BI Norwegian Business School - campus Oslo

GRA 19703

Master Thesis

Thesis Master of Science

Is the Philosopher's Stone Green? An Event Study on Market Reactions to Green Initiatives in Listed Companies

Navn: Jesper Nordgaard Hembre, Chrisander Kollen Engen

Start: 15.01.2021 09.00

Finish: 01.07.2021 12.00

Is the Philosopher's Stone Green? An Event Study on Market Reactions to Green Initiatives in Listed Companies

Master Thesis

by
Jesper N. Hembre and Chrisander K. Engen
MSc in Finance

Supervisor
Samuli Knüpfer

Oslo, 1 July 2021

ABSTRACT

This thesis aims to investigate whether shifts in ESG can be driven by financial incentives rather than sustainability itself by employing an event study methodology. We first analyse abnormal trading volumes and find that the shifts carry a significant weight. Next, we investigate if the shifts carry any price impact by analysing abnormal returns. We find that only the more material shifts carry a price impact from the abnormal return analysis. Lastly, we find that companies are improving their sustainable performance by analysing the development in ESG ratings and pillar scores following the shifts. We conclude that shifts in ESG are not driven by the hypothesised financial incentives, and that companies are improving their sustainable performance after the announced shift.

Acknowledgements

We would like to thank NFI Professor of Finance, Samuli Knüpfer, for his constructive feedback and suggestions throughout our thesis process. In addition to the support, he inspired us to conduct our research in the field of sustainable finance. This inspiration will also be taken with us into our working lifes. Further, we would also like to thank family and friends that have provided us with continuous support throughout our Masters degree. We never would have been in the position we are in today without your support.

Contents

List of Abbreviations

$\overline{AR}$	Average Abnormal Return
$\overline{AV}$	Average Abnormal Trading Volume
APT	Arbitrage Pricing Theory
AR	Abnormal Returns
AV	Abnormal Trading Volume
BLUE	Best Linear Unbiased Estimators
$\overline{CAR}$	Cumulative Average Abnormal Return
CAPM	Capital Asset Pricing Model
CAR	Cumulative Abnormal Returns
CFP	Corporate Financial Performance
CSP	Corporate Social Performance
CSR	Corporate Social Responsibility
EMH	Efficient Market Hypothesis
ESG	Environmental, Social, and Governance
FF3	Fama-French Three Factor Model
KPI	Key Performance Indicators
MM	Market Model
OLS	Ordinary Least Squares

List of Tables

List of Figures

1 Introduction

In recent years, the importance of Environmental, Social, and Governance (ESG) among publicly listed companies has substantially increased. International treaties like the Paris Agreement and the forthcoming European Union Taxonomy Regulation are both causing material changes in corporate behaviour. Corporations are gradually distancing from the famous Friedman doctrine (1970): *The Social Responsibility of Business is to Increase Its Profits*, and instead give precedence to all of its stakeholders (Business Roundtable, 2019). Differences in ESG have shown to affect the investor base of companies, and subsequently key factors like the cost of capital. Hence, corporations that are classified as sustainable might be able to raise more favourable financing as a result of a wider investor base. This hypothesised incentive leads up to the question: do corporations exploit the benefits of sustainability by catering to investor sentiment[1], as supported in Baker & Wurgler (2007)? We will investigate this question by analysing stock market reactions to sustainable initiatives, in a similar manner to Cooper et al. (2000) on the "dotcom" effect in listed companies.

1.1 Motivation

The motivation behind this thesis is to add valuable insights to the ever-expanding research revolved around ESG within corporations. We attempt to provide insight into the underlying incentives of sustainable initiatives, instead of solely focusing on the link between ESG and asset prices. This is done in order to investigate whether or not the companies joining the *"green wave"* are attempting to capture benefits from doing so. Empirical evidence suggests that such benefits exist, like the aforementioned effect on cost of capital. We will employ an event study methodology in order to investigate if two types of

[1] Catering to investor sentiment is the process of companies adapting to specific investor preferences in order to appear more attractive.

ESG-related shifts generate abnormal returns following the announcement. If we find evidence of positive abnormal returns, we argue that this can work as a financial incentive for the companies in question. This hypothesised incentive is what inspired our symbolic link to the philosopher's stone: can corporations create value by going green, similar to the philosopher's stone that could turn any base metal into gold?

We will also investigate both abnormal trading volumes and subsequent ESG ratings following the shift. This analysis is done in order to investigate whether the shifts generates market reactions, in addition to see if the companies actually improve their ESG ratings. We consider this to be an important aspect of the study, as companies could both capture financial gains from the shift while also improving their performance related to sustainability. We deem this to be a "win-win" scenario, where corporations can capture benefits from reducing their negative externalities.

However, we hypothesise that a "win-win" scenario is unlikely, and that companies still mainly focus on maximising value for shareholders. Thus, companies could attempt to convey a message of being sustainable while not actually improving their sustainable performance if an underlying financial incentive exists. We argue that such behavior can be deemed as a form of large scale "greenwashing"[2], where companies attempt to convey a false impression of sustainability in order to cater to investor preferences.

Such behaviour amongst companies can lead to investors being misguided by ESG initiatives, which could directly reduce the positive effects of sustainable investing. We believe this notion is what drives the importance of this study, considering that sustainable investing can help to create positive externalities to the society as a whole.

[2]Greenwashing is the process of conveying a false impression or providing misleading information about how a company's products are more environmentally sound.

2 Literature Review

We have observed a greater emphasis on companies taking more responsibility for their environmental and societal externalities over the past decade. This responsibility is referred to as Corporate Social Responsibility (CSR), which is the predecessor of the more recent term ESG. ESG is simplified to be CSR plus the corporate governance performance (Gerard, 2019). The research in sustainable finance has significantly increased over the past decade, with over 12,000 released articles in Web of Science when searching for ESG and CSR.

To understand why greenwashing is incentivised among companies, we must first understand how greenwashing is made possible in the first place. Greenwashing is directly encouraged by limited and imperfect information about firm environmental performance, and further reinforced by limitations of regulatory punishment (Delmas & Cuerel Burbano, 2011). The authors acknowledge our concern of the consequences of misleading information, in terms of maintaining investor confidence and supporting environmentally friendly behaviour across firms.

We will discuss the initial motivation behind our research, before focusing on how the current literature share different views on the relation between sustainability and firm valuation. The event study conducted by Cooper et al. (2000) investigates how corporate name changes to Internet related dot-com names generated abnormal returns in subsequent periods. This study documents a striking positive stock price reaction to the announcement of a corporate name change. The more interesting is that the effect is similar across all firms - regardless of the firm's level of involvement with the Internet. The conclusion of Cooper et al. (2000) is that a mere association with the Internet seems enough to provide a firm with a large and permanent value increase. This finding has been the main motivation behind conducting our research, and particularly the inspiration behind our rebranding sample. However, we have found contradicting results on corporate name changes in earlier studies by Bosch & Hirschey (1989) and Karpoff & Rankine (1994). Both studies finds insignificant abnormal returns around the announcement date, with a positive pre-announcement drift followed by a negative post-announcement drift. All three studies are in line with our research where we apply a similar methodology for investigating both the valuation effect of a sustainable rebranding, but also the value impact of improvements in the sustainability reporting framework.

Recent literature argues in favour of sustainability in terms of a firm's cost of capital, its investor base, and the litigation risk (Bolton & Kacperczyk, 2020; Hartzmark & Sussman, 2019; Hong & Kacperczyk, 2009). We find it important to also consider potential commercial damages regarding corporate brand and recruitment capabilities, with the litigation risk increasing alongside the regulatory framework. One way to assess how investors are valuing sustainability is to examine event studies of suitable events. Krüger (2014) run event studies on stock market reactions to both positive and negative events concerning a firm's corporate social responsibility. Regarding our own research, we find it valuable that offsetting CSR[3] seems to be valued by investors. When excluding offsetting CSR, the author finds a negative relation between CSR and stock market reactions. This finding is consistent with Fisher-Vanden & Thorburn (2011) who investigate stock price reactions to announcements on voluntary corporate green initiatives. Whether the negative market reaction is explained by agency conflicts or the belief of green initiatives destroying shareholder value is a subject for further research.

Besides the short-term valuation measures of stock price reactions in event studies, we also find it interesting to study long-term valuation links with ESG. Clark et al. (2015) review studies that investigate the effects of

[3] Offsetting CSR is the event where firms with history of weak stakeholder relations present positive news related to their social responsibilities.

sustainable practices on metrics like operational performance, cost of capital, and stock price. The study concludes that strategies integrating ESG issues tends to outperform comparable non-ESG strategies. The outperformance is proven to involve both an increased operational performance and less risk. A similar study investigating the relationship between voluntary integration of environmental- and social issues and the following financial performance is conducted by Eccles et al. (2014). Their findings show that high sustainability firms have outperformed the low sustainability firms in the period 1993–2010.

The meta-analysis by Friede et al. (2015) combines more than 3,700 results from over 2,200 unique empirical studies on the link between ESG and financial performance. The analysis finds that approximately 95% of the studies have a positive relation between ESG and financial performance, where the majority of the relations were proven to be stable over time. Although the majority of current research concludes on a positive link between ESG and financial performance, there still exists conflicting evidence on this topic in current literature.

One empirical study that finds a negative relation between company value and various "green" activities is Nyilasy et al. (2014). This study finds evidence of a negative effect on brand attitudes and purchase intentions in the presence of green advertising. The findings support both scenarios where firms have a low and high environmental performance. The explanation is argued to lie in attribution theory where consumers may become sceptical and form negative attributions about the ulterior motive of the company. This theory could be an argument if our research were to find investors devaluing a structural shift within a company; consumers form attributions that the shift is being motivated by financial motives rather than sustainability. The conclusion of Nyilasy et al. (2014) is aligned with Ferrell et al. (2016), who argues that the effect of CSR on firm value will be dependent upon the incentives of the management. Further support of Nyilasy et al. (2014) is found in Brammer et al.

(2006), who investigates the relation between Corporate Social Performance (CSP) and financial performance using stock returns of UK quoted companies. The main finding is that firms with higher CSP tend to be outperformed by firms with the lowest CSP. However, the investors' ability to disvalue a structural shift within a firm must be assumed to be conditional on the respective firm's close competitors, or the ability of screening out its product line.

The empirical evidence concerning the association of corporate social actions and financial performance is found to be non-significant, despite the frequently claimed causal impacts. Baron et al. (2011) examines the interrelations among Corporate Financial Performance (CFP), CSP, and social pressure using a large sample of firms with social engagement in the period 1996–2004. The authors of the study find CFP and CSP to be largely unrelated using the full dataset. This finding is consistent with the theories in which the social market line[4] is horizontal or that CSP provides product differentiation. However, the authors find an interesting relation with social pressure: greater social pressure is associated with a greater CSP, contrary to being associated with a weaker CFP. The latter can be rationalised by social pressure reflecting the pressure on firms' brand equities and productivity.

Besides the relation between financial performance and various sustainable initiatives, we also investigate how the daily trading volume is likely to be impacted by the sustainable events. We have drawn inspiration from Beaver (1968), Foster (1973), Morse (1981), and Bamber (1986) on abnormal trading volume around earnings announcements using an event study methodology. The major objective in these studies is to assess the extent to which the trading volume around the event is abnormal, which all four studies found to be positive and significant.

[4]The social market line represent the equilibrium relation between the CFP and CSP for a respective firm.

3 Theory

Two key theories to help explain the relation between sustainable initiatives and financial performance are the shareholder- and stakeholder theory. The shareholder theory argues the only responsibility of a corporation is to maximise the profits accruing to its shareholders. On the contrary, the stakeholder theory argue that corporations have a greater responsibility to create value for all stakeholders, not just its shareholders. Both theories are unanimous of their respective views being the most value-creating and/or costs-reducing way to manage a corporation.

