BI

WISEflow
Europe/Oslo(CEST)
29 Jun 2022

Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	16-01-2022 09:00	**Termin:**	202210
Sluttdato:	01-07-2022 12:00	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202210||10936||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Julie Høiestad Stenberg og Julie Bergh Andersen

Informasjon fra deltaker

Tittel *:	The Exclusionary Effect on The Stock Performance
Navn på veileder *:	Leon Bogdan Stacescu

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	228
Andre medlemmer i gruppen:	

THE EXCLUSIONARY EFFECT ON THE STOCK PERFORMANCE

An Empirical Study of The Exclusionary Effect on The Stock Performance of Companies Excluded by The Norwegian Government Pension Fund Global.

Norwegian
Business School

MASTER THESIS

by

Julie Bergh Andersen & Julie Høiestad Stenberg

Supervised by Leon Bogdan Stacescu

Oslo, June 29. 2022

Master of Science in Business, Major in Finance

BI Norwegian Business School

ABSTRACT

This study examines the direct effect on stock performance following exclusion announcements made by the Norwegian Government Pension Fund Global within a short- and long-term horizon. Through an event study, we provide evidence that market participants perceive the exclusion from the pension fund negatively, as the cumulative abnormal returns (CAR) are significantly less than zero. To further elaborate on our results, we include mean-differences tests between three subsample splits: developed- and emerging markets, product- and conduct-based exclusion criteria, and small and large firms' market capitalisation. Our findings show that all mean-differences tests give significant results, suggesting that extraneous variables have individual effects on CAAR. Additionally, we have separately estimated the change in systematic- and unsystematic risk to examine the exclusion effects in the long term. Our findings show no significant increase or decrease in risk after an exclusion. Since we use change in risk as an estimator for change in rate of return, we conclude that exclusion has no detrimental impact on stock performance in the long run.

ACKNOWLEDGEMENTS

First and foremost, we would like to thank our supervisor, Leon Bogdan Stacescu, at BI Norwegian Business School. During the whole process, he has been incredibly helpful, contributing with his expertise in the field beyond all expectations. Additionally, we want to thank Matthew Bowler for his great input on economic models. Lastly, we highly appreciate all who have been supportive throughout our years at BI and have contributed with interesting perspectives for this master thesis to be realised.

CONTENTS

LIST OF ABBREVIATIONS

AAR	Average Abnormal Returns
APT	Arbitrage Pricing Theory
AR	Abnormal Returns
AUM	Assets Under Management
CAAR	Cumulative Average Abnormal Returns
CAPM	Capital Asset Pricing Model
CAR	Cumulative Abnormal Returns
CL	Climate Leaders
CSR	Corporate Social Responsibility
EMH	Efficient Market Hypothesis
EPS	Earnings Per Share
ESG	Environment, Social, Governance
ESM	Event Study Methodology
GSIA	Global Sustainable Investment Alliance
GPFG	Government Pension Fund Global
HML	High Minus Low
ISO	International Organization for Standardisation
M&A	Merger and Acquisition
MPT	Modern Portfolio Theory
NGPFG	Norwegian Government Pension Fund Global
S&P 500	Standard and Poor's 500
SIC	Standard Industrial Classification
SMB	Small Minus Big
SRI	Socially Responsible Investments
SWF	Sovereign Wealth Fund

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION & MOTIVATION

During the past decades, academia and practitioners have turned their attention toward sustainability and socially responsible investments (SRI). As a result, SRI has gained considerable momentum in financial markets. The Global Sustainable Investment Alliance (GSIA) discloses that global SRI had a total value of USD 35.3 trillion in five major markets[1] in early 2020, representing a 15% increase over two years (2018-2020). Global sustainable assets under management (AUM) made up 35.9% of the total AUM, a 33.4% rise over the same period (GSIA, 2021). Given its growing prominence, it is surprising that investors still have no clear consensus on what constitutes sustainable assets. For clarity, we employ an inclusive definition of SRI used by Renneboog et al. (2008a, b) and Scholtens & Sievänen (2013). They describe SRI as an approach considering environmental, social, and governance (ESG) factors and ethical criteria in the management and selection of portfolios. However, ESG is not a one-size-fits-all approach, and how to define and measure ESG is still a matter of debate.

This rapid development within sustainable investments coincides with the idea that investors are indirectly accountable for the businesses they fund and their potential wrongdoings. Such obligations especially apply to public asset owners like Sovereign Wealth Funds (SWFs), as they invest significant amounts of state-owned assets to benefit the public (Richardson, 2011). Moreover, SWFs' relation with their beneficiaries differs substantially from mutual funds. Beneficiaries of public asset owners do not have the same opportunity to exit if they disagree with the fund's objectives (Clark, 2004), making SWFs as an investor group subject to investigation. Both the Norwegian Government Pension Fund Global (NGPFG) and the Swedish AP-funds have made headlines on several occasions. One prominent example is the attack of the NGPFG for owning shares in the mining company POSCO, accused of human rights violations in India (Hoepner & Schopohl, 2018).

[1] Five major markets: Europe, the US, Japan, Canada, and Australasia (Australia and New Zealand).

A common method among large institutional investors to avoid public disapproval is employing exclusionary screenings, one of the most used SRI approaches (USD 15.9 trillion). This strategy restricts investors' investment universe as they divest from companies deemed unethical or unsustainable (GSIA, 2021). In the NGPFG's work to maintain its role as a pioneer within SRI, the fund publishes its exclusion list with complete information regarding their decisions. Additionally, the chief executive of the fund, Nicolai Tangen, told the parliament of Norway in May 2021 that they would screen all 500-600 companies added to their reference index each year (Milne, 2021). As one of the largest SWFs worldwide and a forerunner within SRI, the NGPFG's choices strongly affect global asset owners. Hence, exclusion announcements by the fund are typically followed by several other investors, leading to an even larger negative demand shock for these stocks in the market (Bengtsson, 2008; Scholtens & Sievänen, 2013). Such domino effects are likely not seen for SRI mutual funds, interpreting the Norwegian GPFG's exclusions of greater importance to the financial markets and the excluded firms themselves.

As sustainable investing has entered mainstream asset management, the link between financial success and societal good is of great importance. With this thesis, we wish to contribute to this discussion. What especially caught our interest is the responsibility that large institutional investors like SWFs carry. We also find the strong signalling effect an exclusion decision by the NGPFG has on other investors as thought-provoking. As a result, our study attempts to answer the research question: *How does the exclusion announcement by the NGPFG affect the targeted firm's stock performance within a short- and long horizon?*

The research question is addressed through an event study by first and foremost examining the excluded firms' cumulative abnormal return (CAR) around the announcement date. The abnormal return is the difference between a stock's actual and expected return, where the latter is obtained through the capital asset pricing model (CAPM). Second, we study the potential change in systematic risk by measuring the beta estimates before and after the exclusion. This is captured using an interaction term between the market excess return and a dummy variable, taking the value of 1 on- and after the event date and 0 if no event occurs. A similar approach is applied for the final analysis, as we compare the equality of the regression residual variances before and after the exclusion. This approach allows us to examine whether the unsystematic risk has increased due to exclusion.

It is noteworthy that extensive research has been done on the topic. Nevertheless, no studies have our exact angle. Most of the published literature on exclusionary screening primarily concentrates on fund performance. For instance, Hoepner & Schopohl (2018) argue that the approach neither worsens nor improves the fund's financial returns. Other studies find that sin stocks (tobacco, alcohol, gaming) outperform comparables and that exclusionary screening harms the fund performance (Hong & Kacperczyk, 2009; Statman & Glushkov, 2009). Furthermore, several studies only emphasise the sector-based criteria in their analysis of exclusionary screening, which also applies to the recognised study by Hong & Kacperczyk (2009).

However, we want to study the direct effect of the exclusion announcements on the stock performance and how the market interprets this new information. This includes examining the short-term abnormal returns due to stock price fluctuations after the announcement. For the long-term effects, we study the potential change in systematic and unsystematic risk resulting from lower demand and limited risk-sharing. Additionally, we do not only want to examine sector-based exclusions but also include norm-based exclusions in our estimations. Studying both these exclusion reasons is a natural extension as we use the NGPFG's exclusion list. Our original sample included all (151) firms on the Norwegian GPFG's blacklist from 2005 to 2022. However, to enhance the validity of our analysis, we omit firms that did not apply to our study, leaving us with a total of 128 excluded companies.

In line with other event studies on similar topics (e.g., Cañón-de-Francia & Garcés-Ayerbe, 2009; Fisher-Vanden & Thorburn, 2011), we find the exclusion announcements lead to a significantly negative CAAR. More specifically, -1.42% and -2.27% over the event windows $[-1,1]$ and $[-2,2]$, respectively. These results suggest that investors interpret the exclusion as unfortunate news for the companies excluded. We also analyse the mean differences of CAAR across three subsamples: developed- and emerging markets, product- and conduct-based exclusion criteria, and small and large firms' market capitalisation. Our findings show that all mean-differences tests give significant results, suggesting that extraneous variables have individual effects on CAAR.

We anticipated that systematic and unsystematic risk would increase for the long-term effects due to exclusion. However, both tests gave insignificant results. For the systematic risk, we observe a slightly negative change in the beta estimate, thus deviating from our predictions. Following Hong & Kacperczyk (2009), there should be a decrease in demand for sinful stocks, as numerous other investors would follow the NGPFG's exclusion decision. This would then lead to lower prices and thus increase the required rate of return. Their results are, among others, based on Merton's recognised market segmentation theory (1987), examining the price implications of limited risk-sharing induced by SRI.

Additionally, we predicted that firm-specific risk factors would increase due to the exclusion of the unsystematic risk. For example, Ilhan et al. (2021) prove that uncertainty related to climate policy is priced into the options market and that it is more expensive for carbon-intense business models to protect against downside tail risks. However, there is no significant change in sigma estimates when evaluating the firms excluded by the NGPFG. Even though the results show a slight difference in both systematic and unsystematic risk, the high p-values show that random variations besides the exclusion can be an explanation.

For the rest of the study, the structure is as follows: "Literature Review", which provides an overview of previous research on the effects of exclusionary screening. The "Theory & Hypotheses" section addresses relevant financial theories regarding our research question and the development of our hypotheses. The "Methodology" section describes the empirical methodology, while the "Data" section describes the necessary inputs in our model and how we collected them. Under "Results & Analysis", we present the empirical results and a discussion. Lastly, our main findings are summarised in the "Conclusion" section.

2.0 LITERATURE REVIEW

In light of the growing interest in sustainable finance, several studies investigate SRI and its effect on fund performance and security prices. This section reviews important studies related to this topic and our research question.

2.1 EXCLUSIONARY SCREENING

SRI includes different approaches. Though, this study is restricted to examining the exclusionary screening strategy. The strategy has evolved from initially consisting of sector-based exclusions (e.g., alcohol, tobacco, gaming) to also considering norm-based exclusions (e.g., violations of human rights, corruption) (*Norm-Based Exclusions*, 2012). Therefore, published literature mainly focuses on sector-based exclusions, including the recognised paper by Hong & Kacperczyk (2009). They prove that sectors related to alcohol, gambling, and tobacco, so-called sin stocks, are less likely to be held by norm-constrained investors. Their findings show that US sin stocks outperform otherwise comparable stocks, conclusively leading institutional investors to pay a financial cost by excluding these sectors. Other studies have agreed or disagreed with their findings by including additional sectors in the screening. E.g., Kempf & Osthoff (2007) find insignificant outperformance of sin stocks relative to comparables, while Statman & Glushkov (2009) find similar results to Hong & Kacperczyk (2009). On the other hand, more recent studies like Hoepner & Schopohl (2018) review both sector- and norm-based exclusions, which is more aligned with our study on the NGPFG's exclusions. They conclude that excluded firms neither under -nor outperform relative to the fund benchmark.

