 to 30.6.2020 and it consists of daily adjusted closing prices of the ETFs contained in this study. The selected timeframe will cover the early beginning of COVID-19, the rapid collapse phase, and the recovery phase - therefore, the research splits the timeframe into three different sections to observe ETFs' price movement trends. Figure 8 below illustrates the examined phases.

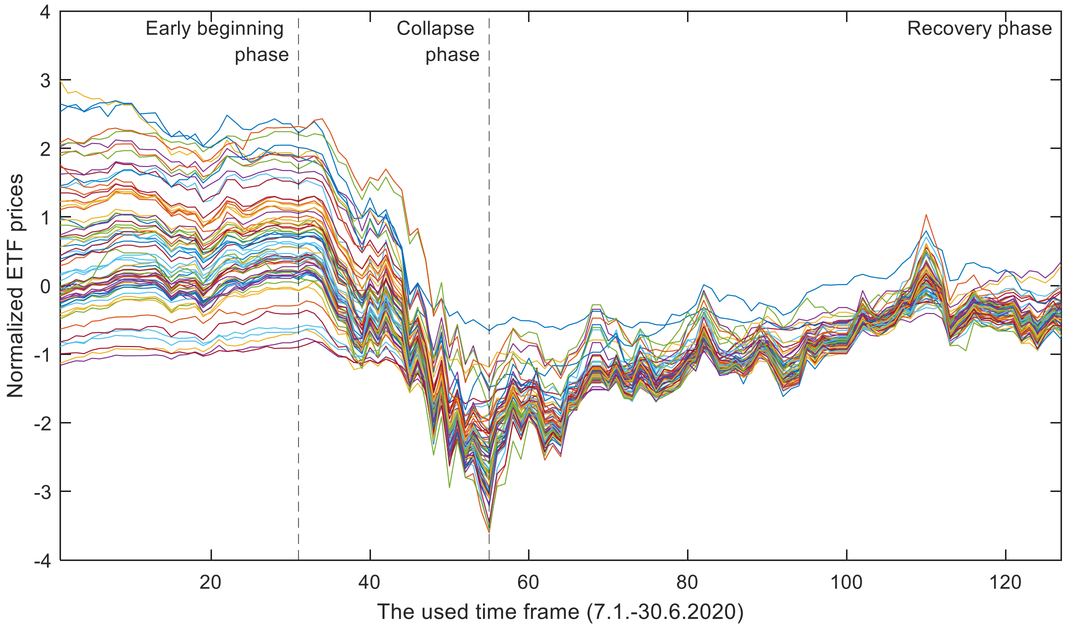


Figure 8. Illustration of the data and examined phases

The phases are described as 1) the early beginning phase: from the 7th of January to the 18th of February 2020, and 2) the collapse phase: from the 19th of February to the 23rd of March 2020. Lastly, 3) the recovery phase: from the 24th of March to the 30th of June of 2020. The two first phases used in this thesis follow the timeframe defined by Pástor and Vorsatz (2020). Data processing and calculations are executed in MATLAB and Microsoft Excel.

4 Methodology

The objective of this research is to identify price movement trends in financial time series data in different phases of the COVID-19 crisis. The methodology of this study is based on unsupervised learning; clustering analysis is performed in order to observe possible similarities with respect to price movements of different ETFs. The methodology of this thesis is constructed in the following way: the two methods of clustering, k-means and Ward's method are presented as well as the used distance measure followed by a clustering validation measure. Euclidean distance is used as the distance measure and the optimal number of clusters is selected using the silhouette coefficient maximization criteria. Finally, different financial characteristics are presented to detect differences between the resulting clusters in the examined phases.

4.1 Partitional clustering approach

In this thesis a partitional clustering approach, k-means clustering (MacQueen 1967) is selected. It is a popular approach due to its simplicity and was frequently used for clustering stocks and ETFs (see e.g. Nanda, Mahanty and Tiwari 2010; Buszko, Orzeszko and Stawarz 2021) The k-means clustering algorithm is a hard clustering algorithm, meaning the observation can only belong to one cluster (Liao 2005).

The k-means method is based on minimizing the objective function, which purpose is to minimize the distance of each data point from the cluster centroid to which the data point belongs (Halkidi, Batistakis and Vazirgiannis 2001). According to Dougherty (2013), King (2015) and Halkidi, Batistakis and Vazirgiannis (2001) first, the algorithm starts with a number of cluster centers, which are pre-determined as the initial partition of the data points. Second, the algorithm assigns observations to their closest centroid using the distance function. Thirdly, the centroids are updated based on the algorithm's new calculations based on the partition obtained in the previous phase. This partitioning process and cluster centers update is repeated until the cluster centers do not move. (Dougherty 2013; King 2015; Paparrizos and Gravano 2017)

The k-means method determines clusters in such a way that the variability among the observations is minimal with respect to the cluster centroids. This is calculated by using the within-cluster sum of squared distances (WCSS) as Equation 2 shows.

$$\text{WCSS} = \sum_{k=1}^{K} \sum_{i:x_i \in C_k} \|x_i - \mu_k\|^2 \qquad (2)$$

K describes the number of clusters whereas clusters are denoted $C_k (k = 1,2 \dots, K)$. The centroids are presented by μ_k. (Buszko, Orzeszko and Stawarz 2021)

4.2 Hierarchical clustering approach

The second chosen clustering method is Ward's method, which is an agglomerative hierarchical clustering approach by Ward (1963). The hierarchical clustering method can be illustrated as a tree of clusters, a dendrogram, which describes the hierarchical structure of clusters. The observations are initially considered as single clusters, which are merged in succession. Thus, larger, and larger clusters are formed by merging the observations one by one. The method's advantage is that the number of clusters does not have to be predefined at first. (Großwendt, Röglin and Schmidt 2019; Gagolewski, Bartoszuk and Cena 2021; Liao 2005; Dougherty 2013)

The Ward's minimum variance algorithm is based on the sum-of-squares variance, like the k-means algorithm presented in the previous subsection, thus, they have the same objective function (Dougherty 2013). Ward method starts by assigning all data points as their own cluster. The sum-of-squares variance is calculated for all clusters in every step of the clustering process. The clusters, which have resulted in the smallest increase of the sum-of-squares variance are merged. Finally, the algorithm stops when there is only one cluster left. (Liao 2005; Gubu, Rosadi and Abdurakhman 2019; Buszko, Orzeszko and Stawarz 2021; Großwendt, Röglin and Schmidt 2019)

Since the number of clusters does not have to be predetermined by default as the algorithm suggests the cluster number which has resulted in the smallest WSCC as the optimal partition, Ward's method is used beside the k-means algorithm. According to Murtagh and

Legendre (2014), the clustering results identified by Ward's method can be used to improve the k-means as the starting approximation of partitioning. Therefore, Ward's methods suggested partition is considered in this thesis when defining the initial partitions for the k-means algorithm.

4.3 Distance measure

The distance measure, also referred to as a similarity measure is defined by Irani, Pise and Phatak (2016) as the distance between different data points. The distance measure in this thesis is the Euclidean distance. The Euclidean distance illustrates the ordinary distance between two observations and is widely used in clustering problems (Irani, Phatak and Pise, 2016). The Euclidean distance *d(T1,T2)*, between two time-series $T1$=($T1_1, T1_2, \dots, T1_n$) and $T2$=($T2_1, T2_2, \dots, T2_n$) is expressed in Equation 3. (Javed, Lee and Rizzo 2020)

$$d(T1, T2) = \sqrt{\sum_{i}^{n} T1_i - T2_i{}^2} \tag{3}$$

The Euclidean distance measures the dissimilarities/similarities of time series objectives based on a fixed order. Thus, Euclidean distance cannot take into account time shifts that might occur in time series data. This means that the time series data points might form similarities considering their shapes in different time points which are not measured in succession. (Izakian, Pedrycz and Jamal 2015) Different distance measures consider the time shifts, for example, the DTW and the k-Shape. However, existing time shifts are not taken into consideration due to the chosen distance measure. In this study, the relationship between the progression of the crisis and ETFs is examined from the perspective of three selected time periods.

4.3.1 Cluster validation

The evaluation of the quality of the partitions produced for the underlying data is one of the most important issues in cluster analysis (Halkidi, Batistakis and Vazirgiannis 2001). After the clustering results are achieved using the chosen clustering method, validation takes place. These measures are used to assess the quality of partitioning the data based on the optimal number of clusters proposed (Nanda, Mahanty and Tiwari 2010). The validation measure in this study is the silhouette coefficient.

The aim of clustering is to group similar data points together to form homogenous groups. The data points within the cluster should be as similar as possible compared to the data points in the other clusters. (Liao 2005) The silhouette coefficient (SC) by Rousseeuw (1987) can be calculated from silhouette values, which determine the similarity of each data point to its own cluster compared to data points in other clusters. To calculate the SC, at first the silhouette values are calculated (Equation 4), where the mean intra-cluster distance a and the mean nearest cluster distance b for each point x_i are determined for number of clusters obtained by *K*.

$$s_{i,k} = \frac{b_i - a_i}{\max{(a_i, b_i)}} \qquad (4)$$

The s_i takes values from a range of [-1,1], where a higher number indicates that point x_i fits better in its specific cluster than the closest other cluster. If the obtained value is 0, it indicates that the point x_i is close to the decision boundary between two neighboring clusters, while a negative value can be interpreted as the point being assigned to a wrong cluster. (Buszko, Orzeszko and Stawarz, 2021)

Then using Equation 5, the SC are calculated. Since the individual points are measured using s_i, evaluating of all obtained clusters *K*, the average silhouette is calculated using Equation (5):

$$\bar{s}_K = \frac{1}{N}\sum_{i=1}^{N} s_i, \qquad (5)$$

where N is indicating the number of data points. The value of $\bar{s}_K$ is the silhouette coefficient, which represents the overall clustering quality. (Buszko, Orzeszko and Stawarz, 2021) Kaufman and Rousseew (2005) have created specific thresholds to interpret the obtained silhouette coefficients, which are used in this study to evaluate the clustering results:

SC ≤ 0.25 indicates that no substantial structure have been found,

0.26 ≤ SC ≤ 0.50 indicates that the structure is weak and could be artificial,

0.51 ≤ SC ≤ 0.70 indicates that a reasonable structure has been found,

0.71 ≤ SC ≤ 1 indicates that there has been found a strong structure.

The partition which results the highest silhouette coefficient is considered to be the optimal solution (Rodriguez et al. 2019).

4.3.2 Financial characteristics

The differences between obtained clusters in the three phases of the COVID-19 crisis are examined using annualized volatility, average, minimum and maximum daily and annualized returns. First, ETFs' daily returns are calculated using natural logarithms. According to Wells (2004) natural logarithms are more accurate estimations compared to simple percentage change, since the simple percentage change is more sensitive to the arithmetic anomaly bias. The arithmetic anomaly bias considers that the average of stock return changes must be zero, therefore in the case stock price has increased on day one and then the next day the price decreases to the same level as on day one, the change should be zero. The returns are calculated using the following Equation (6):

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right), \tag{6}$$

where r_t indicates the daily return, P_t is the day's adjusted closing price and P_{t-1} considers the previous day's ETFs adjusted closing price. The volatility is measured based on the standard deviation of the daily natural logarithmic returns of an ETF for the examined phases. The volatility, denoted by σ is calculated using the following Equation (7):

$$\sigma_{annualized} = \sigma_{daily}\sqrt{252}, \tag{7}$$

where $\sigma_{annually}$ is determined by taking the square root of number of trading days (252) and multiplying it with the daily volatility. Lastly, the annualized return is calculated using Equation (8):

$$r_{annualized} = ((1 + return)^{252}) - 1, \tag{8}$$

where similarly to Equation 7, trading days are used to annualize the ETFs daily logarithmic returns, which makes the cluster characteristics comparison more generalizable.

5 Results

In this chapter, the results of clustering ESG, conventional and all ETFs are presented for different phases of the COVID-19 pandemic. The following subsections are divided into the three examined phases, where clustering results, their validation and the obtained clusters' financial characteristics are discussed and lastly the results are analyzed.

5.1 Examination of different time frames

Different time frames for the whole dataset are illustrated in Figure 9 below. The phases are described as 1) the early beginning phase: from the 7th of January to the 18th of February 2020, and 2) the collapse phase: from the 19th of February to the 23rd of March 2020. Lastly, 3) the recovery phase: from the 24th of March to the 30th of June of 2020.

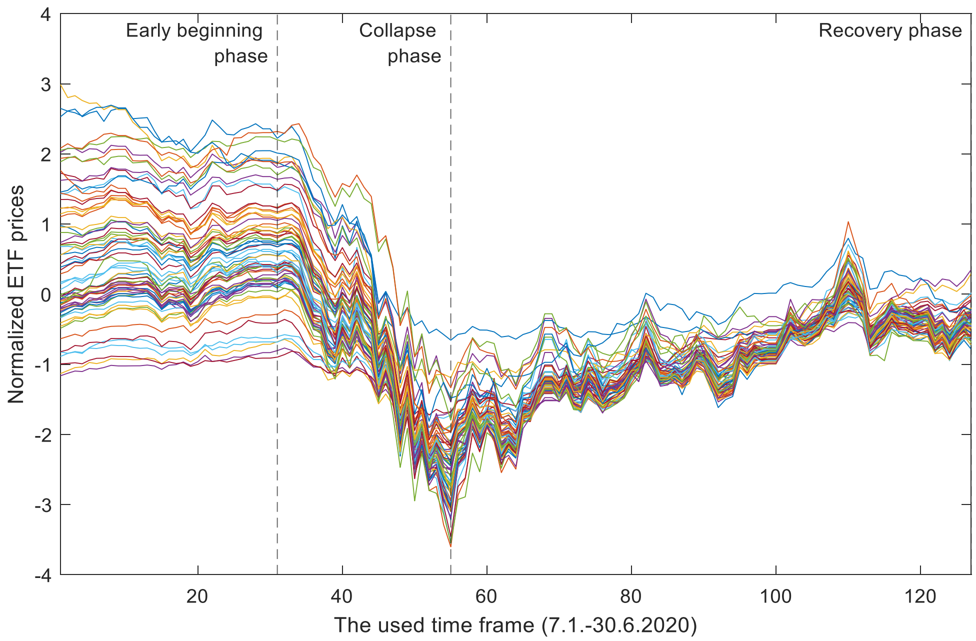


Figure 9. Illustration of the examined phases

Two clustering methods are applied to the data to analyze different ETFs' price movements in multiple phases of the COVID-19 crisis. The used methods included a partitional method, k-means, and a hierarchical method, Ward's method. The Euclidean distance is applied for both methods as the distance measure. To select the appropriate number of clusters, the criterion of maximization of the mean silhouette coefficient was used, since it can be used as a measure of clustering quality as performed for example in the research of Buszko, Orzeszko and Stawarz (2021) and D'Urso, De Giovanni and Massari (2021).

