Tianqi Wang

An In-Depth Analysis of the Impact of ESG Investing on Returns Using Large-Scale News Data

Submission Date:  August 26th 2019

Adviser        Prof. Peter Bühlmann
Supervisor:    Dr. Ryohei Hisano
I would first like to express my deepest gratitude to my thesis advisor Dr. Ryohei Hisano for introducing me to this interesting topic and providing me with support and guidance throughout this process. I thank Prof. Sornette Didier, Dr. Ke Wu, and Chair of Entrepreneurial Risk for giving me the opportunity to work on this thesis here and providing me with additional data required for the analysis. Furthermore, I thank Prof. Peter Bühlmann for co-supervising this master thesis. Last but not the least, I want to express my gratitude to my parents and boyfriend for their support and encouragement all the time.
Abstract

By using the Dow Jones Adverse Media Entity dataset that covers 2012 to 2019, this thesis finds that environmental, social, and governance (ESG) risks generate negative short-run stock returns with a negative cumulative abnormal return of -0.32% for a 21-day event window. This discovery demonstrates that the market fails to fully incorporate value-relevant ESG information into stock prices in the short-run, and thus implies potentially profitable investment strategies. An underperformance of controversial firms in long-run with a monthly alpha of -4.37% is also reported. This result is robust to outliers and different risk model specifications, but not to industry or certain combinations of firm characteristics. This shows that the underperformance may stem from underperforming industries and common characteristics of controversial firms rather than ESG incidents. This thesis also examines the financial performance of several ESG investment strategies, including negative and relative screening. The results show that while a equal-weighted portfolio of worst ESG-performing stocks underperform the best ones by 0.54% and a value-weighted portfolio overperforms S&P 500 portfolio by 0.2%, negatively screened portfolios do not systematically overperform or underperform the benchmarks.
Contents

1 Introduction 1

2 Background 5
   2.1 Theoretical Considerations of ESG investings 5
   2.2 Literature Review 6
      2.2.1 Sin Stocks 7
      2.2.2 Information Voluntarily Disclosed by the Company 7
      2.2.3 Information from Rating Agencies 8
      2.2.4 News Information 9

3 Data Description 11

4 Event Study 19
   4.1 Methodology 19
      4.1.1 Expected Return Models 19
      4.1.2 Abnormal Return and Cumulative Abnormal Return 21
      4.1.3 Pros and Cons of Event Study 22
   4.2 Results 23
   4.3 Longevity Study 24

5 Jensen’s Alpha Approach 31
   5.1 Methodology 31
   5.2 Introduction to Jensen’s Alpha Approach 31
   5.3 Improvements and Limitations 32
   5.4 Results 33
   5.5 Robustness Checks 33
      5.5.1 Outliers 34
      5.5.2 Misspecified Risk Factors 34
      5.5.3 Industry 35
      5.5.4 Firm Characteristics 37

6 Cross-sectional Time-Series Regression 41
   6.1 Methodology 41
      6.1.1 Model Specification 41
      6.1.2 Fama-MacBeth Regression 42
   6.2 Results 42

7 Portfolio Construction 47
CONTENTS

7.1 Methodology ................................................. 48
7.2 Relative Screening ........................................ 48
7.3 Negative Screening ......................................... 51

8 Conclusion and Future Work ................................. 55

Bibliography ...................................................... 57

A Heteroskedasticity and Autocorrelation Consistent Standard Errors 61
A.1 Ordinary Least Square (OLS) ............................... 61
A.2 Heteroskedasticity and Serial-Correlation .................. 63
   A.2.1 Correlogram and Residual Plots ....................... 63
   A.2.2 Tests for Heteroskedasticity and Serial-Correlation .... 65
   A.2.3 Estimators Addressing Heteroskedasticity and Autocorrelation 66
A.3 Cluster-robust Inference .................................... 68
   A.3.1 One-way Clustering .................................. 68
   A.3.2 Multi-way Clustering ................................ 69

B Hypothesis Testing with Sharpe Ratio Test .................. 71

C Code .................................................................. 73
   C.1 Event Study .............................................. 73
   C.2 Jensen’s Alpha ............................................ 77
   C.3 Panel Regression ......................................... 88
   C.4 Relative Screening ....................................... 93
   C.5 Negative Screening ....................................... 96
List of Figures

3.1 Distribution of number of ESG incidents per subcategory ........................................ 12
3.2 Stock prices of Celsion ............................................................................................... 13
3.3 Trend of ESG incidents per year from 2012 to 2018 .................................................. 13
3.4 Distribution of number of ESG incidents per company ............................................. 15
3.5 Distribution of number of days between ESG incidents .............................................. 15
3.6 Median market capitalization per year from 2012 to 2018 ....................................... 16
3.7 Distribution of industries of controversial firms, compared with the universe of all US public traded stocks ................................................................................. 17
4.1 Timeline of event study .............................................................................................. 19
4.2 Cumulative abnormal return(CAR) within the 21-day event window ....................... 24
4.3 Cumulative abnormal return(CAR) for each event subcategory within the 21-day event window ....................................................................................... 26
4.4 Cumulative abnormal return(CAR) and confidence interval for each event subcategory at day 21. The confidence interval is constructed at 90% confidence level. ................................................................. 27
4.5 Cumulative abnormal return(CAR) plot used for longevity study ......................... 27
4.6 Cumulative abnormal return(CAR) plot for each event category, used for longevity study ........................................................................................................ 29
5.1 Monthly performance of several industries from 2010 to 2019(Fidelity Investment 2019) ........................................................................................................... 37
5.2 Illustration of the procedures to form 125-fractile portfolios .................................... 38
7.1 Profit and loss graph for equal-weighted portfolio of worst 30 stocks, compared with SP500, all US, and best 30 stocks ............................................................ 49
7.2 Profit and loss graph for value-weighted portfolio of worst 30 stocks, compared with SP500, all US, and best 30 stocks ............................................................ 50
7.3 Profit and loss graph for equal-weighted negative screened portfolio on all US stocks ................................................................................................................. 52
7.4 Profit and loss graph for equal-weighted negative screened portfolio on S&P 500 stocks ........................................................................................................... 52
7.5 Profit and loss graph for value-weighted negative screened portfolio on all US stocks ................................................................................................................. 53
7.6 Profit and loss graph for value-weighted negative screened portfolio on S&P 500 stocks ........................................................................................................... 53
A.1 Correlogram of time-series of portfolio return consisted of controversial firms over the risk-free-rate ........................................ 64
A.2 Residual plot of portfolio return consisted of controversial firms over the risk-free-rate .................................................... 64
## List of Tables

<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Number of ESG events per event subcategory</td>
<td>12</td>
</tr>
<tr>
<td>3.2</td>
<td>Number of total ESG incidents and median market capitalization by year from 2012 to 2018</td>
<td>14</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary statistics of number of ESG incidents per company and number of days between ESG incidents</td>
<td>15</td>
</tr>
<tr>
<td>3.4</td>
<td>Industry composition of controversial firms and all US public traded stocks</td>
<td>18</td>
</tr>
<tr>
<td>4.1</td>
<td>Values of cumulative abnormal return (CAR) within the 21-day event window</td>
<td>25</td>
</tr>
<tr>
<td>5.1</td>
<td>Coefficients for portfolio regression, where four factor-model is used</td>
<td>33</td>
</tr>
<tr>
<td>5.2</td>
<td>Alpha and p-value for each event subcategory</td>
<td>34</td>
</tr>
<tr>
<td>5.3</td>
<td>Alpha and p-value for each event category</td>
<td>34</td>
</tr>
<tr>
<td>5.4</td>
<td>Coefficients for portfolio regression, after removing outliers</td>
<td>34</td>
</tr>
<tr>
<td>5.5</td>
<td>Coefficients for portfolio regression with three-factor model</td>
<td>35</td>
</tr>
<tr>
<td>5.6</td>
<td>Coefficients for portfolio regression with five-factor model</td>
<td>35</td>
</tr>
<tr>
<td>5.7</td>
<td>Coefficients for portfolio regression on industry-matched returns</td>
<td>36</td>
</tr>
<tr>
<td>5.8</td>
<td>Alphas and p-values of portfolio regressions for each industry</td>
<td>37</td>
</tr>
<tr>
<td>5.9</td>
<td>Coefficients for portfolio regressions with characteristics-matched portfolio returns</td>
<td>39</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary statistics of control variables</td>
<td>43</td>
</tr>
<tr>
<td>6.2</td>
<td>Results of cross-sectional time-series regression, where the event indicator is a count of the number of events</td>
<td>43</td>
</tr>
<tr>
<td>6.3</td>
<td>Results of cross-sectional time-series regression, where the event indicator is a dummy variable</td>
<td>44</td>
</tr>
<tr>
<td>6.4</td>
<td>Results of cross-sectional time-series regression, where the number of events is counted for each subcategory</td>
<td>45</td>
</tr>
<tr>
<td>7.1</td>
<td>Coefficients for equal-weighted portfolio consisted of worst 30 stocks, compared with all US stocks, SP500 stocks, and best 30 stocks</td>
<td>49</td>
</tr>
<tr>
<td>7.2</td>
<td>Results for Sharpe Ratio test, where the worst 30 stocks are compared with best 30, SP500, and all US stocks</td>
<td>49</td>
</tr>
<tr>
<td>7.3</td>
<td>Coefficients for value-weighted portfolio consisted of worst 30 stocks, compared with all US stocks, SP500 stocks, and best 30 stocks</td>
<td>50</td>
</tr>
<tr>
<td>7.4</td>
<td>Results for Sharpe Ratio test, where the a portfolio of value-weighted worst 30 stocks are compared with best 30, SP500, and all US stocks</td>
<td>50</td>
</tr>
<tr>
<td>7.5</td>
<td>Coefficients for equal-weighted negatively screened portfolio, compared with benchmarks</td>
<td>51</td>
</tr>
</tbody>
</table>
7.6 Results for Sharpe Ratio test, where equal-weighted portfolios are compared 51
7.7 Coefficients for value-weighted negatively screened portfolio, compared with benchmarks .......................... 51
7.8 Results for Sharpe Ratio test, where value-weighted portfolios are compared 52
Chapter 1

Introduction

Socially responsible investing (SRI) is broadly defined as an investment process that involves identifying companies based on their corporate social responsibility profiles (CSR). ESG investing belongs to the class of SRI and specifically considers integrating environmental, social, and governance factors into investment decisions (Renneboog, Horst, and Zhang 2008). It reflects investors’ wish of ‘doing well by doing good.’ In other words, in addition to the financial utilities, ESG investors also strive for non-financial utilities by holding portfolios that are consistent with social norms and values.