Several existing studies evaluate other approaches besides exclusions, and there is a consensus within SRI literature that exclusionary screening is an outdated technique (Hoepner & Schopohl, 2018). For instance, Sparkes & Cowton (2004) argue that SRI has adopted a more sophisticated strategy through active ownership and positive screening. This statement is in line with Chen et al. (2020), studying how institutional shareholders use active ownership and monitoring to influence the Corporate Social Responsibility (CSR) of firms included in their portfolios. The study evaluates the level of institutional ownership and the degree of attention investors provide to affect CSR policies. However, exclusionary screening is still one of the most employed SRI approaches by large institutional investors, which favours the need for additional research on this specific strategy.

2.2 FUND- & STOCK PERFORMANCE

The documented effect of SRI on mutual fund performance is diverse. One notable example is the analysis by Renneboog et al. (2008a), evaluating how SRI and stakeholder governance impacts risk-adjusted return and whether SRI funds obtain a superior return in the market. The study measures the risk and returns characteristics of nearly all SRI mutual funds globally and how they perform in relation to domestic benchmark portfolios and conventional funds. They find that SRI funds strongly underperform and, on average, accumulate a risk-adjusted return between –2.2% and –6.5% per annum. These results imply that firms included in SRI funds, satisfying high ethical standards and strict stakeholder governance criteria, are overpriced, i.e., investors paying a fee for SRI. A drawback of this study is the SRI fund categorisation. Given the wide variety in screening criteria, the fund classification can be conflicting. El Ghoul & Karoiu (2017) solved this by ignoring distinct fund categories. By sampling 2,168 US equity funds without considering SRI, they found similar results as Renneboog et al. (2008b), arguably that high-CSR funds achieve a weaker return than their low-CSR counterparts.

These results are aligned with findings from Hong & Kacperczyk (2009), who argue that US sin stocks tend to offer superior financial performance. Due to the exclusion from SRI funds, sin stocks are considerably under-priced. Consequently, investors can forgo profitable investment opportunities by limiting their investment universe. Following this theory, it is expected that the Norwegian GPFG would see a decrease in return by excluding sin stocks from their portfolio. Contradictory, Renneboog et al. (2008a) argue that the risk-adjusted return of SRI funds does not significantly differ from conventional funds. This argument is backed up by Hoepner & Schopohl's (2018) study on the trade-off between SRI and financial gains.

Most studies apply screening criteria to a predefined investment universe and can thus freely reallocate or exclude companies from the portfolio. However, employ the actual NGPFG's announced exclusion dates and historical data on stock prices. The screening criteria method is used in, e.g., Chen et al. (2020) and Statman & Glushkov (2009), who measure the performance based on a fictive portfolio. Even though the approach allows for isolating and testing the impact of exclusion, it is debatable if the fictive circumstances would apply to a real-world scenario (Hoepner & Schopohl, 2018). Adamsson & Hoepner (2015) argue that the significant excess return of sin stocks found in previous studies would vanish if only stocks that are liquid enough to be considered by large institutional investors were included. In their research, they try to re-examine whether the sin stock premium recorded by Hong & Kacperczyk (2009) can be achieved by only including stocks in global equity index benchmarks.

2.3 EVENT STUDIES, ENVIRONMENTAL PERFORMANCE

Our paper differs from many studies on exclusionary screening in terms of methodology. Since we also want to capture the short-term effect, we have conducted an event study. However, this method has been applied in research on the opposite scenario, such as companies joining different "green initiatives" or achieving different certifications. A standard prediction in these studies is that investors reward information suggesting a future increase in firm value and punish the contrary (Fisher-Vanden & Thorburn, 2011). This expectation is consistent with a study by Klassen & McLaughlin (1996). They find that improved environmental performance news results in positive abnormal returns, and environmental crisis news results in negative abnormal returns. However, other event studies argue the opposite. Cañón-de-Francia & Garcés-Ayerbe (2009) finds that firms that announce their ISO[2] 14001 certification experience a drop in the stock price. One would believe that this is a positive thing for the company. Regardless, the authors explain this as investors perceiving the resources required to comply with the certification as costly to the firm, with little or no offsetting benefit.

Fisher-Vanden & Thorburn (2011), with their event study on Voluntary Corporate Environmental Initiatives and Shareholder Wealth, found mixed results. They estimate the cumulative abnormal stock returns for firms announcing participation in one of two voluntary environmental programs. The companies announcing membership in Climate Leaders (CL) experience a significant drop in stock price, while companies that join Ceres experience insignificant changes. They explain this by different degrees of difficulty in assessing what the memberships imply for the firm's cost structure. While the Ceres program has a broader scope and general principles, the costs related to CL are more visible. Thus, investors interpret the membership in CL as imposing a high cost, like Cañon-de-Francia & Garcés-Ayerbe's findings.

[2] International Organization for Standardization (ISO)

2.4 CONCLUDING POINTS

After examining a considerable amount of published research on SRI in general, exclusionary screening in specific, and event studies on environmental issues, we conclude that there are gaps in the existing literature. The most differentiating aspect of our study is evaluating the effect of exclusion on the stock itself rather than the fund performance. In addition, we contribute to the current literature by analysing the exclusion effect through different time frames. Not only do we look at the immediate market reaction to an exclusion announcement, but we extend our research to include a long-term horizon. We do this by performing an event study using the Norwegian Government Pension Fund Global exclusion dates. There is literature on these aspects separately, but we have yet to find a study combining all the above factors.

Most studies only investigate sector-based criteria when studying the exclusionary screening strategy (i.e., Hong & Kacperczyk, 2009; Kempf & Osthoff, 2007; Statman & Glushkov, 2009). We wish to extend this research by including norm-based standards. Further, we find that existing literature focuses on the investor's performance (e.g., mutual funds, constructed portfolios), such as Renneboog et al. (2008a) and Hoepner & Schopohl (2018). In contrast, we aim to evaluate how an exclusion announcement affects the targeted company by studying the stock price pre, -during, -and post exclusion. Following the studies by Adamsson & Hoepner (2015) and Hong & Kacperczyk (2009), sin stocks are frequently excluded by large norm-constrained investors, significantly affecting the company's cost of capital. Due to the additional risk, sin stocks have a higher expected return than comparables. These findings will be used to establish our hypotheses.

3.0 THEORY & HYPOTHESES

This section examines relevant financial theories in the context of our study. Based on these frameworks and previous research assessed in section 2, we develop the hypotheses that must be tested to answer our research questions.

3.1 THEORETICAL FRAMEWORKS

3.1.1 Efficient Market Hypothesis

The event study methodology assumes efficient markets, making the Efficient Market Hypothesis (EMH) a central theory in our study. Fama (1970) defines efficient markets as all available information is fully reflected in the security prices and that prices adjust immediately to new information without delay. The EMH thus emphasises that it is impossible for an investor to consistently "beat the market". This assertion means that additional fundamental or technical analysis will not contribute to achieving greater returns than other randomly selected portfolios (Malkiel, 2003). The theory has evolved over time, and a distinction has been introduced between three forms of efficient markets based on different information sets: weak, semi-strong, and strong. The weak form indicates that security prices reflect all historical information but leaves open that fundamental analysis can contribute to yielding abnormal returns. The semi-strong form implies that historical and public information is reflected in the price, precluding that any additional analysis contributes to abnormal profits. The strong form states that public -and inside information are reflected in the prices, meaning that not even market participants with privileged information can exploit it (Malkiel, 1989).

3.1.2 Modern Portfolio Theory

In our study on exclusionary screening, we cannot avoid discussing the portfolio allocation in pension funds. In general, the aim is to combine returns with an acceptable level of risk to generate long-term value. The Modern Portfolio Theory (MPT) by economist Harry Markowitz provides a framework to create a portfolio that maximises returns given the investor's risk preference. Markowitz argues that the key is diversification. MPT states that a stock's risk and return should not be evaluated individually but by how it affects the combined portfolio. By constructing a well-diversified portfolio with unrelated securities, the investor can obtain greater returns without increasing the exposure to risk (Markowitz, 1952).

A core assumption in Markowitz (1952) is homogenous investors with mean-variance preferences, entailing that all investors have the same expectations and make identical decisions in each scenario. In violations of MPT, exclusionary screening causes investors to be less homogenous. Due to the restricted investment universe, exclusionary investors face increased risk in their portfolios, backed up by the efficient frontier, a cornerstone in MPT. The efficient frontier indicates the perfect portfolio combination with high expected returns at low levels of risk. With an exclusionary strategy, and reduced opportunity to diversify, the efficient frontier subsides, harming financial performance.

Following Hong & Kacperczyk (2009), a collective exclusion of sin stocks leads to decreased demand for these firms, thus limiting risk-sharing and lowering the stock price. Besides the stock under-pricing, this can pose an investment opportunity for risk-neutral investors since a higher cost of equity contributes to a risk premium on the excluded stocks. MPT assumes risk-averse investors with a preference for less risky assets given a certain level of return, or vice versa. This relation implies that aversion should correspond to investing in multiple asset classes. However, norm-constrained investors consider and value other factors besides maximising profits.

3.1.3 Downward Sloping Demand Curve & Market Segmentation

Excluded companies must compensate investors for limited risk-sharing in the form of a stock premium. As we assume a long-term drop in demand after an exclusion, this study aims to test potential price adjustments by relating our findings to calibrations by Petajisto (2009). He incorporates theory from Merton (1987) on neglected stocks and market segmentation. In our case, there is a segmented market consisting of socially responsible investors practising negative screening (the Norwegian GPFG) and neutral investors who will invest regardless of ESG norms.

Merton's market segmentation theory concludes that excluded stocks should have a lower price due to limited risk-sharing and carry a higher expected rate of return (Merton, 1987). Consequently, shunned stocks are anticipated to outperform comparables. We can use the same approach as Petajisto (2009) to measure the price effect of exclusions. This model analyses the price effect on a company by being included in the S&P 500[3] instead of excluded. In the study, he argues that

[3] Standard & Poor's 500 (S&P 500). A stock market index that follows 500 publicly traded US firms.

management fees almost entirely determine the slope of the demand curve and not the level of risk aversion. Using realistic fees charged by mutual funds, Petajisto (2009) measured the change in stock price due to a supply shock created by an index inclusion. Findings show evidence for a downward sloping demand curve for stocks, similar to our thesis on a drop in demand post exclusion. This finding contradicts traditional asset pricing models, such as CAPM, which assumes that the demand curve for a stock is nearly perfectly horizontal. Petajisto (2009) argues that this is due to financial intermediates, as money managers' risk preferences determine the slope of the demand curve for a stock.

Conclusively, Petajisto finds a negative supply shock due to an S&P 500 inclusion since passive index funds suddenly buy stocks of the newly added company. This boosted demand will, in turn, increase the stock price, given realistic fees and volatility (Hong & Kacperczyk, 2009). This effect is relatable to our thesis, except that we anticipate the opposite effect; a negative demand shock as the Norwegian GPFG excludes sin stocks and drops their shares onto the market.