5.2 Early beginning phase

The first examined phase of this study is the early beginning phase (7.1.-18.2.2020). The conventional and ESG ETFs' normalized prices are presented below (Figure 10). It can be observed from Figure 10, that the price movements of both ETFs were fairly small and there are no drastic changes within the data. However, a slight difference can be seen from the ESG ETFs where most of the ETFs appear to be concentrated more around zero, which

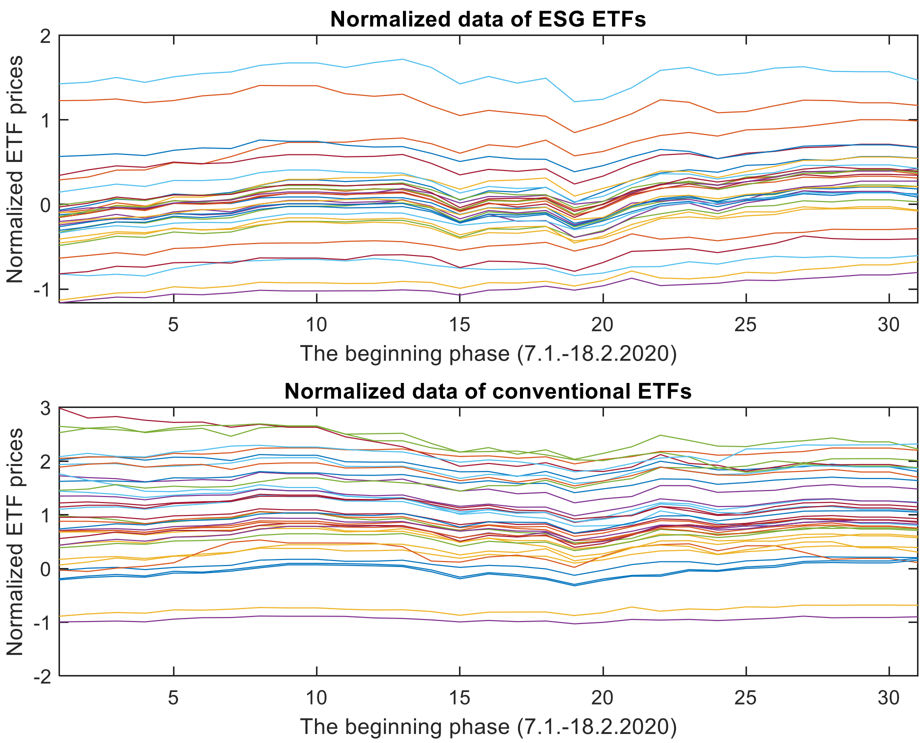


Figure 10. Illustration of the early beginning phase

indicates that the ESG ETFs' normalized prices are closer to their mean. Conventional ETFs are more evenly spread upwards the scale during the first examined phase, except for two ETFs at around -1 indicating that their daily price movements are below their mean with respect to conventional ETF population.

5.2.1 Clustering results of the early beginning phase

The first clustering method applied was Ward's method followed by k-means. The calculated number of clusters ranged from 2 to 15, however, clusters remained stable and therefore the range of 2 to 8 is presented. In Figure 11, the average silhouette coefficient suggested that the optimal number of *k* (number of clusters) for the ESG ETFs is 3 using the k-means method as well as Ward's method. For the conventional ETFs, the optimal number of *k* increased to 4. For all ETFs in the early beginning phase, the optimal number of *k* remained 4 using the SC. Both selected clustering methods were consistent and suggested the same number of *k*.

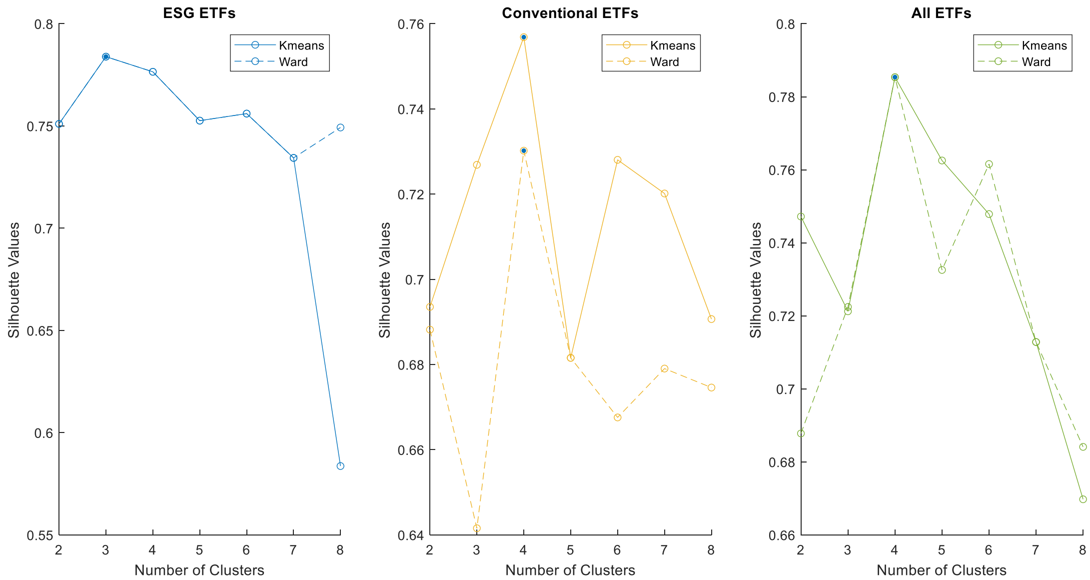


Figure 11. The silhouette values of the early beginning phase

To assess the quality of all *k* clusters, the obtained average silhouette coefficients are used, and they are presented in Figure 11 as discussed above. They are derived from silhouette values, which indicates how well-separated the resulting clusters are (Appendix 3). According to Kaufman and Rousseeuw (2005), the structures obtained from clustering the ETFs can be interpreted as strong, since SC in all obtained clusters exceeds the 0.71 threshold. The k-means have resulted in slightly higher SC in the case of the conventional ETFs, however the number of suggested *k* remains the same using both methods.

5.2.2 Cluster characteristics of the early beginning phase

The cluster characteristics can be seen in Table 2 in the next page, where the ETFs' tickers are color-coded depending on whether the ETF is conventional (blue) or ESG (green). The cluster characteristics are the results of using the k-means method, as Ward's method has been used to improve the k-means method's initial approximation of partitioning, however, both methods have been fairly consistent in terms of the number of clusters obtained. Considering all ETFs, conventional and ESG ETFs are identified within every obtained cluster. Cluster two is the biggest, as it holds the largest number of members, which are mostly consisting of the ESG ETFs. This particular cluster has resulted in the lowest volatility (12.03%); on the contrary, the highest annualized volatility was obtained in cluster three (20.75%), mostly composed of conventional ETFs. Interestingly, clusters two and four have outperformed clusters one and three, indicating that the clusters containing relatively more ESG ETFs positively impact the clusters' returns and generally lower the volatilities. In addition, the minimum daily return belongs to cluster three, where all ETFs are conventional ETFs except one, DURA. From Appendix 1, the Morningstar sustainability rating for DURA is 2 globes, expressing that DURA belongs to the lower rating category.

Clustering the ESG ETFs resulted in three clusters, from which cluster number three is the biggest. It positions itself between clusters one and two regarding of the returns and volatility. The clustering results of obtained cluster three are logical, since 17 members out of 21 follow the same index "Morningstar US Large-Mid TR USD". Cluster two reached the highest returns and the highest volatility (15.18%) and this cluster consists of ESG ETFs whose exposure to the technology sector is the largest. Cluster one was the most stable considering the examined phase's volatility was closest to zero. The average Morningstar sustainability ratings (MSR) show that cluster three has obtained the highest average sustainability rating (4.1) while simultaneously it has a significantly lower annualized volatility than cluster two, in which MSR was 3.2. However, interestingly the lowest annualized volatility and the lowest MSR were identified for cluster one.

Table 2. Cluster characteristics of the early beginning phase

	Cluster	No of members	Members (ticker)	Min r	Max r	$\bar{r}$	Annualized r	Annualized σ
All ETFs	Cluster 1	22	QLV, SHE, ESML, ISMD, RZG, OMFS, UWM, CSML, JPSE, RZV, PSCT, GSSC, EES, XSVM, FYX, VTWV, PRFZ, PFM, GSEW, EPS, DEUS, EWMC	-4.14 %	4.12 %	0.02 %	4.45 %	14.84 %
	Cluster 2	27	SUSL, SNPE, IQSU, HLAL, SPUS, TDV, VEGN, NULC, SUSA, ESGU, PHO, BIBL, DSI, NACP, ESGV, WOMN, ACSI, ETHO, KRMA, CATH, CHGX, FYC, DWAS, IHE, FCOM, VOX, VIG	-2.46 %	2.12 %	0.18 %	58.76 %	12.03 %
	Cluster 3	12	DURA, QABA, FDM, SMLV, DPST, XES, KIE, XSLV, MLPA, PEJ, KRE, KBWY	-5.64 %	6.76 %	-0.11 %	-24.73 %	20.75 %
	Cluster 4	7	CLOU, WCLD, PAWZ, AIQ, BOSS, XITK, IBUY	-2.70 %	3.00 %	0.29 %	106.40 %	16.07 %

	Cluster	No of members	Members (ticker)	MSR	Min r	Max r	$\bar{r}$	Annualized r	Annualized σ
ESG ETFs	Cluster 1	5	QLV, DURA, SHE, ESML, ISMD	3	-2.30 %	1.96 %	0.08 %	23.17 %	11.35 %
	Cluster 2	5	CLOU, WCLD, PAWZ, AIQ, BOSS	3.2	-2.70 %	2.86 %	0.31 %	120.55 %	15.18 %
	Cluster 3	21	SUSL, SNPE, IQSU, HLAL, SPUS, TDV, VEGN, NULC, SUSA, ESGU, PHO, BIBL, DSI, NACP, ESGV, WOMN, ACSI, ETHO, KRMA, CATH, CHGX	4.1	-2.46 %	2.12 %	0.19 %	63.18 %	11.90 %

	Cluster	No of members	Members (ticker)	Min r	Max r	$\bar{r}$	Annualized r	Annualized σ
Conventional ETFs	Cluster 1	16	RZG, OMFS, UWM, CSML, JPSE, RZV, GSSC, EES, XSVM, FYX, VTWV, PRFZ, PFM, DEUS, EPS, EWMC	-4.14 %	4.12 %	-0.01 %	-2.49 %	15.64 %
	Cluster 2	11	QABA, FDM, SMLV, DPST, XES, KIE, XSLV, MLPA, PEJ, KRE, KBWY	-5.64 %	6.76 %	-0.10 %	-23.23 %	20.54 %
	Cluster 3	2	XITK, IBUY	-2.22 %	3.00 %	0.22 %	74.85 %	18.02 %
	Cluster 4	8	FYC, DWAS, IHE, PSCT, FCOM, VOX, VIG, GSEW	-2.56 %	1.74 %	0.12 %	36.77 %	13.05 %

MSR – Morningstar Sustainability Rating (ETFs' average globes),

Min r – ETFs minimum daily return in a cluster,
Max r – ETFs maximum daily return in a cluster,
$\bar{r}$ – average daily return of ETFs' in a cluster during the examined phase,
Annualized r – cluster's daily average return transformed into a yearly metric,
Annualized σ – examined phase's standard deviation of the returns transformed into a yearly metric.

The conventional ETFs have been clustered into four clusters. Cluster one is the largest cluster containing 16 ETFs from which 12 follow a small cap related index, such as "Morningstar US Small Cap Ext TR USD ", and "Morningstar US Small Brd Val Ext TR USD". The average returns of cluster one are the smallest, while the volatility is the second lowest. Cluster number two is showing the highest max daily returns as well as minimum daily returns, and intuitively its volatility is the highest compared to the other obtained clusters (20.54%) in the examined phase. Cluster three is the smallest, only containing two members: XITK (index focused mostly on technology sector) and IBUY (index focused mostly on consumer cyclical sector). Its volatility is the second highest and average returns and annualized returns in the early beginning phase are the highest compared to the other clusters formed of the conventional ETFs. The fourth cluster contains ETFs, which follow indexes that are mostly focused on the technology, healthcare, and communication services sectors. The cluster has the lowest volatility among all conventional ETF clusters and the second highest average and annualized returns.

The overall clustering results of the early beginning phase have shown that there are few general patterns that ETFs tend to follow and there are similarities between the clusters obtained in both groups (ESG and conventional). In both groups, there are clusters with lower volatility and clusters with much higher volatility. Generally, ESG ETFs have resulted in lower volatilities than conventional ETFs. From the ESG ETFs, cluster two has the highest volatility (15.18%) consisting of ETFs mostly focused on the technology sector. While in the case of conventional ETFs, cluster two has resulted in the highest volatility, in which two particular ETFs (XES and DPST) have increased the cluster's volatility significantly. XES (index following mainly energy sector) individual annualized volatility in the examined phase is 33.92% and DPST (index following technology and energy sectors) resulted in 53.74%. All obtained clusters of ESG ETFs have resulted in positive average and annualized returns, whereas the conventional ETFs have shown both, positive and negative returns. From the ESG ETFs, cluster two has resulted in the minimum, maximum, and highest average, annualized returns and also the highest volatility. In the case of the conventional ETFs, cluster two reached the highest volatility, minimum and maximum returns while the annualized return was negative. The movement of cluster memberships through the different phases is discussed in subsection 5.5.1, once all the phases considered have been covered.

5.3 Collapse phase

The second examined phase is the collapse phase (19.2.-23.3.2020). This phase can be seen as the breakdown phase of the COVID-19 crisis, when the WHO declared the disease a pandemic. In Figure 12 below, the movement of normalized prices for conventional ETFs and ESG ETFs are illustrated.