In modern era, SRI became more formalized with the growth of mutual fund industry. As the fund management industry grew, activists started to take SRI as an opportunity for shareholders to influence corporate behaviours. For instance, between the 1970s and 90s, pressures on fund managers to avoid investing in South African companies contributed to the ending of Apartheid, an institutionalized race segregation in South Africa. In the 1980s and beyond, the awareness of climate change and global warming, along with industrial disasters such as Bhopal gas leak and Exxon Valdez oil spill, have brought up concerns regarding environmental issues and drives investors away from these controversial firms.

Nowadays, SRI has grown into a major force in the global market. As revealed in The Global Sustainable Investment Review (GSIR) released by Global Sustainable Investment Alliance, SRI assets in five major markets (Europe, the United States, Japan, Canada, Australia, and New Zealand) approximated $30 trillions at the beginning of 2018, featuring a 34% growth in two years. They constituted more than 50% of total professionally managed assets in Canada, Austria, and New Zealand, almost half in Europe, 26% in the United States, and 18% in Japan.

Common ESG strategies used by individuals or institutional investors include negative screening, besting in-class-screening, active ownership, ESG integration, and themed in-
Negative screening, which entails excluding certain companies or sectors from a fund or portfolio based on ESG criteria, is by far the most widely used strategy. It is also unarguably the most extensively studied strategy in the field of finance.

As reported by Global Sustainable Investment Alliance, there are two main motivations for fund managers to apply ESG investing. The first one is to “achieve social and environmental benefit or to fulfill their firms’ missions.” The second one is to “minimize risk and improve financial performance over time.” While the non-financial utility of ESG investing is unquestionable, its financial performance remains unclear. Therefore, the conversations about ESG investing has shifted from the social responsibility issues to questions about the impact of ESG incidents on stock price and the methods for integrating non-financial information into investment theories.

In past decades, many empirical studies have been conducted in ESG investing, but they show inconclusive and mixed results regarding its financial performance. While many studies find a zero-effect (Bauer, Koedijk, and Otten 2005) or a negative effect (Renneboog et al. 2008; Hong and Kacperczyk 2009), some studies find evidence that SRI improves investment performance (Edmans 2011). The source of controversial results could come from many aspects, such as the exact subject of study, sample period, country, and methodologies used.

One important source of uncertainty is the lack of clear definitions and metrics for ESG responsibility, which forms an obstacle to assess the value of the investment. To be more specific, a traditional way of defining low ESG responsibility stocks is by the notion of ‘sin stocks’. Sin stocks is defined by the industrial sector of the firm and often includes alcohol, tobacco, gambling, sex-related industries and weapon manufacturers (Hong and Kacperczyk 2009). However, this methods defines ESG responsibility broadly on an industry-level and ignore characteristics of individual firms. Furthermore, it does not allow for researchers to study the changes in stock prices around ESG shocks.

Alternatively, many recent studies start to use ESG ratings as a measure for ESG performance (Glosner 2017), with the emergence of specialized ESG rating agencies. Several examples include RepRisk, Sustainalytics, and MSCI (formerly named KID). Typically, information sources on the predefined ESG issues are screened and then used to calculate the ESG score on a monthly basis. A decrease in the rating of certain amount is usually seen as a negative shock in the ESG issues. However, as the rating is reported on a monthly basis, it introduces time latencies in event reporting.

Different from these two approaches, this thesis takes an event study approach by utilizing a large dataset on negative ESG-related news curated by Dow Jones for professional use. Our event study approach allows us to identify ESG shocks directly and to separate
controversial stocks from well-behaving ones by using the occurrences of negative news as a proxy of bad behavior. Compared with the broad industry classification, we check the ESG issues at an individual firm level, and thus making it easier to measure the impact of negative news in a more direct manner. In contrast to the ESG rating approach, our data allows us to study the short-term effect in a timely manner.

This thesis is consisted of four parts. In the first part, we introduce some theoretical foundations related to returns of ESG investing and review related literature. In the second part, we briefly explain the datasets used throughout the thesis and also provide some basic descriptive analysis of the data. In the third part, we investigate the relationship between companies’ ESG performance and stock returns in both the short and long run, using methodologies of event study, Jensen’s alpha, and cross-sectionally time-series regression. In the last part, we compare the ESG-based investment strategies, specifically negative screening and relative screening, with passive investment in order to study whether investors can profit from ESG-investing.

The thesis contributes to the literature on ESG investing in the following ways. First of all, by utilizing a large dataset on ESG related news, we perform a comprehensive analysis of the impact of ESG related news on stock returns. As would be detailed in the next section, our ESG related news dataset categorizes news events into 17 categories. These categories enable us to measure the impact as well as confidence intervals of the impact of ESG related news on stock returns. Secondly, we perform an in-depth analysis of the impact of ESG related news events on stock returns paying particular attention to the robustness of the results. We show that negative impact of ESG related news event seems to be robust to long-run impact, outliers, different risk model specification, but not to industry classification and firm characteristics. Thirdly, we also examine the financial performance of several ESG investment strategies, including both negative and relative screening.
Chapter 2

Background

In this section, we briefly summarize the theoretical considerations to ESG investing and provide a review of related empirical literature.

2.1 Theoretical Considerations of ESG investings

Current research on ESG investing shows both evidence of overperformance and underperformance of stocks with high ESG profiles. In the following, we summarize their main arguments.

Some studies show that ESG investing can improve investment performance relative to passive investing (Edmans 2011). A theoretical explanation is that the stock market undervalues ESG risks. In other words, the value effects of ESG events is not fully captured by the stock market. To further explain, while negative ESG events can almost always destroy shareholder value, as the firms have to pay severe penalties and bear the damages to their intangibles such as reputations, positive events can increase the shareholder value in the opposite way. However, investors may ignore ESG information because they are either obsessed with the short-term financial metric and too myopic and to incorporate long-term fundamentals (Porter 1992), or they fail to recognize the importance of ESG. Moreover, the impact on intangible assets is often underestimated by the market and is only reflected when manifested in some tangible outcomes. Evidence of underreaction to intangibles includes R & D costs, patent citations, advertising, and software development costs.

A second reason for the outperformance of high ESG stocks is that ESG investing has become more popular over time. Therefore, a growing demand for high ESG stocks can push up their prices, and a decreasing demand for low ESG stocks can result in lower stock
prices.

Demand effect is also one reason that some researchers argue that high ESG stocks underperforms relative to low ESG stocks (Hong and Kacperczyk 2009) or the market. To be more specific, when ESG investors ignore certain controversial firms, their stock prices can become undervalued. While these firms have low initial returns, they may yield high subsequent returns relative to high ESG stocks. Another reason is that firms in industries shunned by ESG investors, such as alcohol and tobacco, have incentives for conservative accounting because of more scrutiny from regulators. This will lead to underreaction and thus high subsequent returns.

Other studies show that there is no underperformance or overperformance of high ESG stocks because of the perfectly elastic demand curve (Bauer et al. 2005). To further explain, ESG investing might not be sufficiently widespread to affect prices. Moreover, arbitrageurs can offset any effect of ESG investors by trading in the opposite direction.

2.2 Literature Review

A vast body of literature is devoted to studying the financial impact of ESG investing. The literature can be split into different categories according to the source of information they used. A traditional way of screening ESG stocks is to avoid the so-called “sin stocks”. This screening method is rooted in Islamic religious principles and has thousands of years of history. However, in modern times, with the formalization of financial reporting procedures and the era of big data, ESG researchers are accessible to more information sources. Rather than sin stocks screening, modern literature approaches can be split into three main categories: i) information voluntarily disclosed by the company, ii) information from a rating agency, and iii) news information.

Compared with (i) and (ii), (iii) is a more neutral and reliable approach. On one hand, (i) can be cooked, which is evident in the Enron Scandal. Enron had fooled regulators for long with fake holding and off-the-books accounting, until the problem was discovered and investigated and the company collapsed. On the other hand, (ii) can be influenced by a conflict of interest. Efing and Hau (2013) study 6500 structured debt ratings produced by Standard and Poor’s, Moody’s, and Fitch, and found that ratings by agencies are “biased in favor of issuer clients that provide the agencies with more rating business.” This phenomenon is evident in the subprime crisis.

Hvidkjaer (2017) provides a nice overview of ESG literature.
2.2 Literature Review

2.2.1 Sin Stocks

Hong and Kacperczyk (2009) is one of the most cited articles addressing the return effects of investing in sins stocks. They found that a classical Triumvirate of sin stocks (consisted of tobacco, alcohol, and gambling companies) of 156 US stocks outperform their industrial comparable stocks during 1965-2006. To be more specific, the result shows that applying Carhart four-factor model to the portfolio of returns of sin stocks net of their industry comparables yields a monthly alpha of 26 bps at 5% significance level. Moreover, a cross-sectional returns model finds a monthly excess performance of 30 bps at 5% significance level.

Rather than defining sin stocks according to the broad industry classification, Trinks and Scholtens (2017) selects stocks at individual level based on fourteen potentially controversial issues. They show that sin stocks exhibit high returns during 1991-2012 in 94 international markets. Specifically, they show that the value-weighted TotalSin portfolio, which is consisted of over 1,600 stocks, outperforms the global Fama-French four-factor benchmark by 91-104 bps per month, which is significant at 1% level. Also, note that they apply a median regression to address the non-normality of excess returns.