3.1 Shareholder Theory

The shareholder theory, also known as the Friedman doctrine, was introduced by the American economist Milton Friedman in 1970. Through his essay, *The Social Responsibility of Business is to Increase Its Profits*, he argued that a company has no social responsibility to the society; its only responsibility is to its shareholders. Friedman justified his view by shareholders being the rightful owners of the company: *"A corporate executive is an employee of the owners of the business. He has direct responsibility to his employers. That responsibility is to conduct the business in accordance with their desires".*

Further, Friedman assessed that an executive spending company money on social causes was in fact spending somebody else's money for their own purposes. Shareholders should decide for themselves if and how they wish to contribute to society by using their private wealth, generated from the company. The shareholder theory is supported by Elaine Sternberg, a philosopher specializing in business ethics and corporate governance. She argues that the stakeholder theory undermines two of the most fundamental features that characterise modern society: private property and the duties that agents owe to principals (Sternberg, 1997). The undermining of private property is explained with the stakeholder theory denying owners the right to determine how their property will be used.

However, one common mistake of the Friedman doctrine is the belief that profit-maximisation is encouraged at any cost. From Friedman (1962), *Capitalism and Freedom*, there is quoted that the social responsibility of business is to *"engage in activities designed to increase its profits so long as it engages in open and free competition without deception or fraud"*. It is argued that the critique of the shareholder theory often stems from the misinterpretation of Friedman encouraging illegal and unethical behaviour. In fact, the shareholder theory is encouraging allocation of funds to socially responsible activities — given it is the most value-creating action to make (Smith, 2003).

3.2 Stakeholder Theory

The stakeholder theory was introduced by the American philosopher R. Edward Freeman in 1984. His publication, *Strategic Management: A Stakeholder Approach*, was a response to the existing shareholder theory. We present the following definition from Freeman (1984): *"Stakeholder theory is a view of capitalism that stresses the interconnected relationships between a business and its customers, suppliers, employees, investors, communities and others having a stake in the organization"*. The theory argues that a firm should create value for all stakeholders, not just its direct owners.

Supporters of the stakeholder theory share many different views on how improved stakeholder relations may increase financial performance. Pfeffer (1995) argues that investing in human resources can increase efficiency and earn competitive advantages. Other views, of more relevance with the current regulatory circumstances, are the environmental aspects we find in Dechant et al. (1994) and Shrivastava & Hart (1995). The authors of both studies argue that investments in environmental activities can be cost reducing if such proactiveness would bring the companies in advance of present or future regulations.

Further, they argue that competitive advantages can arise from eco-friendly products through enhanced reputation, or loyalty from stakeholder relations like governments and customers. Excellent community relations might provide incentives for competition-enhancing tax breaks, or reduced regulation, thereby reducing costs to the firm and improving the bottom line (Waddock & Graves, 1997).

We need to establish the *causation* in order to evaluate the relation between sustainability integration and financial performance. The causal relation is argued to be a two-way relationship: firms with high sustainability rating leads to enhanced financial performance through improved stakeholder relations resulting in reduced agency costs. Alternatively, the causality implies that firms with strong financial performance will improve their sustainability integration, by having more resources available (Waddock & Graves, 1997). Our thesis will be focusing on the former alternative where sustainability impacts financial performance.

3.3 Efficient Market Hypothesis

The primary role of the capital market is allocation of ownership of the economy's capital stock (Fama, 1970). The ideal is a market in which investors can choose among the securities that represent ownership of firms' activities under the assumption that security prices at any time fully reflect all available information. Such a market is what theory defines as *"efficient"*. Further, the Efficient Market Hypothesis (EMH) is separated into three different informational subsets.

3.3.1 Strong Form Efficiency

Strong form efficiency is the most stringent version of EMH investment theory. It is required that security prices, at all times, fully reflect all information available — both public and private. This form of efficiency implies that neither technical analysis, equity research, nor inside information can provide investor advantages. With strong form efficiency, the only way to maximise profits would be to follow a buy-and-hold strategy (Malkiel, 2003).

3.3.2 Semi-Strong Efficiency

Semi-strong efficiency implies that all stock prices will immediately, and efficiently, adjust to new publicly available information (Fama, 1970). This form of efficiency is considered to be the most practical of all EMH hypotheses. We will rely on this form of efficiency when conducting our statistical analysis. Thus, we assume that all existing information is already factored into the price of the security before the announcement of the respective events.

3.3.3 Weak Form Efficiency

Weak form efficiency states that all current information is reflected into the stock price, and is the least stringent version of EMH. Additionally, future prices are random and not influenced by historical information. This form of EMH is also known as the random walk theory and has found support in the study of Van Horne & Parker (1967). On the contrary, Doan & Lo (1988) rejects the random walk hypothesis of weekly stock returns using a simple volatility-based specification. However, the authors are unable conclude on the causality of inefficiency in stock-price formations.

3.4 Hypotheses

We have developed the following testable hypotheses, from the aforementioned theories, which are presented in this subsection. The hypotheses are structured in a specific order for which they are connected, but must not be confused with dependency.

3.4.1 Abnormal Trading Volumes

An impact of a company addressing a change towards an improved sustainable profile should be observed in the trading volume. By referring to the views of the shareholder- and stakeholder theory, investors have fundamentally different opinions concerning the value of sustainability. This difference of opinion should result in positive abnormal trading volumes, where investors of the two theories acts in line with their personal beliefs.

Hypothesis 1: Shifts in ESG generates positive abnormal trading volumes

3.4.2 Abnormal Returns

Following the first hypothesis of positive abnormal trading volumes, it is of interest to address the price impact of the increased trading activity. This impact is further linked to the shareholder- and stakeholder theory: We should observe negative abnormal returns following the shift in ESG if the market consensus is in favour of the shareholder theory. Oppositely, we should observe positive abnormal returns following the shift if the consensus is alligned with the stakeholder view of sustainability. The sign of the abnormal return can be considered as a horserace between the investor preference of whether ESG is value-creating or value-destroying. This hypothesis is closely connected to the study by Cooper et al. (2000), where the authors investigated the effect of company name changes related to the boom that internet-related stocks experienced at the time.

Hypothesis 2: Shifts in ESG generates abnormal returns

3.4.3 ESG Ratings

Lastly, it is also of interest to analyse the development in the ESG ratings following the shift to determine whether companies are attempting to cater to specific investor preferences. We only conduct this analysis on the more material changes in the sustainable policy of the firms, in order to avoid a mix-up between cause and result. This subject is further discussed in section 4.6. We should observe no increase in ESG ratings if companies are indeed focusing solely on maximising shareholder value, and trying to convey a false notion of becoming more sustainable. Oppositely, we should observe improved ESG ratings following the shift if stakeholder relations is what drives the decision-making of the companies. This hypothesis also includes a separate analysis of the individual pillar scores in order to assess the drivers of the potential change in the overall ESG rating.

Hypothesis 3: Rebranding does not cause changes in ESG Ratings

4 Methodology

The methodology for conducting the analysis is based upon the *deductive* research approach. A deductive approach involves exploring a known theory or phenomenon and test if that theory is valid in given circumstances. We will formulate hypotheses and subject these to testing during the research process, contrasting to the *inductive* research approach that develops hypotheses based on the findings. The motivation behind the deductive approach is to better explain causal relationships between concepts and variables, in addition to being able to measure concepts quantitatively. The theoretical framework and structure of analysis adopted in order to conduct the event study is inspired by MacKinlay (1997).

Event Studies

We have adopted the event study methodology in order to examine the effect of sustainable initiatives on a company's stock price and trading volume. The foundation for conducting modern event studies is commonly associated with Ball & Brown (1968) and Fama et al. (1969), while the first publishment of an event study is assumed to be Dolley (1933) (MacKinlay, 1997; Bowman, 2006).

Event studies rely on the assumption of market efficiency, which is thoroughly discussed in section 3.3. The benefits of adopting an event study are the possibilities of drawing conclusions of investors' general consensus regarding firm affection, in addition to the economic value attributed to the event (MacKinlay, 1997). Following are the stepwise general flow of analysis when conducting an event study.

4.1 Event Identification

The sustainable events examined through the event study, over the period 2004 to 2019, contains 37 different companies in the rebranding sample, and

297 different companies in the sustainability reporting sample. Companies included in the rebranding sample must fulfill at least one of the three conditions: strategic name changes, commitment to sustainability-linked frameworks, or material operational changes. For companies to be included in the sustanability reporting sample, their reporting routine towards sustainability must be significantly improved through a clear shift. A more comprehensive overview of event identification is provided in section 5.2. The event dates ($t = 0$) are determined using the first day of publication of the company announcement, which is in line with MacKinlay (1997) and Flammer (2015) among others. In order to capture whether information about the event has been leaked to the stock market, the event window is lagged with 5 days ($t - 5$). Similarly, the market might need additional time to absorb the informational content of the event. Subsequently, the event window is extended with 120 days ($t+120$). Resultingly, we define the event window [-5,+120] as the time period over which the security prices are analysed (Brooks, 2019).

4.2 Return Measurements

Appraisal of the event's impact requires a measure of Abnormal Returns (AR). The abnormal return is the actual observed return of the security over the event window minus the normal return of the firm over the event window. The normal return is defined as the expected return without conditioning on the event taking place. The expected return is found using the estimation window prior to the event window. Using the definition from MacKinlay (1997), the AR for firm i and date t can be expressed as

$$AR_{i,t} = R_{i,t} - E(R_{i,t}|X_t) \tag{4.1}$$

where $AR_{i,t}$, $R_{i,t}$, and $E(R_{i,t}|X_t)$ are the abnormal, actual, and normal returns respectively for the time period t. X_t is the conditioning information for the normal return model.

There are numerous approaches available for calculating the normal return of a given security. However, the available approaches can be roughly grouped into two categorical models — statistical and economic.

4.2.1 Economic Models

Economic models rely on the assumptions concerning investors' behaviour and are not solely based on statistical assumptions. These models can be cast as restrictions on the statistical models to provide more constrained normal return models (MacKinlay, 1997). Two common economic models are the Capital Asset Pricing Model (CAPM) and Arbitrage Pricing Theory (APT).

The CAPM was established by Treynor (1961, 1962), Sharpe (1964), Lintner (1965), and Mossin (1966) and is a single factor model for expected return. The CAPM is based on the Modern Portfolio Theory by Markowitz (1952) where the market exposure (β) of the respective asset is the only factor explaining the expected return. The reason why the CAPM has remained popular given the origination of more modern approaches, such as the APT, is due to its simplicity and utility in a variety of situations (Hall & Asteriou, 2015).

The APT is an asset pricing theory where the expected return of a given security is a linear combination of multiple risk factors (Ross, 1976). A general finding is that the most important factor of the APT behaves like a market factor, and that the additional factors marginally affects the explanatory power of the model. Thus, the potential gains from using the APT Model rather than the statistical Market Model (MM) are small (Brown & Weinstein, 1985).

4.2.2 Statistical Models

Statistical models differentiate from the economic models by following statistical assumptions concerning the behaviour of assets returns, and do not depend on economic arguments. The most persistent models discussed in MacKinlay (1997) are the *Constant Mean Return Model*, the *Market Model*, and *Multifactor Models*.

The Constant Mean Return Model assumes that the expected return is being constant over time (Campbell & Wasley, 1993), whereas the Market Model assumes a stable linear relationship between the expected return of security i and the market return (Bodie et al., 2019). However, the Market Model represents a potential improvement of the Constant Mean Return Model: the variance of the abnormal return is reduced by removing the portion of the return related to variation in the market's return (MacKinlay, 1997).

The third statistical model, the Multifactor Model, is motivated by the benefits of reducing the variance of the abnormal return by explaining more of the variation in the normal return. Generally, the benefits from employing multifactor models for event studies are limited. This limitation is explained by the empirical fact that the marginal explanatory power of additional factors to the market factor is small. This in turn limits the reduction in variance of the abnormal returns.

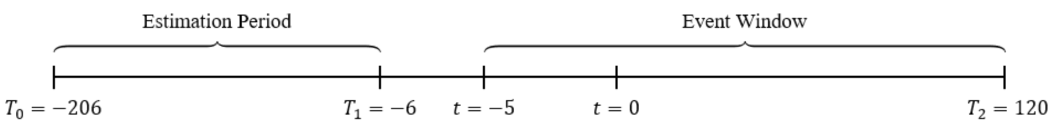
4.2.3 Model Selection

The potential benefit from using economic models based on the APT is to eliminate the biases introduced by the CAPM. However, since these biases also are eliminated by using the statistically motivated models, the economic models are overshadowed by statistical in event studies (MacKinlay, 1997). All things considered, we have chosen to estimate the normal returns using two statistical models: The Market Model, followed by Fama-French Three Factor Model (FF3) to ensure robustness.