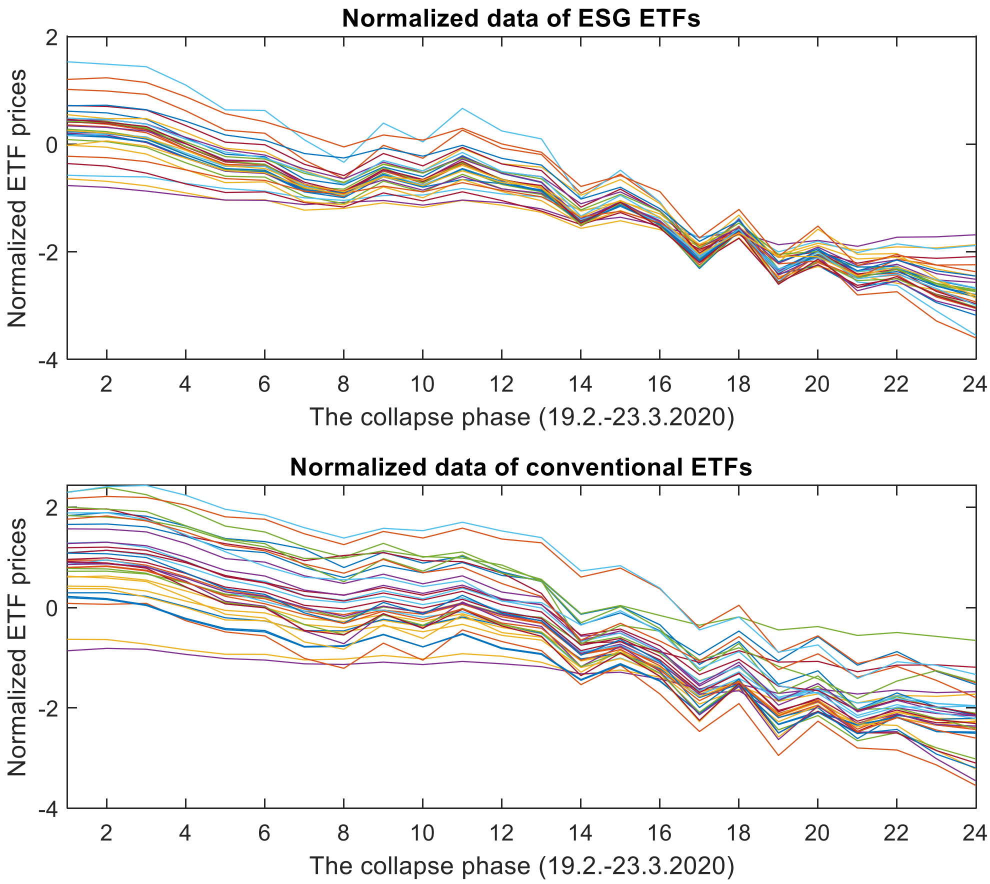


Figure 12. Illustration of the collapse phase

From Figure 12, at first, both groups (conventional and ESG) seem to move differently, as ESG ETFs' prices decrease more similarly than the conventional ETFs' prices. In the middle, the prices of both ETFs move similarly as they both appear to have similarities in upward and downward trends. Towards the end of the examined phase, ETFs continue to behave differently. They are eventually settling down at the end, however, the conventional ETFs seem to have more dispersion in the price movement.

5.3.1 Clustering results of the collapse phase

Similarly, the k-means and Ward's method are applied to examine possible price movement behavior in the collapse phase of the COVID-19 crisis. The calculated range of clusters was 2 up to 15 clusters, however, the result 2 to 8 is shown since the number of optimal clusters remained stable. In Figure 13, the average silhouette coefficient and the suggested number of optimal clusters are shown. For ESG ETFs, the optimal number of *k* based on the SC differs. The k-means suggests 3 and Ward's method states the *k* to be 2. The number of *k* is selected to be 3 as the k-means method suggests, since it is providing higher SC compared if the *k* would be 2. In the case of the conventional ETFs and all ETFs combined, both methods are consistent and they suggest the optimal number of clusters to be 2. It can be seen from Figure 13 that the SC are decreasing considerably after the suggested number of optimal clusters.

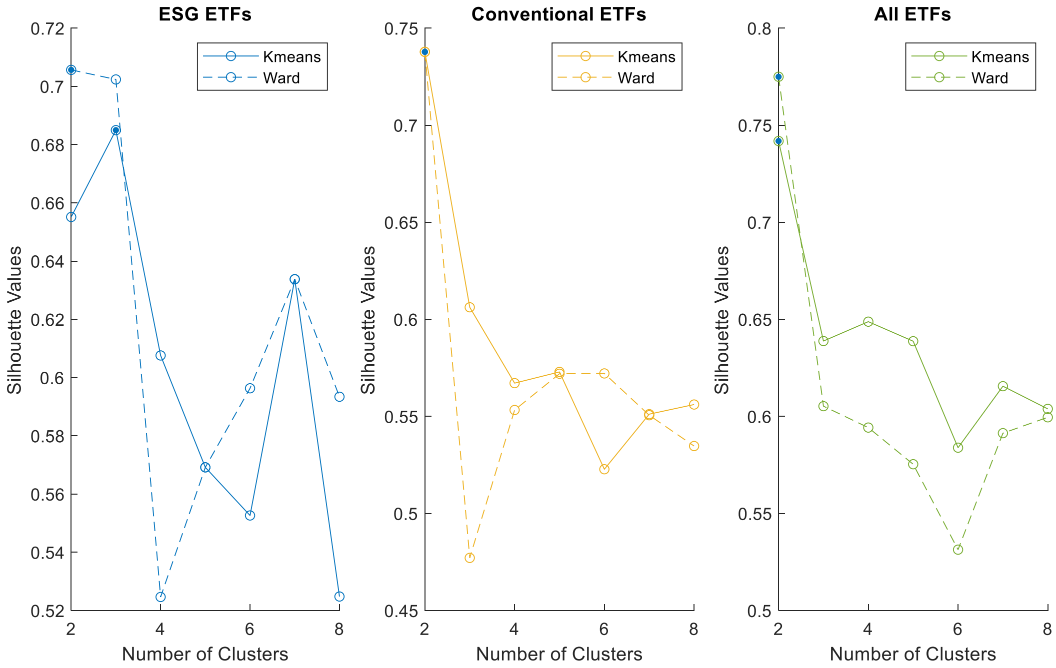


Figure 13. The silhouette values of the collapse phase

To assess the quality of all obtained *k* clusters, the average silhouette coefficients are used. They can be seen in Figure 13 above. The silhouette values from which the SC are derived are presented in Appendix 4. In the case of all ETFs and conventional ETFs, both methods suggested the same number of clusters (*k*=2) maximizing the silhouette coefficient criterion. For ESG ETFs, the k-means and Ward's method based on the silhouette coefficient differ slightly, from which the higher number of clusters are used based on the k-means higher SC for *k*=3 clusters than *k*=2 clusters. Kaufman and Rousseeuw (2005) developed general thresholds for interpreting the silhouette coefficients. According to them, the structures obtained from clustering the ETFs can be interpreted as strong for all ETFs and conventional ETFs. However, the ESG ETFs silhouette coefficient for *k*=3 is slightly under the 0.71 threshold which indicates that the structure of the data is reasonable.

5.3.2 Cluster characteristics of the collapse phase

The cluster characteristics can be seen in Table 3, where the ETFs' tickers are color-coded depending on whether the ETF is conventional (blue) or ESG (green). Clustering results of all ETFs concluded two clusters. The first cluster contains 18 members of which only two are ESG ETFs. From Appendix 1, the Morningstar sustainability rating for ISMD is 1 and for DURA it is 2 globes, which expresses that DURA and ISMD belong to the lower category of ETFs concerning the sustainability rating. The cluster one has significantly higher volatility and the minimum and maximum returns, which are the results of specific ETFs DPST and XES. DPST follows an index containing the technology and energy sector and XES follows an index that heavily focuses on the energy sector. In addition to DPST, UWM follows the same index "Morningstar US Market TR USD" as DPST. Cluster two contains the rest of the ESG ETFs, while the conventional ETFs are presenting slightly less than half of the ETFs in cluster two. Of the two obtained clusters, cluster two is less volatile and results in lower maximum, minimum and average as well as annualized returns.

Clustering ESG ETFs resulted in three clusters, from which cluster one and two contains five members and cluster three contains 21. Overall, clusters have rather similar characteristics with slight modifications. Cluster one has the smallest maximum return and the lowest volatility, whereas cluster two results in the lowest average return and highest volatility. The clusters containing only ESG ETFs have obtained the lowest volatilities among all considered ETFs in the collapse phase. Interestingly, cluster two has resulted in

Table 3. Cluster characteristics of the collapse phase

	Cluster	No of members	Members (ticker)	MSR	Min r	Max r	$\bar{r}$	Annualized r	Annualized σ
All ETFs	Cluster 1	18	DURA, ISMD, QABA, FDM, UWM, CSML, DPST, XES, KIE, EES, XSLV, VTWV, MLPA, PEJ, KRE, XSVM, SMLV, KBWY		-61.50 %	26.70 %	-2.84 %	-99.93 %	127.52 %
	Cluster 2	50	SUSL, SNPE, CLOU, WCLD, IQSU, PAWZ, AIQ, HLAL, QLV, SPUS, TDV, VEGN, NULC, SUSA, ESGU,PHO, BIBL, DSI, NACP, SHE, BOSS, WOMN, ACSI, ETHO, KRMA, CATH, ESML, CHGX, RZG, XITK, OMFS, JPSE, FYC, IBUY, RZV, DWAS, IHE, PSCT, GSSC, FCOM, FYX, PRFZ, VOX, VIG, PFM, GSEW, EPS, DEUS, EWMC		-17.36 %	12.76 %	-1.47 %	-97.61 %	83.19 %
	Cluster	**No of members**	**Members (ticker)**	**MSR**	**Min r**	**Max r**	$\bar{r}$	**Annualized r**	**Annualized σ**
ESG ETFs	Cluster 1	5	CLOU, WCLD, PAWZ, AIQ, BOSS	3.2	-14.61 %	9.60 %	-1.17 %	-94.88 %	75.55 %
	Cluster 2	5	QLV, DURA, SHE, ESML, ISMD	3	-15.33 %	10.20 %	-1.58 %	-98.18 %	87.07 %
	Cluster 3	21	SUSL, SNPE, IQSU, HLAL, SPUS, TDV, VEGN, NULC, SUSA, ESGU, PHO, BIBL, DSI, NACP, ESGV, WOMN, ACSI, ETHO, KRMA, CATH, CHGX	4.1	-14.29 %	10.45 %	-1.35 %	-96.76 %	80.27 %
	Cluster	**No of members**	**Members (ticker)**		**Min r**	**Max r**	$\bar{r}$	**Annualized r**	**Annualized σ**
Conventional ETFs	Cluster 1	26	RZG, XITK, OMFS, UWM, CSML, JPSE, FYC, IBUY, RZV, DWAS, IHE, PSCT, GSSC, EES, XSVM, FCOM, FYX, VTWV, PRFZ, VOX, VIG, PFM, GSEW, EPS, DEUS, EWMC		-30.90 %	17.37 %	-3.20 %	-99.97 %	92.59 %
	Cluster 2	11	QABA, FDM, SMLV, DPST, XES, KIE, XSLV, MLPA, PEJ, KRE, KBWY		-61.50 %	26.70 %	-1.82 %	-99.03 %	140.78 %

MSR – Morningstar Sustainability Rating (ETFs' average globes),
Min r – ETFs minimum daily return in a cluster,
Max r – ETFs maximum daily return in a cluster,
$\bar{r}$ – average daily return of ETFs' in a cluster during the examined phase,
Annualized r – cluster's daily average return transformed into a yearly metric,
Annualized σ – examined phase's standard deviation of the returns transformed into a yearly metric.

the lowest MSR while simultaneously it has resulted in the highest volatility and the lowest average return among the obtained ESG ETFs clusters. Cluster three obtained the highest MSR (4.1), in which maximum return was the highest and minimum return the smallest. In addition, it resulted in lower volatility than cluster two, which may indicate that it may be more risk-tolerant than cluster two.

The conventional ETFs were clustered into two very different clusters, from which cluster one contains 26 members and cluster two 11. Cluster one has resulted in overall smaller returns as well as the specific phase volatility and the annualized volatility. Cluster two has significantly higher volatilities, in addition, it includes DPST and XES ETFs which contribute to cluster performance, as seen also in the cluster characteristics of all ETFs (cluster 1).

The collapse phase resulted in all obtained clusters' volatilities being very high as well as their annualized returns being negative. Compared to the previously examined phase, all ESG ETFs' average returns were positive and two out of four obtained clusters of conventional ETFs were positive as well. Whereas in the collapse phase, the ESG ETFs resulted in lower negative returns whereas minimum and maximum returns were relatively smaller compared to conventional ETFs. Generally, in the collapse phase as well as in the early beginning phase the ESG ETF clusters obtained lower volatilities and in addition, ESG ETFs lowered the volatilities when all ETFs were considered. The movement of cluster memberships during different phases are discussed in subsection 5.5.1.

5.4 Recovery phase

The recovery phase (24.3.-30.6.2020) is the third and final examined phase in this thesis. The COVID-19 crisis market collapse phase was investigated in the previous phase whereas the recovery and possible smoothing of securities markets are considered next. The recovery phase is the longest time frame examined. Figure 14 shows the price movement of ESG and conventional ETFs in the examined phase. It can be observed, that ESG ETFs have developed more consistently, although there have been small differences between the beginning and the end of the period. In the case of conventional ETFs, the price movement from the beginning to halfway differs from ESG ETFs. Both groups of ETFs however show a continuing upward trend reaching a stable point at the end of the examined phase.

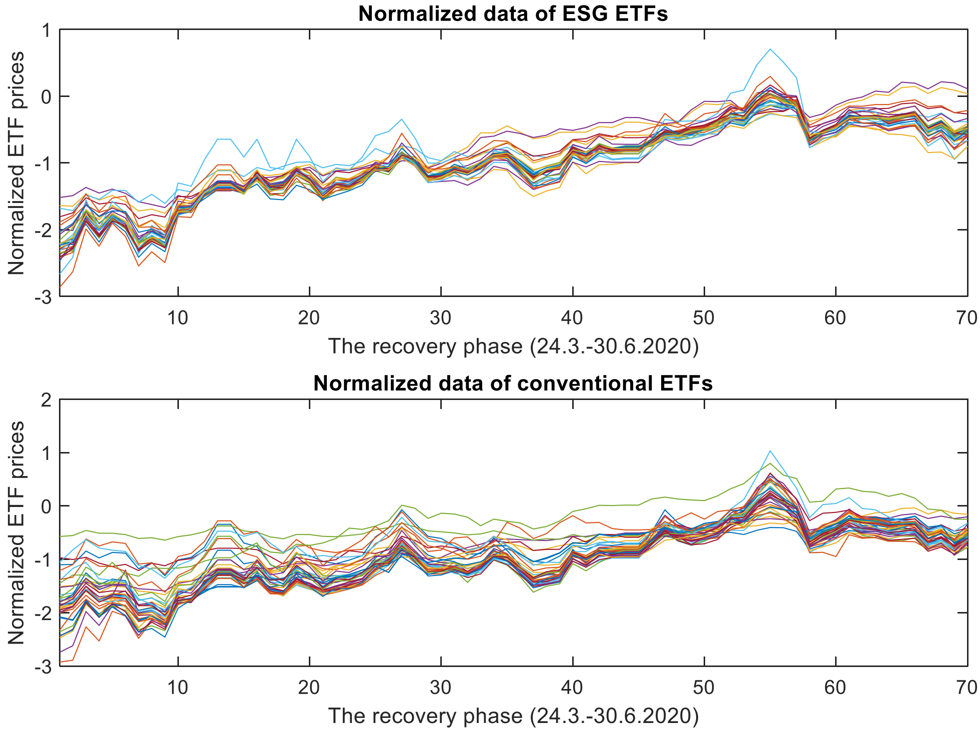


Figure 14. Illustration of the recovery phase

5.4.1 Clustering results of the recovery phase

As in the previous phases, the k-means and Ward methods are applied to the data to examine the recovery phase. The calculated range of clusters was 2 up to 15 clusters, however, the result 2 to 8 is shown since the number of optimal clusters remained unchangeable. The average silhouette coefficients and suggested number of optimal clusters are shown in Figure 15 and the silhouette values are presented in Appendix 5. For ESG ETFs, k-means and Ward's method suggested the same k=3. For the conventional ETFs and all ETFs, both methods remained stable and they suggest k to be 2.