Similar to the idea of sin stocks, Edmans (2011) conducts social screens by selecting firms with high employee satisfaction, which is defined as being on the list of “100 best companies to work for in America”. He shows that firms with high employee satisfaction exhibit high future stock returns. As the paper noted, a value-weighted portfolio of the listed firms earned an annual alpha of 3.5% from 1984 to 2009 and it is 2.1% above industrial benchmarks. The results are robust to controls for firm characteristics, different weighted methodologies, and removal of outliers. In addition to the industry-matched returns mentioned in Hong and Kacperczyk (2009), Edmans also uses characteristic-adjusted returns. This controls for interaction effects, which may potentially cause the model to misprice certain stocks. Moreover, Fama-Macbeth cross-sectional regressions confirm the results of the factor regressions.

2.2.2 Information Voluntarily Disclosed by the Company

Jacobs, Singhal, and Subramanian (2010) apply event study methodology to study the stock market response to the announcement of various types of corporate environmental initiatives, such as environmental business strategies, eco-friendly products, renewable energy, and recycling. The announcements include both Corporate Environmental Initiatives (CEI), which is the self-reported corporate endeavors to avoid environmental impacts, and Environmental Awards and Certifications (EACs), which is about recognition granted
by a third party. The sample includes 780 announcements between 2004 and 2006. They find mostly insignificant results, except for significantly negative returns for voluntary emission reduction and positive returns for environmental philanthropy and ISO14001 certification.

Fisher-Vanden and Thorburn (2011) use the same methodology to study the abnormal returns around announcements that firms have joined two environmental programs, namely the U.S. Environmental Protection Agency’s Climate Leaders program and Ceres. The analysis is based on a sample of 117 announcements by large U.S. firms during the period 1993 to 2008. The results show that the market values decrease by around 1% when firms join the Climate Leaders program, which is statistically significant. By contrast, abnormal returns are insignificant when firms join Ceres.

2.2.3 Information from Rating Agencies

Kempf and Osthof (2007) use KLD rating data to construct value-weighted portfolios for the period 1992-2004 based on three strategies, namely negative, positive, and best-in-class screenings. Their first finding is that high ESG-rated portfolio performs better than low-rated portfolio for certain qualitative criteria. Specifically, a long-short strategy yields a positive alpha of up to 8.4% per year using Carhart four-factor model. Furthermore, the maximum alpha is obtained by the best-in-class screening, choosing stocks with extreme ratings and applying several screenings at the same time. However, negative screening alone does not lead to any significant performance effect. Statman and Glushkov (2009) generally confirm the results of Kempf and Osthof (2007) based on 1992-2007 KLD rating data. Both of the studies also show that portfolios formed on community and employee relations criteria result in high excess returns, while those formed on diversity, environment, products and human rights do not have significant results.

Humphrey and Tan (2014) use KLD data to construct portfolios using positive and negative screening strategies, during the sample period 1996-2010. They show that the ESG portfolio returns are lower than, but not significantly different from the benchmarks. They also compare the Sharpe ratios and found that none of the difference is significant.

Kruger (2015) uses event study methodology to study the stock market response to both positive and negative corporate ESG events, identified by KLD data around seven ESG issues. He finds a significantly negative response to negative ESG events and a slightly negative reaction to positive ESG events. To be more specific, the mean 21-day cumulative abnormal return is approximately -111 bps for negative events and -47 bps for positive events.
Halbritter and Dorfleitner (2015) shows that the outperformance of ESG portfolio disappears after the initial sample period in Kempf and Osthof (2007). Moreover, they investigate the returns to ESG ratings of two other agencies, namely ASSET4, which covers a sample period from 2003 to 2012, and Bloomberg, which covers 2006 to 2012. Alphas are generally also insignificant, confirming the KLD results. However, the Fama-Macbeth cross-sectional regression shows a different result: while the excess returns to KLD ratings remain insignificant, those obtained from ASSET4 and Bloomberg are statistically significant.

2.2.4 News Information

Flammer (2013) identifies eco-harmful and eco-friendly events based on word search of the Wall Street Journal (WSJ) articles obtained from Factiva, which is one of the major newspaper databases. The search provides 117 eco-friendly and 156 eco-harmful events, ranging from 1980 to 1999. By applying event study methodology, Flammer (2013) finds a negative market reaction to eco-harmful events and a positive reaction to eco-friendly events. She also examines the trend of market reaction over decades and shows that the reward for eco-friendly initiatives has decreased over time while the penalty for eco-harmful events has increased.
Chapter 3

Data Description

The primary dataset we used in this data is purchased from Dow Jones. The dataset is called Dow Jones Adverse Media Entity List, and it is said to be created by curating news from about 10,000 sources all over the world in the Factiva dataset. The data covers the negative ESG-related events from January 2012 to January 2019. There are 2,967 firms and 16,380 events involved during this period.

For the robustness study, we collected additional data of company industry, market capacity, dividend yield, and book-to-market ratio for all controversial firms from the Thomson Reuters database. In order to compare the controversial firms with other firms, we also obtained these data for all US public stocks traded on NASDAQ, NYSE, and AMEX, which include around 6,900 stocks, as well as S&P 500 firms. To convert stock symbols from ticker format to RIC format for queries on Thomson Reuter, we applied the inherent function of Eikon. we compared the matched company names and manually checked entries where names differ in order to ensure that the matching is correct.

In our dataset, the negative ESG-related news are divided into four main categories and more subcategories, namely competitive/financial(information, anti-competitive, ownership, management, financial, association), environmental/production(service, supply, environmental), social/labour(discrimination, workplace, workforce, human), and regulatory (regulatory, fraud, corruption, sanctions). The distribution of number of negative news per category is as figure 3.1. The most frequent subcategories appeared are product/service, regulatory, and information, and they altogether account for 55% of the incidents. Few observations lie in the subcategory of association, sanction, corruption, human, and financial.
A visualization of the stock returns around event date can help us better understand the impact of ESG incidents. Figure 3.2 shows the stock prices of Celsion, a biopharmaceutical company. The failure of its experimental liver cancer treatment ThermoBox in the Phase 3 heat trial in January 2013 causes a plunge of 80% in its stock price.
An inspection of the yearly trend shows that the number of incidents nearly doubles, from 1,225 events in the year 2012 to 2,420 events in 2018. This partly reflects an expanding coverage universe of the dataset and an increasing attention to ESG incidents, and partly indicates that companies become more negligent in their ESG obligations.
<table>
<thead>
<tr>
<th>year</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
<th>2017</th>
<th>2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>total number of incidents</td>
<td>1255</td>
<td>2210</td>
<td>2648</td>
<td>2561</td>
<td>2802</td>
<td>2372</td>
<td>2420</td>
</tr>
<tr>
<td>median market capitalization(bn $)</td>
<td>1.85</td>
<td>2.27</td>
<td>2.74</td>
<td>2.63</td>
<td>2.28</td>
<td>2.73</td>
<td>3.12</td>
</tr>
</tbody>
</table>

Table 3.2: Number of total ESG incidents and median market capitalization by year from 2012 to 2018

The distribution of number of negative news per company is shown in figure 3.4. Around 57% of the companies have recurring negative news, up to February 2019. The median time interval between two events is 52 days. Figure 3.5 reveals that as the number of previous incidents for a company increases, the time interval between incidents decreases. It shows that negative events are persistent, which is another reason why investors should pay attention to ESG incidents.
Figure 3.4: Distribution of number of ESG incidents per company

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std</th>
<th>min</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of events per company</td>
<td>5.56</td>
<td>14.65</td>
<td>1.00</td>
<td>1.00</td>
<td>2.00</td>
<td>4.00</td>
<td>297.00</td>
</tr>
<tr>
<td>number of days between incidents</td>
<td>15.20</td>
<td>23.88</td>
<td>0.00</td>
<td>5.00</td>
<td>9.00</td>
<td>17.00</td>
<td>271.50</td>
</tr>
</tbody>
</table>

Table 3.3: Summary statistics of number of ESG incidents per company and number of days between ESG incidents

Figure 3.5: Distribution of number of days between ESG incidents

The median market capitalization for controversial companies in our dataset increases from $1.8 billion as of the year 2012 to $3.1 billion in 2018. Most of our controversial
firms are small to medium cap. This is a contrast to RepRisk data, which shows a median market cap of around $49 billion for controversial firms.