4.3 Estimation Window

The estimation window provides the information needed to estimate the normal return, and represent the time period prior to the event. Typically, the estimation- and event window do not overlap, ensuring that the parameters for the normal return model are not influenced by the returns around the event (MacKinlay, 1997). The length of the estimation window is usually ranging from 100 to 300 days (Peterson, 1989; Armitage, 1995), while MacKinlay (1997) suggest an optimal window between 180 and 250 days. The key consideration when deciding on the estimation length is that the window is long enough to give a precise representation of the volatility in the given security. We have adopted a 200-day estimation window ranging from $[-206, -6]$, which is consistent with Flammer (2015) whose study is similar to ours. There is no gap between the estimation- and the event window, explained by the possibility of capturing whether information has been leaked to the market before the specified event date ($t = 0$).

Figure 1: Timeline of Event Study



4.4 Testing Framework

This section provides the testing framework for the abnormal returns. All return variables are in the logarithmic form in the following calculations:

$$R_{i,t} = ln\left(\frac{P_{i,t}}{P_{i,t-1}}\right) \tag{4.2}$$

Where P represents the closing price of security i at time t. The logarithmic form of returns is empirically proven to enhance the normality of the return distribution (Henderson, 1990). Logarithmic returns also simplifies the summing process when estimating cumulative returns, compared to the arithmetic form.

4.4.1 The Market Model

The Market Model is a one factor model that relates the return of any given security to the return of the market portfolio. The linear specification of the model follows from the assumption of joint normality across asset returns and can be expressed as:

$$R_{i,t} = \alpha_i + \beta_i R_{M,t} + \epsilon_{i,t} \tag{4.3}$$

$$E(\epsilon_{i,t}) = 0 \qquad var\left(\epsilon_{i,t}\right) = \sigma^2_{\epsilon_i}$$

where

$R_{i,t}$ = Period t return on security i

α_i = Intercept of security i

β_i = Slope coefficient of the systematic risk of security i

$R_{M,t}$ = Period t return on the market portfolio M

$\epsilon_{i,t}$ = Zero mean disturbance term

Under general conditions, the Market Model parameters α_i, β_i, and σ_{ϵ_i} are estimated using the Ordinary Least Squares (OLS) procedure. The estimators of the model parameters will be Best Linear Unbiased Estimators (BLUE) if the following assumptions hold (Brooks, 2019):

(1) $E[u_t] = 0$	The errors have zero mean
(2) $Var[u_t] = \sigma^2 < \infty$	The variance of the errors is constant and finite
(3) $Cov[u_i, u_j] = 0$	The errors are linearly independent of one another
(4) $Cov[u_t, x_t] = 0$	There is no relationship between the error and the corresponding x variable(s)

If we include a constant term to our regression model, the first assumption is valid. The second assumption assumes that the error terms are homoscedastic, while the third assumption assumes no autocorrelation, implying independence across the error terms. We will adopt the *Newey-West Heteroscedasticity-and Autocorrelation-Corrected Standard Errors* (HAC) in order to ensure that assumptions (2) and (3) are not violated. The last assumption assumes that the error terms are non-stochastic, implying they are not correlated with the regressors of the regression model.

4.4.2 Multifactor Model

The motivation of adopting a multifactor model is to reduce the variance of the abnormal return by explaining more of the variation in the normal return by introducing additional factors. One way to develop explaining factors is to study common market anomalies and conclude on their robustness. The study done by Fama & French (1993) identifies five common risk factors in the returns on stocks and bonds, whereas three are related to the stock-market. These three common risk factors include an overall market factor and factors related to firm size and book-to-market equity. The size effect refers to the phenomenon where the performance of relatively small stocks tends to exceed that of larger stocks, measured by market cap. The value effect is another phenomenon where stocks with high book-to-market ratio, known as value stocks, earn higher average returns compared to stocks with a lower book-to-market ratio, so-called growth stocks (Bodie et al., 2019). This finding is what constitutes to the FF3, which we adopt in order to supplement the statistical inferences drawn from the Market Model.

Similar to the Market Model, the linear specification of our multi-factor model can be expressed as:

$$R_{i,t} = \alpha_i + \beta_{i,M} R_{M,t} + \beta_{i,SMB} SMB_t + \beta_{i,HML} HML_t + \epsilon_{i,t} \tag{4.4}$$

where

$\beta_{i,SMB}$ = Slope coefficient of the size factor for security i

SMB_t = Return on the size factor in period t

$\beta_{i,HML}$ = Slope coefficient of the value factor for security i

HML_t = Return on the value factor in period t

4.4.3 Statistical Properties

We are able to define and analyse the abnormal returns using the parameter estimates of the two models. We denote the abnormal return in the event window $t = T_1 + 1, ..., T_2$ for security i as:

$$AR_{i,t} = R_{i,t} - \widehat{\alpha}_i - \widehat{\beta}_{i,M} R_{M,t} \tag{4.5}$$

$$AR_{i,t} = R_{i,t} - \widehat{\alpha}_i - \widehat{\beta}_{i,M} R_{M,t} - \widehat{\beta}_{i,SMB} SMB_t - \widehat{\beta}_{i,HML} HML_t \tag{4.6}$$

Abnormal Return

The abnormal return is the disturbance term of both the Market Model, and the multifactor model, calculated on an out-of-sample basis (MacKinlay, 1997). We hypothesise that the abnormal returns, conditional on the event window returns, will be jointly normally distributed with a zero conditional mean and conditional variance, $\sigma^2(AR_{i,t})$, expressed as:

$$\sigma^2(AR_{i,t}) = \sigma^2_{\epsilon_i} + \frac{1}{L_1}\left[1 + \frac{\left(R_{M,t} - \overline{R}_M\right)^2}{\widehat{\sigma}^2_M}\right] \tag{4.7}$$

The conditional variance has two components. One component is the disturbance variance, $\sigma^2_{\epsilon_i}$, and the other component is additional variance caused by the sampling error in α_i and β_i. However, the second component approaches zero as the sampling error of the two parameters diminishes when the length of the estimation window (L_1) increases. This implies that the variance of the abnormal return will be $\sigma^2_{\epsilon_i}$, and the observations will become independent over time (MacKinlay, 1997).

Cumulative Abnormal Return

The abnormal return observations must be aggregated in order to draw inferences of the respective events. The observations are aggregated both over time and across securities. The Cumulative Abnormal Returns (CAR) from t_1 to t_2 is the sum of the included abnormal returns of security i:

$$CAR_{i,(t_1,t_2)} = \sum_{t=t_1}^{t_2} AR_{i,t} \tag{4.8}$$

Following the discussion from equation 4.7, the variance of CAR can be defined as:

$$\sigma^2_{i,(t_1,t_2)} = (t_2 - t_1 + 1)\,\sigma^2_{\epsilon_i} \tag{4.9}$$

Cumulative Average Abnormal Returns

It is necessary to include not only one event observation in order to achieve any valuable results from various tests. This problem is solved by aggregating abnormal return observations for the event window, assuming there is no overlap between the respective securities. We arrive at Cumulative Average Abnormal Return ($\overline{CAR}$) by aggregating the Average Abnormal Return ($\overline{AR}$), similar to equation 4.8.

$$\overline{CAR}_{(t_1,t_2)} = \sum_{t=t_1}^{t_2} \overline{AR}_t \tag{4.10}$$

4.5 Abnormal Trading Volume

The shifts can also carry a significant impact on trading volumes as a results of increased investor attention. This hypothesised impact closely follows the discussion in section 3.4.1 related to the conflicting theories related to investors' perception of sustainable initiatives. We utilise the methodology laid out by Ajinkya & Jain (1989) in order to estimate and test the significance of the Abnormal Trading Volume (AV).

4.5.1 Abnormal Volume Estimation

The abnormal trading volume estimation closely follows the abnormal return estimation where we subtract the normal trading volume in order to determine any abnormal patterns.

$$AV_{i,t} = V_{i,t} - E(V_{i,t}|X_t) \tag{4.11}$$

Where $AV_{i,t}$, $V_{i,t}$, and $E(V_{i,t}|X_t)$ represents the abnormal, actual and normal trading volume for each company i at time t.

Several previous studies have shown that raw trading volume does not follow a normal distribution (Ajinkya & Jain, 1989; Cready & Ramanan, 1991). However, Campbell & Wasley (1996) states that a log transformation yields a trading volume measure that is approximately normally distributed. We will employ the following measure for trading volume in order to obtain the desired statistical features required for significance testing:

$$V_{i,t} = ln\left(\frac{n_{i,t}}{S_{i,t}} \times 100\right) \tag{4.12}$$

Where $n_{i,t}$ and $S_{i,t}$ represents the number of shares traded and number of shares outstanding for company i at time t.

We employ a mean-adjusted model for the normal trading volume when estimating the abnormal trading volume, defined as:

$$AV_{i,t} = V_{i,t} - \overline{V}_i \tag{4.13}$$

Where

$$\overline{V}_i = \frac{1}{L_1} \sum_{t=T_0}^{T_1} V_{i,t} \tag{4.14}$$

L_1 denotes the number of days in the estimation window. We employ the same number of days as in the abnormal return estimation, which is 200 days prior to the announcement.

4.5.2 Significance Testing of Abnormal Trading Volume

We employ a parametric t-test in order to test the significance of the abnormal trading volume, similar to Ajinkya & Jain (1989). The t-test can be conducted as a result of the log-transformation described in the previous section, as the normality assumption is now likely to be fulfilled.

The test statistics of the t-test can be computed as follows:

$$t_{\overline{AV},t} = \frac{\overline{AV}_t}{S_{\overline{AV}_t}} \tag{4.15}$$

Where

$$\overline{AV}_t = \frac{1}{N}\sum_{i=1}^{N} AV_{i,t} \tag{4.16}$$

And

$$S_{\overline{AV}_t} = \sqrt{\frac{1}{L_1 - 1}\sum_{t=1}^{L_1}\left(\overline{AV}_t - \overline{\overline{AV}}\right)} \tag{4.17}$$

$\overline{AV}_t$ corresponds to the cross-sectional average at time t, and $\overline{\overline{AV}} = \frac{1}{L_1}\sum_{t=1}^{L_1} \overline{AV}_t$.

4.6 Change in ESG Ratings and Pillar Scores

Another impact that can be seen following the shift is the subsequent change in a company's ESG rating. We believe this impact to be especially important for the rebranding sample, where companies have publicly stated that the strategic direction of the company will be changed in order to obtain a more sustainable profile.

We do not conduct this analysis on the sustainability reporting sample, as reporting carries a weight in the estimation of the retrieved ESG ratings. As we implicitly would mix the cause with result, we believe that there will exist a causality issue. Subsequently, effects seen in the ESG ratings could solely reflect the initiation of reporting and not an improvement in the actual performance of the company in question.

It is therefore of interest to see if there is any significant shift in the ESG ratings of the companies included in our rebranding sample. Pillar scores are also analysed in order to identify the drivers of the potential change in the overall rating. We will employ a paired t-test in order to see if there is any significant change in the ESG ratings, accompanied by a Wilcoxon Signed Rank Test. From hypothesis 3, we hypothesise that we should not see an improvement in the ratings for the companies following the shift.

We will use the last reported ESG ratings and pillar scores prior to the announcement as a baseline (t), and test the significance of the relative difference in the scores received following the announcement.

We are not able to test whether one can observe significant changes in longer time horizons than 2 years following the shift ($t+2$) for the t-test, due to data limitations affecting our sample sizes. However, we believe this topic to be highly relevant for future research when more data becomes available. Additionally, we have tested the significance up to $t+5$ using the non-parametric test, as it can provide some insight even with smaller samples.

5 Data

This section presents the sampling procedure used in the event study to assess the market reactions to corporate sustainable initiatives. We will disclose both the data collection process, and the criteria used in the data screening procedure.

5.1 Data Collection

The motivation of the data collection is to secure a large and exhaustive sample of all sustainable shifts that have taken place in the current century within the larger European companies. We choose to focus on the benchmark indices in order to ensure the larger companies within each country is included. Further, since we adopt the FF3 in the estimation of abnormal returns in the event window, we are dependent on each country in the sample also being included in the Fama French Daily European Three Factors, which we retrieve from Kenneth R. French' data library (Fama & French, 2021). This criterion has resulted in 16 countries being included in the initial data collection phase. The respective countries are presented in table 1 below.