3.2 HYPOTHESIS DEVELOPMENT

3.2.1 The Short-Term Effect

Firstly, we want to test the short-term effects on excluded firms' stock performance. In other words, the market reaction to the exclusions made by the Norwegian GPFG. In developing a hypothesis, we turn to the EMH and published literature applying the event study method on environmental topics (i.e., Fisher-Vanden & Thorburn, 2011; Cañón-de-Francia & Garcés-Ayerbe, 2009). Under the EMH (assuming semi-strong form), there should be an immediate reaction on the event date but no further reaction on subsequent trading days. The magnitude of a potential change in the stock price then provides a measure of the unexpected impact the exclusion has on shareholder value. Based on predictions about negative and positive news elaborated in section 2, an exclusion from the NGPFG should impose an adverse reaction to the stock price. However, turning to Fisher-Vanden & Thorburn's (2011) findings on investors being indifferent to companies joining vague "green" initiatives may suggest less significant results. Based on this, we formulate the following hypothesis:

Hypothesis 1) *The market will adversely react to an exclusion announcement made by the Norwegian GPFG, resulting in negative announcement effects.*

3.2.2 The Long-Term Effect

Secondly, we want to measure the long-term effects of exclusion. Since we expect to see a decrease in demand following exclusion from the Norwegian GPFG, the excess stock quantity in the market should push the price down. Our objective is to examine whether this predicted decrease in demand also contributes to an increase in risk-adjusted return. This matter is answered by evaluating the stock's long-term change in systematic and unsystematic risk.

Following the CAPM theory, the expected return is only determined by the firm's exposure to systematic risk, i.e., the risk inherent to the market. Aligned with Hong & Kacperczyk (2009), we expect the long-term change in systematic risk to be greater than zero since sin stocks are anticipated to be more volatile than the overall market. On the other hand, Merton's market segmentation theory concludes that CAPM no longer holds. Due to neglection and increased litigation risk, return is not just a product of systematic risk (beta) but also unsystematic risk. As excluded stocks are exposed to limited risk-sharing, increased risk for each investor contributes to a higher expected rate of return (Merton, 1987).

Our hypotheses can be linked to the paper by Ilhan et al. (2021) regarding the cost of options protection for firms with carbon-intense business models. It is hard to predict how future climate regulations will affect this sector in terms of decreased stock prices and increased volatility. Hence, the additional risk leads to a greater cost of options protection for these firms. This effect can be translated onto sin stocks, as uncertainty in future ESG policies might harm their performance. In turn, uncertainty contributes to higher risk, resulting in a higher cost of equity to compensate investors. Based on this, we are formulating the following two hypotheses to draw an overall conclusion:

Hypothesis 2) *An exclusion announcement by the Norwegian GPFG leads to an increase in the systematic risk of the excluded company's stock in the long term.*

Hypothesis 3) *An exclusion announcement by the Norwegian GPFG leads to an increase in the unsystematic risk of the excluded company's stock in the long term*.

4.0 METHODOLOGY

This section outlines the employed methodology and a formal statement of the developed hypotheses. The event study analyses the potential market reaction following the exclusion announcement. The variable of interest for the short-term effects is the cumulative abnormal returns, examined over the defined event windows. To analyse the long-term changes in systematic risk, we include an interaction term on a dummy variable and the market excess return to isolate the effect of exclusion before and after the announcement. We test the equality of the regression residual variances before and after the event for the unsystematic risk.

4.1 EVENT STUDY METHODOLOGY

The event study methodology (ESM) is a well-known and frequently used tool in financial research, and the approach has a long history. It was initially introduced to a broad audience of economists in the two landmark papers by Ball & Brown and Fama et al. published in 1968 and 1969, respectively (Corrado, 2011). Several modifications have been made in applying this methodology; accordingly, there is no standard in the published research. However, the core elements of event studies are presented in these early papers.

The ESM is usually applied when the objective is to review the market's response to a specific event and its effect on a financial variable. These events are often related to announcements of new information to market participants, such as mergers & acquisitions (M&A), stock splits, or earnings announcements through corporate releases or financial press (Peterson, 1989). In our case, the event is the exclusion decision, and we assess the announcement's impact on the excluded company's stock price by determining the risk-adjusted performance. The abnormal return is the primary variable in our analysis as it is standard practice in ESM to measure a stock's price reaction by comparing predicted and actual returns. The predicted returns are obtained through the Capital Asset Pricing Model (CAPM).

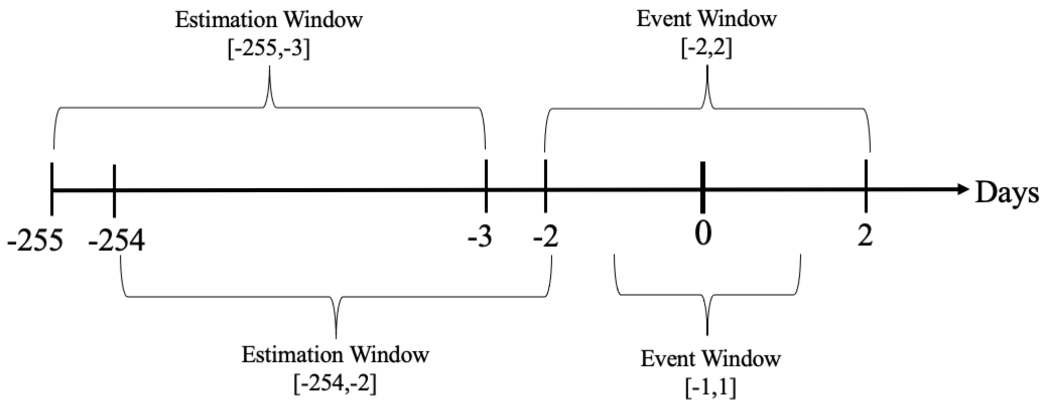
4.1.1 Identification of Event & Time Parameters

The initial step in constructing the event study is defining the event of interest and the corresponding event window. This is the period where the firms' stock prices are examined. The identified event in our study is the exclusion announcement made by the Norwegian GPFG. All information regarding the exclusion, such as exclusion criteria and the exact exclusion date, is published on the Norwegian Central Bank's website (*Observation and Exclusion of Companies*, 2019) and through other financial publications. The event date is denoted as day 0 ($t = 0$), and we evaluate 128 events ($N = 128$) after firms with incomplete information are omitted (Appendix A) from the dataset. Moreover, MacKinlay (1997) underpins the importance of the choice of data frequency, and we use daily stock returns to observe the potential change around the event. The paper argues that the power of EMS in detecting abnormal performance is greater when daily observations are employed than weekly or monthly observations, which means that the same power can be achieved with a smaller N.

Using a standardised event window across all observations is one remarkable advantage of the ESM. MacKinlay (1997) highlights that the event window is typically longer than the event date ($t = 0$) since this will permit investigation of the time surrounding the event. However, it is crucial to be aware that broader event windows can be inconsistent with the EMH as this can add noise to the estimation of how the markets interpret the news (McWilliams & Siegel, 1997). To be in line with the EMH and still be able to examine the time around the exact date, we set the event windows to one and two days before and after the exclusion date. These are defined as $[-1, 1]$ and $[-2, 2]$. The additional days prior to the announcement are to capture any information leakage, and the extra subsequent days are to capture an information delay or an incremental update of the stock price.

In the same matter, the estimation period for the expected returns must be defined. These are the returns that are expected to be observed if no event occurs. The typical lengths of such estimation periods typically range from 100 to 300 days for studies using daily observations and 24-60 months for monthly studies. However, there is a trade-off between the costs and benefits of more extended estimation periods. The cost is related to parameter instability, and the benefit is linked to an improved prediction model (Peterson, 1989). Under this trade-off, the estimation period for the short-term test is set to one trading year (252 days) and 24 months for the long-term. Moreover, MacKinlay (1997) states that the estimation period should not overlap with the event window, as this can lead to the event returns affecting the expected returns if no event occurs. Therefore, the estimation windows are defined as $[-254, -2]$ and $[-255, -3]$ for the event windows $[-1, 1]$ and $[-2, 2]$, respectively. This is illustrated in Figure 1.

Figure 1: Estimation- and Event Window



This figure demonstrates the defined estimation windows $[-254,-2]$ and $[-255,-3]$, and the event windows $[-1,1]$ and $[-2,2]$.

4.1.2 Expected Stock Returns

We need an estimation of the expected returns to compute the abnormal returns if no event occurs. This can be obtained in several ways, but we have employed the CAPM for this study. This model measures the stock performance relative to a market benchmark by linearly relating a security's excess return to the market's excess return. We have applied a market risk premium corresponding to the stock's benchmark index where the firm originated. An interesting question is whether α (alpha) should be included in the estimation period of the return. The alpha can be exceptionally high (low) due to unrelated incidents, and this pushes up (down) the expected return (Brooks, 2014). Since the stocks we are studying might be more volatile than other comparables, we believe it is preferable to assume an alpha with an expected value of zero and thus exclude it from the computation. The CAPM can be expressed as:

$$E[\mathrm{R_{i,t}}] - R_{f,t} = \beta_i(R_{m,t} - R_{f,t}) + \epsilon_{\mathrm{i,t}} \qquad (\mathbf{1})$$

where $E[\mathrm{R_{i,t}}]$ is the logarithmic return of stock i on day t, $R_{f,t}$ is the 3-month US Treasury Bill rate on day t, serving as a proxy for the risk-free rate[4]. Furthermore, $R_{m,t}$ is the logarithmic return of stock i's market benchmark and the beta of stock i is denoted as β_i, capturing the stock's systematic risk exposure. Lastly, $\epsilon_{\mathrm{i,t}}$ is the independent disturbance term, which is assumed to be equal to zero.

[4] We transform the 3-month U.S. Treasury bill rate into a daily continuously compounded risk-free rate using the following formula: $R_{f,t} = \ln\left[1 + TB_{3m} * \frac{90}{360}\right]^{\frac{1}{90}}$, where $R_{f,t}$ is the continuously compounded daily rate and TB_{3m} is the stated 3-month Treasury Bill Rate.

4.1.3 Abnormal Stock Returns

Abnormal return (AR) is a common measure to assess an event's effect on stock prices (MacKinlay, 1997). This is defined as the difference between an individual stock's actual and expected return over the predefined event windows (Peterson, 1989). The null hypothesis in event studies typically assumes no abnormal performance, meaning that abnormal returns will have an expected value of zero. The abnormal return for stock i on day t is estimated in the following way:

$$\widehat{AR}_{i,t} = R_{i,t} - \hat{\beta}_i R_{m,t} \quad (\mathbf{2})$$

where $R_{i,t}$ is the excess return of the stock i on day t, $R_{m,t}$ is the excess market return on day t, and $\hat{\beta}_i$ are the estimated coefficients from the CAPM **(1)** over the defined estimation windows $[-255, -3]$ and $[-254, -2]$.