Figure 3.6: Median market capitalization per year from 2012 to 2018

Figure 3.7 is the industry distribution for companies that experiences ESG incidents, compared with the universe of all public US-traded stocks. The controversial firms are mostly concentrated in capital goods, pharmaceutical, and banking industries. Their industry compositions are mostly similar to that of all firms, but more concentrated in retailing, utilities, and material industries, while less concentrated in pharmaceutical and banking industries.
Figure 3.7: Distribution of industries of controversial firms, compared with the universe of all US public traded stocks
<table>
<thead>
<tr>
<th>Industry</th>
<th>Controversial Firms</th>
<th>All US Traded Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food, Staples Retailing</td>
<td>0.011</td>
<td>0.006</td>
</tr>
<tr>
<td>Household Products</td>
<td>0.012</td>
<td>0.006</td>
</tr>
<tr>
<td>Automobiles</td>
<td>0.016</td>
<td>0.008</td>
</tr>
<tr>
<td>Consumer Durables</td>
<td>0.019</td>
<td>0.023</td>
</tr>
<tr>
<td>Telecommunication</td>
<td>0.022</td>
<td>0.009</td>
</tr>
<tr>
<td>Transportation</td>
<td>0.027</td>
<td>0.019</td>
</tr>
<tr>
<td>Food, Beverage, Tobacco</td>
<td>0.027</td>
<td>0.019</td>
</tr>
<tr>
<td>Insurance</td>
<td>0.031</td>
<td>0.022</td>
</tr>
<tr>
<td>Semiconductors</td>
<td>0.031</td>
<td>0.029</td>
</tr>
<tr>
<td>Real Estate</td>
<td>0.033</td>
<td>0.026</td>
</tr>
<tr>
<td>Commercial Services</td>
<td>0.035</td>
<td>0.026</td>
</tr>
<tr>
<td>Consumer Services</td>
<td>0.037</td>
<td>0.023</td>
</tr>
<tr>
<td>Entertainment</td>
<td>0.038</td>
<td>0.040</td>
</tr>
<tr>
<td>Utilities</td>
<td>0.038</td>
<td>0.010</td>
</tr>
<tr>
<td>Technology Equipment</td>
<td>0.048</td>
<td>0.052</td>
</tr>
<tr>
<td>Diversified Financials</td>
<td>0.049</td>
<td>0.060</td>
</tr>
<tr>
<td>Retailing</td>
<td>0.051</td>
<td>0.032</td>
</tr>
<tr>
<td>Materials</td>
<td>0.056</td>
<td>0.028</td>
</tr>
<tr>
<td>Software and Services</td>
<td>0.057</td>
<td>0.068</td>
</tr>
<tr>
<td>Healthcare</td>
<td>0.061</td>
<td>0.071</td>
</tr>
<tr>
<td>Energy</td>
<td>0.067</td>
<td>0.037</td>
</tr>
<tr>
<td>Banks</td>
<td>0.072</td>
<td>0.134</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>0.080</td>
<td>0.198</td>
</tr>
<tr>
<td>Capital Goods</td>
<td>0.081</td>
<td>0.048</td>
</tr>
</tbody>
</table>

Table 3.4: Industry composition of controversial firms and all US public traded stocks
Chapter 4

Event Study

4.1 Methodology

Event study is a statistical method for examining the impact of an event on the value of a firm. It is widely used to investigate the stock market responses to macroeconomic events such as market shock, as well as to corporate events, such as earning announcement, merge and acquisition, and dividend announcement. In this study, we use event study methodology to study the stock market response to ESG-related news in the short-run. This addresses the problem as to how the stock market perceives the value of an ESG shock.

![Timeline of event study](image)

Figure 4.1: Timeline of event study

4.1.1 Expected Return Models

We proceed with an event study framework following Mackinlay (1997). For every event, we use a pre-event window to estimate the expected stock returns.

\[ R_{it} = E[R_{it} | X_t] + e_{it} \]  

(4.1.1.1)
where $R_{it}$ is the return for stock $i$ at time $t$, $X_t$ is the conditioning information, and $\epsilon_{it}$ is the abnormal return.

There are a variety of expected return models, such as market model, capital asset pricing model (Treynor 1961, Treynor 1962, Sharpe 1964, Lintner 1965, Mossin 1966), Fama-French three factor model (Fama and R. French 1992), and Carhart four factor model (Carhart 1997). We introduce each of them below. These asset pricing models are also used in the portfolio regressions in the next section.

These estimated coefficients are used to compute the expected returns during the event window as follows.

$$E[R_{it} | X_t] = \alpha_i + \beta_i R_{it}$$

(4.1.1.2)

Market Model

Market model is also called single index model. It is a common specification that models the return of individual security or portfolio on the market portfolio. It is written as:

$$R_{it} = \alpha_{it} + \beta_{1i} R_{MKT} + \epsilon_t$$

(4.1.1.3)

, where $R_{MKT}$ is the return of market portfolio. As it is impossible to collect returns of all the stocks, it is a common practice to use S&P 500 index as the market return.

Capital Asset Pricing Model (CAPM)

CAPM was introduced by Treynor (1961), Sharpe (1964), Lintner (1965), and Mossin (1966). It builds on the earlier work of Markowitz (1952) on diversification and modern portfolio theory. It assumes that the expected return of a security is based on risk-free-rate plus a risk premium of the market return over the risk-free-rate. The risk-free-rate usually used is the one-month treasury bill rate.

$$R_{it} - Rf_t = \alpha_{it} + \beta_{1i}(R_{MKT} - Rf_t) + \epsilon_t$$

(4.1.1.4)

, where the additional parameter $Rf_t$ is the risk-free-rate.
4.1 Methodology

Fama-French Three-factor Model

Fama and French (Fama and R.French 1992) three-factor model expands on CAPM by adding size and value risk factors. This model considers the fact that value stocks outperform growth stocks and that small cap stocks outperform big cap stocks on a regular basis. By including these two additional factors, the model adjusts for this outperforming tendency.

\[ R_{it} - R_{ft} = \alpha_{it} + \beta_1(R_{MKT_i} - R_{ft}) + \beta_2SMB_t + \beta_3HML_t + \epsilon_t \]  

(4.1.5)

where SMB is the size premium, measured by the average return on three small portfolios minus three big portfolios, formed in terms of market capacity. HML is the value premium, computed as the average return on two value portfolio minus two growth portfolios formed according to book-to-value ratio.

Carhart Four-factor model

Carhart four-factor model (Carhart 1997) extends Fama and French three-factor model by adding momentum as an additional risk factor, which is defined as a tendency for the stock price to continue rising if it is in an upward trend and to continue declining if it is in a downward trend.

\[ R_t - R_{ft} = \alpha + \beta_1MKT_t + \beta_2HML_t + \beta_3SMB_t + \beta_4MOM_t + \epsilon_t \]  

(4.1.6)

where MOM is the momentum premium, calculated by subtracting the average return on lowest-performing firms to highest-performing firms lagged one month.

4.1.2 Abnormal Return and Cumulative Abnormal Return

The abnormal return (AR) is defined as the difference between the actual and expected return during the event window. If AR is systematically different from 0, this suggests that the market misprices the ESG incidents and implies a profitable trading rule. The cumulative abnormal return (CAR) is a computed as time-series aggregation of AR.

\[ AR_{it} = R_{it} - E[R_{ist}|X_t] \]  

(4.1.2.1)
\[ CAR(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_t \]  

(4.1.2.2)

The cross-sectional mean abnormal return (AAR) and cumulative abnormal return (CAAR) for any is defined as follows:

\[ AAR_t = \frac{1}{n} AR_t \]  

(4.1.2.3)

\[ CAAR(t_1, t_2) = \sum_{t=t_1}^{t_2} AR_t \]  

(4.1.2.4)

A t-test can be used for testing the following null hypothesis. The null hypothesis of a zero cumulative abnormal return (CAR) is tested against the alternative hypothesis of a nonzero CAR.

\[ H_0 : CAR_{it} = 0 \]  

(4.1.2.5)

\[ H_1 : CAR_{i} \neq 0 \]  

(4.1.2.6)

The test statistic is computed as follows:

\[ s^2_{AR_i} = \frac{1}{M_i - 2} \sum_{t=t_0}^{t_1} e^2_{i,t} \]  

(4.1.2.7)

\[ s^2_{CAR_i} = Ls^2_{AR_i} \]  

(4.1.2.8)

\[ t(CAR) = \frac{CAR_i}{s_{CAR}} \]  

(4.1.2.9)

, where M is the estimation window and L is the event window.

4.1.3 Pros and Cons of Event Study

A strength of event study is that the volatility of returns over a short time window is low, and hence the test power would be high and it would be easier to establish the statistical
relationship. Another advantage is that the expected return component is typically neglig-ible. Therefore, it rules out much of the probability that the abnormal return can be caused by some latent variable.

Shedding light to the criticisms of event study, some people argue that t-test might be misspecified. To further explain, the assumptions for t-test are that the abnormal returns are normally distributed, the variances of abnormal returns are equal across firms, and there is no cross-correlation in abnormal returns. If abnormal returns exhibit either heteroskedasticity or cross-sectional dependence, t-test can be inappropriate. However, J.Brown and B.Warner (1985) rebutted this criticism and conclude from their specification test that “although daily excess returns are so highly non-normal, there is evidence that the mean excess return in a cross-section of securities converges to normality as the number of sample securities increases. Standard parametric test are well specified.”

Another criticism of event study is that other events may occur within the event window, which could confound the result. For example, a positive ESG event could have been a response to some negative news about the firm. In that case, even if the ESG event is value-creating, the stock market may respond negatively because of the previous negative news. Evidence supporting this theory is found in Kruger (2015): Kruger finds that the stock market has a larger response to a positive ESG event if KLD has registered an ESG concern for the company within the last year.

4.2 Results

We applied a pre-event window ranging from 200 to 20 trading days prior to the event, and an event window covering 20 trading days before and after the event.

We chose a simple market model to compute the expected stock return. The first reason is that small error in risk adjustment does not make much difference in a short window. Empirically, J.Brown and B.Warner (1985) test the market model in an event study simulation and find out that it incurs very small type one error. Another reason is that Fama-French and Carhart models are often used to compute monthly abnormal over the long-run. Even though they can be adapted to daily data and short-run windows, their statistical properties as an underlying model for short-term event studies has not been confirmed.

The resulting cross-sectional mean for the CAR is -0.32%. Given that the median market capitalization for stocks in our dataset is around $2.6 billion, the loss caused by a single ESG incident is approximated at $8 million. Applying a two-tailed t-test for CAR, we
conclude a nonzero CAR at 1% significance level. This roughly aligns with the result of Glosner (2017) which finds a CAR of 0.4%.

Figure 4.1 reveals an increase in the cumulative abnormal return in the controversial stock, even though most of it happens prior to the reporting date of negative news. The possible reasons could be the following: (i) the incident is reflected in other aspects of the company and the price is partly adjusted beforehand, (ii) lag between the date on which ESG incident occurred and the news reported, (iii) investors react before the news released based on their speculations, (iv) insider trading, (v) confounding event.

Figure 4.2: Cumulative abnormal return (CAR) within the 21-day event window

The CAR plots by incident subcategory is shown in figure 4.2 and 4.3. The plot shows that many of the event categories have a negative CAR and is significant at 90% confidence level, including product/service, workforce, environmental, and financial. Event categories such as association and corruption have a large increase in the magnitude in CAR, but a conclusion cannot be drawn because of the insufficient amount of data in these categories.