Table 1: Countries Included in the Sample

Norway	Sweden	Denmark	Finland
Italy	France	Switzerland	Germany
Spain	Austria	Belgium	Ireland
Netherlands	United Kingdom	Greece	Portugal

The rationale of collecting data from the relatively larger companies is the availability of public data, more precisely the following three categories:

5.1.1 Stock Price Data

All stock price data is gathered from the Compustat database, accessed through the Wharton Research Data Services (WRDS). The time series data for the respective indices are retrieved from Bloomberg. The data is cleaned and adjusted for both dividends and stock splits in order to calculate the daily returns, and further used in the estimation of the abnormal returns. We ensure that Compustat can deliver stock price information both pre- and post the sustainable initiative of interest by focusing on the relatively larger companies.

5.1.2 ESG Data

We have collected numerical ESG ratings and pillar scores from Refinitiv in order to conduct our analysis on the development in sustainable performance. This procedure is dependent upon the ESG ratings and pillar scores both being reliable, and having historical data available. We have experienced that larger companies tend to have more ESG data available, which can be used for hypothesis testing.

5.1.3 Trading Volume Data

We retrieve the daily trading volume and shares outstanding from the Compustat database for each company in our samples. This data is necessary for conducting the analysis of abnormal trading volume around the time of the events. Like the stock price data and the ESG data, the relatively larger companies tend to have better information in Compustat.

5.2 Screening Criteria

The sample companies included in the event study are imposed two key selection criteria: the listing criterion followed by the sustainability criterion. The listing criterion involves the requirement of being listed on the benchmark index in one of the countries included in the Fama French European Three Factors. The background for this requirement is that the FF3 is adopted in order to test the robustness of the inferences drawn from the Market Model estimates.

We deem the sustainability criterion to be the most essential, as the sustainable initiatives are the background of this study. This criterion is separated into two different criteria that represents our two final samples: the rebranding sample and the sustainability reporting sample.

5.2.1 Rebranding

The rebranding criterion constitutes to the rebranding sample and is the most stringent criterion of the two. With the term rebranding, we are referring to company announcements signaling a material change towards the company's strategic direction or vision. Companies included in this sample must fulfill at least one of the three conditions listed below:

1. Strategic name changes

We assume that companies with its brand closely linked to carbon-intensive industries announces a corporate name change to signal its improved sustainable vision. For example, this involves companies with a current name including the words *"Petroleum"* or *"Oil"*, which then announces to replace these parts with *"Energy"* or *"Renewables"*. The sample companies captured by a strategic name change are explicitly operating in the energy sector, and more closely the oil and gas industry.

2. Commitment to sustainability-linked frameworks

Another condition of the rebranding criterion are companies that publicly announces their commitment to sustainability-linked framework and guidelines. We assume that such a commitment is to indicate that a company has chosen to redirect its vision into a greater sustainable manner.

3. Material operational changes

The last condition has similarities to the strategic name change criteria in that it is mainly directed towards the carbon-intensive sectors. With material operational changes, we are referring to carbon-intensive companies that publicly announces a commitment to zero-emission or carbon neutral operations within a specific timeframe. The sample companies captured by this sub-criterion was all, at the time of the announcement, operating with relatively high carbon emission in industries with high exposure to climate risk.

We were left with 712 companies from the benchmark indices of the 16 countries listed in the Fama French European Three Factors after applying the listing criterion. Our final rebranding sample contains 37 companies across 13 countries after imposing the rebranding conditions.

5.2.2 Sustainability Reporting

The reporting criterion constitutes to our sustainability reporting sample and is less stringent than the rebranding criteria. Companies included in this sample have all experienced a material change in their sustainability reporting routine. The complication of this criterion is to avoid subjective inconsistency in the decision of whether a change is material or not. This subjectiveness is avoided by setting clear guidelines on how a change in the reporting framework of a company should be interpreted. These guidelines includes that all sustainability reporting within a company must contain Key Performance Indicators (KPI) for tracking the progress, a clear timeframe for which the target measures should be fulfilled, and detailed company-specific information describing the current situation of the firm.

The main procedure is to examine the exact date a company publishes their first ever sustainability report. With sustainability, we are also referring to CSR- and environmental reports. However, companies might have been documenting their sustainable efforts in previous annual reports. We must therefore analyse if this reporting was done in a proper manner before concluding with the date of the sustainability report. The company will not be included in our sample if the first sustainability report does not represent a material change from previous reporting standards.

The second procedure is the examination of a company's integrated reporting standards, which is the most exposed to subjective inconsistency. This procedure is undertaken for companies that do not publish a stand-alone sustainability report, but rather integrate sustainable efforts into their annual report. Consequently, all annual reports are carefully investigated in order to conclude whether the company has experienced a material change in their sustainability reporting standards. However, we have chosen to be conservative in the decision of including these into the sample, since such a sustainable shift will be recorded at the date of the release of the annual report. We require an extensive increase in both the length and the quality of their integrated report, with several relevant KPI's as a necessity in order to be included. This requirement is naturally explained by wanting to limit all sources of noise stemming from coinciding events in our event window before running the statistical analysis.

Our final sustainability reporting sample contains 297 companies across all 16 countries after imposing the reporting criterion.

6 Empirical Findings

We will provide the empirical results from our analysis in this section. First, we show that both samples exhibit positive significant abnormal trading volume in the days surrounding the event date.

Next, we show that the only shift generating significant abnormal returns is the rebranding scenarios. We observe the returns being initially positive in the days following the announcement, followed by a reversion shortly after the event.

Lastly, we show that the companies included in the rebranding sample experiences a significant positive change in their ESG ratings and pillar scores following the announcement. This finding contradicts our initial hypothesis that companies attempt to cater to specific investor preferences regarding sustainability in listed companies.

6.1 Abnormal Trading Volume

According to hypothesis 1, we expect to find positive Average Abnormal Trading Volume ($\overline{AV}$) surrounding the event dates for our two samples. This abnormal volume comes as a result of the hypothesised increased attention stemming from the respective shifts.

The results in table 2 show that the announcement generates positive abnormal trading volume on day 0 for the rebranding sample. The abnormal trading volume is significant at the 5% level, and in line with hypothesis 1. The same can also be seen in the reporting sample, with significant positive trading volume in the days surrounding the event date. However, it should also be noted that this sample is more prone to coinciding events in the event window, which could contribute to the increased trading in these companies.

Additionally, the log transformation successfully satisfied the normality assumption for the rebranding sample, while we have evidence supporting a positively skewed and platykurtic distribution in the reporting sample.

However, the distribution should be approximately normal according to the Central Limit Theorem taking sample size into consideration (Brooks, 2019).

Table 2: Abnormal Trading Volume

	Reporting N = 297		Rebranding N = 37	
T	$\overline{AV}$	t-test	$\overline{AV}$	t-test
-5	0.059	0.84	-0.031	-0.31
-4	0.092	1.30	-0.116	-1.17
-3	0.100	1.41	0.051	0.51
-2	0.124	1.75*	0.032	0.32
-1	0.169	2.38**	0.082	0.83
0	0.258	3.64***	0.231	2.33*
1	0.261	3.68***	0.061	0.61
2	0.188	2.65***	0.085	0.86
3	0.135	1.90*	0.143	1.44
4	0.099	1.39	0.152	1.54
5	0.091	1.28	0.051	0.52
6	0.120	1.70	0.057	0.57
7	0.077	1.08	0.109	1.10
8	0.115	1.62	0.046	0.47
9	0.096	1.35	-0.006	-0.06
10	0.119	1.67*	0.049	0.49
11	0.064	0.91	-0.014	-0.14
12	0.114	1.61	0.115	1.17
13	0.134	1.89*	0.142	1.44
14	0.144	2.03**	0.127	1.29
15	0.116	1.63	-0.041	-0.41
Avg. Skewness		0.282**		0.477
		(0.141)		(0.393)
Avg. Kurtosis		-1.461***		-1.051
		(0.282)		(0.770)

The table represents the $\overline{AV}s$ and corresponding t-statistics. The average skewness and kurtosis for the samples are also reported in order to check the normality assumption with corresponding standard errors in parentheses. The kurtosis coefficients reflects the subtraction of 3.0, which is the expected value under normality. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10.

Overall, it appears that the companies in our samples are experiencing abnormal trading volumes surrounding the event date, implying that the shifts increases the liquidity of the shares for the companies. Thus, the evidence suggests that the shifts increases investor attention.

6.2 Abnormal Returns

In order to determine whether the different events have an abnormal impact on the returns of the companies in our samples, it is of interest to interpret the build-up of $\overline{CAR}$ surrounding the event dates. A significant positive CAR could work as a financial incentive for the companies in question, following our initial hypothesis.

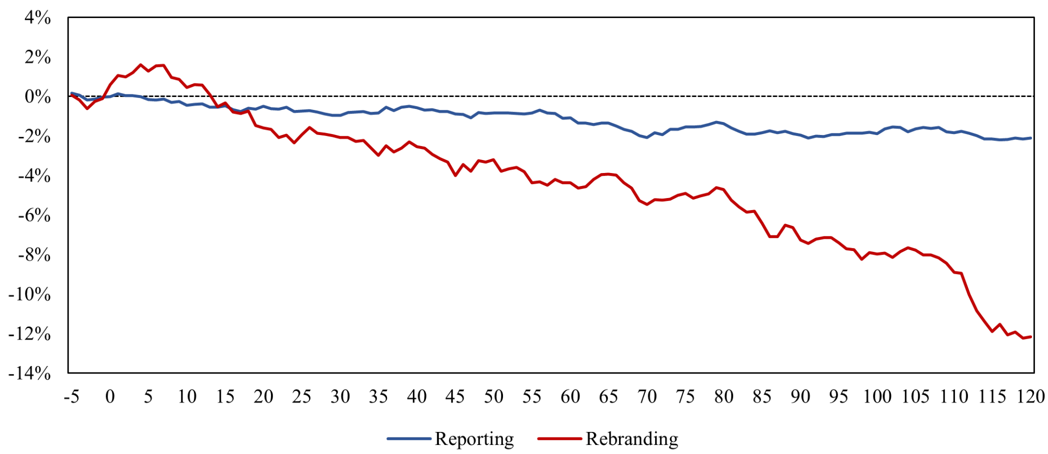
Thus, the first step is to test the following hypothesis for both of our samples:

$$H_0 : \overline{CAR} = 0 \quad vs. \quad H_1 : \overline{CAR} \neq 0$$

We applied the event study methodology laid out in section 4 in order to test this hypothesis. We estimate the $\overline{CAR}s$ separately for both of our different events in order to isolate the effect of the different shifts.

In figure 2 we present the $\overline{CAR}$ for both of the event types in our study. From the plot, sustainability reporting seems to carry no significant effect in terms of abnormal returns. However, the rebranding cases seems to generate abnormal returns following the event.

Figure 2: Cumulative Average Abnormal Returns



We observe a clear distinction between the two events already from the visual analysis of the $\overline{CARs}$. There seems to be evidence of abnormal returns in the rebranding sample surrounding the event date, while no distinct effects can be seen in the reporting sample.

Another feature of the rebranding sample is that there seem to be a reversion and negative post-announcement drift taking place. We see an initial increase in the immediate days following the announcement, followed by a continuous decrease as the event window increases. We argue that the initial increase in value for the companies in the sample are not permanent, but rather that we see an initial overreaction and that investors slowly incorporate the new information into the stock price. Such market reactions contradicts the EMH, and can seen in several previous studies on other types of shifts like the well-known earnings announcement drift (Bernard & Thomas, 1989, 1990).

6.2.1 Significance of Abnormal Returns

We are interested in interpreting the significance of returns generated by the event following the preliminary analysis. This analysis is done both for the $\overline{AR}$ and the $\overline{CAR}$, which enables us to interpret both the cumulative- and the single day effect carried by the event. Hypothesis 2 state that the event should generate abnormal returns around the announcement date.