4.1.4 Aggregation of Abnormal Returns

The test requires the estimation of AR for a sample of stocks and must be aggregated through time and across securities to draw inferences. It is thus necessary to study the cumulative abnormal return (CAR) when the event window consists of multiple days. The estimated $\widehat{CAR}_i$ for stock i is defined as the sum of the included AR in the event window $[t_1, t_2]$ (MacKinlay, 1997):

$$\widehat{CAR}_{i,[t_1,t_2]} = \sum_{t=\mathrm{t}_1}^{\mathrm{t}_2} \widehat{AR}_{i,t} \quad (\mathbf{3})$$

Since the test includes more than one event observation, the abnormal returns must, in addition to time, be aggregated across the event observations. So, the average abnormal return, given N events for a given day t, is the arithmetic mean of the $\widehat{AR}_{i,t}$ for a specific day t (Peterson, 1989):

$$\widehat{AAR}_{N,t} = \frac{1}{N} \sum_{i=1}^{\mathrm{N}} \widehat{AR}_{i,t} \quad (\mathbf{4})$$

The same approach used when calculating CAR for each stock i can also be applied when aggregating the average abnormal returns over the event window. Hence, the cumulative average abnormal return $\widehat{CAAR}$ across the event window $[t_1, t_2]$ is given by (MacKinlay, 1997):

$$\widehat{CAAR}_{[t_1,t_2]} = \sum_{t=t_1}^{t_2} \widehat{AAR}_{N,t} \quad (\mathbf{5})$$

4.2 LONG-TERM STUDY

The long-term effects are evaluated by measuring the volatility of the excluded firm before and after the exclusion announcement. The total risk can be divided into two main components: systematic- and unsystematic.

4.2.1 Systematic Risk

Following the CAPM, we study the potential change in systematic risk by measuring the beta of stock i. We create a dummy variable and an interaction term to isolate the excess market return before and after the exclusion date. The dummy variable D takes the value of 0 for observations prior to the event and 1 for all observations both on and after the event date.

The interaction term is defined as $D * R_m$, where the dummy variable is multiplied by the excess return in the period after the exclusion date. This captures the effects on the systematic risk after the exclusion. The interaction term of the dummy and the excess market returns allows for different slopes. In this way, we can test the potential change in beta after the event, β_2, for statistical significance by inferring whether the beta estimate prior is statistically significantly different to the beta estimate after. To make the equation cleaner, we set the excess stock return and excess market return to $R_e = (R_{i,t} - R_{f,t})$ and $R_m = (R_{m,t} - R_{f,t})$, respectively. Hence, the regression, including the interaction term of the dummy variable, is:

$$R_e = \beta_0 + \beta_1 R_m + \beta_2 R_m * D + \varepsilon_{i,t} \qquad (6)$$

where β_1 is the beta if no event occurs, and β_2 equals the difference in beta after the event. This regression can now be broken into two separate regressions:

$$R_e = \begin{cases} \beta_0 + \beta_1 R_m + \varepsilon_{i,t}\,, & when\ D = 0 \\ \beta_0 + (\beta_1 + \beta_2) R_m + \varepsilon_{i,t}, & when\ D = 1 \end{cases} \qquad (7)$$

4.2.2 Unsystematic Risk

To test for a potential change in the unsystematic risk after an exclusion, we employ a similar method as for the systematic risk. However, instead of comparing the beta estimates, we compare the residual variances of the daily stock returns before and after the event date. This is done through running two separate regressions, and also here we include the interaction term.

Otherwise, we will include the change in systematic risk as part of it. After that, we perform an F-test, as we operate with two different sample populations. The F-distribution arises in tests of hypotheses concerning whether two population variances are equal and whether three or more population means are equal. If the variance of the regression residuals has increased, we can conclude on greater unsystematic risk.

4.3 ALTERNATIVE METHODOLOGIES

Besides economic models like the CAPM or the Arbitrage Pricing Theory (APT), there are numerous ways the expected returns can be obtained. MacKinlay (1997) groups the approaches into two main groups: statistical and economical. The statistical models do not depend on economic arguments and only include statistical assumptions. E.g., they assume returns are jointly multivariate and follow a normal independently distribution through time. The most common within this category are the single-index model (also known as the constant mean model) and the market model (MacKinlay, 1997). The constant mean return model is perhaps the simplest one, assuming that a security's mean return is constant through time. The market model is one of the most used models in event studies, and it linearly relates a security's return to the market portfolio's return (Ma et al., 2009).

The CAPM, on the other hand, relies on investor behaviour, even though it is necessary to include statistical assumptions (MacKinlay, 1997). This model assumes that the expected stock return is a linear function of its covariance with the market portfolio's return. (Ma et al., 2009). The CAPM also requires a risk-free return in estimating the normal returns, while the statistical market model assumes this to be constant. Stapleton & Subrahmanyam (1983) argue that the linear market model is sufficient for deriving the CAPM relationship between market beta and expected returns. However, they are debating that characterisation theory suggests that the difference between the assumptions of multivariate normality and the market model linearity is not great. One can also include additional factors in the CAPM, such as momentum, size (SMB[5]), and book-to-market (HML[6]). However, MacKinlay (1997) argues that using more elaborate pricing models typically makes an insignificant difference when short windows are applied.

---

[5] Small Minus Big (SMB) – Small-cap companies over large-cap companies.
[6] High Minus Low (HML) – Value stocks over growth stocks.

4.4 FORMAL HYPOTHESIS STATEMENTS

Hypothesis 1) *The market will adversely react to an exclusion announcement made by the Norwegian GPFG, resulting in negative announcement effects.*

$$\boldsymbol{H_0}:\ \widehat{CAAR}_{[t_1,t_2]} = 0\,, \qquad \boldsymbol{H_A}:\ \widehat{CAAR}_{[t_1,t_2]} \neq 0 \qquad (\mathbf{8})$$

The first hypothesis is tested by examining the Cumulative Average Abnormal Returns (CAAR) for the two defined event windows and then checking if they are significantly different from zero. If the null hypothesis is rejected by observing CAARs different from zero, we then check whether they are positive or negative. If the CAARs are negative, it supports our prediction of exclusion by the Norwegian GPFG being interpreted as bad news for the company.

Hypothesis 2) *An exclusion announcement by the Norwegian GPFG leads to an increase in the systematic risk of the excluded company's stock in the long term.*

$$\boldsymbol{H_0}: \beta = 0\,, \qquad \boldsymbol{H_A}:\ \beta > 0 \qquad (\mathbf{9})$$

To test the second hypothesis, we perform a one-tailed t-test to check if the potential change in beta pre- and post-exclusion is statistically significant. A rejection of the null hypothesis implies an increase in systematic risk, i.e., the excess stock return will increase accordingly.

Hypothesis 3) *An exclusion announcement by the Norwegian GPFG leads to an increase in the unsystematic risk of the excluded company's stock in the long term.*

$$\boldsymbol{H_0}: \sigma^2_{pre} = \sigma^2_{post}\,, \qquad \boldsymbol{H_A}:\ \sigma^2_{pre} < \sigma^2_{post} \qquad (\mathbf{10})$$

For the third hypothesis, an F-test is conducted. This test compares the equality of the variances of the regression residuals before and after the event date. If the null hypothesis is rejected, we conclude that the unsystematic risk has increased.

5.0 DATA

This section describes how we selected and gathered the necessary data to test the hypotheses. Furthermore, it presents the descriptive statistics of the sample and its distribution, allowing us to get a first assessment of the results.

5.1 SAMPLE SELECTION

5.1.1 Data Collection

As this study analyses the change in stock price for companies excluded by the Norwegian GPFG, their official exclusion list was the starting point when determining the sample. Hoepner & Schopohl (2018) have used a similar approach but have included several funds like the Swedish AP-funds. The complete exclusion list by the Norwegian GPFG is publicly available and transparent by disclosing the company name, category (conduct-/product based), exclusion criteria, decision (exclusion/observation), and the exact publishing date of the exclusion. A piece of such detailed information on exclusions made by private (SRI) mutual funds is most probably not available to market participants (Hoepner & Schopohl, 2018).

Contradictory to most reviewed research, which collects a sample based on specific criteria (e.g., Standard Industrial Classification (SIC) codes, country of origin), our dataset was to some degree predefined. Still, we had to clean the data and identify the relevant firms. We started with all current firms included in the blacklist with exclusion dates from 05/01/2005 to 08/04/2022, equalling 173 events. Since we are studying the effect of an actual exclusion, we removed 22 firms marked under "observation", leaving us with 151 events. Not all applied to our dataset, and we had to apply requirements to avoid errors, resulting in omitting 15 firms (Appendix A). The main reasons were delisting and currency conversion errors.

In total, 136 companies were suitable for data collection. The next step consisted of retrieving historical closing prices for the excluded companies. Since we use each company's listing on the national stock exchange, prices are quoted in local currency. For consistency, we convert all values into USD. Additionally, we collected spot prices for the corresponding benchmark index (e.g., CAC 40 for France) to calculate the market return. All historical prices are retrieved from the Bloomberg database (Appendix B). To compute excess return through CAPM, we used the 3-month US Treasure Bill as a proxy for the risk-free rate. Lastly, we collect the historical market capitalisation for each firm during the last quarter before exclusion. The differentiation between small and large-cap is used as subsamples in mean difference tests described in section 5.2.3.

5.1.2 Data Processing & Requirements

After cleaning the data in Microsoft Excel, four separate datasets were imported into the statistical tool, R Studio, to conduct the event study: risk-free rate, stock prices, corresponding market index prices, and the exclusion dates. We matched the tickers across sheets with a one-to-one relationship to link the files in R. Further, we sorted data from exclusion date and stock return based on the ticker. When matching these columns, we discovered a missing link between ticker and exclusion dates for six firms (Appendix A). This lack of correspondence would create N/A terms further down the line. Thus, these companies were removed, leaving us with 130 firms to analyse. Additionally, the exclusion date for each company was linked to daily stock returns to pinpoint day 0. It is worth mentioning that all exclusions are on a trading day. If there was a non-trading day within the event window, we used the closest trading day after the announcement to calculate CAAR.

The statistical tests depend on having sufficient historical data on each company. Since our estimation period is 252 trading days, the data must contain approximately one year of stock prices before the exclusion. This information was not available for two of the firms and was removed from the dataset (Appendix A). These were not publicly listed a year before the exclusion and therefore did not have enough data points. After considering all factors that might significantly affect the regression, we have a total of 128 events to analyse. For these, we converted historical stock prices and the indices into daily, continuously compounded returns[7].

[7] The daily continuously compounded returns are expressed as: $R_{i,t} = \ln\left[\frac{P_t}{P_{t-1}}\right]$, where $R_{i,t}$ is the continuously compounded stock return for day t, P_t is the closing stock price of company i at day t, and P_{t-1} is the closing stock price of company i at the previous day $t-1$.

5.1.3 Treatment of Outliers

Stock return anomalies must be detected and analysed to avoid inaccurate results. To check for abnormal values in our dataset, we calculate the mean and standard deviation of the stock returns. Outliers have the potential to distort the regression, thereby making the hypothesis test less reliable. Even though historical stock return data are legitimate observations, significant deviations from the mean can be caused by other extraordinary situations besides the exclusion, e.g., M&A announcements. Since this paper evaluates stock returns resulting from regular price changes over time and not a few numbers of extreme values, we adjust the dataset for outliers. Following Osborne & Overbay (2004), data points below and above three standard deviations from the mean can disproportionally impact the regression. Therefore, we only include values within this range (Appendix C).