4.3 Longevity Study

In addition to the magnitude of abnormal returns, the longevity is also an interesting subject. If excess returns result from mispricing of ESG incidents rather than risk, we would expect the drift to decline over time for two reasons. First, ESG incidents are persistent but not permanent. Put differently, the value of the intangible asset ignored by market is lower overtime, so there should be less mispricing. Secondly, even for firms that persist, the mispricing may be corrected over time as the market slowly learns about their
We proceed longevity study similar to Edmans (2011). The longevity of underperformance can be calculated with the cumulative abnormal return (CAR) up to time $t$. The result shows that the magnitude of CARs continue to grow through around day 777, but starts to decline after that. This suggests that it takes around 26 months for the abnormal return to decline. The result is consistent with our hypothesis and implies that the reduction in drift could occur because the market gradually learns how to price intangibles. Note that our event window is around 4 years, so we are only focusing events before the beginning of 2015.

<table>
<thead>
<tr>
<th>day</th>
<th>CAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td>0.000268</td>
</tr>
<tr>
<td>-9</td>
<td>-0.000045</td>
</tr>
<tr>
<td>-8</td>
<td>-0.000339</td>
</tr>
<tr>
<td>-7</td>
<td>-0.000246</td>
</tr>
<tr>
<td>-6</td>
<td>-0.000519</td>
</tr>
<tr>
<td>-5</td>
<td>-0.000841</td>
</tr>
<tr>
<td>-4</td>
<td>-0.000979</td>
</tr>
<tr>
<td>-3</td>
<td>-0.001410</td>
</tr>
<tr>
<td>-2</td>
<td>-0.001414</td>
</tr>
<tr>
<td>-1</td>
<td>-0.001960</td>
</tr>
<tr>
<td>0</td>
<td>-0.002695</td>
</tr>
<tr>
<td>1</td>
<td>-0.002693</td>
</tr>
<tr>
<td>2</td>
<td>-0.003080</td>
</tr>
<tr>
<td>3</td>
<td>-0.002884</td>
</tr>
<tr>
<td>4</td>
<td>-0.002804</td>
</tr>
<tr>
<td>5</td>
<td>-0.003094</td>
</tr>
<tr>
<td>6</td>
<td>-0.002636</td>
</tr>
<tr>
<td>7</td>
<td>-0.002638</td>
</tr>
<tr>
<td>8</td>
<td>-0.002867</td>
</tr>
<tr>
<td>9</td>
<td>-0.003029</td>
</tr>
<tr>
<td>10</td>
<td>-0.003186</td>
</tr>
</tbody>
</table>

Table 4.1: Values of cumulative abnormal return (CAR) within the 21-day event window value.
Figure 4.3: Cumulative abnormal return (CAR) for each event subcategory within the 21-day event window.
4.3 Longevity Study

Figure 4.4: Cumulative abnormal return (CAR) and confidence interval for each event subcategory at day 21. The confidence interval is constructed at 90% confidence level.

Figure 4.5: Cumulative abnormal return (CAR) plot used for longevity study.
We also study the longevity of events by category. The result shows that while drift for categories social/label, environmental/production, competitive/financial start to decrease between month 21 to 26, the cumulative abnormal return continues to grow for regulatory. This implies that regulatory incidents persist longer or it takes longer for the market to learn to price regulatory incidents.

Edmans (2011) continue to confirm this hypothesis by showing that the abnormal return die down even in firms for which the event is permanent. Because of the scope of study, we will not further explore it in this paper.
Figure 4.6: Cumulative abnormal return (CAR) plot for each event category, used for longevity study.
Chapter 5

Jensen’s Alpha Approach

5.1 Methodology

It is possible that the short-term event study returns substantially underestimate the relationship, as the market fails to fully incorporate intangible price in the short-term. Therefore, we extend the previous study and investigate the impact of ESG incidents on the stock returns in the long run.

For the long-term study, the event study method introduced in the previous section is no longer appropriate. One reason is that the long-horizontal error in risk adjustment accumulates, and thus cannot be neglected anymore as in the short-run. Another reason is that it is unclear which expected return model is correct, so estimates of abnormal returns over long horizons are very sensitive to model choice. Instead, Jensen’s alpha approach is often an alternative choice for event study in the long-run.

5.2 Introduction to Jensen’s Alpha Approach

Jensen’s alpha approach, also called calendar-time portfolios abnormal returns, can be summarized with two steps. The first step involves constructing a portfolio that consists of all firms experiencing the event of interest. The second step is to compute the abnormal return of the portfolio. This is done by a time-series regression of actual monthly returns net of the risk-free-rate(or the benchmark return) against some predefined risk factors that explains the systematic risk. Fama-French three-factor model and Carhart four-factor model are often used in this case.

The intercept, also called alpha, provides the estimated average monthly abnormal return of the portfolio of controversial firms over the market. An alpha significantly different
Jensen’s Alpha Approach

from zero implies that the risk factors fail to fully capture the excess return. R-squared is a statistical measure that represents the percentage of a fund or portfolio’s movement that is explained by the factors in the regression. It can be used to assess how specific ESG investments correlated with the movements of the broader market.

In our study, we investigate the stock performance over 1 year following the ESG incident. Given that the sample period starts in January 2012, we create the portfolio for the first time in January 2013. We reform the portfolio every month and all companies experiencing the event within previous 12 months are added to the portfolio. This procedure results in a time-series of monthly portfolio returns from January 2013 to January 2019. The risk model that we choose is the Carhart four-factor model. The factors are obtained from Fama French’s homepage.

5.3 Improvements and Limitations

The time-series of portfolios returns often suffers autocorrelation and heteroskedasticity in error terms. To address the problem, we estimated the standard error using Newey-West method, the detail of which is further explained in the appendix.

There are several limitations to Jensen’s Alpha approach. First, it can only be computed at a portfolio rather than individual firm level, and therefore they do not produce separate measures of abnormal return for each event. As we want to measure it for each event category, we need to group firms into several portfolios according to the event category and then measure a separate alpha for each portfolio. Moreover, there is no consensus on the appropriate kind and number of risk factors.

Another criticism is that Jensen’s alpha suffers from a low statistical power (Loughran and Ritter 2000), which implies a high Type 2 error. To further explain, it will more likely to find insignificant results that are consistent with market efficiency. The rationale is that corporate executives exploit mispricing by timing corporate events to coincide with the misvaluation, but the Jensen’s alpha approach underweights the timing decisions. As is mentioned in Loughran and Ritter (2000), “if there are time-varying misvaluations that firms capitalize on by taking some action, there will be more events involving larger misvaluations in some periods than in others...... In general, tests that weight firms equally should have more power than tests that weight each time period equally.” Since Jensen’s alpha approach weights each time period equally, it has lower power to detect abnormal returns if managers time corporate events. To address this problem, a commonly used technique is the weighted least square (WLS) regression, which provides greater weight to months with more firms in portfolio. As a common practice, the weighted vector is the
square root of the number of controversial firms in the relevant month. Since we use news
data rather than self-reported corporate events, it is less likely that the events will be
subject to behavioral timing considerations. Therefore, we will not further explore this
and cons of Jensen’s alpha approach.

5.4 Results

Our portfolio contains 341 firms on average. The resulting equal-weighted monthly alpha is
-4.32% and it is statistically significant at level 1%, which implies an abnormal performance
of stock price in the long-term. A value-weighted portfolio return is also computed. It
results in a monthly alpha of -4.37% and is also statistically significant. These results
provide evidence that an US portfolio of controversial firms underperforms the market by
about 4% per month.

<table>
<thead>
<tr>
<th></th>
<th>equal-weighted</th>
<th>value-weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td>0.0014</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>HML</td>
<td>0.0071 *</td>
<td>0.0060</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>HML</td>
<td>0.0119 ***</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>MKT</td>
<td>0.0029</td>
<td>0.0038</td>
</tr>
<tr>
<td>MOM</td>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>CONST</td>
<td>-0.0432 *</td>
<td>-0.0437 *</td>
</tr>
<tr>
<td></td>
<td>(0.022)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>R²</td>
<td>0.322</td>
<td>0.052</td>
</tr>
</tbody>
</table>

Table 5.1: Coefficients for portfolio regression, where four factor-model is used

We also examine events by categories and subcategories. The alphas are significant at 10%
significance level for all of the event categories, except for those with too few observations
that are excluded from our study.

5.5 Robustness Checks

This section is intended to provide further evidence that controversial firms are associated
with negative risk-adjusted stock returns. We run several robustness tests, aiming to
rule out the possibility that the negative alpha stems from outliers, industries, or firm
characteristics.
Table 5.2: Alpha and p-value for each event subcategory

<table>
<thead>
<tr>
<th>Event Category</th>
<th>Alpha</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product/Service</td>
<td>-0.044</td>
<td>0.050</td>
</tr>
<tr>
<td>Regulatory</td>
<td>-0.044</td>
<td>0.0548</td>
</tr>
<tr>
<td>Information</td>
<td>-0.042</td>
<td>0.060</td>
</tr>
<tr>
<td>Discrimination/Workforce</td>
<td>-0.042</td>
<td>0.073</td>
</tr>
<tr>
<td>Anti-Competitive</td>
<td>-0.044</td>
<td>0.048</td>
</tr>
<tr>
<td>Fraud</td>
<td>-0.044</td>
<td>0.046</td>
</tr>
<tr>
<td>Ownership</td>
<td>-0.040</td>
<td>0.084</td>
</tr>
<tr>
<td>Workplace</td>
<td>-0.044</td>
<td>0.051</td>
</tr>
</tbody>
</table>

Table 5.3: Alpha and p-value for each event category

<table>
<thead>
<tr>
<th>Event Category</th>
<th>Alpha</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment/Production</td>
<td>-0.044</td>
<td>0.046</td>
</tr>
<tr>
<td>Competitive/Financial</td>
<td>-0.043</td>
<td>0.053</td>
</tr>
<tr>
<td>Regulatory</td>
<td>-0.043</td>
<td>0.051</td>
</tr>
<tr>
<td>Social/Labour</td>
<td>-0.044</td>
<td>0.055</td>
</tr>
</tbody>
</table>

5.5.1 Outliers

Our first concern is that outliers drive the results. We test this by winsorizing the stock returns of the controversial firms at 1% and 5% before creating the portfolio. The resulting alphas are still statistically significant.