From table 3, we see that sustainability reporting seems to generate some significant abnormal returns in the days surrounding the event, but not at the actual event day. As we observe significant abnormal trading volumes, but no abnormal returns, we interpret this finding as a sign of disagreement amongst investors regarding the value impact of sustainability reporting. Additionally, there are no conclusive results from our three tests on any given day in the event window. Thus, the significance can be caused by violated statistical assumptions made in each respective test, thereby making the robustness of these findings questionable. Subsequently, we are not able to conclude that the event generates any conclusive significant abnormal return.

Table 3: Average Abnormal Returns

	Reporting N = 297				Rebranding N = 37			
T	$\overline{AR}$	t-test	Patell Z	Wilcox.	$\overline{AR}$	t-test	Patell Z	Wilcox.
-5	0.15 %	1.34	1.44	0.50	-0.09 %	-0.41	-0.58	-0.78
-4	-0.11 %	-0.91	-0.88	-0.75	-0.15 %	-0.68	-0.38	-0.88
-3	-0.24 %	-1.66*	-2.45**	-1.10	-0.19 %	-0.99	-0.94	-1.55
-2	0.07 %	0.58	0.78	0.71	0.31 %	1.64	1.15	1.21
-1	0.09 %	0.76	0.18	0.35	0.05 %	0.24	0.48	0.02
0	0.03 %	0.15	0.80	-0.18	0.67 %	2.32**	2.80***	2.27**
1	0.14 %	0.91	1.78*	0.78	0.33 %	1.05	1.59	1.02
2	-0.11 %	-0.99	-0.49	-1.09	0.02 %	0.09	0.31	-0.20
3	0.00 %	0.05	0.03	-0.14	0.23 %	0.95	1.37	1.05
4	-0.05 %	-0.41	-0.22	0.25	0.44 %	1.61	1.51	1.67
5	-0.15 %	-1.16	-1.72*	-1.06	-0.17 %	-0.72	-0.15	-0.67
6	0.00 %	-0.05	-0.81	-0.88	0.26 %	1.24	1.56	1.17
7	0.04 %	0.33	-0.16	-0.65	-0.01 %	-0.03	0.51	0.02
8	-0.16 %	-1.21	-1.98**	-0.11	-0.45 %	-2.36**	-2.27**	-2.30**
9	0.04 %	0.37	0.87	0.97	-0.14 %	-0.78	-0.43	-0.66
10	-0.19 %	-1.47	-1.58	-1.26	-0.21 %	-0.93	-0.64	-1.09
11	0.04 %	0.34	0.15	0.24	0.11 %	0.67	0.50	0.66
12	0.02 %	0.14	0.11	0.13	-0.06 %	-0.27	-0.31	-0.64
13	-0.16 %	-1.24	-2.30*	-2.49**	-0.32 %	-0.70	-0.53	0.07
14	-0.01 %	-0.05	1.21	0.54	-0.60 %	-0.74	-2.63***	0.87
15	0.07 %	0.51	0.67	0.23	0.11 %	0.41	0.03	0.75

The table presents the $\overline{AR}s$ surrounding the announcement of the reporting- and rebranding sample from 5 days prior to the announcement to 15 days post-announcement [-5,15]. Abnormal returns are estimated using the Market Model. We include a t-test, Patell Z test, and Wilcoxon Signed Rank Test in order to assess whether or not the $\overline{AR}s$ are statistically different from zero. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10. See appendix A.1 for the estimation techniques for each of the different test statistics.

However, there are some interesting findings in the rebranding sample. Rebranding seems to cause significant abnormal returns on the event date, with an $\overline{AR}$ of 0.67% being significant at the 5% level across all three tests. On the other hand, the events seem to carry a relatively small impact, with the $\overline{ARs}$ lying in the interval [-0.60%, 0.67%]. Thus, the impact at $t = 0$ is significant, but the economic significance of the shift seems to be rather small. We also have significant negative $\overline{AR}$ of -0.45% on day 8, marking the starting point of the reversion that we saw in the graphical illustration of the $\overline{CAR}$s in figure 2.

We see similar patterns in the sustainability reporting sample when interpreting the results from our analysis on the $\overline{CARs}$ in table 4. Reporting seems to carry no significant impact in terms of CAR when taking conclusiveness of the tests into account. We see that the Patell Z test yields significant findings in the longer event windows, namely [-5,+30], [-5,+60], and [-5,+120]. However, for longer event windows, it is also recommended to rely on non-parametric tests as they do not depend as heavily on strict assumptions related to the statistical features of the abnormal return series.

We see similarities for the rebranding sample as well. We find some significant positive build-up of $\overline{CAR}$ in the days following the event, with a reversion taking place as the event window increases. Especially in the longest event window, [-5,+120], we find a $\overline{CAR}$ of -10.05% that is significant across all of our tests at a 5% level. The finding is in line with the shareholder theory. That is, shareholders might believe that ESG initiatives ultimately destroy value for the shareholders as the companies gradually move away from solely focusing on maximising shareholder value.

Another explanation could lie in attribution theory, where investors are sceptical of the ulterior motive of the companies. If investors believe that an ulterior motive behind the rebranding exists, it might result in lower valuations similar to the findings by Nyilasy et al. (2014).

Table 4: Cumulative Average Abnormal Returns

	Reporting N = 297				Rebranding N = 37			
Window	$\overline{CAR}$	t-test	Patell Z	Wilcox.	$\overline{CAR}$	t-test	Patell Z	Wilcox.
[-5,-5]	0.15 %	1.32	1.44	0.50	-0.09 %	-0.41	-0.58	-0.78
[-5,-4]	0.04 %	0.27	0.40	-0.46	-0.24 %	-0.74	-0.67	-1.27
[-5,-3]	-0.20 %	-0.99	-1.09	-1.02	-0.43 %	-1.14	-1.09	-1.88*
[-5,-2]	-0.13 %	-0.56	-0.55	-0.50	-0.12 %	-0.32	-0.37	-0.87
[-5,-1]	-0.04 %	-0.14	-0.41	-0.15	-0.07 %	-0.16	-0.12	-0.88
[-5,0]	-0.01 %	-0.04	-0.05	0.03	0.60 %	0.97	1.03	0.28
[-5,+1]	0.13 %	0.37	0.63	0.30	0.93 %	1.38	1.56	1.15
[-5,+2]	0.02 %	0.06	0.41	0.28	0.95 %	1.25	1.57	0.82
[-5,+3]	0.03 %	0.07	0.40	0.42	1.18 %	1.44	1.93*	1.12
[-5,+4]	-0.02 %	-0.06	0.31	0.20	1.62 %	1.94*	2.31**	1.47
[-5,+5]	-0.17 %	-0.37	-0.23	0.01	1.45 %	1.56	2.16**	1.02
[-5,+6]	-0.18 %	-0.37	-0.45	-0.45	1.71 %	1.90*	2.52**	1.23
[-5,+7]	-0.14 %	-0.28	-0.48	-0.31	1.70 %	1.79*	2.56**	1.21
[-5,+8]	-0.30 %	-0.58	-0.99	-0.29	1.25 %	1.30	1.86*	0.94
[-5,+9]	-0.26 %	-0.48	-0.73	0.08	1.11 %	1.07	1.69*	0.78
[-5,+10]	-0.45 %	-0.81	-1.11	-0.19	0.90 %	0.86	1.47	0.60
[-5,+11]	-0.41 %	-0.73	-1.04	-0.18	1.01 %	0.94	1.55	0.75
[-5,+12]	-0.39 %	-0.70	-0.98	-0.07	0.95 %	0.85	1.43	0.66
[-5,+13]	-0.55 %	-0.99	-1.48	-0.69	0.63 %	0.51	1.27	0.64
[-5,+14]	-0.55 %	-0.99	-1.17	-0.69	0.03 %	0.02	0.65	0.34
[-5,+15]	-0.48 %	-0.84	-1.00	-0.48	0.14 %	0.10	0.64	0.60
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+30]	-0.97 %	-1.36	-2.50**	-1.54	-1.11 %	-0.56	-0.34	-0.13
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+60]	-1.10 %	-1.13	-2.75***	-1.79*	-3.28 %	-1.43	-1.48	-1.41
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+120]	-2.11 %	-1.27	-3.31***	-1.62	-10.05 %	-2.74***	-3.05***	-2.56**

The table presents the $\overline{CAR}s$ surrounding the announcement of the reporting- and rebranding sample from 5 days prior to the announcement to 120 days post-announcement [-5,+120]. Abnormal returns are estimated using the Market Model. We include a t-test, Patell Z test, and Wilcoxon Signed Rank Test in order to assess whether or not the $\overline{CAR}s$ are statistically different from zero. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10. See appendix A.1 for the estimation techniques for each of the different test statistics.

The rebranding sample could also provide support for the theory laid out by Asness (2017) related to the lower expected returns achieved when investing in companies not affected by negative screening. A wider investor base would lower the expected returns, compared to companies classified as "sin stocks"[5], where investors gets rewarded with a higher expected return for holding such securities.

Overall, our results from the reporting sample indicate that initiation of sustainability reporting does not generate any significant abnormal returns. The same does not seem to be true for the rebranding sample, where we find evidence suggesting that a positive effect can be seen in the days following the event. We also have evidence of a reversion and a significant negative post-announcement drift over the longer event windows.

For robustness we have estimated and tested the significance of $\overline{AR}$ and $\overline{CAR}$ using the Fama-French Three Factor Model (see Appendix A.2). Model selection only marginally affect our results, with strong similarities in both estimates and significance based on our two statistical models for the normal return, consistent with the theory presented in section 4.2.2.

6.3 Change in ESG Ratings Following the Announcement

This study also provides an analysis on the actual development of the reported ESG ratings of the companies included in our rebranding sample. This development is closely aligned with hypothesis 3, where we want to see whether or not the companies actually improve their sustainable profiles. The reason behind why the reporting sample is not included can be found in section 4.6. We utilise a paired t-test for differences in means in order to assess the development. However, we are only able to test the development up to 2 years following the shift due to sample size limitations.

[5] Sin stocks in this setting refers to listed companies that a specific group of investors exclude from their investment sets due to unwanted characteristics.

From table 5, we see that there exists a significant initial positive change in the ESG ratings in the first subsequent rating received by the companies. However, this effect is not present at $t = 2$. We believe this to be a sign of the rating agency being overly optimistic in the first rating, followed by a re-adjustment in $t = 2$. This belief is also accompanied by the fact that the shifts are relatively large, meaning that the companies need time to adapt their new strategies.

Table 5: t-test for Change in ESG Ratings and Pillar Scores

T	0	1	2
Avg. $ESG\ Score_T$	61.36	64.69	64.22
$\overline{\Delta}^{ESG}_{[0,T]}$		3.33***	0.90
		(2.79)	(0.29)
Avg. $E\ Score_T$	59.06	67.74	70.06
$\overline{\Delta}^{E}_{[0,T]}$		8.68**	6.92*
		(2.59)	(1.92)
Avg. $S\ Score_T$	62.31	67.67	66.15
$\overline{\Delta}^{S}_{[0,T]}$		5.36**	2.10
		(2.49)	(0.54)
Avg. $G\ Score_T$	59.69	60.46	61.40
$\overline{\Delta}^{G}_{[0,T]}$		0.77	-0.23
		(0.42)	(-0.05)
N		29	25

The table represents the results from the paired t-test for differences in means on the change in ESG ratings relative to the last reported ESG rating prior to the shift ($t = 0$). The same analysis is also conducted on the pillar scores for the respective ESG rating. $\overline{\Delta}_{[0,T]}$ corresponds to the mean change relative to $t = 0$ for each subsequent reported rating. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10 and corresponds to a two-sided test for a change in means.

We find similar results when looking at the pillar scores. Both the environmental and the social pillar score experiences a significant increase in the first rating received following the shift, while we see no significant change in the governance score. This is an interesting finding, as it provides evidence that the increase in the overall ESG rating is driven by environmental- and/or social matters. We argue that the companies manage to improve their performance following the shift, as the rebranding cases focuses on environmental- and social matters. If we were to see significant increases in the governance scores, we argue that the increase could be driven by internal matters and ultimately not affect the negative externalities produced by the company.

An important aspect that could undermine these findings is the non-normality feature of the respective scores, in addition to the limited sample size. Following this, we have conducted a complimentary analysis using a non-parametric test, namely the Wilcoxon Signed Rank Test. This methodology enables us to test the robustness of the findings from the t-test. Additionally, we are able to test longer time horizons with smaller sample sizes as the test does not make any assumptions regarding the distribution of the scores. This analysis supported the findings in table 5, and can be found in appendix A.3.