5.2 VARIABLE DESCRIPTIONS

5.2.1 Dependent Variable

The next step is determining if the exclusion or other independent variables have a causal effect on the dependent variable, CAAR. CAAR is the total sum of abnormal returns across companies and time, and the variable is used to measure the effect of the exclusion in our model. As the abnormal return is the difference between actual and predicted returns (equation 2), it allows us to determine the stock's risk-adjusted performance compared to the benchmark indices after the exclusion announcement. This is because the estimated expected returns are derived from the CAPM formula, which accounts for the market risk premium.

5.2.2 Independent Variables

The primary explanatory variable in determining the cumulative abnormal returns is the exclusion date from the Norwegian GPFG's exclusion list. The variable is expressed as a dummy variable (D), indicating whether the stock returns observations are before or after the exclusion date. By comparing the observed stock returns after the exclusion with the estimated stock returns if no event occurs, we measure the magnitude of CAAR over the defined event windows. To assess the long-term effect of exclusion, we create an interaction term between the dummy variable and the excess market return ($D * R_{m,t}$) to isolate excess market return before and after the Norwegian GPFG announcement.

The dummy variable follows the Bernoulli distribution, i.e., takes value 1 with probability p and value 0 with probability $1-p$. This equals a dummy that takes the value 0 for observations before the exclusion and 1 for observations on and after the exclusion. In turn, the dummy is included in the CAPM regression to determine whether the coefficient of interest, β_2, is statistically significant.

5.2.3 Subsample Splits & Variables

Besides the exclusion, we control for other underlying factors that might affect the dependent variable, CAAR. To determine the significance of extraneous variables, we use subsample splits as robustness tests. This is conducted by running a series of mean difference tests on the given subsamples. We inspected six subsamples and tested their mean differences to measure their effect on CAAR. It is worth mentioning that dividing the sample into partial samples will reduce the number of observations in each group. Regardless, we deem each of these subsamples splits large enough to return valid results.

Adamsson & Hoepner (2015) state that the interpretation of social norms is not homogenous across markets and may vary over industries and regions. Therefore, we check if the firms' location, divided into emerging and developed markets, significantly affects CAAR. Preferably, we would like to extend our analysis by performing a mean difference test based on, for instance, continents. However, we discarded this split due to few observations in each subsample, potentially providing inaccurate results. Alternatively, we could compare the two largest continents in terms of excluded companies: North America and Asia. Since these subsamples mainly consist of the United States, Canada, China, and India, we expect the result to be approximately equal to testing emerging versus developed markets.

Research by Hoepner & Schopohl (2018) proves that exclusion criteria can be explanatory for the impact an exclusion has on stock performance. To determine if similar results are found with the Norwegian GPFG announcements, we create a subsample divided into the categories of conduct- and product-based exclusions. Like countries, we cannot expand our test to include exclusion criteria due to the low individual sample size.

Lastly, it is worth mentioning that even though CAPM does not account for size- and value risk, these factors should be controlled to avoid inaccurate conclusions. Regularly, small-cap firms tend to outperform the market. Adamsson & Hoepner (2015) argue that this is also a bias in the sin stock paper by Hong & Kacperczyk. To check if this outperformance tendency is present in our model, we create a subsample split with market capitalisation divided into upper and lower 50th percentile.

Table 1: Variable- and Subsample Descriptions

Dependent Variables	
$\widehat{CAAR}_{[-1,1]}$	Cumulative average abnormal return over a 3-day event window around the exclusion date.
$\widehat{CAAR}_{[-2,2]}$	Cumulative average abnormal return over a 5-day event window around the exclusion date.
Independent Variables & Subsamples	
Exclusion Date (D)	Dummy variable that takes the value 1 on and after the exclusion date and 0 before the exclusion date.
$D \cdot E[R_{m,t}]$	Interaction term between the market excess return and the dummy variable.
Ln Market Cap	The logarithm of the historical market capitalisation of the excluded company in USD.
Emerging	Subsample split for firms excluded from the NGPFG the country of origin is defined as an emerging market.
Developed	Subsample split for firms excluded from the NGPFG the country of origin is defined as a developed market.
Conduct Based	Subsample split for firms excluded from the NGPFG due to conduct-based criteria.
Product Based	Subsample split for firms excluded from the NGPFG due to product-based criteria.
Large	Subsample split for firms in the upper 50th percentile measured by market capitalisation.
Small	Subsample split for firms in the lower 50th percentile measured by market capitalisation.

This table summarises the variables and subsample splits used in the analysis with a description.

5.3 DESCRIPTIVE STATISTICS

The descriptive summary statistics of CAAR for the two defined event windows are provided in Table 2, allowing for a preliminary analysis of the results. Panel A reports different summary statistics for the entire sample, while Panel B focuses on the mean values of CAAR sorted by exclusionary screens. The CAARs are primarily negative, which is expected. However, the highest observation across the full sample has a positive value. Nevertheless, these are relatively small in absolute terms compared to the minimum values. Examining the mean CAARs by exclusionary screens from Panel B of Table 2, we observe slight variations between the reasons for exclusions. The largest (absolute) value is for tobacco, indicating a negative CAAR of -0.0197, and the smallest (absolute) is observed for controversial weapons, -0.0059.

Table 2: Summary Statistics CAAR

	$\widehat{CAAR}_{-1,1}$	$\widehat{CAAR}_{-2,2}$
Panel A: Summary Statistics on CAAR		
Mean	-0.0142	-0.0227
SD	0.0026	0.0034
Min	-0.1554	-0.1954
Max	0.0525	0.0803
Observations	128	128
Panel B: CAAR by Exclusionary Screens		
Conduct Based	-0.0131	-0.0129
Environmental Issues	-0.0134	-0.0194
Violation of Human Rights	-0.0102	-0.0121
Product Based	-0.0146	-0.0263
Tobacco	-0.0197	-0.0193
Controversial Weapons	-0.0047	-0.0141
Coal Based Energy	-0.0059	-0.0145

This table reports the descriptive statistics of the CAARs of the firms excluded by the NGPFG. Panel A focuses on the CAAR for the full sample. Panel B reports CAAR sorted by reason of exclusion.

Table 3 presents a yearly overview of new exclusions from 2006 to 2022 (cumulative summations of yearly exclusions are provided in Appendix D). It is essential to note that this is an overview of the dataset used in our analysis. Thus, we acknowledge that the original list is more comprehensive than the one presented in this paper. However, to get a first assessment and explanation of our results, we have chosen to describe our dataset in this section, not the entire exclusion list.

Table 3: Descriptive Statistics Exclusion List

Panel A: Characteristics of new exclusions	Number of Exclusions (per year)			
	Mean	Min	Max	Median
Total New Exclusions (per year)	8	0	53	4

	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	Total
Panel B: New exclusions per year																	
Total Exclusions	4	-	2	2	12	1	-	8	1	4	53	10	8	2	14	7	128
Thereof excluded due to																	
Conduct Based	-	-	-	1	-	-	-	3	1	4	2	1	5	-	10	7	34
Environmental Issues	-	-	-	1	-	-	-	3	1	4	1	1	-	-	6	5	22
Violation of Human Rights	-	-	-	-	-	-	-	-	-	-	1	-	5	-	4	2	12
Product Based	4	-	2	1	12	1	-	5	-	-	51	9	3	2	4	-	94
Tobacco	-	-	-	-	12	1	-	2	-	-	-	-	-	-	-	-	15
Controversial Weapons	4	-	2	1	-	-	-	3	-	-	-	-	3	-	-	-	13
Coal or Coal-Based Energy	-	-	-	-	-	-	-	-	-	-	51	9	-	2	4	-	66
Thereof located in the following regions																	
Africa	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-	2	3
Asia	-	-	-	-	5	1	-	4	1	4	29	6	4	-	3	7	64
Australia	-	-	-	-	-	-	-	-	-	-	2	-	-	1	1	-	4
Europe	1	-	1	-	3	-	-	-	-	-	2	2	1	-	2	-	12
North America	3	-	1	2	4	-	-	4	-	-	17	1	2	1	4	-	39
South America	-	-	-	-	-	-	-	-	-	-	2	1	1	-	2	-	6
Thereof part of the following markets																	
Developed Markets	4	-	2	2	7	1	-	5	-	-	28	2	3	2	7	5	68
Emerging Markets	-	-	-	-	5	-	-	3	1	4	25	8	5	-	7	2	60

This table is a summary of all new annual exclusion announcements per year by the Norwegian GPFG, sorted by chosen characteristics. The sample consists of announced exclusion from 2006 to 2021. Companies omitted from the list are described in section 5.

Panel A of Table 3 provides the descriptive statistics of the new exclusions of the Norwegian GPFG's blacklist. On average, the fund excludes nine new companies annually, and the list compromises a mean of 51 exclusions per year (Appendix D). The maximum number of new exclusions observed is in 2016, with 53 exclusions, while 2007 and 2012 had zero new exclusions. The magnitude of exclusionary screening seems small when comparing the number to the total number of firms the fund is invested in. They currently have a small stake in about 9300 companies worldwide (*The Fund*, 2017), meaning that the excluded companies only represent 1.38% (128/9300) of these.

Panel B of Table 3 reports the number of new exclusions per year for the entire sample and chosen characteristics, such as exclusion criteria and region of origin. This overview enables us to draw rough conclusions about our results and the sample's skewness. In addition, it lets us observe the patterns and development in the fund's exclusionary approaches. The first section in Panel B of Table 3 shows the trend in the total number of new exclusions each year. From these numbers, we see an increasing trend of exclusions. The yearly distribution of our dataset is thus skewed to the right (Appendix E). We also get an insight into where the excluded companies are located and whether their country of origin belongs to developed or emerging markets from Panel B. Asia and North America represent the overweight, while only a few are in Africa, Australia, and South America.

Additionally, we observe clusters of exclusions. Some years that stand out are 2010, 2016, and 2020, with 12, 53, and 14 new exclusions per year, respectively. In 2010, 12 out of 12 exclusions were excluded due to tobacco production, which can be explained by the press release by the Ministry of Finance of the fund adding tobacco companies to its exclusion criteria (Ministry of Finance, 2010). In 2016, 51 out of 53 exclusions were related to coal and coal-based energy. This cluster of exclusions is also detailed in a press release by Norges Bank Investment Management (NBIM) in April 2016. This exclusion criterion affected companies with more than 30% of sales activity in coal (Holter, 2016).

6.0 RESULTS & ANALYSIS

The following section describes the conducted hypothesis tests and analyses. We first provide the results of the main model specifications in how the exclusion affects the stock performance in the short and long term. Multiple tests are conducted primarily on cumulative abnormal returns for the additional analyses. This is done through subsamples, allowing us to isolate individual effects. This is favoured to multivariate regression due to our limited sample size.

6.1 THE SHORT-TERM EFFECTS

The CAAR for both event windows and the corresponding standard deviation, t-statistics, and p-value, are presented in Table 4. The results show that the exclusion announcement effect is negative and statistically significantly different from zero for all relevant significance levels $(1\%, 5\%, 10\%)$. Hence, we reject the null hypotheses of CAAR being equal to zero for the event windows $[-1,1]$ and $[-2,2]$, as all p-values are below 0.01 (1%). This is not in line with Fisher-Vanden & Thorburn's (2011) findings on the Ceres Membership announcements, as they observed insignificant abnormal returns. Furthermore, the univariate test indicates that the overall exclusion announcements are adversely perceived as they are negative. This supports our first hypothesis regarding the stock market responding negatively to the exclusion announcements, which also is in line with findings from Klassen & McLaughlin (1996).