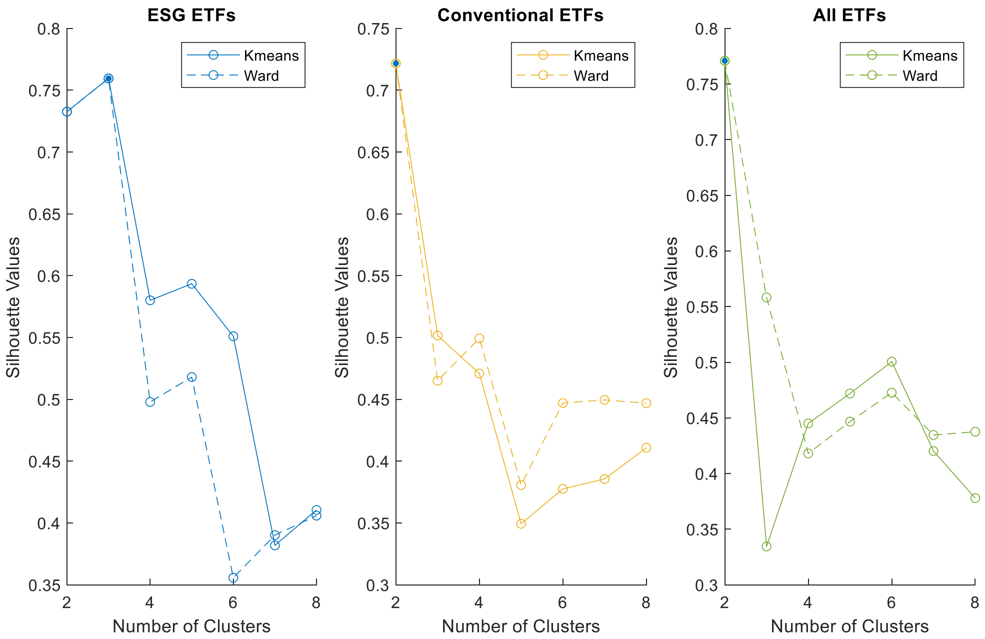


Figure 15. The silhouette values of the recovery phase

In the case of all ETFs, ESG ETFs and conventional ETFs, both clustering methods suggested the same number of optimal clusters utilizing the maximizing the silhouette coefficient criterion. According to Kaufman and Rousseeuw (2005), the silhouette coefficients for all ETFs, conventional ETFs, and ESG ETFs exceeded the threshold SC≤0.71 indicating that a strong structure in the data is found.

5.4.2 Cluster characteristics of the recovery phase

The cluster characteristics can be seen in Table 4, where the ETFs' tickers are color-coded depending on whether the ETF is conventional (blue) or ESG (green). The clustering results of all ETFs concluded two clusters, where cluster one has significantly more members: 61 members compared to the cluster two, which has seven members. Cluster one contains all considered ESG ETFs and it has resulted in substantially lower volatility and higher annualized returns than cluster two which consists entirely of conventional ETFs.

Table 4. Cluster characteristics of the recovery phase

	Cluster	No of members	Members (ticker)	MSR	Min r	Max r	$\bar{r}$	Annualized r	Annualized σ
All ETFs	Cluster 1	61	SUSL, SNPE, CLOU, WCLD, IQSU, PAWZ, AIQ, HLAL, QLV, SPUS, TDV, VEGN, DURA, NULC, SUSA, ESGU, PHO, BIBL, DSI, NACP, SHE, ESGV, BOSS, WOMN, ACSI, ETHO, KRMA, CATH, ESML, ISMD, CHGX, RZG, XITK, FDM, OMFS, UWM, CSML, SMLV, JPSE, FYC, IBUY, RZV, DWAS, IHE, PSCT, GSSC, KIE, EES, XSVM, FCOM, FYX, VTWV, PEJ, PRFZ, VOX, VIG, PFM, GSEW, EPS, DEUS, EWMC		-16.43 %	14.05 %	0.40 %	170.83 %	41.67 %
	Cluster 2	7	QABA, DPST, XES, XSLV, MLPA, KRE, KBWY		-33.27 %	26.03 %	0.28 %	101.32 %	90.46 %
	Cluster	**No of members**	**Members (ticker)**	**MSR**	**Min r**	**Max r**	$\bar{r}$	**Annualized r**	**Annualized σ**
ESG ETFs	Cluster 1	25	SUSL, SNPE, IQSU, HLAL, QLV, SPUS, TDV, VEGN, NULC, SUSA, ESGU, PHO, BIBL, DSI, NACP, SHE, ESGV, ACSI, ETHO, KRMA, CATH, ESML, ISMD, CHGX	3.96	-8.00 %	8.30 %	0.36 %	150.42 %	35.44 %
	Cluster 2	5	CLOU, WCLD, PAWZ, AIQ, BOSS	3.2	-6.47 %	7.54 %	0.59 %	338.15 %	35.63 %
	Cluster 3	1	DURA	2	-6.53 %	7.05 %	0.22 %	75.79 %	36.56 %
	Cluster	**No of members**	**Members (ticker)**		**Min r**	**Max r**	$\bar{r}$	**Annualized r**	**Annualized σ**
Conventional ETFs	Cluster 1	30	RZG, XITK, FDM, OMFS, UWM, CSML, SMLV, JPSE, FYC, IBUY, RZV, DWAS, IHE, PSCT, GSSC, KIE, EES, XSVM, FCOM, FYX, VTWV, PEJ, PRFZ, VOX, VIG, PFM, GSEW, EPS, DEUS, EWMC		-16.43 %	14.05 %	0.40 %	170.68 %	48.04 %
	Cluster 2	7	QABA, DPST, XES, XSLV, MLPA, KRE, KBWY		-33.27 %	26.03 %	0.28 %	101.32 %	90.46 %

MSR – Morningstar Sustainability Rating (ETFs' average globes),
Min r – ETFs minimum daily return in a cluster,
Max r – ETFs maximum daily return in a cluster,
$\bar{r}$ – average daily return of ETFs' in a cluster during the examined phase,
Annualized r – cluster's daily average return transformed into a yearly metric,
Annualized σ – examined phase's standard deviation of the returns transformed into a yearly metric.

Cluster two has obtained higher values in minimum and maximum returns as well as phase volatility and annualized volatility, which are results of the ETF DPST. In cluster one ETF UWM has resulted in the highest maximum return as well as the minimum return. In addition, the presence of ESG ETFs in cluster one balances the volatility significantly.

The ESG ETFs have formed three clusters, each of which is very different in size. Cluster two contains five members and cluster three has only one member, while cluster one consists of 25 ETFs. Cluster one has obtained the highest minimum return resulting from ETF ISMD and the highest maximum return, resulting from ETF ETHO. In this particular cluster, there are multiple indexes that the ETFs follow in contrast to cluster two where all ETFs followed indexes that focus heavily on the technology sector and, three members out of five follow the same index “Morningstar US Technology TR USD”. Cluster three consists of only ETF DURA, which Morningstar sustainability rating is 2 and this particular cluster have resulted in the lowest annualized return while annualized volatility is only slightly higher than in other obtained clusters. Cluster one has significantly more members which have higher MSR, however this has not shown especially better risk tolerance measured by volatility in the examined phase. The best performing cluster measured by the annualized return is cluster two, which MSR is 3.2.

Clustering the conventional ETFs resulted in two clusters, where cluster one contains 30 ETFs out of 37 and cluster two contains seven ETFs. If we examine the obtained volatilities, cluster one’s reaction in the recovery phase has been much more stable than cluster two’s. Cluster two includes ETFs that follow indices focusing heavily on financial, technology, and energy sectors, and cluster two contains the aforementioned ETFs (XES and DPST), which have been associated with the largest price movements in the previously examined phases as well.

5.5 Analysis of the results

The overall clustering results for the early beginning phase indicate that the phase considered led to fairly similar cluster characteristics in terms of annualized volatilities. ETFs were clustered into three (ESG ETFs) and four (all ETFs and traditional ETFs) clusters. However, the conventional ETFs resulted in a bit higher volatility compared to the ESG ETFs, and the maximum daily returns were much higher. Nevertheless, two out of four obtained clusters of conventional ETFs' annualized returns stayed negative while all obtained clusters of ESG ETFs resulted in positive values. The returns used in the analysis are not risk-adjusted, thus this needs to be taken into consideration.

In the collapse phase, all ETFs were clustered into two clusters, ESG ETFs into three clusters and conventional ETFs into two clusters. As illustrated in Table 3, the results of volatilities between the conventional and ESG ETFs were quite different, where the conventional ETFs obtained generally higher or notably higher volatilities. These observations of return volatilities are in line with Albuquerque et al. (2020) findings. In addition, similarities with the study by Omura, Roca and Nakai (2021) were found concerning the average returns. Results showed that ESG ETFs did not suffer the most in the market crash period, which indicates that ESG might offer some form of protection for investors in market downturns, which was also acknowledged by Pisani and Russo (2021).

The third and last examined phase considered is the recovery phase of the COVID-19-related market crash. As in the previous phases examined, the return volatilities were lower for ESG ETFs compared to the conventional ETFs, and ESG clusters seem to recover faster concluded also by Pisani and Russo (2021), which is reflected in the cluster characteristics as the min and max, as well as the volatilities have resulted in more stable values. The characteristics of conventional ETF clusters, especially higher volatilities, may indicate that these specific ETFs have recovered from market crash more slowly.

5.5.1 Cluster memberships during the examined phases

The cluster memberships in the three examined phases are discussed next. Figure 16, Figure 17, and Figure 18 present the transition of considered ETFs' during the different phases. Colors in the figures are explaining the transitions, where every cluster has its unique color. The early beginning phase is used as a benchmark and clusters in this phase are color-coded in the lightest color. When proceeding to the collapse phase, the ETFs that have previously formed a cluster in the early beginning phase now have a darker color as it indicates how a specific ETF or group of ETFs has transitioned from the beginning phase to the next phase and eventually to the third phase. For example, the collapse phase cluster one (Figure 16) contains members from the early beginning phase of clusters one and three. Further on, in the recovery phase cluster one consists of all original cluster members, the only exception is the original red cluster which has split into two separate clusters.

All examined phases and their clustering results from clustering all ETFs are illustrated in Figure 16. As mentioned earlier, the early beginning phase resulted in four clusters, which consist of both ESG and conventional ETFs. As the pandemic progresses, ETF price movements are closer together. Thus, the obtained number of clusters decreases, and two clusters are obtained in both next proceeding phases. Cluster number two in the collapse phase includes all members of cluster two from the early beginning phase, which the darker yellow color is indicating. Similarly, both red and blue color-coded clusters' ETFs have wholly transitioned to clusters one and two without the separation of original cluster members from the beginning phase to the collapse phase. In addition, the red clusters' members have been transitioned into two different clusters at the last studied phase. Interestingly only cluster one's (green) members have been transitioning into two separate clusters, and in the last examined phase they have re-clustered again. The cluster members color-coded as blue have always remained in the same cluster despite the different phases of COVID-19, which may be a result of all the ETFs examined, these ETFs are the most concentrated in the technology sector. The exception is IBUY, which mainly focuses on the consumer cyclical sector.

All ETFs

Phase	Cluster	Members
The early beginning phase	Cluster 1	QLV SHE ESML ISMD RZG OMFS UWM CSML JPSE RZV PSCT GSSC EES XSVM FYX VTWV PRFZ PFM GSEW EPS DEUS EWMC
	Cluster 2	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP ESGV WOMN ACSI ETHO KRMA CATH CHGX FYC DWAS IHE FCOM VOX VIG
	Cluster 3	DURA QABA FDM SMLV DPST XES KIE XSLV MLPA PEJ KRE KBWY
	Cluster 4	CLOU WCLD PAWZ AIQ BOSS XITK IBUY
The collapse phase	Cluster 1	DURA QABA FDM SMLV DPST XES KIE XSLV MLPA PEJ KRE KBWY ISMD UWM CSML EES VTWV XSVM
	Cluster 2	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP WOMN ACSI ETHO KRMA CATH CHGX ESGV FYC DWAS IHE FCOM VOX VIG CLOU WCLD PAWZ AIQ BOSS XITK IBUY QLV SHE ESML RZG OMFS JPSE RZV PSCT GSSC FYX PRFZ PFM GSEW EPS DEUS EWMC
The recovery phase	Cluster 1	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP ESGV WOMN ACSI ETHO KRMA CATH CHGX FYC DWAS IHE FCOM VOX VIG CLOU WCLD PAWZ AIQ BOSS XITK IBUY QLV SHE ESML ISMD RZG OMFS UWM CSML JPSE RZV PSCT GSSC EES XSVM VTWV FYX PRFZ PFM GSEW EPS DEUS EWMC DURA FDM SMLV KIE PEJ
	Cluster 2	QABA DPST XES XSLV MLPA KRE KBWY

Figure 16 illustrates how clusters' memberships are moving during the examined phases. The early beginning phase shows the original cluster membership of ETFs as the lightest color as a benchmark. The color darkens in the next phase and shows how memberships have moved as the pandemic

Figure 16. Movement of all ETFs cluster members at the examined phases

Based on Figure 17 below, the clustering results obtained in the three phases of COVID-19 studied for ESG ETFs are shown. Clustering resulted in that certain ETFs being clustered together in all examined phases. These ETFs (blue) included in particular ESG ETFs: CLOU, WCLD, PAWZ, AIQ, and BOSS. As mentioned in earlier, the aforementioned ETFs follow indices that are heavily focused on the technology sector. In addition, obtained clusters in the early beginning phase and the collapse phase have remained unchanged.