Thus, our findings contradicts hypothesis 3 that the companies cater to investor preferences regarding sustainability, but rather improves their sustainable performance following the shift.

7 Conclusion

This thesis analyses how stock markets react to companies announcing shifts related to ESG, and how the sustainable performance of these companies are affected following the shift. We measure market reactions by analysing abnormal trading volume surrounding the event date, followed by an analysis focusing on abnormal returns generated by the shifts. We use ESG ratings and pillar scores as proxies for sustainable performance, and use these as inputs in significance testing. This analysis aims to investigate the development relative to the rating prior to the announcement of the shift. The broad aim of this thesis is to explore whether shifts in ESG can be driven by financial motives rather than sustainability itself, and is closely linked to the title of the thesis.

We first observe significant positive abnormal trading volume for both of our samples in the days surrounding the event date. This finding supports our hypothesis that these shifts leads to increased investor attention, proxied by the increased liquidity in the shares of the sample companies. Following this, we conclude that the shifts generates market reactions in terms of increased trading.

We find that our rebranding scenarios generates positive abnormal returns at the event date. On the other hand, there seems to only be some minor effects observed in the $\overline{CAR}$. Our evidence suggests a positive buildup of $\overline{CAR}$ in the days following the event, but that a reversion materialised shortly after the event date. We interpret this finding to be a sign of an initial overreaction by investors, and that the shift does not generate a permanent value increase. We find negative $\overline{CARs}$ when looking at the longer event windows, providing similarities with a negative post-announcement drift where investors are slowly incorporating the new information. We fail to find any significant evidence suggesting that sustainability reporting carries any price impact, even though we find significant abnormal trading volumes. We interpret these findings as a sign of disagreement between investors regarding the actual value impact of the initiation of sustainability reporting.

Our results are in line with findings in existing literature, such as Krüger (2014) and Fisher-Vanden & Thorburn (2011), who finds a negative relation between ESG initiatives and stock price reactions. This thesis contributes to the literature by showing that ESG initiatives fails to generate positive abnormal returns, and subsequently provides evidence that the green shift amongst listed companies are not driven by financial incentives, proxied by abnormal returns. An extension to this thesis is to dive deeper into other possible effects that these shifts can carry, which could be a subject for future research. One idea could be to look closer at the cost of capital effects to see if these could potentially provide another form of financial incentive.

We argue that the long term effects seen in the rebranding sample can be the result of the shareholder theory. Shareholders may believe that ESG initiatives destroy shareholder value, resulting in lower valuations following the announcement. Another explanation could lie in attribution theory. That is, investors might become sceptical and form negative attributions about the ulterior motive of the company in question. We argue that investors might deem these shifts as an attempt to capture the hypothesised financial motive of becoming more sustainable, catering to investor sentiment regarding sustainability. However, this hypothesis is contradicting to the findings from our analysis of the development in ESG ratings and pillar scores.

We find evidence suggesting a significant increase in the overall ESG ratings from our analysis on the development in the ESG ratings and pillar scores for the companies in our rebranding sample. We also find that this increase is driven by significant increases in the environmental- and social pillar scores. We therefore suggest that the companies are indeed improving their sustainable performance, proxied by the ratings and the pillar scores.

In conclusion, this study shows that shifts related to ESG generates a market reaction, seen from the abnormal trading volume. The shifts also carries a price impact for the rebranding sample, while the price impact for the reporting scenarios are ambiguous. We fail to find any evidence suggesting that the rebranding scenarios are an attempt to cater to investor sentiment, seen from the subsequent significant improvement in the sustainable performance of the companies in question. In other words, we believe that the colour of the philosopher's stone in modern day stock markets is still up for debate.

A Appendix

A.1 Statistical Tests

We have relied on several different statistical tests in order to test the significance of the abnormal- and cumulative abnormal returns. These are more closely discussed in the following subsections. The reasoning behind this methodology is that the different tests are designed to handle different problems that may occur following the assumptions laid out by each framework.

A.1.1 Cross-Sectional t-test

We want to test whether the $\overline{AR}s$ and $\overline{CAR}s$ are significantly different from zero, following the methodology laid out by MacKinlay (1997). We employ two-sided cross-sectional t-tests under the assumption of no overlap in the event window of the different observations in each sample. The tests will have the following statistical features:

$$\overline{AR}_t \sim N\big[0, var(\overline{AR}_t)\big] \quad \text{(A.1)}$$

$$\overline{CAR}_{(t_1,t_2)} \sim N\big[0, var(\overline{CAR}_{(t_1,t_2)})\big] \quad \text{(A.2)}$$

We are required to compute $var(\overline{AR})$ in order to estimate $var(\overline{CAR})$, following the relationship below.

$$var(\overline{CAR}_{(t_1,t_2)}) = \sum_{t=T_1}^{T_2} var(\overline{AR}_t) \quad \text{(A.3)}$$

$$var(\overline{AR}_t) = \frac{1}{N^2}\sum_{i=1}^{N} \sigma^2_{\epsilon_i} \quad \text{(A.4)}$$

The residual variance required to estimate $var(\overline{AR})$, and subsequently $var(\overline{CAR})$, is unknown. Patell (1976) showed that the residual variance from the model used to estimate the expected returns is unbiased. MacKinlay (1997) further proposed that the residual variance, $\sigma^2_{\epsilon_i}$, is an appropriate estimator when estimating equation A.4. We therefore use this estimator when computing the variance of the $\overline{AR}$.

Under the null hypothesis we have that $\overline{CAR} = 0$ and can be tested with the following test statistic:

$$\theta_1 = \frac{\overline{CAR}_{(t_1,t_2)}}{\sqrt{var(\overline{CAR}_{(t_1,t_2)}}} \sim N(0,1) \quad \text{(A.5)}$$

Similarly, we can also test $\overline{AR} = 0$ using the following formula:

$$\theta_1 = \frac{\overline{AR}_t}{\sqrt{var(\overline{AR}_t)}} \sim N(0,1) \quad \text{(A.6)}$$

The test statistics' distributional results are asymptotic, both in terms of number of securities (N), and the length of the estimation window. However, as the convergence to the asymptotic distribution is reasonably quick for the test statistic, this is usually not a large problem in event studies (MacKinlay, 1997). One modification that can improve the power of the test is to standardise each abnormal return. This standardisation is utilised in the Patell test, and discussed in the next subsection.

A.1.2 Patell Z Test

Patell (1976) presents a modification to the standard cross-sectional tests that can lead to more powerful tests. The approach is centered around standardising the abnormal returns by dividing the AR_i with the forecast-error corrected standard deviation.

$$SAR_{i,t} = \frac{AR_{i,t}}{S_{AR_{i,t}}} \quad \text{(A.7)}$$

Where $SAR_{i,t}$ is the standardised abnormal returns, and $S_{AR_{i,t}}$ is the standard deviation of the abnormal returns in the estimation window.

Patell (1976) adjust the standard error by the forecast error as the event window abnormal returns are out-of-sample predictions. The forecast-error corrected standard error is estimated with the following formula:

$$S_{AR_{i,t}} = \sqrt{S^2_{AR_i}\left(1 + \frac{1}{L_1} + \frac{\left(R_{M_t} - \overline{R}_M\right)^2}{\sum_{t=T_0}^{T_1}\left(R_{M_t} - \overline{R}_M\right)^2}\right)} \tag{A.8}$$

Where $S^2_{AR_i}$ is the unadjusted standard error, L_1 is the number of days in the estimation period, R_{M_t} is the return of the relevant market index at time t, and $\overline{R}_M$ is the mean market return in the estimation window.

Further, the test statistic for testing $\overline{AR} = 0$ is computed as:

$$Z_{Patell,t} = \frac{ASAR_t}{S_{ASAR_t}} \tag{A.9}$$

Where

$$ASAR_t = \sum_{i=1}^{N} SAR_{i,t} \tag{A.10}$$

With an expected value of zero and variance

$$S^2_{ASAR_t} = \sum_{i=1}^{N} \frac{L_1 - 2}{L_1 - 4} \tag{A.11}$$

Similarly, the test statistic for $\overline{CAR} = 0$ is given by

$$Z_{Patell} = \frac{1}{\sqrt{N}} \sum_{t=t_1+1}^{t_2} \frac{CSAR_i}{S_{CSAR_i}} \tag{A.12}$$

Where $CSAR_i$ is the cumulative $SAR_{i,t}$ over the event window for company i. $S^2_{CSAR_i}$ is computed using the following formula:

$$S^2_{CSAR_i} = L_2 \frac{L_1 - 2}{L_1 - 4} \tag{A.13}$$

With L_2 corresponding to the number of days in the event window.

Under the assumptions laid out by Patell (1976), Z_{Patell} will follow a normal distribution.

A.1.3 Wilcoxon Signed Ranked Test

The Wilcoxon Signed Rank Test is a non-parametric statistical hypothesise test for examining the difference across repeated measurements across a single sample (Wilcoxon, 1945). The benefit of adopting a non-parametric approach is to avoid specific assumptions concerning the distribution of returns (MacKinlay, 1997). For the purposes of this study, it was tested whether our two distinctive samples contained abnormal returns significantly different from the normal returns predicted using the Market Model.

The Wilcoxon Signed Rank Test considers both the magnitude and the sign of abnormal returns (Dutta, 2014). Further, the test assumes that none of the absolute values are equal and non-zero. Let

$$W_t = \sum_{i=1}^{N} rank(AR_{i,t})^+ \tag{A.14}$$

Where $rank(AR_{i,t})^+$ is the positive rank of the absolute value of the abnormal return $AR_{i,t}$ at point t for firm i.

The hypotheses of the testing framework are the following

$$H_0\colon \overline{AR} = 0 \quad vs. \quad H_1\colon \overline{AR} \neq 0$$
$$H_0\colon \overline{CAR} = 0 \quad vs. \quad H_1\colon \overline{CAR} \neq 0$$

for the $\overline{AR}$ and $\overline{CAR}$ respectively. The test statistic is defined as

$$Z_{Wilcoxon,t} = \frac{W - \mu_W}{\sigma_W} \tag{A.15}$$

Where μ_W is the expected value of the W statistic, and computed as

$$\mu_W = \frac{n(n+1)}{4} \tag{A.16}$$

The standard deviation of the W statistic is given by

$$\sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}} \tag{A.17}$$

A.2 Robustness: Market Model vs. Fama-French

We also need to determine whether the model used to estimate the expected return series would affect our statistical inferences, in the interest of robustness. We have conducted both a paired t-test for differences in means between the two statistical models for both abnormal- and cumulative abnormal returns.

Additionally, we have performed significance testing on the $\overline{CARs}$ using and OLS framework to see if model selection affects the significance.

A.2.1 Paired t-test: MM and FF3

First, we utilise a paired t-test in order to determine whether or not there exists a significant difference between the two estimation techniques for the abnormal returns. A paired testing framework is used as the observations corresponds to the same company on the same day in both samples. Our hypothesis is the following:

$$H_0 \colon \mu_d = 0 \quad vs. \quad H_1 \colon \mu_d \neq 0$$

Where

$$\mu_d = \frac{\sum_{i=1}^{N} \left(AR_{MM_{i,t}} - AR_{FF3_{i,t}} \right)}{N} \tag{A.18}$$

The test statistic is given by:

$$t = \frac{\overline{x}_d}{\frac{\sigma_d}{\sqrt{N}}} \tag{A.19}$$

$\overline{x}_d$ corresponds to the average difference between the two estimated abnormal returns at time t, and σ_d is the standard deviation of the differences.

The same test is also applied to the $CARs$ with the equivalent procedure. Results for the abnormal- and cumulative abnormal returns for both of our samples are reported in table 6, and 7 respectively.