<table>
<thead>
<tr>
<th>1%</th>
<th>5%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td></td>
</tr>
<tr>
<td>-0.0004</td>
<td>-0.0027 *</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>HML</td>
<td></td>
</tr>
<tr>
<td>0.0065</td>
<td>0.0041</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>MKT</td>
<td></td>
</tr>
<tr>
<td>0.0107 ***</td>
<td>0.008 ***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>MOM</td>
<td></td>
</tr>
<tr>
<td>0.0036</td>
<td>0.0043</td>
</tr>
<tr>
<td>(0.004)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>CONST</td>
<td></td>
</tr>
<tr>
<td>-0.0424 *</td>
<td>-0.0409 *</td>
</tr>
<tr>
<td>(0.023)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.256</td>
</tr>
</tbody>
</table>

Table 5.4: Coefficients for portfolio regression, after removing outliers

5.5.2 Misspecified Risk Factors

When estimating abnormal returns, one important concern is a misspecified factor model. We address this problem by additionally estimating the three-factor and five-factor (Fama and R.French 2015) model. The specification for the Fama-French five-factor model is as
follows:

\[ R_{it} - R_{ft} = \alpha_i + \beta_1(R_{MKT_i} - R_{ft}) + \beta_2SMB_i + \beta_3HML_i + \beta_4RMW_t + \beta_5CMA_t + \epsilon_t \]  

(5.5.2.1)

, where RMW is the return spread of the most profitable firms minus the least profitable, and CMA is the return spread of firms that invest conservatively minus aggressively.

These two models both result in statistically significant negative alpha.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td>0.0013</td>
<td>0.002</td>
</tr>
<tr>
<td>HML</td>
<td>0.0051**</td>
<td>0.003</td>
</tr>
<tr>
<td>MKT-RF</td>
<td>0.0116***</td>
<td>0.002</td>
</tr>
<tr>
<td>CONST</td>
<td>-0.0421*</td>
<td>0.022</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.326</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.5: Coefficients for portfolio regression with three-factor model

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Value</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>HML</td>
<td>0.0057</td>
<td>0.005</td>
</tr>
<tr>
<td>MKT-RF</td>
<td>0.0114***</td>
<td>0.002</td>
</tr>
<tr>
<td>RMW</td>
<td>0.0029</td>
<td>0.004</td>
</tr>
<tr>
<td>CMA</td>
<td>-0.0023</td>
<td>0.007</td>
</tr>
<tr>
<td>CONST</td>
<td>-0.0423*</td>
<td>0.022</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.329</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.6: Coefficients for portfolio regression with five-factor model

5.5.3 Industry

It is possible that the portfolio of controversial firms has a negative alpha because they come from some underperforming industries. We investigate this problem by creating
an industry-matched portfolio, using the 12-industry portfolios from Kenneth French’s website.

According to Fama and R.French (1997), each NYSE, AMEX, and NASDAQ stocks are assigned to one of the 12 industries based on the last four digits of company Compustat SIC codes. 12 industrial portfolios are constructed according to this industrial classification and value-weighted returns are computed. As we do not have the access to the SIC data, we retrieve the Global Industry Classification Standard (GICS) industrial groups data for each company and map companies from the 24 industry groups to the 12 industries of the Fama-French industrial portfolio.

For each stock in the portfolio, we subtract the value-weighted returns of the corresponding industry portfolio from the stocks’ raw returns. Then we calculate the risk-adjusted stock returns of the industry-matched portfolio.

\[
\begin{array}{lll}
\text{SMB} & -0.0004 \quad *** \\
         & (0.000) \\
\text{HML} & -0.0002 \\
         & (0.000) \\
\text{MKT} & -0.0001 \\
         & (0.000) \\
\text{MOM} & 0.0002 \\
         & (0.000) \\
\text{CONST} & -0.0002 \\
         & (0.000) \\
\end{array}
\]

\[R^2 = 0.102\]

Table 5.7: Coefficients for portfolio regression on industry-matched returns

The result shows that an alpha of -0.0002 with p-value of 0.534, which is not statistically significant. In other words, the controversial stocks do not significantly underperform their industry comparable. This suggests that the underperformance of controversial stocks which was discovered previously may stem from the fact that these companies come from underperforming industries. Recall from chapter 2 that we found the controversial firms are more concentrated in energy, utility, materials, telecommunications, automobiles, while they are less concentrated on healthcare, pharmaceutical, banks, and software industries. This coincides with the fact that the former industries underperform the latter ones within the recent 10 years, shown in the figure below. One objection of this conclusion is that many companies fail because of firm-specific events instead of widespread industry features. For example, recall from Chapter 2 that the stock price of Celsion drops sharply because of its unsuccessful drug test, while the average return in pharmaceutical industry is positive at the time. Our explanation is that such firm-specific cases with extreme stock returns
are so rare that they do not make statistically significant impact.

Figure 5.1: Monthly performance of several industries from 2010 to 2019 (Fidelity Investment 2019)

We proceed to investigate the excess return for each industry by running separate industrial portfolio regressions, where the industry benchmark return is subtracted from the raw stock return. The result shown below indicates that the alphas for many industries are statistically significant, including money, non-durables, health, shops, energy, and others.

<table>
<thead>
<tr>
<th>industry</th>
<th>alpha</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money</td>
<td>-0.004630</td>
<td>0.000064</td>
</tr>
<tr>
<td>Other</td>
<td>0.003169</td>
<td>0.024338</td>
</tr>
<tr>
<td>BusEq</td>
<td>0.000186</td>
<td>0.938074</td>
</tr>
<tr>
<td>Manuf</td>
<td>0.000441</td>
<td>0.819017</td>
</tr>
<tr>
<td>Hlth</td>
<td>0.007502</td>
<td>0.006547</td>
</tr>
<tr>
<td>Shops</td>
<td>0.004938</td>
<td>0.051794</td>
</tr>
<tr>
<td>NoDur</td>
<td>-0.002271</td>
<td>0.086337</td>
</tr>
<tr>
<td>Enrgy</td>
<td>0.013829</td>
<td>0.001568</td>
</tr>
<tr>
<td>Utils</td>
<td>-0.001846</td>
<td>0.195401</td>
</tr>
<tr>
<td>Telm</td>
<td>-0.005850</td>
<td>0.123074</td>
</tr>
<tr>
<td>Chems</td>
<td>0.001402</td>
<td>0.730204</td>
</tr>
<tr>
<td>Durbl</td>
<td>-0.006831</td>
<td>0.149209</td>
</tr>
</tbody>
</table>

Table 5.8: Alphas and p-values of portfolio regressions for each industry

5.5.4 Firm Characteristics

To rule out the possibility that the negative alpha stems from firm characteristics, We took the approach from Wermers (2003), which matches each stock to a portfolio of stocks with similar market capitalization, book-to-market ratio, and momentum.
Market capitalization is the firm’s market value, computed as the outstanding shares by the price of a single share. It is one of the basic determinants of various characteristics including risk. The smaller the market capitalization, the higher the investment risk.

Book-to-market ratio is defined as the dividend of a firm’s book value to market value. Book value can be calculated by subtracting total liabilities, preferred shares, and intangible assets from the total assets of the company. It represents how much a company would have left in assets if it went out of business. If a firm’s book-to-market ratio is large, it means that its market value is trading higher than its book value per share and that the company is overvalued.

Momentum here is defined as the sum of 12-months past returns of the stock. It considers the trend that the stock price is likely to continue to rise if it has a high past-year return.

Considering the characteristics-matched portfolio, specifically we first rank all stocks traded on NYSE, NASDAQ, and AMEX stocks by their market capitalization at the end of each June. Quintile portfolios are formed according to the rank, and each quintile portfolio is further divided into book-to-market ratio quintiles, based on their most recently available data. Finally, each of the resulting 25-group portfolios is further divided into quintiles based on the 12-month past returns of stocks. This three-way ranking procedure results in 125-group portfolios, each having different combinations of size, book-to-market ratio, and momentum. The portfolio is turned over every year. Each controversial firm is mapped to one of these portfolios based on their characteristics and the equal-weighted portfolio return is subtracted from the company’s raw stock return.

We are left with 42 stocks in each fractile. The result shows an alpha that is not statistically
significant. This suggests firm characteristics, specifically size, book-to-market-ratio, and momentum, is another source for the underperformance of controversial firms.

We proceed to examine which characteristics specifically changes the result and leads to a statistically insignificant alpha. We achieve this by removing one factor from our portfolio-forming procedure at a time. This results in three 25-group portfolios, formed on i) book-to-market ratio and yearly return, ii) market capitalization and yearly return, iii) market capitalization and book-to-market ratio. The result shows that the alpha for portfolio specification ii) remains insignificant, while for i) and iii) are negative and statistically significant. This implies that that the excess return is not robust to the characteristics combination of market capitalization and book-to-market ratio.