ESG ETFs

	Cluster	MSR	Members (ticker)
The early beginning phase	Cluster 1	3	QLV DURA SHE ESML ISMD
	Cluster 2	3.2	CLOU WCLD PAWZ AIQ BOSS
	Cluster 3	4.1	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP ESGV WOMN ACSI ETHO KRMA CATH CHGX
The collapse phase	Cluster 1	3.2	CLOU WCLD PAWZ AIQ BOSS
	Cluster 2	3	QLV DURA SHE ESML ISMD
	Cluster 3	4.1	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP ESGV WOMN ACSI ETHO KRMA CATH CHGX
The recovery phase	Cluster 1	3.96	SUSL SNPE IQSU HLAL SPUS TDV VEGN NULC SUSA ESGU PHO BIBL DSI NACP ESGV ACSI ETHO KRMA CATH CHGX WOMN ESML ISMD QLV SHE
	Cluster 2	3.2	CLOU WCLD PAWZ AIQ BOSS
	Cluster 3	2	DURA

Figure 17 illustrates how clusters' memberships are moving during the examined phases. The early beginning phase shows the original cluster membership of ETFs as the lightest color as a benchmark. The color darkens in the next phase and shows how memberships have moved as the pandemic proceeds. MSR – Morningstar Sustainability Rating (ETFs' average globes)

Figure 17. Movement of ESG ETFs cluster members at the examined phases

Some industries are more volatile than others in different economic cycles, as mentioned by Haroon and Rizvi (2020); Buszko, Orzeszko and Stawarz (2021); Mazur, Dang and Vega (2021). However, this study’s objective is to find clusters of ETFs with similar behaviour in terms of daily price movements, not focusing on too much on the question of relationships between sectors’ performance. Nevertheless, the results are intuitive since similar movements are identified with respect to ETFs in the same sectors. Lastly, in the recovery phase, the cluster (yellow) which has remained unchangeable in two previously examined phases was in the end transformed into two separate clusters. Differences between the transitioned members are ETF DURA's index focuses mostly on financial services and healthcare while other ETF indices focus on technology and industrials which may influence their price movements differently in the recovery phase.

The conventional ETFs' clustering results concerning cluster memberships are illustrated in Figure 18 below. The beginning phase clustering resulted in four clusters, which then decreased into two clusters for the collapse and recovery phase. As seen in the case of ESG

Conventional ETFs

Phase	Cluster	Members (ticker)										
The early beginning phase	Cluster 1	RZG	OMFS	UWM	CSML	JPSE	RZV	GSSC	EES	XSVM	FYX	VTWV
		PRFZ	PFM	DEUS	EPS	EWMC						
	Cluster 2	QABA	FDM	SMLV	DPST	XES	KIE	XSLV	MLPA	PEJ	KRE	KBWY
	Cluster 3	XITK	IBUY									
	Cluster 4	FYC	DWAS	IHE	PSCT	FCOM	VOX	VIG	GSEW			
The collapse phase	Cluster 1	RZG	OMFS	UWM	CSML	JPSE	RZV	GSSC	EES	XSVM	FYX	VTWV
		PRFZ	PFM	DEUS	EPS	EWMC	XITK	IBUY	FYC	DWAS	IHE	PSCT
		FCOM	VOX	VIG	GSEW							
	Cluster 2	QABA	FDM	SMLV	DPST	XES	KIE	XSLV	MLPA	PEJ	KRE	KBWY
The recovery phase	Cluster 1	RZG	OMFS	UWM	CSML	JPSE	RZV	GSSC	EES	XSVM	FYX	VTWV
		PRFZ	PFM	EPS	DEUS	EWMC	FDM	SMLV	PEJ	KIE	XITK	IBUY
		FYC	DWAS	IHE	PSCT	FCOM	VOX	VIG	GSEW			
	Cluster 2	QABA	DPST	XES	XSLV	MLPA	KRE	KBWY				

Figure 18 illustrates how clusters' memberships are moving during the examined phases. The early beginning phase shows the original cluster membership of ETFs as the lightest color as a benchmark. The color darkens in the next phase and shows how memberships have moved as the pandemic proceeds.

Figure 18. Movement of conventional ETFs cluster members at the examined phases

ETFs, also in conventional ETFs the beginning phase and the collapse phase have resulted in clusters, in which each member has transitioned into a new cluster wholly. As seen in the initial yellow, red, and blue clusters, their members have formed a new bigger cluster number one in the collapse phase as well as recovery phase. Only the green cluster’s members have transitioned into two separate clusters in the last examined phase, which might be the result of a different mix of indices focusing on the energy, real estate, financial services, and technology sector (cluster two) whilst cluster one's ETFs are focusing on financial services, industrials, and consumer cyclical sectors.

Overall, there have been more transitions of cluster members in the conventional ETFs and all ETFs compared to ESG ETFs during the examined phases. The original number of clusters decreased from four to two in the case of all ETFs and conventional ETFs, however clustering ESG ETFs resulted in always the same number of clusters, which in addition contained the same members during the phases except for the recovery phase.

5.5.2 Silhouette coefficients during the examined phases

The homogeneity of the obtained clusters is analyzed based on the average silhouette coefficient. Observing the examined phases' average silhouette coefficients, the highest SC was reached in the early beginning phase followed by the recovery phase and lastly in the collapse phase in the case of all ETFs and ESG ETFs. Only clusters consisting of conventional ETFs resulted in higher SC in the collapse phase compared to the recovery phase. However, the SC for all ETFs, conventional ETFs and ESG ETFs generally have resulted in values between $0.71 \leq SC \leq 1$ in all examined phases (excluding the collapse phase SC for ESG ETFs) indicating a good clustering result from the used methods and the distance measure. It is worth noting that when all ETFs were clustered together, the SC gave higher values than when they were clustered separately (ESG and conventional) in every examined phase.

The bigger number of obtained clusters in the early beginning phase may indicate that the price movements of the ETFs were more diversified (Buszko, Orzeszko and Stawarz, 2021). The highest SC was obtained for the early beginning phase, which supports the findings of Buszko, Orzeszko and Stawarz (2021), however, the difference between phases SC is small. In the case of the collapse phase and the recovery phase all ETFs resulted in two clusters. Therefore, the price movements of all ETFs and conventional ETFs showed that they consist of two larger groups, and no specific ETFs were found that stood out from the groups, which would have resulted in more clusters.

5.5.3 Sustainability ratings during the examined phases

Examined phases have resulted in interesting observations of Morningstar's sustainability rating, volatilities, and performance among clusters, which are presented in Table 5 below.

Table 5. Obtained clusters' information in the examined phases

		The early beginning phase			The collapse phase			The recovery phase		
	Cluster	MSR	Annualized r	Annualized σ	MSR	Annualized r	Annualized σ	MSR	Annualized r	Annualized σ
ESG ETFs	1	3	23.17 %	11.35 %	3.2	-94.88 %	75.55 %	3.96	150.42 %	35.44 %
	2	3.2	130.55 %	15.18 %	3	-98.18 %	87.07 %	3.2	338.15 %	35.63 %
	3	4.1	63.18 %	11.90 %	4.1	-96.76 %	80.27 %	2	75.79 %	36.56 %
Conventional ETFs	1		-2.49 %	15.64 %		-99.97 %	92.59 %		170.68 %	48.04 %
	2		-23.23 %	20.54 %		-99.03 %	140.78 %		101.32 %	90.46 %
	3		74.85 %	18.02 %						
	4		36.77 %	13.05 %						

MSR – Morningstar Sustainability Rating (ETFs' average globes),
Annualized r – cluster's daily average return transformed into a yearly metric,
Annualized σ – examined phase's standard deviation of the returns transformed into a yearly metric.

In the early beginning phase, the obtained clusters' average Morningstar sustainability rating shows that a higher MSR would not indicate a better risk tolerance measured by volatilities, which is also acknowledged by Folger-Laronde et al. (2020) and Hartzmark and Sussman (2019). This is observed as ESG cluster 1: average of 3 globes, and ESG cluster 3: average of 4.1 globes does not have a distinct difference among annualized volatilities.

In the case of the collapse phase, MSR indicates that a lower rating would result in higher volatility. Interestingly, ESG cluster two is not too far away from conventional ETF cluster one's annualized volatility (92.59 %), however, ESG ETFs have resulted in lower volatilities and higher returns, which is in line with Pástor and Vorsatz (2020). In addition, Pisani and Russo (2021) state a high ESG rating is leading to better performance compared to lower-rated ESG funds during COVID-19, which the results of this study are also implying as ESG cluster 2 (lowest MSR) and cluster 3 (highest MSR) have quite a significant difference between returns and volatilities.

In the recovery phase, observing the ESG ETFs' cluster characteristics, MSR has changed for the first time during the examined phases. Although there are no significant differences among the obtained cluster characteristics in the recovery phase with respect to the volatilities, MSR gives a slight indication that a higher rating is associated with lower volatility and higher return, which are also acknowledged by Pavlova and de Boyrie (2022) and Broadstock et al. (2021).

6 Conclusions

Crises occur irregularly and their consequences are impossible to anticipate in all aspects. The COVID-19 pandemic has been one of its kind, it has not only affected people’s health globally but has also caused significant restrictions to the socio-economic activities of countries all over the world. The COVID-19 pandemic and the resulting crisis have created an interesting possibility to explore the relationship between socially responsible investing and conventional investing in unordinary market conditions to which this thesis aims to contribute.

This thesis used unsupervised machine learning to examine the stock market reaction between conventional and ESG ETFs during the first six months of the COVID-19 pandemic. The objective was to identify possible trends among the chosen ETFs' normalized prices using clustering analysis. In addition, the six-month time frame was divided into three different phases to capture occurring changes within the pandemic. The used clustering algorithms were k-means and Ward’s method using the Euclidean distance measure and the obtained number of clusters was validated by maximizing the silhouette coefficient criterion.

6.1 Answers to the research questions

The first main research question is related to the market crises, while simultaneously investigating the reaction of both conventional and ESG investing during uncertain market environment. Thus, the first main question was:

“What kind of performance did socially responsible investing instruments show in market crises according to the literature compared to conventional instruments?”

The technology bubble burst, the global financial crisis as well as COVID-19 crisis were examined based on previous literature. Based on the results of this study, the COVID-19 crisis environment showed that ESG ETFs generally had resulted in higher or similar returns to conventional ETFs. According to the previous studies concerning crises, investors have been able to protect their investments' returns and outperformed conventional funds during market downturns (Nofsinger and Varma 2014; Becchetti et al. 2015; Albuquerque et al. 2020). However, some findings suggest that ESG investing has the opposite effect on the returns (Demers et al. 2020). Based on the results of this study ESG ETFs increased investments' short-term returns during COVID-19, as well as ESG investments have performed generally better compared to its conventional counterparts. Overall, the average returns of ESG ETFs have not decreased as sharply during the COVID-19 crisis as conventional ETFs, which strengthens the conclusion that investors may benefit from using ESG investing in their investing strategy.

Concerning the volatilities during the first six months of the COVID-19 pandemic, ESG ETFs have generally resulted in lower return volatility and higher average returns than conventional ETFs. However, most of the ESG ETFs indices have a large exposure to the technology sector, which would indicate that in the COVID-19 pandemic this might result in mitigating the impact of the crash on the ESG ETFs, which is the opposite of Becchetti et al. (2015) findings as they concluded that during the dot-com crisis socially responsible funds did not perform well due to the high exposure to high-tech stocks. However, to mention again each crisis has its unique setting and in the COVID-19 crisis He et al. (2020) found that the technology sector benefitted from the crisis compared to more traditional sectors. In addition, Albuquerque et al. (2020) find that firms in the energy industry have resulted high volatilities as well as lower returns compared to other firms. This also is supported by the results of this study, where particular conventional ETFs XES and MLPA have resulted in the highest volatilities and lowest average returns during the examined phases of COVID-19. Consistent with the results of this study concerning average returns and return volatilities indicate that researched ESG ETFs' reaction to the COVID-19 crisis has been less volatile compared to conventional investments.

The second main research question was:

"Can clustering analysis identify different kinds of behavior with respect to price movements between ESG and conventional ETFs in different phases of the COVID-19 crisis?"

The research was divided into three phases, where clustering analysis was implemented: the early beginning phase, the market collapse phase, and the recovery phase. The clustering analysis implemented indeed was able to identify clusters with distinct returns and volatility during the examined phases. Results of clustering ETFs in the early beginning phase showed that the COVID-19 crisis progressed fairly smoothly at first, as indicated by the moderate level of annualized volatilities of the resulting clusters. Obtained ESG and conventional ETF clusters did not have drastic differences, however, the conventional ETF clusters have generally slightly higher volatilities and lower average returns compared to ESG clusters. Interestingly Morningstar's sustainability rating during the examined phase did not indicate that a cluster with a higher sustainability rating (ESG cluster 3: average of 4.1 globes) would perform better than a cluster that has resulted in a lower level of rating (ESG cluster 1: average of 3 globes). Folger-Laronde et al. (2020) and Hartzmark and Sussman (2019) have suggested similar conclusions, as they have found that a higher level of sustainability rating does not add resilience to investments' returns during COVID-19. However, compared to conventional ETF clusters in this study, ESG ETFs have been more stable concerning returns and volatilities during the beginning of COVID-19.

The market collapse phase resulted in fewer clusters, indicating that the price movements of ETFs moved more closely together than in the beginning phase. This phase's clusters have resulted in very high volatilities and negative annualized returns for all obtained clusters. ESG ETFs have resulted in three clusters, which can be characterized as less risky clusters compared to obtained conventional ETF clusters, in terms of returns as well as volatilities. The two clusters formed of conventional ETFs have resulted in significantly higher volatilities and lower minimum returns compared to ESG clusters. However, differences among the conventional clusters are discovered as cluster two has notably different characteristics compared to cluster one, indicating the difference between the selected conventional ETFs during the market collapse phase, possibly explained by ETFs focusing on specific industries. This study's results also support the conclusion by Pisani and Russo (2021) concerning sustainability ratings, indicating that a higher rating would slightly soften the price crash in the collapse phase. In addition, Pisani and Russo (2021), Broadstock et al. (2021) and Albuquerque et al. (2020) state that ESG funds seem to have a better reaction to an unexpected COVID-19 pandemic which can also be seen from the obtained ESG clusters.

The recovery phase resulted in the same number of clusters as in the collapse phase, indicating that the readjustment of the considered ETFs was gradual. ESG ETFs resulted in three clusters with quite similar characteristics, whereas conventional ETF clusters concerning their volatilities were notably different. However, both ESG and conventional ETFs resulted in positive average returns after the collapse phase and showed signs of market stabilization. Interestingly, MSR resulted in different values than in previous phases. The cluster with the highest MSR (average of globes 3.96) resulted in the lowest volatility while the cluster with the lowest MSR (2 globes) resulted in the highest volatility. Similarly in Pavlova and de Boyrie (2022) and Broadstock et al. (2021) studies, the highest MSR has been associated with low volatilities and a high sustainability rating has been connected with better risk tolerance measured by short-term returns. During the two previously examined phases there has been a rather minor difference between the medium/average MSR obtained cluster and high MSR obtained cluster and their financial characteristics.