Table 6: Paired t-test $\overline{AR}$ – MM vs. FF3

	Reporting N = 297			Rebranding N = 37		
T	$\overline{AR}_{MM}$	$\overline{AR}_{FF3}$	t-test	$\overline{AR}_{MM}$	$\overline{AR}_{FF3}$	t-test
-5	0.15 %	0.14 %	0.49	-0.09 %	-0.13 %	0.40
-4	-0.11 %	-0.12 %	0.19	-0.15 %	-0.08 %	-1.03
-3	-0.24 %	-0.21 %	-1.22	-0.19 %	-0.17 %	-0.34
-2	0.07 %	0.09 %	-0.67	0.31 %	0.29 %	0.26
-1	0.09 %	0.07 %	1.14	0.05 %	0.09 %	-0.82
0	0.03 %	0.02 %	0.14	0.67 %	0.69 %	-0.25
1	0.14 %	0.22 %	-2.04**	0.33 %	0.27 %	0.95
2	-0.11 %	-0.13 %	0.61	0.02 %	0.02 %	0.07
3	0.00 %	-0.01 %	0.63	0.23 %	0.29 %	-0.78
4	-0.05 %	-0.01 %	-1.11	0.44 %	0.29 %	2.33**
5	-0.15 %	-0.16 %	0.38	-0.17 %	-0.12 %	-0.83
6	0.00 %	0.02 %	-1.12	0.26 %	0.22 %	0.58
7	0.04 %	0.04 %	-0.32	-0.01 %	-0.08 %	1.03
8	-0.16 %	-0.17 %	0.66	-0.45 %	-0.42 %	-0.72
9	0.04 %	0.00 %	1.44	-0.14 %	-0.19 %	0.89
10	-0.19 %	-0.19 %	0.08	-0.21 %	-0.05 %	-2.02*
11	0.04 %	0.06 %	-0.74	0.11 %	0.14 %	-0.38
12	0.02 %	0.00 %	0.93	-0.06 %	0.04 %	-1.90*
13	-0.16 %	-0.16 %	0.40	-0.32 %	-0.27 %	-0.79
14	-0.01 %	-0.03 %	0.87	-0.60 %	-0.58 %	-0.12
15	0.07 %	-0.19 %	0.97	0.11 %	0.06 %	0.57

This table reports the $\overline{AR}$ for both the Market Model and the Fama-French Three Factor Model for our two samples. The t-statistic is estimated using a paired t-test for differences in means. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10.

From table 6, we see that there are some days where we have a significant difference between the two estimated $\overline{AR}s$ surrounding the event date. However, from table 7, we see that no significant difference exists in the estimated $\overline{CAR}s$.

Table 7: Paired t-test $\overline{CAR}$ – MM vs. FF3

	Reporting N = 297			Rebranding N = 37		
Window	$\overline{CAR}_{MM}$	$\overline{CAR}_{FF3}$	t-test	$\overline{CAR}_{MM}$	$\overline{CAR}_{FF3}$	t-test
[-5,-5]	0.15 %	0.14 %	0.49	-0.09 %	-0.13 %	0.40
[-5,-4]	0.04 %	0.02 %	0.47	-0.24 %	-0.20 %	-0.33
[-5,-3]	-0.20 %	-0.19 %	-0.18	-0.43 %	-0.37 %	-0.69
[-5,-2]	-0.13 %	-0.10 %	-0.41	-0.12 %	-0.08 %	-0.36
[-5,-1]	-0.04 %	-0.04 %	0.01	-0.07 %	0.01 %	-0.64
[-5,0]	-0.01 %	-0.02 %	0.06	0.60 %	0.70 %	-0.67
[-5,+1]	0.13 %	0.20 %	-0.81	0.93 %	0.97 %	-0.27
[-5,+2]	0.02 %	0.07 %	-0.58	0.95 %	0.99 %	-0.25
[-5,+3]	0.03 %	0.06 %	-0.39	1.18 %	1.28 %	-0.50
[-5,+4]	-0.02 %	0.05 %	-0.71	1.62 %	1.57 %	0.34
[-5,+5]	-0.17 %	-0.11 %	-0.59	1.45 %	1.44 %	0.04
[-5,+6]	-0.18 %	-0.09 %	-0.79	1.71 %	1.66 %	0.21
[-5,+7]	-0.14 %	-0.04 %	-0.80	1.70 %	1.59 %	0.48
[-5,+8]	-0.30 %	-0.22 %	-0.61	1.25 %	1.17 %	0.34
[-5,+9]	-0.26 %	-0.22 %	-0.29	1.11 %	0.98 %	0.50
[-5,+10]	-0.45 %	-0.41 %	-0.27	0.90 %	0.93 %	-0.13
[-5,+11]	-0.41 %	-0.35 %	-0.37	1.01 %	1.07 %	-0.22
[-5,+12]	-0.39 %	-0.35 %	-0.23	0.95 %	1.12 %	-0.55
[-5,+13]	-0.55 %	-0.52 %	-0.18	0.63 %	0.85 %	-0.78
[-5,+14]	-0.55 %	-0.54 %	-0.07	0.03 %	0.27 %	-0.74
[-5,+15]	-0.48 %	-0.74 %	0.78	0.14 %	0.32 %	-0.62
⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+30]	-0.97 %	-1.51 %	1.23	-1.11 %	-0.65 %	-0.97
⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+60]	-1.10 %	0.61 %	-0.57	-3.28 %	-1.89 %	-1.55
⋮	⋮	⋮	⋮	⋮	⋮	⋮
[-5,+120]	-2.36 %	-0.91 %	-0.48	-10.05 %	-7.14 %	-1.57

This table reports the $\overline{CAR}$ for both the Market Model and the Fama-French Three Factor Model for our two samples from 5 days prior to 120 days after the announcement [-5,+120]. The t-statistic is estimated using a paired t-test for differences in means. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10.

However, following differences in standard errors, it still might be differences in significance based on the model specification. Please see next subsection for a comparison of significance between the Market Model and the Fama-French Three Factor Model.

A.2.2 Comparison of Significance of $\overline{CAR}s$ – MM vs. FF3

As we have based our significance tests in section 6 on the $\overline{CAR}s$ estimated using the Market Model, it is also of interest to compare the significance of the $\overline{CAR}s$ estimated using our other normal return model.

We run OLS regressions without any explanatory variables for both of our samples, and both of our normal return models in order to compare the significance. By definition, this methodology coincides with a standard t-test with $\overline{CAR} = 0$ under the null hypothesis. The $\overline{CAR}s$ for each sample are tested for different event windows in order to see whether there exists any significant differences with longer windows. The probability of seeing different conclusions due to the accumulation of abnormal returns increases alongside the event window.

Results are reported in table 8. First, we see different conclusions in the [-5,+30] event window for the reporting sample. The Market Model fails to reject the null, while the Fama-French Three Factor Model provides significant $\overline{CAR}s$ at the 10% level. In our rebranding sample, the only difference in conclusions is that the Market Model provides significant $\overline{CAR}$ in the [-5,+120] event window at the 1% level, while the Fama-French Model based estimate is only significant at the 5% level.

Following this comparison, we feel confident that our statistical inference in section 6 based on the Market Model provide us with robust conclusions. This argument is also based on the different significance tests that were conducted based on the Market Model estimates.

Table 8: OLS Significance Comparison – MM vs. FF3

Panel A: Reporting

	Market Model					**Fama-French 3 Factors**				
	$\overline{CAR}_{[-5,+1]}$	$\overline{CAR}_{[-5,+5]}$	$\overline{CAR}_{[-5,+30]}$	$\overline{CAR}_{[-5,+60]}$	$\overline{CAR}_{[-5,+120]}$	$\overline{CAR}_{[-5,+1]}$	$\overline{CAR}_{[-5,+5]}$	$\overline{CAR}_{[-5,+30]}$	$\overline{CAR}_{[-5,+60]}$	$\overline{CAR}_{[-5,+120]}$
Constant	0.13	-0.17	-0.97	-1.10	-2.36	0.20	-0.11	-1.51*	0.61	-0.91
	(0.38)	(-0.37)	(-1.38)	(-1.14)	(-1.45)	(0.59)	(-0.25)	(-1.96)	(0.19)	(-0.27)
N	297	297	297	297	296	297	297	297	297	296

Panel B: Rebranding

	Market Model					**Fama-French 3 Factors**				
	$\overline{CAR}_{[-5,+1]}$	$\overline{CAR}_{[-5,+5]}$	$\overline{CAR}_{[-5,+30]}$	$\overline{CAR}_{[-5,+60]}$	$\overline{CAR}_{[-5,+120]}$	$\overline{CAR}_{[-5,+1]}$	$\overline{CAR}_{[-5,+5]}$	$\overline{CAR}_{[-5,+30]}$	$\overline{CAR}_{[-5,+60]}$	$\overline{CAR}_{[-5,+120]}$
Constant	0.93	1.45	-1.11	-3.28	-10.05***	0.97	1.44	-0.65	-1.89	-7.14**
	(1.38)	(1.56)	(-0.56)	(-1.43)	(-2.74)	(1.5)	(1.61)	(-0.36)	(-0.88)	(-2.07)
N	37	37	37	37	37	37	37	37	37	37

This table presents the results from regressing $\overline{CAR}s$ only on constants in order to conduct a t-test for significance. By definition, constants corresponds to the percentage $\overline{CAR}s$ across the event windows. The $\overline{CAR}s$ included are in the range of 5 days prior to the announcement, to 120 days after. Number of companies (N) are reported for each of the event windows. One observation is dropped in the longest event window in our reporting sample due to insufficient data. The t-statistics are reported in parentheses, and significance is denoted as ***p<0.01, **p<0.05, and *p<0.10.

A.3 Wilcoxon Signed Rank Test for Changes in ESG Ratings and Pillar Scores

We also utilise a non-parametric test on the changes in the ESG ratings and the related pillar scores due to the aforementioned non-normality feature and limited sample sizes. We employ a Wilcoxon Signed Rank Test to test for any differences in the median following the shift. Another desirable feature of this test is that we are able to test the median differences, compared to the average differences that we used in the paired t-test in section 6.3. We utilise the effect size developed by Cohen (1988), where effect sizes above 0.3 corresponds to a moderate change in order to assess the impact of the increase.

From table 9, we see that we obtain similar results as the t-test in the first two subsequent ESG ratings after the shift, with significant increases in the first year and insignificant changes in $t = 2$. Following $t = 2$ our sample size becomes limited, but we observe significant positive changes in the ESG ratings. This finding is interpreted as the companies starting to adopt and employ their new strategic directions.

Similar patterns are also seen in the pillar scores, where both the environmental- and social scores increases after the shift. As with the t-test, no significant changes can be seen in the governance scores.

Thus, the non-parametric test supports our findings from the t-test in section 6.3. The longer time horizons also show signs of a continuous increase from $t = 2$. We believe this to be a subject for future research when more data becomes available in order to obtain more robust findings, due to the very limited sample sizes.

Table 9: Wilcoxon Signed Rank Test for Change in ESG Ratings and Pillar Scores

T	1	2	3	4	5
			ESG Rating		
$\widetilde{\Delta}^{ESG}_{[0,T]}$	1.59***	1.38	5.93***	8.43**	7.99***
	(2.36)	(1.18)	(2.47)	(1.89)	(1.83)
r^{ESG}_T	0.31	0.17	0.45	0.37	0.41
			Environmental Pillar Score		
$\widetilde{\Delta}^{E}_{[0,T]}$	1.88**	1.36	6.64*	4.57	5.57
	(2.41)	(1.12)	(1.87)	(0.94)	(0.87)
r^{E}_T	0.32	0.16	0.34	0.19	0.19
			Social Pillar Score		
$\widetilde{\Delta}^{S}_{[0,T]}$	2.81**	3.16	10.12**	10.41**	7.37*
	(2.07)	(0.93)	(2.16)	(2.20)	(1.68)
r^{S}_T	0.27	0.13	0.39	0.43	0.38
			Governance Pillar Score		
$\widetilde{\Delta}^{G}_{[0,T]}$	0.97	-0.02	4.47	9.25	4.34
	(0.49)	(-0.12)	(0.80)	(1.08)	(1.48)
r^{G}_T	0.06	0.02	0.15	0.21	0.33
N	29	25	15	13	10

The table represents the results from the Wilcoxon Signed Rank Test on the change in ESG ratings relative to the last reported ESG rating prior to the shift ($t = 0$). The same analysis is also conducted on the pillar scores for the respective ESG ratings. $\widetilde{\Delta}_{[0,T]}$ corresponds to the median change relative to $t = 0$ for each subsequent reported rating. Significance is denoted as ***p<0.01, **p<0.05, and *p<0.10 and corresponds to a two-sided test for change in the median difference. We utilise an exact estimation technique for the p-values, due to the non-normality feature of the ESG ratings. Z-statistics are reported in parentheses, and used to estimate the effect size, r, for each respective rating. Effect size is estimated as $r = Z/\sqrt{N}$, where N corresponds to the number of observations.