Seen in relation to the Efficient Market Hypothesis's (EMH) different forms, these results contradict the strong form but support the semi-strong form. In a strong form of efficient markets, this information would already be priced into the stock price of the excluded firms, as it assumes that public and private information is accounted for. However, our results prove that there is an immediate reaction among the market participants, suggesting that the markets are of a semi-strong form.

Table 4: Cumulative Average Abnormal Returns

	Mean	SD	t-statistics	p-value	Obs.
$\widehat{CAAR}_{[-1,1]}$	-0.0142	0.0026	5.4	0.00^{***}	128
$\widehat{CAAR}_{[-2,2]}$	-0.0227	0.0034	6.7	0.00^{***}	128

This table provides the CAAR over two event windows for the full sample. The table also reports the SD, t-statistics, and the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

6.2 THE LONG-TERM EFFECTS

The long-term effect is measured between each company's exclusion date and the subsequent year, i.e., 252 trading days. Our study compares the long-term effects on the excluded company's total risk, given by the variance of the abnormal returns.

6.2.1 Change in Systematic Risk

By performing a t-test, we test the hypothesis that change in systematic risk before and after the announcement date is significantly different from zero. Since we expect the change in beta to be positive, we conduct a one-sided test. The estimated beta is calculated using each company's abnormal return one year after exclusion. The t-statistic[8] is the ratio of the parameter estimate to its standard error. Depending on the degrees of freedom, we can determine whether the exclusion significantly affects systematic risk in the long run. Based on CAPM, a change in the stock's beta will change the excess return accordingly. Therefore, we are testing for a structural break in the systematic risk factor (beta) after the exclusion date.

Following our second hypothesis, we expect to see a positive beta estimate. Contradictory, Table 5 shows that the change in estimated beta is slightly negative. This implies a decrease in systematic risk, i.e., the stock return will decrease after an exclusion. Given $N = 128$ observations, we have $128 - 1 = 127$ degrees of freedom. By conducting a one-sided test, we see that $t_{crit} > t_{stat}$ for all relevant significance levels (1%, 5%, 10%). Therefore, we cannot reject H_0, and the results are not statistically significant. Additionally, we obtain a high p-value. This indicates that the minor change in beta is caused by random variations and not by the Norwegian GPFG exclusion.

Table 5: Change in Systematic Risk

Coefficient	$SE(\hat{\beta})$	t-statistics	p-value	Obs.
-0.0008	0.0085	0.1	0.46	128

This table shows the output from the interaction dummy coefficient. The table also reports the SD, t-statistics, and the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

[8]The t-statistic is derived in the following way: $t = \frac{\hat{\beta}}{SE(\hat{\beta})}$, where β is the average beta estimate and $SE(\hat{\beta})$ is the standard error of the beta estimate.

6.2.2 Change in Unsystematic Risk

To conclude on the third hypothesis related to change in variance for the regression residuals, we perform a lower-tailed F-test. As stated by the null and alternative hypothesis, we predict that the variance will increase after the exclusion, i.e., the unsystematic risk will increase. To compare the variances prior- and post-exclusion, we need the F value, given by the formula:

$$F = \frac{\sigma_1^2}{\sigma_2^2} \quad (\mathbf{11})$$

where σ_1^2 is the variance in the sample prior to the exclusion and σ_2^2 is the variance in the sample on and after the exclusion.

The F-test assumes that these two samples are normally distributed and independent of each other. In addition, the F-test also requires two degrees of freedom, which in both cases are equal to 127. As shown in Table 6, the resulting p-value is above all relevant significance levels, i.e., the change in sigma is insignificant. Therefore, the null hypothesis is not rejected, and we conclude that there is insufficient evidence to state that unsystematic risk will increase after an exclusion.

Table 6: Change in Unsystematic Risk

$\boldsymbol{\sigma_{pre}}$	$\boldsymbol{\sigma_{post}}$	**F-statistics**	**p-value**	**Obs.**
0.0015	0.0014	1.078	0.66	128

This table shows the output from the unsystematic risk. The table also reports the SD, F-statistics, and the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

When deciding if our results are significant, the F-value should always be used along with the p-value. Given a significance level of 0.05, the critical value is 1.34. Since the F-statistic is less than the critical value at the required level of significance, we get the same results as above, namely, not rejecting H_0.

6.3 ADDITIONAL TESTS & ANALYSES

This section provides additional tests of our results from the primary analysis to obtain more robust answers. First, we extend the event windows by an additional day prior to and after the event to confirm if the results are more significant further away from the event day. Second, we analyse the individual days after the event date. We do this to study whether we observe other trends than previously. Lastly, we conduct three mean-differences tests through subsample analyses.

6.3.1 Extended Event Window

Since we have significant abnormal returns for both event windows and the $\widehat{CAAR}_{[-2,2]} > \widehat{CAAR}_{[-1,1]}$, it is interesting to see whether this trend continues. If the CAAR is getting more significant for days further from the event date, it may suggest that the markets are not efficient, as we concluded they were in the primary analysis. Table 7 presents the results for the event window $[-3,3]$, showing that it is still significantly different from zero. Following the EMH, these results are contradictory, as the abnormal performance should disappear after some time and normalise. However, the t-statistics tell us that the results are starting to get less significant compared to the event window $[-2,2]$, which had a t-statistics of 6.7.

Table 7: Cumulative Average Abnormal Return

	Mean	SD	t-statistics	p-value	Obs.
$\widehat{CAAR}_{[-3,3]}$	-0.0185	0.0030	6.1	0.00^{***}	128

This table shows the CAAR for a 7-day event window. The table reports the SD, t-statistics, and the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

6.3.2 Average Abnormal Returns

When examining the results of the CAAR within the two defined event windows, it suggests they are supportive of the EMH if the market is of a semi-strong form. Both observations show abnormal returns in the time surrounding the event, and the t-statistics are significant. However, the results do not explain the development in the abnormal returns from day-to-day post exclusion. Nor does it explain whether there are more significant results on the event day ($t = 0$) than on the following one, two, or three days. The EMH, assuming semi-strong form, does not only emphasise that there should be abnormal returns on the day of the unanticipated announcement, but there should be insignificant t-statistics on the subsequent trading days (Kritzman, 1994).

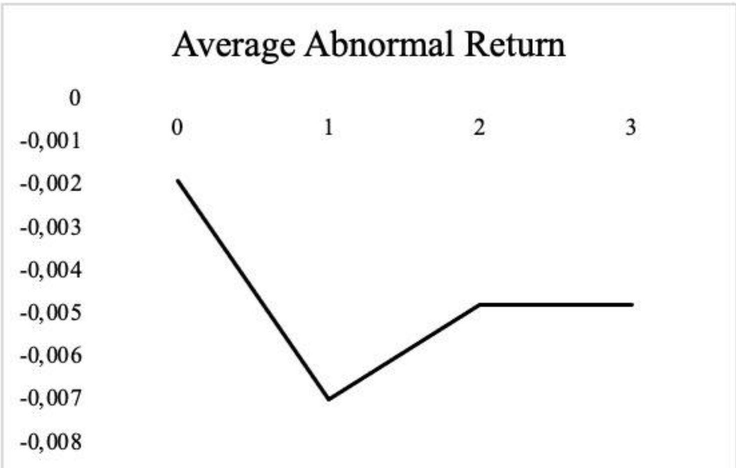
Table 8: Average Abnormal Returns

	Mean	SD	t-statistics	p-value	Obs.
$\widehat{AAR}_0$	-0.0019	0.0015	1.3	0.102	128
$\widehat{AAR}_1$	-0.0070	0.0015	4.7	0.000^{***}	128
$\widehat{AAR}_2$	-0.0048	0.0015	3.2	0.001^{***}	128
$\widehat{AAR}_3$	-0.0048	0.0015	3.2	0.001^{***}	128

This table shows the AAR for four days. The table also reports the SD, t-statistics, and the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

To conclude, we want to study the individual abnormal returns instead of the aggregated abnormal return over the previously defined event windows. Therefore, we have calculated the average abnormal return each day from the event date ($t = 0$) and the subsequent three days. For the EMH to hold, the t-statistic should be insignificant in the following dates (Kritzman, 1994). The results in Table 8 show that this is not the case, as they are all still significant. However, they are getting less significant each day. Another interesting observation is the AAR for the event date. This is not significant for any relevant significance levels, suggesting that it takes some time before the new information reaches the market. Figure 2 graphically clearly shows this trend in AAR.

Figure 2: Average Abnormal Return



This figure shows a graph of the average abnormal return on and after the event day.

6.3.3 Subsample Analysis – Mean Differences Tests

The subsample splits allow us to examine if the exclusion effect differs across the chosen characteristics. We can then check whether our findings in the primary model specification result from individual effects and thus extend the analysis of the relationship between exclusion and abnormal returns. As the variables may have different significance when the whole sample is included compared to sample splits, we perform t-tests in partial subsamples created for all defined dummy variables. One major weakness of this method is fewer observations in each test, resulting in reduced robustness and increased difficulty in detecting whether the results are significant even when they are. This is known as Type II errors.

Due to our limited sample size, investigating subsamples is favoured to control for variables in multivariate regression. A concern in multivariate regressions with interacted effects is that we would not have sufficient power to statistically identify any effect due to the correlation between variables of interest. The subsamples we want to study are developed markets, emerging markets, conduct-based criteria, product-based criteria, large firms, and small firms.

<u>Developed Markets vs Emerging Markets</u>

Table 9: CAAR of Developed & Emerging

	$\widehat{CAAR}_{-1,1}$	$\widehat{CAAR}_{-2,2}$
Developed Markets $N = 68$	-0.0194 (0.00^{***})	-0.0287 (0.00^{***})
Emerging Markets $N = 60$	-0.0082 (0.02^{**})	-0.0159 (0.00^{***})
Mean Differences	-0,0112 (0.00^{***})	-0,0128 (0.00^{***})

This table shows the results from t-tests of the mean differences of the CAAR for Developed Markets and Emerging Markets. The table also reports the p-values, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

Table 9 shows the excluded firms divided into two subsamples: developed and emerging markets. The number of firms is approximately the same in both subsamples, 68 and 60, respectively. The results show that an excluded company operating within a developed market will obtain a negative CAAR for both event windows at all relevant significance levels. On the other hand, the CAAR for an excluded company within the emerging market subsample is less significant. These results are only significant if the significance level is 5% or above. If we chose an $\alpha = 0.01$, the exclusion has no significant negative effect on emerging markets in the event window [-2, 2]. Nevertheless, the fact that CAAR is negatively affected by exclusion aligns with our short-term findings.

The mean difference is significant for both intervals at all relevant significance levels. Therefore, we can conclude that an exclusion will have a slightly more negative effect on CAAR for companies in developed markets than in emerging. These findings can be linked to several explanations. First and foremost, one assumption is that developing countries do not have sufficient incentives or means to implement strict ESG and ethical standards (costs may exceed the benefits). Hence, emerging markets might not consider sustainable investing at the same level as developed markets (Dasgupta et al., 2001). In turn, this can be related to access to information, which is not as readily available in emerging markets. Secondly, the reviewed studies mainly focus on stock performance in developed markets. Renneboog et al. (2008a) state that this is due to a lack of SRI funds in emerging markets, i.e., a small number of observations. This indicates a gap in existing literature, and more research on the connection between emerging markets, SRI and financial performance is needed.