Considering the cluster memberships, the transitioning of both ESG and conventional ETF members has resulted in a different number of clusters during the examined phases. For ESG ETFs, the transition between different phases has been more subtle as only the recovery phase resulted in slight adjustments to the original cluster memberships. Compared to ESG ETF clusters, conventional ETFs have had more movement concerning cluster memberships. The members of ESG ETFs clusters from the early beginning phase to the collapse phase have remained the same. The changes observed during the recovery phase compared to the two previously studied phases is that in one cluster (cluster three) there is only one member, DURA, which has the lowest MSR. As mentioned earlier, the ESG ETFs are quite technology sector-focused, which may result in similarities in their behavior during COVID-19. The transition from the early beginning phase to the collapse phase has resulted in the original number of conventional ETF clusters being reduced to half. The three original clusters have formed a new single cluster (cluster one) as their prices have moved closer together while cluster two's members have remained unchangeable. Finally, in the recovery phase cluster two, which has so far contained the same ETF members, was partly separated into cluster one.

Overall, the obtained clusters of ESG ETFs have had a less volatile reaction to the COVID-19 crisis in every examined phase compared to the conventional ETFs, excluding one conventional ETF’s cluster number four in the early beginning phase. In addition, obtained ESG clusters resulted in generally higher annualized returns compared to the conventional ETFs clusters during the early beginning phase as well as the collapse phase. In the last examined phase the average returns of conventional and ESG clusters were at quite similar levels, however, the highest return was obtained by one of the ESG ETF clusters. Clustering all ETFs and conventional ETFs has resulted in very different kinds of clusters within the examined phases compared to ESG ETFs, which have generally resulted in more similar clusters. A higher sustainability rating did not protect in a market downturn, however, ESG ETFs did not result in worse performance than conventional ETFs. Therefore, investors who seek to maximize their portfolio performance might find ESG investing as a strengthening option during uncertain market situations.

6.2 Limitations of the study and research criticality

The COVID-19 is still an ongoing pandemic; therefore, criticism should be addressed regarding the used time frame in the study. The study examines the pandemic using three different phases; however, the phases altogether cover only the first six months of the crisis. The limitations of this study consider the challenges of the time series clustering methodology and the artificial split into the three examined phases, which may have an impact on the resulting clusters.

Limitations also apply to the market area, since this study only considers the U.S. stock market, which is the largest in the world. Therefore, the obtained results cannot be directly generalized to other market areas. In addition, this study only considered the COVID-19 crisis, thus the results cannot be necessarily applied to other crises and different types of crises (e.g., financial crises).

Another limitation that needs to be taken into consideration is the size of the sample, the number of used ETFs. Using 68 ETFs can be seen as a relatively small sample, which may impact the reliability of the results and performance of the utilized machine learning algorithms. However, the ETFs used represent a total AUM of 136 Bn$. In addition, the study used only normalized prices for clustering while many studies also utilized other variables such as financial ratios. Despite that, there are also studies focusing purely on time series data concerning normalized prices, however, this might affect the clustering performance.

6.3 Suggestion for further research

Further research suggestions have emerged from this study. Firstly, it would be very interesting to extend the analysis for longer time periods in order to compare different crisis and non-crisis environments as Nofsinger and Varma (2014) and Becchetti et al. (2015) included the global financial crisis and the technology bubble burst in their study. This would add not only different aspects concerning crisis and socially responsible investing performance but their differences in different market conditions. In addition, a broader time frame and using a market benchmark would give more insights into the crises environment and the market behaviour, and further on, dividing the time frames into short and long ones as examined by Buszko, Orzeszko and Stawarz (2021) may be useful for investors' decision-making. While the researchers have traditionally proceeded with factor models focusing on strictly performance comparison, it would be interesting to alter the research with unsupervised machine learning and investigate different characteristics of obtained clusters, for example, sustainability rating relationships might bring more insights into the behaviour of clusters in different market environments.

7 References

Aghabozorgi, S., Shirkhorshidi, A. S. & Teh, Y. W. 2015. Time-series clustering – A decade review. Information Systems. Vol. 53, pp. 16–38.

Albuquerque, R., Koskinen, Y. & Zhang, C. 2019. Corporate Social Responsibility and Firm Risk: Theory and Empirical Evidence. Management Science. Vol. 65(10), pp. 4451–4469.

Albuquerque, R., Koskinen, Y., Yang, S. & Zhang, C. 2020. Resiliency of Environmental and Social Stocks : An Analysis of the Exogenous COVID-19 Market Crash, Review of Corporate Finance Studies, Vol. 9, pp. 593–621.

Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A. & Aljaaf, A. J. 2020. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science. Springer Nature Switzerland AG.

Almeida, V., Barrios, S., Christl, M., De Poli, S., Tumino, A. & van der Wielen, W. 2021. The impact of COVID-19 on households ´ income in the EU. The Journal of Economic Inequality. Vol. 19, pp. 413–431.

Alpaydin, E. 2014. Introduction to machine learning. Cambridge, MA. The MIT Press.

Alvarez, M. & Rodriguez, J. 2015. Water-related mutual funds: Investment performance and social role. Social Responsibility Journal. Vol. 11, pp. 502–512.

Bandyopadhyay, S. & Saha, S. 2013. Unsupervised Classification Similarity Measures, Classical and Metaheuristic Approaches, and Applications. Berlin, Heidelberg, Springer.

Barko, T., Cremers, M. & Renneboog, L. 2021. Shareholder Engagement on Environmental, Social, and Governance Performance. Journal of Business Ethics. Vol. 180, pp. 777-812.

Barnett, M. L. & Salomon, R. M. 2006. Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. Strategic Management Journal. Vol. 27(11), pp. 1101–1122.

Beal, D. & Goyen, M. 1998. Putting your money where your mouth is, a profile of ethical investors. Financial Services Review. Vol. 7, pp. 129–143.

Becchetti, L., Ciciretti. R., Dalo, A. & Herzel, S. 2015. Socially responsible and conventional investment funds: performance comparison and the global financial crisis, Applied Economics. Vol. 47(25), pp. 2541–2562.

Bello, Z. Y. 2005. Socially Responsible Investing and Portfolio Diversification. The Journal of Financial Research. Vol. 28 (1), pp. 41–57.

Berndt, D. & Clifford, J. 1994. Using dynamic time warping to find patterns in time series. Proceedings of the AAAI-94 Workshop, Knowledge Discovery in Databases, pp. 359–370.

Bilbao-Terol, A., Arenas-Parra, M., Canal-Fernandez, V. & Bilbao-Terol, C. 2016. Multi-criteria decision making for choosing socially responsible investment within a behavioral portfolio theory framework: a new way of investing into a crisis environment. Annals of Operations Research. Vol. 247(2), pp. 549–580.

Billio, M., Costola, M., Hristova. I., Latino, C. & Pelizzon, L. 2021. Inside the ESG ratings: (Dis)agreement and performance, Corporate Social Responsibility and Environmental Management. Vol. 28, pp. 1426–1445.

Broadstock, D. C., Chan, K., Cheng, L. & Xiaowei, W. 2021. The role of ESG performance during times of financial crisis: Evidence from COVID-19 in China, Finance Research Letters. Vol. 38.

Buszko, M., Orzeszko, W. & Stawarz, M. 2021. COVID-19 pandemic and stability of stock market — A sectoral approach. PLoS ONE. Vol. 16(5).

Camilleri, M. A. 2021. The market for socially responsible investing: a review of the developments. Social Responsibility Journal. Vol. 17(3), pp. 412–428.

Capelle-Blancard, G. & Monjon, S. 2014. The Performance of Socially Responsible Funds: Does the Screening Process Matter? European Financial Management. Vol. 20(3), pp. 494–520.

Carhart, M. 1997. On Persistence in Mutual Fund Performance. The Journal of Finance. Vol. 52(1), pp. 57-82.

Choi, J. W. 2013. The 2007–2010 U.S. financial crisis: Its origins, progressions, and solutions. The Journal of Economic Asymmetries. Vol. 10(2), pp. 65–77.

Das, N., Chatterje, S., Ruf, B. & Sunder, A. 2018. ESG Ratings and the Performance of Socially Responsible Mutual Funds: A Panel Study. Journal of Finance Issues. Vol. 17(1), pp. 49-57.

Dawkins, C. E. 2018. Elevating the Role of Divestment in Socially Responsible Investing. Journal of Business Ethics. Vol. 153(2), pp. 465–478.

Demers, E., Hendrikse, J., Joos, P. & Lev, B. 2020. ESG Didn't Immunize Stocks Against the Covid-19 crisis, but investments in intangible assets did. Journal of Business Finance & Accounting. Vol. 48(3-4), pp. 433-462.

Derwall, J., Koedijk, K. & Ter Horst, J. 2011. A tale of values-driven and profit-seeking social investors. Journal of Banking and Finance. Vol. 35(8), pp. 2137–2147.

Dixon, M. F., Halperin, I. & Bilokon, P. 2020. Machine Learning in Finance: From Theory to Practice. Springer International Publishing.

Dougherty, G. 2013. Pattern Recognition and Classification. New York, Springer Science Business Media.

Dragut, A. B. 2012. Stock Data Clustering and Multiscale Trend Detection, Methodology and Computing in Applied Probability. Vol. 14(1), pp. 87–105.

Drucker, P.F. 1984. Converting Social Problems into Business Opportunities: The New Meaning of Corporate Social Responsibility. California Management Review. Vol. 26(2), pp. 53-63.

Durán-Rosal, A. M., de la Paz-Marin, M., Gutierrez, P. A. & Hervas-Martinez, C. 2017. Identifying Market Behaviours Using European Stock Index Time Series by a Hybrid Segmentation Algorithm. Neural Processing Letters. Vol. 46(3), pp. 767–790.

D'Urso, P., De Giovanni, L. & Massari, R. 2021. Trimmed fuzzy clustering of financial time series based on dynamic time warping. Annals of Operations Research. Vol. 299(1), pp. 1379–1395.

ETF.com. 2022. Screener & Database. [www-page].[Accessed 25.5.2022] Available at: https://www.etf.com/etfanalytics/etf-finder/?sfilters=eyJ0aWNrZXIiOnsiaW4iOltdfX0=

ETFdatabase. 2022. Socially responsible ETF List. [www-page].[Accessed 26.5.2022] Available at: https://etfdb.com/esg-investing/social-issues/#etfs__returns&sort_name=fiftytwo_week_return&sort_order=desc&page=8

Fama, E. F. 1971. Risk, return, and equilibrium. Journal of Political Economy. Vol. 79(1), pp. 30–55.

Ferriani, F. & Natoli, F. 2020. ESG risks in times of Covid-19. Applied Economics Letters. Vol. 28(18), pp. 1537-1541.

Fidelity. 2022. What's in your ETF? [www-page].[Accessed 12.7.2022] Available at: https://www.fidelity.com/learning-center/investment-products/etf/whats-in-your-etf

Folger-Laronde, Z., Pashang, S., Feor, L. & ElAlfy, A. 2020. ESG ratings and financial performance of exchange-traded funds during the COVID-19 pandemic funds during the COVID-19 pandemic. Journal of Sustainable Finance & Investment. Vol. 12(2), pp 490-496.

Freeman, R., E. 1984. Strategic management: a stakeholder approach. Boston, MA, Pitman.

Gagolewski, M., Bartoszuk, M. & Cena, A. 2021. Are cluster validity measures (in)valid? Information Sciences. Vol. 501, pp. 620–636.

Graves, S., B. & Waddock, S., A. 2000. Beyond built to last…stakeholder relations in built-to-last companies. Business and Society Review. Vol. 105(4), pp. 393–418.

Grewenig, E., Lergetporer, P., Werner, K., Woessmann. & Zierow, L. 2021. COVID-19 and educational inequality: How school closures affect low- and high-achieving students. European Economic Review. Vol. 140, pp. 103920.

Großwendt, A., Röglin, H. & Schmidt, M. 2019. Analysis of Wards Method. Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 2939–2957.

GSIR. 2020. Global Sustainable Investment Review 2020. [Online document].[Accessed 13.4.2022] Available at: http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf

Gubu, L., Rosadi, D. & Abdurakhman. 2019. Classical portfolio selection with cluster analysis: Comparison between hierarchical complete linkage and Ward algorithm. AIP Conference Proceedings. Vol. 2192(1).

Halkidi, M., Batistakis, Y. & Vazirgiannis, M. 2001. On Clustering Validation Techniques. Journal of Intelligent Information Systems. Vol. 17, pp. 107–145.

Han, J., Pei, J. & Kamber, M. 2012. Data mining concepts and techniques. 3rd ed. Burlington, MA, Morgan Kaufmann.

Haroon, O. & Rizvi, S. A. R. 2020. COVID-19: Media coverage and financial markets behavior - A sectoral inquiry. Journal of behavioral and experimental finance. Vol. 27.

Hartzmark, S.M. & Sussman, A.B. 2019. Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. The Journal of Finance. Vol. 74 (6), pp. 2789–2837.

Hassan, S. M. & Riveros Gavilanes, J. M. 2021. First to React Is the Last to Forgive: Evidence from the Stock Market Impact of COVID 19. Journal of Risk and Financial Management. Vol. 14(1), pp. 1-25.

He, H., Chen, J., Jin, H., Chen, SH. 2007. Trading Strategies Based on K-means Clustering and Regression Models. In: Chen, SH., Wang, P.P., Kuo, TW. (eds) Computational Intelligence in Economics and Finance. Berlin, Springer.

He, P., Sun, Y., Zhang, Y. & Li, T. 2020. COVID – 19' s Impact on Stock Prices Across Different Sectors — An Event Study Based on the Chinese Stock Market. Emerging Markets Finance and Trade. Vol. 56(10), pp. 2198–2212.

Henrique, B. M., Sobreiro, V. A. and Kimura, H. 2019. Literature review : Machine learning techniques applied to financial market prediction R, Expert Systems With Applications. Vol. 124, pp. 226–251.

Hoepner, A. G. F. 2010. Portfolio diversification and environmental, social or governance criteria: Must responsible investments really be poorly diversified? Social Science Research Network Working (SSRN) papers.

Hyrske, A., Lönnroth, M., Savilaakso, A. & Sievänen, R. 2012. Vastuullinen sijoittaminen. Tampere, Tammerprint Oy.

Iraci, J. A. & Iraci, A. J. 2020. The Case For Business Resiliency. The RMA Journal. Vol. 102(9), pp. 32–37.

Irani, J., Phatak, M. & Pise, N. 2016. Clustering Techniques and the Similarity Measures used in Clustering: A Survey, International Journal of Computer Applications, Vol. 134(7), pp. 9-14.

Ivanisevic Hernaus, A. 2019. Exploring the strategic variety of socially responsible investment financial performance insights about SRI strategy portfolios. Sustainability accounting, management and policy journal. Vol. 10(3), pp. 545–569.

Izakian, H., Pedrycz, W. & Jamal, I. 2015. Engineering Applications of Artificial Intelligence Fuzzy clustering of time series data using dynamic time warping distance, Engineering Applications of Artificial Intelligence. Vol. 39, pp. 235–244.

James, N. & Menzies, M. 2021. Association between COVID-19 cases and international equity indices. Physica D. Vol. 417.