<table>
<thead>
<tr>
<th></th>
<th>mc, bv, yr</th>
<th>bm, yr</th>
<th>mc, yr</th>
<th>mc, bm</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMB</td>
<td>-0.0004***</td>
<td>-0.0046***</td>
<td>-0.0032***</td>
<td>-0.0033***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>HML</td>
<td>4.374e-05</td>
<td>-0.0003</td>
<td>0.0004</td>
<td>-0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>MKT</td>
<td>0.0025 *</td>
<td>0.0007</td>
<td>0.0003</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.0027***</td>
<td>-0.0003</td>
<td>-0.0012***</td>
<td>-0.0013***</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>CONST</td>
<td>0.0023</td>
<td>-0.0037**</td>
<td>-0.0024*</td>
<td>-0.0016</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.063</td>
<td>0.532</td>
<td>0.556</td>
<td>0.536</td>
</tr>
</tbody>
</table>

Table 5.9: Coefficients for portfolio regressions with characteristics-matched portfolio returns
Chapter 6

Cross-sectional Time-Series Regression

6.1 Methodology

6.1.1 Model Specification

A concern with the previous approaches is that the underperformance of controversial firms stems from unknown omitted variables rather than the unexplained abnormal return. The characteristics-matched approach is aimed to solve the problem, but as the number of variables increase, the number of stocks in each group decreases, and thus the method does not work anymore. In order to address this problem, in this section we apply the cross-sectional time-series regression to control for a wide range of firm characteristics. The regression model is specified as follows:

\[ R_{i,t} = \alpha_0 + \beta_1 X_{i,t-k} + \beta_2 Z_{i,t-k} + \epsilon_{i,t} \]  

(6.1.1.1)

where \( R_{i,t} \) is the return of stock i, \( X_{i,t-k} \) are the factors that control for firm characteristics, and \( Z_{i,t-k} \) is an indicator for ESG incident. We tried three specifications for the indicator: i) a dummy variable, which equals to 1 if there are at least 1 ESG incidents in month i for stock t, and equals to 0 otherwise, ii) a count of number of ESG incidents in month i, and iii) separate counts of number of ESG incidents for each event subcategory. The indicator is lagged 1 month.

Following Brennan, Chordia, and Subrahmanyam (1998) and Edmans (2011), we use firm-characteristics-related factors as control variables, which include the following: “Variable
size” is the logarithms of market capitalization in terms of billions of US dollars. “Variable bm” is the logarithm of book-to-market ratio. “Variable div” is the ratio of dividends paid over the market value. “Variables ret23”, “ret46”, and “ret712” are the logarithms of compounded returns returns over previous months t-2 through t-3, t-4 through t-6, and t-7 through t-12. “Variable prc” is the logarithm of the reciprocal of the stock price. These risk factors are all lagged 1 month.

6.1.2 Fama-MacBeth Regression

We apply the Fama-Macbeth model (Fama and MacBeth 1973) to estimate the betas and risk premia for the risk factors.

To obtain the Fama-MacBeth estimator, we perform T regressions, one for each sample period using all of the entity data. The model is given by:

\[ y_{it} = \beta' x_{it} + \epsilon_{it} \]  

(6.1.2.1)

Denote the estimated model parameters as \( \hat{\beta}_t \). The reported estimator is then:

\[ \hat{\beta} = T^{-1} \sum_{t=1}^{T} \hat{\beta}_t \]  

(6.1.2.2)

\[ \sigma(\hat{\beta}) = \frac{\sigma(\hat{\beta}_t)}{\sqrt{M}} \]  

(6.1.2.3)

One criticism of the Fama-MacBeth regression is that it is robust only against time but not against firm effects. Alternatively, we can apply a pooled-OLS panel regression and double-cluster the standard error to account for cross-sectional and serial correlations. The detail of it is explained in appendix A.

6.2 Results
Table 6.1: Summary statistics of control variables

<table>
<thead>
<tr>
<th></th>
<th>mc</th>
<th>prc</th>
<th>bm</th>
<th>div</th>
<th>ret23</th>
<th>ret46</th>
<th>ret712</th>
<th>ret</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>1.915</td>
<td>-1.301</td>
<td>-0.829</td>
<td>2.372</td>
<td>0.002</td>
<td>0.005</td>
<td>0.009</td>
<td>0.010</td>
</tr>
<tr>
<td>std</td>
<td>1.476</td>
<td>0.198</td>
<td>0.660</td>
<td>1.595</td>
<td>0.020</td>
<td>0.028</td>
<td>0.045</td>
<td>0.070</td>
</tr>
<tr>
<td>min</td>
<td>-2.978</td>
<td>-1.819</td>
<td>-3.217</td>
<td>0.018</td>
<td>-0.132</td>
<td>-0.198</td>
<td>-0.329</td>
<td>-0.3090</td>
</tr>
<tr>
<td>25%</td>
<td>0.872</td>
<td>-1.441</td>
<td>-1.268</td>
<td>1.263</td>
<td>-0.008</td>
<td>-0.009</td>
<td>-0.014</td>
<td>-0.031</td>
</tr>
<tr>
<td>50%</td>
<td>1.951</td>
<td>-1.322</td>
<td>-0.789</td>
<td>2.038</td>
<td>0.003</td>
<td>0.006</td>
<td>0.011</td>
<td>0.011</td>
</tr>
<tr>
<td>75%</td>
<td>3.044</td>
<td>-1.184</td>
<td>-0.350</td>
<td>3.050</td>
<td>0.013</td>
<td>0.020</td>
<td>0.035</td>
<td>0.051</td>
</tr>
<tr>
<td>max</td>
<td>5.307</td>
<td>-0.312</td>
<td>1.049</td>
<td>13.553</td>
<td>0.136</td>
<td>0.188</td>
<td>0.246</td>
<td>0.408</td>
</tr>
</tbody>
</table>

Table 6.2: Results of cross-sectional time-series regression, where the event indicator is a count of the number of events
<table>
<thead>
<tr>
<th></th>
<th>double-clustered</th>
<th>Fama-Macbeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>dummy</td>
<td>5.233e-05</td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.0016)</td>
<td>(0.0013)</td>
</tr>
<tr>
<td>mc</td>
<td>-0.0009</td>
<td>-0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0007)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>prc</td>
<td>0.0151 ***</td>
<td>0.0073 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0025)</td>
</tr>
<tr>
<td>bm</td>
<td>0.0019</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0010)</td>
<td>(0.0009)</td>
</tr>
<tr>
<td>div</td>
<td>-0.0002</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.0005)</td>
</tr>
<tr>
<td>ret23</td>
<td>0.0569</td>
<td>0.0076</td>
</tr>
<tr>
<td></td>
<td>(0.1161)</td>
<td>(0.0411)</td>
</tr>
<tr>
<td>ret46</td>
<td>0.0721</td>
<td>-0.0170</td>
</tr>
<tr>
<td></td>
<td>(0.070)</td>
<td>(0.0339)</td>
</tr>
<tr>
<td>ret712</td>
<td>0.0480</td>
<td>0.0029</td>
</tr>
<tr>
<td></td>
<td>(0.0455)</td>
<td>(0.0217)</td>
</tr>
<tr>
<td>const</td>
<td>0.0322 ***</td>
<td>0.0227 ***</td>
</tr>
<tr>
<td></td>
<td>(0.0089)</td>
<td>(0.0057)</td>
</tr>
</tbody>
</table>

Table 6.3: Results of cross-sectional time-series regression, where the event indicator is a dummy variable

Both results from Fama-MacBeth regression and pooled-OLS with double-clustered standard error show that neither a dummy indicator nor a count of the number of events has a significant effect on the stock return. This result is consistent with our finding from the characteristic-matched portfolio that the excess returns from ESG incidents are not robust to firm characteristics.
### 6.2 Results

<table>
<thead>
<tr>
<th>Category</th>
<th>double-clustered</th>
<th>Fama-Macbeth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association</td>
<td>0.0655</td>
<td>0.0018</td>
</tr>
<tr>
<td>Sanctions</td>
<td>0.0398 *</td>
<td>0.0016</td>
</tr>
<tr>
<td>Financial</td>
<td>-0.0072</td>
<td>0.0016</td>
</tr>
<tr>
<td>Corruption</td>
<td>-0.0067</td>
<td>0.0003</td>
</tr>
<tr>
<td>Information</td>
<td>0.0009</td>
<td>-0.0076 *</td>
</tr>
<tr>
<td>Human</td>
<td>-0.0174</td>
<td>0.0002</td>
</tr>
<tr>
<td>Workplace</td>
<td>0.0008</td>
<td>-0.0090 **</td>
</tr>
<tr>
<td>Production/Supply</td>
<td>0.0013</td>
<td>-0.0022</td>
</tr>
<tr>
<td>Environmental</td>
<td>0.0008</td>
<td>0.0040</td>
</tr>
<tr>
<td>Management</td>
<td>0.0066</td>
<td>0.0055</td>
</tr>
<tr>
<td>Workforce</td>
<td>-0.0078</td>
<td>-0.0112 *</td>
</tr>
<tr>
<td>Regulatory</td>
<td>0.0008</td>
<td>0.0048</td>
</tr>
<tr>
<td>Fraud</td>
<td>-0.0007</td>
<td>-0.0051</td>
</tr>
<tr>
<td>Anti-Competitive</td>
<td>-0.0102</td>
<td>-0.0017</td>
</tr>
<tr>
<td>Ownership</td>
<td>0.0054</td>
<td>-0.0009</td>
</tr>
<tr>
<td>Product/Service</td>
<td>-0.0016</td>
<td>-0.0016</td>
</tr>
<tr>
<td>Discrimination/Workforce</td>
<td>0.0015</td>
<td>0.0036</td>
</tr>
<tr>
<td>mc</td>
<td>-0.0008</td>
<td>-0.0007</td>
</tr>
<tr>
<td>prc</td>
<td>0.0154 ***</td>
<td>0.0077 ***</td>
</tr>
<tr>
<td>bm</td>
<td>0.0018 *</td>
<td>0.0003</td>
</tr>
<tr>
<td>div</td>
<td>-0.0002</td>
<td>-0.0008</td>
</tr>
<tr>
<td>ret23</td>
<td>0.0571</td>
<td>0.0003</td>
</tr>
<tr>
<td>ret46</td>
<td>0.0720</td>
<td>-0.0144</td>
</tr>
<tr>
<td>ret712</td>
<td>0.0471</td>
<td>0.0014</td>
</tr>
<tr>
<td>const</td>
<td>0.0326 ***</td>
<td>0.0232 ***</td>
</tr>
</tbody>
</table>

| const                     | 0.0090           | 0.0057       |

Table 6.4: Results of cross-sectional time-series regression, where the number of events is counted for each subcategory.
We proceed to run the regression with our third specification of event indicator, which separately counts the number of events for each subcategory. The pooled OLS result shows that while for most of the categories the coefficients remain insignificant, the coefficients for management, human, and information are statistically significant at 5%, 10%, and 5% level. To be more specific, while other variables remain the same, companies with one additional management incident have 1.12% lower stock returns, companies with one additional human incident have 0.9% lower stock return, and companies with one additional information incident have 0.76% lower stock return for the month.
Chapter 7