References

Ajinkya, B. B., & Jain, P. C. (1989). The Behavior of Daily Stock Market Trading Volume. *Journal of Accounting and Economics*, *11*(4), 331–359. doi:10.1016/0165-4101(89)90018-9

Armitage, S. (1995). Event Study Methods and Evidence on Their Performance. *Journal of Economic Surveys*, *9*(1), 25–52. doi:https://doi.org/10.1111/j.1467-6419.1995.tb00109.x

Asness, C. (2017, May. 18). *Virtue Is its Own Reward: Or, One Man Ceiling Is Another Man Floor.* Retrieved from `https://www.aqr.com/Insights/Perspectives/Virtue-is-its-Own-Reward-Or-One-Mans-Ceiling-is-Another-Mans-Floor`

Baker, M., & Wurgler, J. (2007). Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, *21*(2), 129–152. doi:10.1257/jep.21.2.129

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, *6*(2), 159–178. doi:10.2307/2490232

Bamber, L. S. (1986). The Information Content of Annual Earnings Releases: A Trading Volume Approach. *Journal of Accounting Research*, *24*(1), 40–56. doi:10.2307/2490803

Baron, D. P., Harjoto, M. A., & Jo, H. (2011). The Economics and Politics of Corporate Social Performance. *Business and Politics*, *13*(2), 1–46. doi:10.2202/1469-3569.1374

Beaver, W. H. (1968). The Information Content of Annual Earnings Announcements. *Journal of Accounting Research*, *6*, 67–92. doi:10.2307/2490070

Bernard, V. L., & Thomas, J. K. (1989). Post-Earnings-Announcement Drift: Delayed Price Response or Risk Premium? *Journal of Accounting Research*, *27*, 1–36. doi:10.2307/2491062

Bernard, V. L., & Thomas, J. K. (1990). Evidence that Stock Prices do not Fully Reflect the Implications of Current Earnings for Future Earnings. *Journal of Accounting and Economics*, *13*(4), 305–340. doi:10.1016/0165-4101(90)90008-R

Bodie, Z., Kane, A., & Marcus, A. J. (2019). *Essentials of Investments* (11th ed.). New York, NY: McGraw-Hill Education.

Bolton, P., & Kacperczyk, M. T. (2020). Do Investors Care about Carbon Risk? *Journal of Financial Economics*. doi:10.2139/ssrn.3398441

Bosch, J. C., & Hirschey, M. (1989). The Valuation Effects of Corporate Name Changes. *Financial Management*, *18*(4), 64–73. doi:10.2307/3665798

Bowman, R. (2006). Understanding and Conducting Event Studies. *Journal of Business Finance & Accounting*, *10*(4), 561–584. doi:10.1111/j.1468-5957.1983.tb00453.x

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate Social Performance and Stock Returns: UK Evidence from Disaggregate Measures. *Financial Management*, *35*(3), 97–116. doi:10.2139/ssrn.739587

Brooks, C. (2019). *Introductory Econometrics for Finance* (4th ed.). Cambridge: Cambridge University Press. doi:https://doi.org/10.1017/9781108524872

Brown, S., & Weinstein, M. I. (1985). Derived Factors in Event Studies. *Journal of Financial Economics*, *14*(3), 491–495. doi:https://doi.org/10.1016/0304-405X(85)90010-8

Business Roundtable. (2019, Aug. 19). *Statement on the Purpose of a Corporation.* Retrieved from `https://www.businessroundtable.org/business-roundtable-redefines-the-purpose-of-a-corporation-to-promote-an-economy-that-serves-all-americans`

Campbell, C. J., & Wasley, C. E. (1993). Measuring Security Price Performance Using Daily NASDAQ Returns. *Journal of Financial Economics*, *33*(1), 73–92. doi:10.1016/0304-405X(93)90025-7

Campbell, C. J., & Wasley, C. E. (1996). Measuring Abnormal Daily Trading Volume for Samples of NYSE/ASE and NASDAQ Securities Using Parametric and Nonparametric Test Statistics. *Review of Quantitative Finance and Accounting*, *6*(3), 309–326. doi:10.1007/BF00245187

Clark, G. L., Feiner, A., & Viehs, M. (2015). From the Stockholder to the Stakeholder: How Sustainability Can Drive Financial Outperformance. *Organizations & Markets: Formal & Informal Structures eJournal*. doi:10.2139/ssrn.2508281

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, N.J: Lawrence Erlbaum Associates. doi:https://doi.org/10.4324/9780203771587

Cooper, M. J., Dimitrov, O., & Rau, P. R. (2000). A Rose.Com by Any Other Name. *Journal of Finance*, *56*(6), 2371–2388. doi:10.2139/ssrn.242376

Cready, W. M., & Ramanan, R. (1991). The Power of Tests Employing Log-Transformed Volume in Detecting Abnormal Trading. *Journal of Accounting and Economics*, *14*(2), 203–214. doi:10.1016/0165-4101(91)90005-9

Dechant, K., Altman, B., Downing, R. M., Keeney, T., Mahoney, M., Swaine, A., ... Post, J. (1994). Environmental Leadership: From Compliance to Competitive Advantage [and Executive Commentary]. *The Academy of Management Executive (1993-2005)*, *8*(3), 7–27. doi:https://doi.org/10.5465/ame.1994.9503101163

Delmas, M. A., & Cuerel Burbano, V. (2011). The Drivers of Greenwashing. *California Management Review*, *54*(1), 64–87. doi:10.1525/cmr.2011.54.1.64

Doan, T., & Lo, A. (1988). Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test. *Review of Financial Studies*, *1*(1), 41–66. doi:https://doi.org/10.1093/rfs/1.1.41

Dolley, J. C. (1933). Open Market Buying as a Stimulant for the Bond Market. *Journal of Political Economy*, *41*(4), 513–529. doi:10.1086/254507

Dutta, A. (2014). Parametric and Nonparametric Event Study Tests: A Review. *International Business Research*, *7*(12), 136–142. doi:10.5539/ibr.v7n12p136

Eccles, R., Ioannou, I., & Serafeim, G. (2014). The Impact of Corporate Sustainability on Organizational Processes and Performance. *Management Science*, *60*(11), 2835–2857. doi:10.1287/mnsc.2014.1984

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383–417. doi:10.2307/2325486

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review*, *10*(1), 1–21. doi:10.2307/2525569

Fama, E. F., & French, K. R. (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, *33*(1), 3–56. doi:10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2021). Europe 3 Factors Daily [Data set]. Retrieved from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Ferrell, A., Liang, H., & Renneboog, L. (2016). Socially Responsible Firms. *Journal of Financial Economics*, *122*(3), 585–606. doi:10.1016/j.jfineco.2015.12.003

Fisher-Vanden, K., & Thorburn, K. (2011). Voluntary Corporate Environmental Initiatives and Shareholder Wealth. *Journal of Environmental Economics and Management*, *62*(3), 430–445. doi:https://doi.org/10.1016/j.jeem.2011.04.003

Flammer, C. (2015). Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach. *Management Science*, *61*(11), 2549–2568. doi:http://dx.doi.org/10.1287/mnsc.2014.2038

Foster, G. (1973). Stock Market Reaction to Estimates of Earnings per Share by Company Officials. *Journal of Accounting Research*, *11*(1), 25–37. doi:10.2307/2490279

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach.* Cambridge: Cambridge University Press. doi:10.1017/CBO9781139192675

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from More than 2000 Empirical Studies. *Journal of Sustainable Finance & Investment*, *5*(4), 210–233. doi:10.1080/20430795.2015.1118917

Friedman, M. (1962). *Capitalism and Freedom.* University of Chicago Press. doi:https://doi.org/10.1086/258820

Friedman, M. (1970, Sept. 13). *A Friedman Doctrine: The Social Responsibility Of Business Is to Increase Its Profits.* Retrieved from https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

Gerard, B. (2019). Esg and Socially Responsible Investment: A Critical Review. *Beta*, *33*(1), 61–83. doi:10.18261/issn.1504-3134-2019-01-05

Hall, S. G., & Asteriou, D. (2015). *Applied Econometrics* (3rd ed ed.). Basingstoke: Palgrave Macmillan.

Hartzmark, S. M., & Sussman, A. B. (2019). Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows. *Journal of Finance*, *74*(6), 2789–2837. doi:10.2139/ssrn.3016092

Henderson, G. V. (1990). Problems and Solutions in Conducting Event Studies. *The Journal of Risk and Insurance*, *57*(2), 282–306. doi:10.2307/253304

Hong, H., & Kacperczyk, M. (2009). The Price of Sin: The Effects of Social Norms on Markets. *Journal of Financial Economics*, *93*(1), 15–36. doi:10.2139/ssrn.766465

Karpoff, J. M., & Rankine, G. (1994). In Search of a Signaling Effect: The Wealth Effects of Corporate Name Changes. *Journal of Banking & Finance*, *18*(6), 1027–1045. doi:10.1016/0378-4266(94)00058-1

Krüger, P. (2014). Corporate Goodness and Shareholder Wealth. *Journal of Financial Economics*, *115*(2), 304–329. doi:10.2139/ssrn.2287089

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. doi:10.2307/1924119

MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, *35*(1), 13–39.

Malkiel, B. G. (2003). The Efficient Market Hypothesis and Its Critics. *Journal of Economic Perspectives*, *17*(1), 59–82. doi:10.1257/089533003321164958

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. doi:10.2307/2975974

Morse, D. (1981). Price and Trading Volume Reaction Surrounding Earnings Announcements: A Closer Examination. *Journal of Accounting Research*, *19*(2), 374–383. doi:10.2307/2490871

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, *34*(4), 768–783. doi:10.2307/1910098

Nyilasy, G., Gangadharbatla, H., & Paladino, A. (2014). Perceived Greenwashing: The Interactive Effects of Green Advertising and Corporate Environmental Performance on Consumer Reactions. *Journal of Business Ethics*, *125*(4), 693–707. doi:10.1007/s10551-013-1944-3

Patell, J. M. (1976). Corporate Forecasts of Earnings Per Share and Stock Price Behavior: Empirical Test. *Journal of Accounting Research*, *14*(2), 246–276. doi:10.2307/2490543

Peterson, P. P. (1989). Event Studies: A Review of Issues and Methodology. *Quarterly Journal of Business and Economics*, *28*(3), 36–66.

Pfeffer, J. (1995). Producing Sustainable Competitive Advantage Through the Effective Management of People. *Academy of Management Perspectives*, *9*(1), 55–72. doi:10.5465/AME.1995.9503133495

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory*, *13*(3), 341–360. doi:10.1016/0022-0531(76)90046-6

Sharpe, W. F. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, *19*(3), 425–442. doi:10.2307/2977928

Shrivastava, P., & Hart, S. (1995). Creating Sustainable Corporations. *Business Strategy and the Environment*, *4*(3), 154–165. doi:https://doi.org/10.1002/bse.3280040307

Smith, N. (2003). Corporate Social Responsibility: Not Whether, But How. *California Management Review*, *45*(4), 52–76. doi:10.2307/41166188

Sternberg, E. (1997). The Defects of Stakeholder Theory. *Corporate Governance: An International Review*, *5*(1), 3–10. doi:https://doi.org/10.1111/1467-8683.00034

Treynor, J. L. (1961). Market Value, Time, and Risk.
doi:10.2139/ssrn.2600356

Treynor, J. L. (1962). Jack Treynor's 'Toward a Theory of Market Value of Risky Assets'.
doi:10.2139/ssrn.628187

Van Horne, J. C., & Parker, G. G. C. (1967). The Random-Walk Theory: An Empirical Test. *Financial Analysts Journal*, *23*(6), 87–92. doi:https://doi.org/10.2469/faj.v23.n6.87

Waddock, S. A., & Graves, S. B. (1997). The Corporate Social Performance-Financial Performance Link. *Strategic Management Journal*, *18*(4), 303–319. doi:https://doi.org/10.1002/(SICI)1097-0266(199704)18:43.0.CO;2-G

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, *1*(6), 80–83. doi:10.2307/3001968