Javed, A., Lee, B. S. & Rizzo, D. M. 2020. A benchmark study on time series clustering. Machine Learning with Applications. Vol. 1, 100001.

Javed, B., Sarwer, A., Soto, E. B. & Mashwani Z. 2020. The coronavirus (COVID-19) pandemic's impact on mental health. International Journal of Health Planning and Management. Vol. 35(5), pp. 993–996.

Jobst, A. 2008. What Is Securitization? Finance & Development. Vol. 45(3), pp. 48-49.

Jung, A. 2022. Machine learning: the basics. 1st ed. Singapore, Springer.

justETF. 2022. Exchange-traded funds. [Online document].[Accessed 18.5.2022] Available at: https://www.justetf.com/en/academy/what-is-an-etf.html

Kanuri, S. 2020. Risk and Return Characteristics of Environmental, Social, and Governance (ESG) Equity ETFs. Journal of Index Investing. Vol. 11(2), pp. 66–76.

Kaufman, L. & Rousseeuw, P. J. 2005. Finding Groups In Data: an Introduction to Cluster Analysis. Hoboken, NJ, Wiley.

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

Keogh, E. & Lin, J. 2005. Clustering of time-series subsequences is meaningless: implications for previous and future research. Knowledge and Information Systems. Vol. 8, pp. 154–177.

Kinder, P. D. 2005. Socially responsible investing: An Evolving Concept in Changing World. KLD Research & Analytics Inc.

King, L. & Gish, E. 2015. Marketizing Social Change: Social Shareholder Activism and Responsible Investing. Sociological Perspectives. Vol. 58(4), pp. 711–730.

King, R. S. 2015. Cluster Analysis and Data Mining: An Introduction. Dulles, VA, Mercury Learning and Information.

Koutroumbas, K. & Theodoridis, S. 2009. Pattern Recognition. 4th ed. Burlington, MA, Elsevier.

Kurtz, L. 1997. No effect or no net effect? Studies on socially responsible investing. Journal of Investing. Vol. 6 (4), pp. 37–49.

Lean, H. H. & Pizzutilo, F. 2021. Performances and risk of socially responsible investments across regions during crisis. International journal of finance and economics. Vol. 26, pp. 3556–3568.

Lee, D. D., Humphrey, J. E., Benson, K. L. & Ahn, J. Y. K. 2010. Socially responsible investment fund performance: the impact of screening intensity. Accounting and Finance. Vol. 50(2), pp. 351–370.

Leite, P. & Cortez, M. 2015. International Review of Financial Analysis Performance of European socially responsible funds during market crises: Evidence from France. International Review of Financial Analysis. Vol. 40, pp. 132–141.

Lewis, A. & Mackenzie, C. 2000a. Support for investor activism among U.K. Ethical investors. Journal of Business Ethics. Vol. 24(3), pp. 215–222.

Lewis, A. & Mackenzie, C. 2000b. Morals, money, ethical investing and economic psychology. Human Relations. Vol. 53(2), pp. 179–191.

Liao, W. T. 2005. Clustering of time series data — a survey, Pattern Recognition. Vol. 38, pp. 1857–1874.

Łuczak, M. 2018. Combining raw and normalized data in multivariate time series classification with dynamic time warping. Journal of Intelligent and Fuzzy Systems: Applications in Engineering and Technology. Vol. 34(1), pp. 373–380.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1, pp. 281–297.

Madhulatha, T. S. 2012. An Overview on clustering methods. IOSR Journal of Engineering. Vol. 2(4), pp. 719–725.

Mahesh, B. 2020. Machine Learning Algorithms - A Review. International Journal of Science and Research. Vol. 9(1), pp. 381-386.

Markowitz, H. 1952. Portfolio selection. Journal of Finance. Vol. 7(1), pp. 77–91.

Marti-Ballester, C-P. 2020. Financial Performance of SDG Mutual Funds Focused on Biotechnology and Healthcare Sectors. Sustainability. Vol. 12(5), pp. 1-15.

Mazur, M., Dang, M. & Vega, M. 2021. COVID-19 and the march 2020 stock market crash. Evidence from S & P1500. Finance Research Letters. Vol. 38.

Meziani, A. S. 2016. Exchange-Traded Funds Investment Practises and Tactical Approaches. London, Palgrave Macmillan.

Mierau, J. O. & Mink, M. 2013. Are stock market crises contagious? The role of crisis definitions. Journal of Banking and Finance. Vol. 37(12), pp. 4765–4776.

Milligan, G. W. & Cooper, M. C. 1985. An examination of procedures for determining the number of clusters in a data set. Psychometrika. Vol. 50, pp. 159–179.

Morningstar. 2021. Morningstar Sustainability Rating Methodology. [Online document]. [Accessed 1.6.2022] Available at: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf

Morningstar. 2022. ETF screener. [Online document].[Accessed 27.5.2022] Available at: https://www.morningstar.com/etfs/screener

MSCI. 2022. ESG Rating. [www-page].[Accessed 27.5.2022] Available at: https://www.msci.com/our-solutions/esg-investing/esg-ratings

Murtagh, F. & Legendre, P. 2014. Wards Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Wards Criterion? Journal of Classification. Vol. 31(3), pp. 274-295.

Murty, M. N. & Devi, V. S. 2015. Introduction to pattern recognition and machine learning. Singapore, World Scientific Publishing Co Pte Ltd.

Naffa, H. & Fain, M. 2020. Performance measurement of ESG-themed megatrend investments in global equity markets using pure factor portfolios methodology. PLoS ONE. Vol. 15(12), pp. 1–35.

Nakagawa, K., Imamura, M. & Yoshida, K. 2019. Stock price prediction using k -medoids clustering with indexing. IEEJ Transactions on Electronics, Information and Systems. Vol. 138(8), pp. 986–991.

Nanda, S. R., Mahanty, B. & Tiwari, M. K. 2010. Clustering Indian stock market data for portfolio management. Expert Systems with Applications. Vol. 37(12), pp. 8793–8798.

Nayak, J., Mishra, M., Naik, B., Swapnarekha, H., Cengiz, K. & Shanmuganathan, V. 2022. An impact study of COVID-19 on six different industries: Automobile, energy and power, agriculture, education, travel and tourism and consumer electronics. Expert Systems. Vol. 39(3).

Nilsson, J. 2009. Segmenting socially responsible mutual fund investors: The influence of financial return and social responsibility. International Journal of Bank Marketing, Vol. 27(1), pp. 5–31.

Nofsinger, J. & Varma, A. 2014. Socially responsible funds and market crises. Journal of Banking and Finance. Vol. 48, pp. 180–193.

Ofek, E. & Richardson, M. 2003. DotCom Mania: The Rise and Fall of Internet Stock Prices. The Journal of Finance. LVIII(3).

Oh, H. C., Park, J-H. & Ghauri, P. N. 2013. Doing right, investing right: Socially responsible investing and shareholder activism in the financial sector. Business Horizons. Vol. 56, pp. 703–714.

Omura, A., Roca, E. & Nakai, M. 2021. Does responsible investing pay during economic downturns: Evidence from the COVID-19 pandemic. Finance Research Letters. Vol. 42.

Ortas, E., Moneva, J. M., Burritt, R. & Tingey-Holyoak, J. 2014. Does Sustainability Investment Provide Adaptive Resilience to Ethical Investors? Evidence from Spain, Journal of Business Ethics. Vol. 124, pp. 297–309

Paparrizos, J. & Gravano, L. 2017. Fast and Accurate Time-Series Clustering. ACM Transactions on Database Systems. Vol. 42(2), pp. 1-49.

Pastor, L. & Vorsatz, B. M. 2020. Mutual Fund Performance and Flows during the Covid-19 Crisis. The Review of Asset Pricing Studies. Vol. 10(4), pp. 791–833.

Pava, M. L. 2007. A Response to "Getting to the Bottom of Triple Bottom Line". Business Ethics Quarterly. Vol. 17(1), pp. 105-110.

Pavlova, I. & de Boyrie, M. E. 2022. ESG ETFs and the COVID-19 stock market crash of 2020: Did clean funds fare better? Finance Research Letters. Vol. 44(1).

Pisani, F. & Russo, G. 2021. Sustainable finance and covid-19: The reaction of esg funds to the 2020 crisis. Sustainability. Vol. 13(23), pp. 1-18.

PRI Association. 2022. Annual report 2021-2022. [Online document].[Accessed 12.6.2022] Available at: https://dwtyzx6upklss.cloudfront.net/Uploads/b/f/m/pri_annual_report_2022_689047.pdf

PRI Association. 2020. Sustainable and inclusive: Covid-19 recovery and reform. [Online document].[Accessed 10.6.2022] Available at: https://www.unpri.org/covid-19-resources/sustainable-and-inclusive-covid-19-recovery-and-reform/6016.article

Puaschunder, J. M. 2016. On the Emergence, Current State and Future Perspectives of Socially Responsible Investment (SRI). Consilience: The Journal of Sustainable Development. Vol. 16(1), pp. 38–63.

Renneboog, L., Ter Horst, J. & Zhang, C. 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. Journal of Banking and Finance. Vol. 32(9), pp. 1723–1742.

Revelli, C. 2017. Socially responsible investing (SRI): From mainstream to margin? Research in International Business and Finance. Vol. 39, pp. 711–717.

Rodriguez, J. & Romero, H. 2019. Exchange-Traded Funds as an International Diversification Tool for Socially Responsible Investors. The Journal of Wealth Management. Vol. 22(3), pp. 98–103.

Rodriguez, M. Z., Comin, H, C., Casanova, D., Bruno, O. M., Amancio, D. R., Costa, L. & Rodrigues, F. A. 2019. Clustering algorithms: A comparative approach. PLoS ONE. Vol. 14(1), pp. 1–34.

Rokach, L. & Maimon, O. 2005. Clustering methods, Data Mining and Knowledge Discovery Handbook. Boston, MA, Springer.

Rompotis, G. G. 2013. Actively vs. Passively Managed Exchange Traded Funds. The IEB International Journal of Finance. Vol. 6, pp. 116–135.

Sandberg, J., Juravle, C., Hedesström, T. M. & Hamilton, I. 2009. The Heterogeneity of Socially Responsible Investment. Journal of Business Ethics. Vol. 87(4), pp. 519–533.

Saxena, A., Prasad, M., Gupta. A., Bharill, N., Patel, O. P., Tiwari, A., Er, M. J., Ding, W. & Lin, C.T. 2017. A review of clustering techniques and developments. Neurocomputing. Vol. 267, pp. 664–681.

Schueth, S. 2003. Socially Responsible Investing in the United States. Journal of Business Ethics. Vol. 43(3), pp. 189–194.

Sharpe, W. 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance. Vol. 19, No. 3, p. 425-442.

Shu, M., Song, R. & Zhu, W. 2021. The "COVID " crash of the 2020 U.S. Stock market. North American Journal of Economics and Finance. Vol. 58.

Silvola, H. & Landau, T. (2021) Sustainable investing: beating the market with ESG. Switzerland, Palgrave Macmillan.

Singh, A. 2020. COVID-19 and safer investment bets. Finance Research Letters. Vol. 36.

Statista. 2020a. Total assets under management (AUM) by exchange traded funds (ETFs) worldwide in 2020, by region. [Online document].[Accessed 18.5.2022] Available at: https://www.statista.com/statistics/1196132/global-etf-assets-under-management-region/

Statista. 2022b. Global ESG ETF assets 2006-2022. [Online document].[Accessed 19.5.2022] Available at: https://www.statista.com/statistics/1297487/assets-of-esg-etfs-worldwide/

Statman, M. & Glushkov, D. 2009. The wages of social responsibility. Financial Analysts Journal. Vol. 65(4), pp. 33–46.

Stephens, C., Benink, H. A., Gordillo, H. L. & Pardo-Guerra, J. P. 2021. A New Measure of Market Inefficiency. Journal of Risk and Financial Management. Vol. 14(6), pp. 1-22.

Sutton, R. S. & Barto, A. G. 2018. Reinforcement Learning: An introduction. 2nd. ed. Cambridge, MA, The MIT Press.

Syed, A. M. 2017 Socially responsible: Are they profitable? Research in International Business and Finance. Vol. 42, pp. 1504–1515.

Takahashi, H. & Yamada, K. 2021. When the Japanese stock market meets COVID-19: Impact of ownership, China and US exposure, and ESG channels. International Review of Financial Analysis. Vol. 74, p. 101670.

Thorbecke, W. 2020. The Impact of the COVID-19 Pandemic on the U.S. Economy: Evidence from the Stock Market. Journal of Risk and Financial Management. Vol. 13(10), pp. 1–32.

US SIF (2022a) The Forum for Sustainable and Responsible Investing. [Online document]. [Accessed 12.4.2022] Available at: www.ussif.org

Verheyden, T., Eccles, R. G. & Feiner, A. 2016. ESG for all? The impact of ESG Screening on Return, Risk and Diversification. Journal of Applied Corporate Finance. Vol. 28(6), pp. 47-55.

von Wallis, M. & Klein, C. 2015. Ethical requirement and financial interest: a literature review on socially responsible investing. Business Research. Vol. 8, pp. 61–98.

Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P. & Keogh, E. 2013. Experimental comparison of representation methods and distance measures for time series data. Data Mining and Knowledge Discovery. Vol. 26(2), pp. 275–309.

Wang, Y., Huang, L., Zheng, Y., Lee, S. & Fu, Y. 2021. Stock Price Trend Analysis Method Based on Change Cycle Clustering. Academic Journal of Computing & Information Science. Vol. 4(8), pp. 42–47.

Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association. Vol. 58(301), pp. 236–244.

Wicks, A. C., Berman, S. L., & Jones, T. M. 1999. The structure of optimal trust: moral and strategic implications. Academy of Management Review. Vol. 24(1), pp. 99-116.

World Health Organization. 2020. WHO Director-Generals opening remarks at the media briefing on COVID-19-11March 2020. [Online document].[Accessed 14.5.2022] Available at: https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020

World Health Organization. 2022. COVID-19 Weekly Epidemiological Update. [Online document][Accessed 14.5.2022] Available at: www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---11-may-2022

Zhang, D., Hu, M. & Ji, Q. 2020. Financial markets under the global pandemic of COVID-19. Finance Research Letters. Vol. 36.

Zhang, Y., Zhao, H., Li, X., Gao, S., Kong, K. & Chen, Y. 2020. Exchange Traded Fund Clustering via Metric Learning. IEEE International Conference on Big Data (Big Data).

Zolhavarieh, S., Aghabozorgi, S. & Teh, Y. W. 2014. A Review of Subsequence Time Series Clustering. The Scientific World Journal. Vol. 2014.