Conduct Based Exclusions vs Product Based Exclusions

Most studies only account for sector-based criteria, but we aim to expand this by including norm-based screenings. Research implies that exclusion criteria can be explanatory of the performance (e.g., Hoepner & Schopohl, 2018). The NGPFG blacklist provides a set of exclusion criteria, which are divided into two main groups: conduct-based and product-based. Since the sample size of each exclusion criteria within conduct-based and product-based exclusions are smaller, we chose to study these two groups to get a more extensive set of observations. The division of excluded companies is presented in Appendix D. It is worth mentioning that the number of observations is relatively small for conduct-based exclusions compared to product-based, 34 versus 94, respectively. This can have a negative impact on the significance of the results, i.e., encountering type II errors.

Table 10 shows that firms in the product-based subsample have significant negative CAARs at all relevant significance levels within both event windows. In contrast, conduct-based exclusions provide less significant results, which is anticipated due to the low sample size. Hence, we acknowledge that a thorough analysis of the performance differences between norm-based and sector-based screening would require a more comprehensive set of events. The mean difference test concludes that product-based exclusions will have a more negative effect on CAAR than conduct-based in the five-day interval. For the three-day event window, we only find marginally significant results. This is consistent with Hoepner & Schopohl (2018), who argue that excluded companies neither under nor outperform relative to the benchmark, even when dividing for the sector- and product-based exclusions.

Table 10: CAAR of Product & Conduct

	$\widehat{CAAR}_{-1,1}$	$\widehat{CAAR}_{-2,2}$
Product Based $N = 94$	-0.0146 (0.00^{***})	-0.0263 (0.00^{***})
Conduct Based $N = 34$	-0.0131 (0.01^{**})	-0.0129 (0.04^{**})
Mean Differences	-0,0015 (0.06^{*})	-0,0392 (0.00^{***})

This table shows the results from t-tests of the mean differences of the CAAR for Product Based & Conduct Based exclusions. The table also reports the p-values in parentheses, where *,**, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

Large Firms vs Small Firms

The CAAR for small and large companies are presented in Table 11. As explained in the previous section, the firm size in market capitalisation is sorted in ascending order and divided into upper and lower 50%-percentile. We do this to see if the results are more significant for larger firms than smaller ones. The means are statistically significantly different from zero for both small and large excluded firms. This means that the results we find of the market having an adverse reaction to the exclusion announcement hold for small and large firms. However, there are variations in the coefficient estimates and t-statistics that shows that the exclusion effect is more significant for larger firms. There is not an obvious explanation for this. However, it makes sense that larger companies get more attention when a large pension fund blacklists them compared to smaller companies. Thus, resulting in more significant abnormal returns in the announcement date period.

Table 11: CAAR of Large & Small

	$\widehat{CAAR}_{-1,1}$	$\widehat{CAAR}_{-2,2}$
Large Firms	-0.0178	-0.030
$N = 64$	(0.00^{***})	(0.00^{***})
Small Firms	-0.0090	-0.0166
$N = 64$	(0.01^{**})	(0.00^{***})
Mean Differences	-0,0088	-0,0134
	(0.00^{***})	(0.00^{***})

This table shows the results from t-tests of the mean differences of the CAAR for Large firms and Small firms. The table also reports the p-values in parentheses, where *, **, *** indicate statistical significance at the 10, 5, and 1 % levels, respectively.

7.0 CONCLUSION

Exclusionary screening is a commonly used SRI approach, and it has gained substantial momentum among large institutional investors, such as public SWFs. New literature (Renneboog et al., 2008a, 2008b; Scholtens & Sievänen, 2013) argue that exclusionary screening is an outdated approach, and there is a debate regarding its implications and value. However, pension funds still apply this method to prevent themselves of allegations of unresponsible investments. Most are also transparent regarding their decisions and have their blacklists available to the public. The NGPFG is among these, announcing their exclusions in detail. As this is one of the largest pension funds in the world, we found it interesting to investigate if its blacklisting leads to any change among other market participants. Stated differently, we are testing if an exclusion announcement by the NGPFG leads to abnormal stock price changes in the short term. In the long run, we test whether the stock is more volatile in terms of systematic and unsystematic risk due to the predicted decrease in demand among investors.

To examine the short-term effects, we estimate the abnormal stock return for the excluded firms through the event study methodology (ESM). According to the efficient market hypothesis (EMH), stock prices reflect new information by immediately being adjusted. Hence, the abnormal stock performance around the exclusion date measures this event's unanticipated impact, and we believe ESM provides a suitable environment for assessing this effect. We conduct several mean differences tests between defined subsamples to conclude whether certain factors have significant individual effects on the cumulative average abnormal return (CAAR). Furthermore, we evaluate if an exclusion has a long-term effect on stock performance by measuring the change in systematic and unsystematic risk. First, we test the changes in systematic risk by comparing the beta estimates before and after the exclusion date. Following Merton (1987), we predicted a decrease in demand after the market is segmented, leading to an increased rate of return due to limited risk-sharing. Secondly, the change in unsystematic risk is measured through a similar approach, where we test the equality of stock variances for the excluded firms pre- and post-exclusion.

Our findings suggest that the exclusion announcement by the NGPFG leads to an immediate reaction among the market participants and that the CAAR is significantly different from zero. Additionally, all CAARs are negative, implying that the market perceives the announcement as bad news for the excluded company. The significance of the announcement can be explained through the EMH when the markets are of a semi-strong form. However, the results for the long-term effects are relatively insignificant, and we cannot reject the null hypotheses, i.e., there is not enough evidence to conclude that the change in systematic- and unsystematic risk is different from zero. In turn, this implies that the exclusion has no significant effect on CAAR in the long term. The fact that there is no noticeable change in risk insinuates that the market does not follow NGPFG's example. These findings are somewhat surprising, as we assumed neglected stock to carry a higher rate of return due to limited risk-sharing. One possible explanation is that investors at present do not care enough for ESG and ethical investing. The implementation and attitude towards sustainable investing is a subject that needs further research.

Our thesis contributes to the existing literature by evaluating how exclusions affect stock performance. By examining exclusionary screens done by a large public asset owner like GPFG, we can determine whether exclusions impact the targeted firm. Expanding our study to include other mutual funds would be interesting in this setting. First and foremost, the number of observations would be larger, potentially increasing the statistical power of our test. Larger sample size would also reveal potential ripple effects across SRI funds. In addition, it is of interest to compare the screening criteria across different funds. An extensive subsample split with adequate observations could establish whether the exclusion criterion is detrimental to changes in CAAR, as proposed by Hoepner & Schopohl (2018). One area that needs additional research is SRI funds' progress in emerging markets. Due to the scope, most reviewed literature accounts for SRI in developed countries. As ESG and ethical investing are becoming more widespread, we expect the need for research on developing countries and sustainability to increase accordingly. Lastly, this thesis measures the long-term effect by risk. Further research can supplement our findings by adding other measures of financial performance, such as earnings per share (EPS).

REFERENCES

Adamsson, H., & Hoepner, A. G. F. (2015). *The 'Price of Sin' Aversion: Ivory Tower Illusion or Real Investable Alpha?* (SSRN Scholarly Paper ID 2659098). Social Science Research Network. https://doi.org/10.2139/ssrn.2659098

Bengtsson, E. (2008). Socially responsible investing in Scandinavia – a comparative analysis. *Sustainable Development*, *16*(3), 155–168. https://doi.org/10.1002/sd.360

Berry, T. C., & Junkus, J. C. (2013). Socially Responsible Investing: An Investor Perspective. *Journal of Business Ethics*, *112*(4), 707–720. https://doi.org/10.1007/s10551-012-1567-0

Brooks, C. (2014). *Introductory Econometrics for Finance* (3rd ed.). Cambridge University Press.

Cañón-de-Francia, J., & Garcés-Ayerbe, C. (2009). ISO 14001 Environmental Certification: A Sign Valued by the Market? *Environmental and Resource Economics*, *44*(2), 245–262. https://doi.org/10.1007/s10640-009-9282-8

Chen, T., Dong, H., & Lin, C. (2020). Institutional shareholders and corporate social responsibility. *Journal of Financial Economics*, *135*(2), 483–504. https://doi.org/10.1016/j.jfineco.2019.06.007

Clark, G. L. (2004). Pension fund governance: Expertise and organisational form. *Journal of Pension Economics & Finance*, *3*(2), 233–253. https://doi.org/10.1017/S1474747204001556

Corrado, C. J. (2011). Event studies: A methodology review. *Accounting & Finance*, *51*(1), 207–234. https://doi.org/10.1111/j.1467-629X.2010.00375.x

Dasgupta, S., Laplante, B., & Mamingi, N. (2001). Pollution and Capital Markets in Developing Countries. *Journal of Environmental Economics and Management*, *42*(3), 310–335. https://doi.org/10.1006/jeem.2000.1161

El Ghoul, S., & Karoui, A. (2017). Does corporate social responsibility affect mutual fund performance and flows? *Journal of Banking & Finance*, *77*, 53–63. https://doi.org/10.1016/j.jbankfin.2016.10.009

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

Fisher-Vanden, K., & Thorburn, K. S. (2011). Voluntary corporate environmental initiatives and shareholder wealth. *Journal of Environmental Economics and Management*, *62*(3), 430–445. https://doi.org/10.1016/j.jeem.2011.04.003

GSIA. (2021). *Global Sustainable Investment Review 2020* (p. 32). Global Sustainable Investment Aliance (GSIA). http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf

Hoepner, A. G. F., & Schopohl, L. (2018). On the Price of Morals in Markets: An Empirical Study of the Swedish AP-Funds and the Norwegian Government Pension Fund. *Journal of Business Ethics*, *151*(3), 665–692. https://doi.org/10.1007/s10551-016-3261-0

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics*, *93*(1), 15–36. https://doi.org/10.1016/j.jfineco.2008.09.001

Ilhan, E., Sautner, Z., & Vilkov, G. (2021). Carbon Tail Risk. *The Review of Financial Studies*, *34*(3), 1540–1571. https://doi.org/10.1093/rfs/hhaa071

Kempf, A., & Osthoff, P. (2007). The Effect of Socially Responsible Investing on Portfolio Performance. *European Financial Management*, *13*(5), 908–922. https://doi.org/10.1111/j.1468-036X.2007.00402.x

Klassen, R. D., & McLaughlin, C. P. (1996). The Impact of Environmental Management on Firm Performance. *Management Science*, *42*(8), 1199–1214.

Kritzman, M. P. (1994). What Practitioners Need to Know about Event Studies. *Financial Analysts Journal*, *50*(6), 17–20.

Ma, J., Pagán, J., & Chu, Y. (2009). Abnormal Returns to Mergers and Acquisitions in Ten Asian Stock Markets. *International Journal of Business*, *14*.

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

Malkiel, B. G. (1989). Efficient Market Hypothesis. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Finance* (pp. 127–134). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-20213-3_13

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

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

McWilliams, A., & Siegel, D. (1997). Event Studies In Management Research: Theoretical And Empirical Issues. *Academy of Management Journal*, *40*(3), 626–657. https://doi.org/10.5465/257056

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, *42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Milne, R. (2021, May 3). Norway's $1.3tn oil fund broadens ESG screening to smaller companies. *Financial Times*. https://www.ft.com/content/f6649a49-03cb-47ef-89ab-2181ae9b30c2

Ministry of Finance. (2010, January 19). *Tobacco producers excluded from Government Pension Fund Global* [Pressemelding]. Government.No; regjeringen.no. https://www.regjeringen.no/en/historical-archive/Stoltenbergs-2nd-Government/Ministry-of-Finance/Nyheter-og-pressemeldinger/pressemeldinger/2010/Tobacco-producers-excluded-from-Government-Pension-Fund-Global/id591449/

Norm-based exclusions. (2012). [Research]. Novethic Research. https://www.novethic.com/fileadmin/user_upload/tx_ausynovethicetudes/pdf_complets/Norm-based_exclusions_EN_20120306.pdf

Observation and exclusion of companies. (2019, March 18). Norges Bank Investment Management. https://www.nbim.no/

Petajisto, A. (2009). Why Do Demand Curves for Stocks Slope Down? *Journal of Financial and Quantitative Analysis*, *44*(5), 1013–1044. https://doi.org/10.1017/S0022109009990317

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

Renneboog, L., Ter Horst, J., & Zhang, C. (2008a). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, *14*(3), 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009

Renneboog, L., Ter Horst, J., & Zhang, C. (2008b). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, *32*(9), 1723–1742. https://doi.org/10.1016/j.jbankfin.2007.12.039

Richardson, B. J. (2011). Sovereign Wealth Funds and the Quest for Sustainability: Insights from Norway and New Zealand. *Nordic Journal of Commercial Law*, *2*, 1–27.

Scholtens, B., & Sievänen, R. (2013). Drivers of Socially Responsible Investing: A Case Study of Four Nordic Countries. *Journal of Business Ethics*, *115*(3), 605–616. https://doi.org/10.1007/s10551-012-1410-7

Sparkes, R., & Cowton, C. J. (2004). The Maturing of Socially Responsible Investment: A Review of the Developing Link with Corporate Social Responsibility. *Journal of Business Ethics*, *52*(1), 45–57. https://doi.org/10.1023/B:BUSI.0000033106.43260.99

Stapleton, R. C., & Subrahmanyam, M. G. (1983). The Market Model and Capital Asset Pricing Theory: A Note. *The Journal of Finance*, *38*(5), 1637–1642. https://doi.org/10.2307/2327592

Statman, M., & Glushkov, D. (2009). The Wages of Social Responsibility. *Financial Analysts Journal*, *65*(4), 33–46.

The fund. (2017, April 28). Norges Bank Investment Management. https://www.nbim.no/

APPENDICES

A OMITTED FIRMS FROM DATASET

Company Name	Bloomberg Ticker Code	Country of Origin	Exclusion Date
Error in currency conversion*			
Ashtrom Group Ltd	ASHG IT	Israel	02.09.2021
Danya Cebus Ltd	DNYA IT	Israel	23.08.2010
Elbit Systems Ltd	ESLT IT	Israel	03.09.2009
Elco Ltd	ELCO IT	Israel	02.09.2021
Electra Ltd	ELTR IT	Israel	02.09.2021
Mivne Real Estate KD Ltd	MVNE IT	Israel	19.05.2021
Shapir Engineering & Industry Ltd	SPEN IT	Israel	19.05.2021
Shikun & Binui Ltd	SKBN IT	Israel	15.06.2021
Delisted*			
Airbus Finance BV	-	Netherlands	02.09.2005
Great River Energy	-	United States	07.03.2017
Tri-State Generation & Transmission Association Inc.	-	United States	10.07.2018
Does not list common stocks, only preferred shares			
PacifiCorp		United States	10.07.2016
Not publicly listed at the time of exclusion			
Consol Energy	CEIX US	United States	14.04.2016
Peabody Energy Corp	BTU US	United States	14.04.2016
Poongsan Corp	-	South Korea	06.12.2006
Missing link between ticker and exclusion date			
MMC Norilsk Nickel PJSC	GMKN RU	Russia	19.11.2009
Halcyon Agri Corp Ltd	HACL SP	Singapore	18.03.2019
British American Tobacco Plc	BATS LN	United States	19.01.2010
Public Power Corporation SA	PPC GA	Greece	14.04.2016
Airbus SE	AIR PA	Netherland	01.01.2006
Volcan Cia Minera SAA	VOLCABC1 LM	Peru	14.10.2013
Lack of historical data			
Zuari Agro Chemicals Ltd	ZUAC IN	India	14.10.2013
Freeport-McMoRan Inc	FXC US	United States	06.06.2006

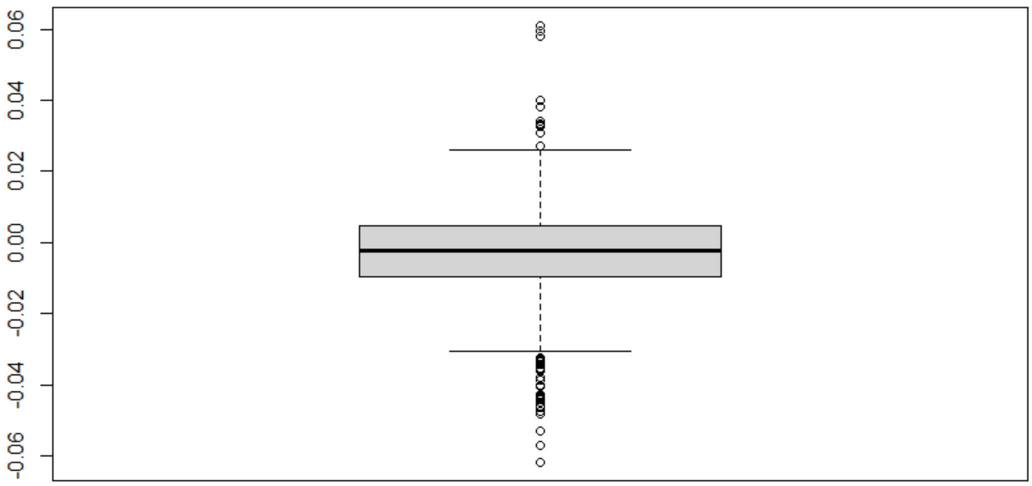
* 8 Israeli companies were removed due to error in currency conversion. The official currency in Israel is Israeli New Shekel (ILS), but Bloomberg reported the Israeli stock prices in agora (ILa). The conversion between ILa and USD was not available through the Bloomberg terminal.

** Bloomberg does not provide the complete historical information needed for the ESM.

B DATA SOURCES

Input	Specified	Source
Exclusion Companies	The Exclusion and Observation List	Norges Bank Investment Management Official Website
Stock Prices	Retrieved using the Bloomberg Data History (BDH)-function in Microsoft Excel.	Bloomberg Database
Market Return	Local Benchmark Index	Bloomberg Database
Risk-Free Rate	3-month US Treasury Bill Rate	Bloomberg Database

C TREATMENT OF OUTLIERS



Boxplot of the outliers in the stock return estimation. Observations outside three standard deviations from the mean are removed from the dataset.

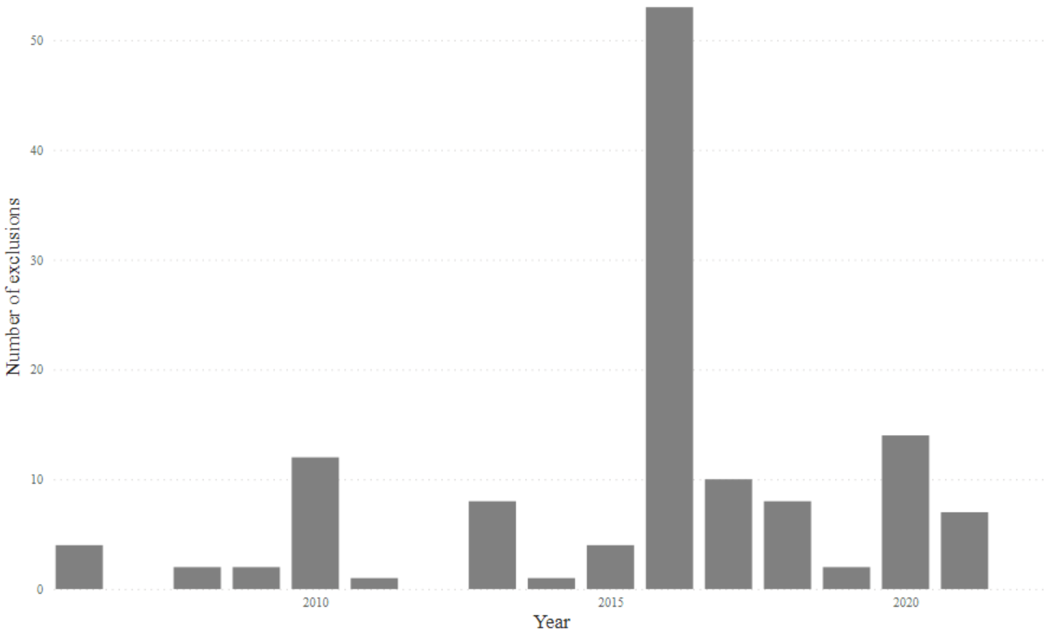
D DESCRIPTIVE STATISTICS (SAMPLE DISTRIBUTION)

	Number of Exclusions (per year)			
Panel A: Characteristics of new exclusions	Mean	Min	Max	Median
Total Exclusions	51	4	128	30

	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	Total
Panel B: Exclusionlist at end of year																	
Total Exclusions	4	4	6	8	20	21	21	29	30	34	87	97	105	107	121	128	128
Thereof excluded due to																	
Conduct Based	-	-	-	1	1	1	1	4	5	9	11	12	17	17	27	34	34
Environmental Issues	-	-	-	1	1	1	1	4	5	9	10	11	11	11	17	22	22
Violation of Human Rights	-	-	-	-	-	-	-	-	-	-	1	1	6	6	10	12	12
Product Based	4	4	6	7	19	20	20	25	25	25	76	85	88	90	94	94	94
Tobacco	-	-	-	-	12	13	13	15	15	15	15	15	15	15	15	15	15
Controversial Weapons	4	4	6	7	7	7	7	10	10	10	10	10	13	13	13	13	13
Coal or Coal-Based Energy	-	-	-	-	-	-	-	-	-	-	51	60	60	62	66	66	66
Thereof located in the following regions																	
Africa	-	-	-	-	-	-	-	-	-	-	1	1	1	1	1	3	3
Asia	-	-	-	-	5	6	6	10	11	15	44	50	54	54	57	64	64
Australia	-	-	-	-	-	-	-	-	-	-	2	2	2	3	4	4	4
Europe	1	1	2	2	5	5	5	5	5	5	7	9	10	10	12	12	12
North America	3	3	4	6	10	10	10	14	14	14	31	32	34	35	39	39	39
South America	-	-	-	-	-	-	-	-	-	-	2	3	4	4	6	6	6
Thereof part of the following markets																	
Developed Markets	4	4	6	8	15	16	16	21	21	21	49	51	54	56	63	68	68
Emerging Markets	-	-	-	-	5	5	5	8	9	13	38	46	51	51	58	60	60

This table is a cumulative summary of all new annual exclusion announcements per year by the Norwegian GPFG, sorted by chosen characteristics. The sample consists of announced exclusion from 2006 to 2021. Companies omitted are described in section 5.

E GRAPH OF YEARLY DISTRIBUTION OF EXCLUSIONS



This figure shows the yearly distribution of exclusion announcements made by the Norwegian GPFG between 2006-2021.