Appendices

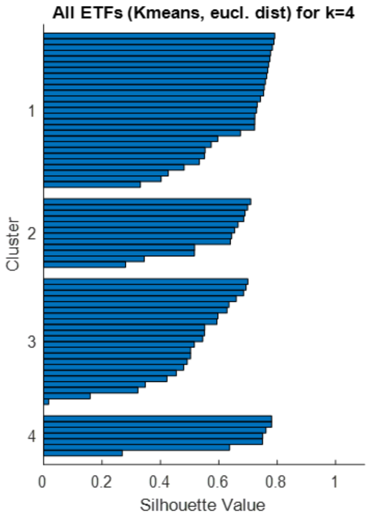
Appendix 1. List of used ESG ETFs

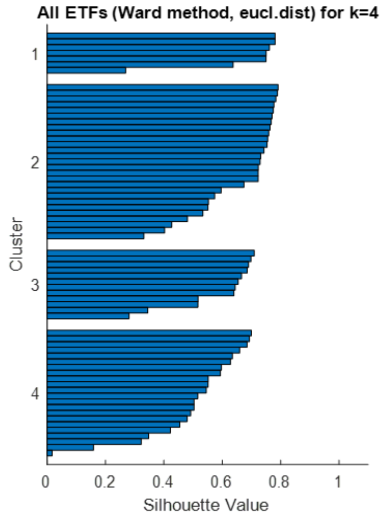
Ticker	ESG ETF	Morningstar Sustainability Rating	Inception	AUM $
BOSS	GLOBAL X FOUNDER- RUN COMPANIES ETF	3	13.2.2017	9 690 000
NULC	NUVEEN ESG LARGE- CAP ETF	5	3.6.2019	25 880 000
WOMN	IMPACT SHARES YWCA WOMENS EMPOWERMENT ETF	4	24.8.2018	32 000 000
NACP	IMPACT SHARES NAACP MINORITY EMPOWERMENT	4	18.7.2018	38 520 000
DURA	VANECK MORNINGSTAR DURABLE DIVIDEND ETF	2	30.10.2018	65 660 000
VEGN	US VEGAN CLIMATE ETF	5	30.10.2018	66 200 000
ACSI	AMERICAN CUSTOMER SATISFACTION ETF	3	1.11.2016	70 810 000
CHGX	AXS CHANGE FINANCE ESG ETF	5	10.10.2017	106 490 000
TDV	S&P TECHNOLOGY DIVIDEND ARISTOCRATS ETF	4	5.11.2019	110 000 000
ISMD	INSPIRE SMALL/MID CAP ESG ETF	1	28.2.2017	116 360 000
SPUS	SP FUND.SP5.SHARIA IND. EXCLUSIONS ETF	3	18.12.2019	146 300 000
QLV	FLEXSHARES US QUALITY LOW VOLATILITY IND. FUND	4	15.7.2019	160 800 000
ETHO	ETHO CLIMATE LEADERSHIP US ETF	5	18.11.2015	163 120 000
HLAL	WAHED FTSE USA SHARIAH ETF	3	16.7.2019	164 100 000
AIQ	GLOBAL X ARTIFICIAL INTELLIGENCE & TECHNOLOGY	4	11.5.2018	169 100 000
PAWZ	PROSHARES PET CARE ETF	3	5.11.2018	188 100 000
SHE	SPDR SSGA GENDER DIVERSITY INDEX ETF	4	7.3.2016	225 090 000
BIBL	INSPIRE 100 ESG ETF	3	30.10.2017	278 470 000
IQSU	IQ CANDRIAM ESG US EQUITY ETF	5	17.12.2019	416 030 000
CATH	GLOBAL X S&P 500 CATHOLIC VALUES ETF	3	18.4.2016	560 600 000
KRMA	GLOBAL X CONSCIOUS COMPANIES ETF	4	11.7.2016	651 390 000
WCLD	WISDOMTREE CLOUD COMPUTING ETF	3	6.10.2019	759 000 000
SNPE	XTRACKERS S&P 500 ESG ETF	4	26.7.2019	794 560 000
CLOU	GLOBAL X CLOUD COMPUTING ETF	3	12.4.2019	814 900 000
ESML	ISHARES ESG AWARE MSCI USA SMALL-CAP ETF	4	10.4.2018	1 500 000 000
PHO	INVESCO WATER RESOURCES ETF	3	6.12.2005	1 650 000 000
SUSL	ISHARES ESG MSCI USA LEADERS ETF	5	7.5.2019	3 496 300 000
SUSA	ISHARES MSCI USA ESG SELECT ETF	5	24.1.2005	3 600 000 000
DSI	ISHARES MSCI KLD 400 SOCIAL ETF	5	14.11.2006	3 830 000 000
ESGV	VANGUARD ESG US STOCK ETF	4	18.9.2018	5 910 000 000
ESGU	ISHARES ESG AWARE MSCI USA ETF	4	1.12.2016	22 950 000 000

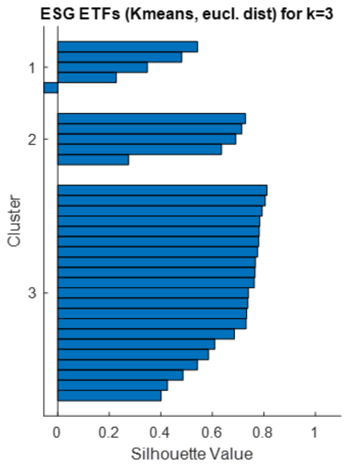
Appendix 2. List of used conventional ETFs

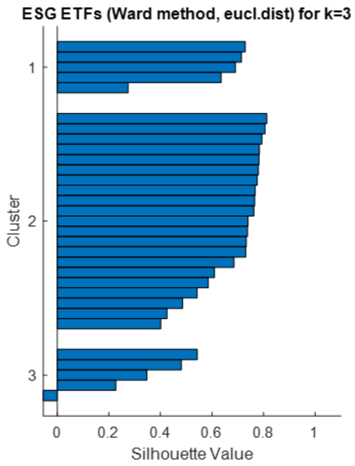
Ticker	Conventional ETFs	Inception	AUM $
QABA	FIRST TRUST NASDAQ ABA COMMUNITY BANK INDEX	29.6.2009	100 760 000
RZG	INVESCO S&P SMALLCAP 600 PURE GROWTH ETF	1.3.2006	109 560 000
EWMC	INVESCO S&P MIDCAP 400 EQUAL WEIGHT ETF	18.12.2010	120 660 000
XITK	SPDR FACTSET INNOVATIVE TECHNOLOGY ETF	13.1.2016	144 400 000
OMFS	INVESCO RUSSELL 2000 DYNAMIC MULTIFACTOR ETF	8.11.2017	148 650 000
DEUS	XTRACKERS RUSSELL U.S. MULTIFACTOR ETF	24.11.2015	153 190 000
FDM	FIRST TRUST DOW JONES SELECT MICROCAP INDEX	27.9.2005	157 040 000
UWM	PROSHARES ULTRA RUSSELL2000	23.1.2007	184 730 000
SMLV	SPDR SSGA US SMALL CAP LOW VOLATILITY INDEX ETF	20.2.2013	199 420 000
FYC	FIRST TRUST SMALL CAP GROWTH ALPHADEX FUND	19.4.2011	208 480 000
JPSE	JPMORGAN DIVERSIFIED RETURN U.S. SMALL CAP EQUITY ETF	15.11.2016	217 770 000
CSML	IQ CHAIKIN U.S. SMALL CAP ETF	16.5.2017	220 310 000
IBUY	AMPLIFY ONLINE RETAIL ETF	20.4.2016	264 930 000
DPST	DIREXION DAILY REGIONAL BANKS BULL 3X SHARES	19.8.2015	265 240 000
KBWY	INVESCO KBW PREMIUM YIELD EQUITY REIT ETF	2.12.2010	297 760 000
RZV	INVESCO S&P SMALLCAP 600 PURE VALUE ETF	1.3.2006	303 770 000
XES	SPDR S&P OIL & GAS EQUIPMENT & SERVICES ETF	19.6.2006	383 130 000
DWAS	INVESCO DWA SMALLCAP MOMENTUM ETF	19.7.2012	387 520 000
GSSC	GOLDMAN SACHS ACTIVEBETA US SMALL CAP EQUITY	7.4.2010	404 230 000
PSCT	INVESCO S&P SMALLCAP INFORMATION TECHNOLOGY ETF	28.6.2017	435 380 000
IHE	ISHARES U.S. PHARMACEUTICALS ETF	1.5.2006	439 330 000
KIE	SPDR S&P INSURANCE ETF	8.11.2005	470 940 000
FCOM	FIDELITY MSCI COMMUNICATION SERVICES INDEX ETF	21.10.2013	600 470 000
EES	WISDOMTREE U.S. SMALLCAP FUND	23.2.2007	607 690 000
EPS	WISDOMTREE U.S. LARGECAP FUND	23.2.2007	659 790 000
GSEW	GOLDMAN SACHS EQUAL WEIGHT U.S. LARGE CAP EQUITY ET	12.9.2017	674 960 000
XSVM	INVESCO S&P SMALLCAP VALUE WITH MOMENTUM ETF	3.3.2005	703 870 000
PFM	INVESCO DIVIDEND ACHIEVERS ETF	15.9.2005	710 050 000
XSLV	INVESCO S&P SMALLCAP LOW VOLATILITY ETF	15.2.2013	759 960 000
FYX	FIRST TRUST SMALL CAP CORE ALPHADEX FUND	8.5.2007	852 560 000
VTWV	VANGUARD RUSSELL 2000 VALUE ETF	20.9.2010	905 050 000
MLPA	GLOBAL X MLP ETF	18.4.2012	1 320 000 000
PEJ	INVESCO DYNAMIC LEISURE & ENTERTAINMENT ETF	23.6.2005	1 320 000 000
PRFZ	INVESCO FTSE RAFI US 1500 SMALL-MID ETF	20.9.2006	1 900 000 000
VOX	VANGUARD COMMUNICATION SERVICES ETF	23.9.2004	3 180 000 000
KRE	SPDR S&P REGIONAL BANKING ETF	19.6.2006	3 820 000 000
VIG	VANGUARD DIVIDEND APPRECIATION ETF	21.4.2006	63 380 000 000

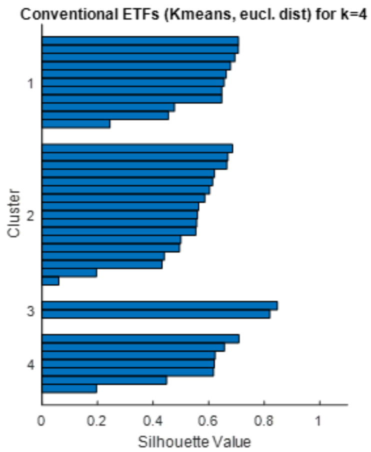
Appendix 3. The silhouette values of the early beginning phase

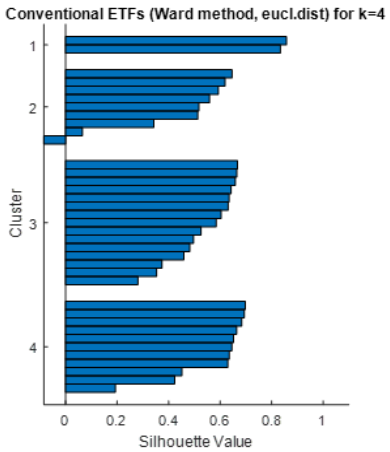




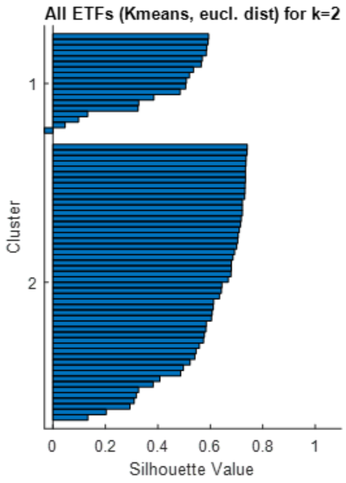


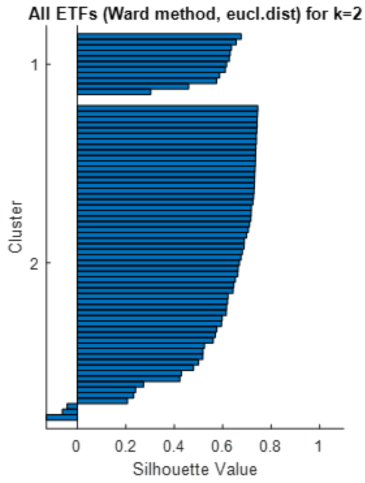


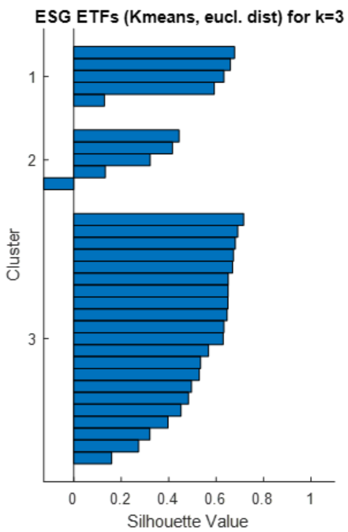


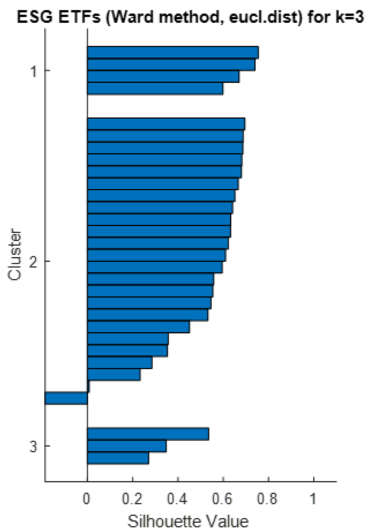


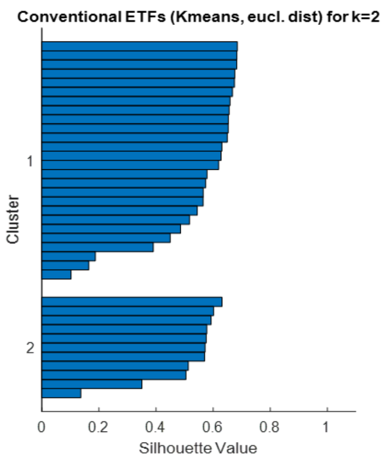
Appendix 4. The silhouette values of the collapse phase

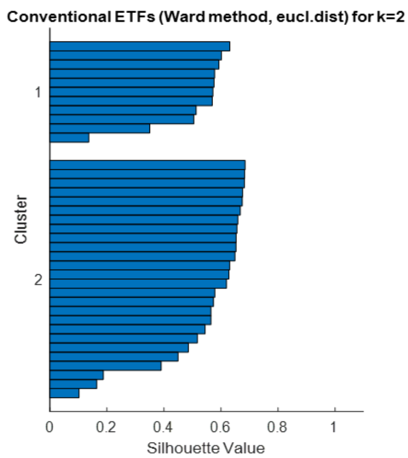












Appendix 5. The silhouette values of the recovery phase