Portfolio Construction

In order to evaluate whether portfolios with low ESG-profile underperform typical passive benchmarks, we compare and construct a series of portfolios, based on several ESG investing strategies. According to the GSIA by Global Sustainable Investment Alliance (2018), ESG investing strategies are divided into the following main categories:

- **Negative screening**: the exclusion of certain companies or sectors from the portfolio based on ESG criteria.
- **Relative/Best-in-class screening**: investment in companies or sectors based on ESG criteria relative to their industry peers.
- **Norm-based screening**: screening of portfolios against some minimum standards of business practice according to international norms, such as those issued by OECD and UN.
- **ESG inclusion**: systematic and explicit incorporation of ESG factors into financial analysis
- **Themed investing**: investment in themes specifically related to ESG considerations
- **Impact investing**: targeted investments aimed at solving ESG problems
- **Corporate engagement**: utilization of shareholder power to influence corporate behaviors, including direct corporate engagement, filing or co-filing shareholder proposals, and proxy voting that is consistent with ESG guidelines.

While investors using themed investing, impact investing, and norm-based screening strategies are more concerned with social responsibility than financial returns, corporate engagement and ESG inclusion are rarely used by individual investors and hard to evaluate. Therefore, in this section we only focus on examining the financial implication of negative
and positive screenings.

7.1 Methodology

Firstly, we construct portfolios based on negative and positive screenings. Then we model the difference between monthly portfolio returns and benchmark returns using four-factor model, which is the same as Jensen’s Alpha approach. Finally, we do hypothesis testing to check if the alpha is statistically significantly different from 0. If not, this implies that ESG portfolios do not significantly underperform or overperform the benchmarks.

A critique of applying Jensen’s Alpha method to investigate the excess returns of positive or negative screening strategy is that this method is primarily concerned with the exposure of market risk and is thus only appropriate for well-diversified portfolios. However, ESG screenings restrict the investment universe. In addition, there is no consensus on the appropriate kind and number of risk factors. Instead, Ledoit and Wolf (2008) suggested comparing the two portfolios by hypothesis testing of Sharpe ratio. Sharpe ratio is defined as the expected return in excess of risk-free-rate per unit of volatility and is widely used to understand the risk-adjusted return potentials for portfolios. We also applied this method and provide the results from both methods below. Details of the hypothesis testing of Sharpe ratio is explained in Appendix B.

7.2 Relative Screening

Relative screening refers to the investment in companies selected based on the criteria of ESG performance relative to industry peers.

In our case, for each month we rank the companies by the number of previous ESG incidents as an approximation for ESG score. Based on this ranking, we select the 30 worst stocks with the most number of previous ESG incidents and form an equal-weighted portfolio. This gives a monthly time-series return from April 2012 to February 2019. Similarly, we also construct a portfolio with the 30 best stocks for companies with the least number of previous ESG incidents, which is used as a benchmark. The other two benchmark portfolios are S&P 500 stocks and all publicly traded US stocks on NASDAQ, NYSE, and AMEX.
7.2 Relative Screening

<table>
<thead>
<tr>
<th></th>
<th>all US stocks</th>
<th>S&amp;P 500</th>
<th>best 30 stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.0003</td>
<td>-0.0010 **</td>
<td>-0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>HML</td>
<td>-0.0010</td>
<td>-0.0014 *</td>
<td>-0.0022 ***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>SMB</td>
<td>-0.0107 ***</td>
<td>-0.0046 ***</td>
<td>-0.0078 ***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.0025 ***</td>
<td>-0.0028 ***</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>const</td>
<td>9.422e-07</td>
<td>-0.0009</td>
<td>-0.0054 **</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.677</td>
<td>0.345</td>
<td>0.414</td>
</tr>
</tbody>
</table>

Table 7.1: Coefficients for equal-weighted portfolio consisted of worst 30 stocks, compared with all US stocks, SP500 stocks, and best 30 stocks

<table>
<thead>
<tr>
<th></th>
<th>worst 30 stocks</th>
<th>best 30 stocks</th>
<th>S&amp;P500 stocks</th>
<th>all US stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe Ratio</td>
<td>0.2342239</td>
<td>0.3217523</td>
<td>0.3264605</td>
<td>0.2260067</td>
</tr>
<tr>
<td>P value</td>
<td>0.3133373</td>
<td>0.2337532</td>
<td>0.2337532</td>
<td>0.9346131</td>
</tr>
</tbody>
</table>

Table 7.2: Results for Sharpe Ratio test, where the worst 30 stocks are compared with best 30, SP500, and all US stocks

Figure 7.1: Profit and loss graph for equal-weighted portfolio of worst 30 stocks, compared with SP500, all US, and best 30 stocks

The result shows that a equal-weighted portfolio of the worst ESG-performing stocks underperforms the best ones with a negative alpha of -0.54% and it is statistically significant at 5% significance level. However, it does not systematically underperform S&P 500 or all US stocks. Moreover, Sharpe ratio test does not show any statistical significance for any
benchmark strategies.

Value-weighted returns are also computed. A value-weighted portfolio of the worst ESG-performing stocks outperforms S&P 500 stocks by 0.29% at 1% significance level, but not systematically outperform all US stocks or underperform best stocks.

<table>
<thead>
<tr>
<th></th>
<th>all US stocks</th>
<th>S&amp;P 500</th>
<th>best 30 stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>0.0020 ***</td>
<td>0.0020 ***</td>
<td>0.0025 *</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>HML</td>
<td>-0.0012</td>
<td>-0.0043 ***</td>
<td>-2.127e-05</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>SMB</td>
<td>-0.0018 **</td>
<td>-0.0008</td>
<td>-0.0048 *</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>MOM</td>
<td>0.0007</td>
<td>0.0006</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>const</td>
<td>0.0003</td>
<td>0.0029 **</td>
<td>-0.0052</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.175</td>
<td>0.344</td>
<td>0.072</td>
</tr>
</tbody>
</table>

Table 7.3: Coefficients for value-weighted portfolio consisted of worst 30 stocks, compared with all US stocks, SP500 stocks, and best 30 stocks

<table>
<thead>
<tr>
<th></th>
<th>worst 30 stocks</th>
<th>best 30 stocks</th>
<th>SP500 stocks</th>
<th>all US stocks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe Ratio</td>
<td>0.4181311</td>
<td>0.3886057</td>
<td>0.3915104</td>
<td>0.4394484</td>
</tr>
<tr>
<td>P value</td>
<td>0.7280544</td>
<td>0.3757508</td>
<td>0.6802639</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4: Results for Sharpe Ratio test, where the a portfolio of value-weighted worst 30 stocks are compared with best 30, SP500, and all US stocks

Figure 7.2: profit and loss graph for value-weighted portfolio of worst 30 stocks, compared with SP500, all US, and best 30 stocks
7.3 Negative Screening

Negative screening entails the exclusion of certain companies from the portfolio based on some ESG criterion. The benchmark portfolios are constructed using both the S&P 500 stocks and all public traded US stocks on NYSE, NASDAQ, and AMEX. In our dataset, approximately 69% of S&P 500 stocks and 19% of all public traded US stocks are controversial firms. For every month, we remove the stocks with historical ESG-related news in the past to construct our comparable ESG portfolio.

<table>
<thead>
<tr>
<th></th>
<th>all US stocks</th>
<th>S&amp;P500</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.0002 ***</td>
<td>-0.0004 ***</td>
</tr>
<tr>
<td></td>
<td>(5.59e-05)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>HML</td>
<td>-3.809e-05</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>SMB</td>
<td>-7.069e-05</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(5.49e-05)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>MOM</td>
<td>-7.324e-05</td>
<td>-9.124e-05</td>
</tr>
<tr>
<td></td>
<td>(9.81e-05)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>CONST</td>
<td>-0.0003</td>
<td>9.811e-05</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.173</td>
<td>0.164</td>
</tr>
</tbody>
</table>

Table 7.5: Coefficients for equal-weighted negatively screened portfolio, compared with benchmarks

<table>
<thead>
<tr>
<th></th>
<th>all public US stocks</th>
<th>negative screened</th>
<th>SP500 stocks</th>
<th>negative screened</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sharpe Ratio</td>
<td>0.2447</td>
<td>0.2333</td>
<td>0.3467</td>
<td>0.3463</td>
</tr>
<tr>
<td>P value</td>
<td>0.0856</td>
<td></td>
<td>0.9650</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.6: Results for Sharpe Ratio test, where equal-weighted portfolios are compared

<table>
<thead>
<tr>
<th></th>
<th>all US stocks</th>
<th>S&amp;P500</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.0007 ***</td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>HML</td>
<td>0.0002</td>
<td>-4.873e-05</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>SMB</td>
<td>-0.0003</td>
<td>-0.0004 **</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>MOM</td>
<td>-0.0007</td>
<td>-0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>CONST</td>
<td>-0.0005</td>
<td>-0.0003</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.217</td>
<td>0.089</td>
</tr>
</tbody>
</table>

Table 7.7: Coefficients for value-weighted negatively screened portfolio, compared with benchmarks
The portfolio regression results show that negatively screened portfolios slightly underper-
7.3 Negative Screening

Figure 7.5: Profit and loss graph for value-weighted negative screened portfolio on all US stocks

Figure 7.6: Profit and loss graph for value-weighted negative screened portfolio on S&P 500 stocks

form the benchmarks, but none of them is statistically significant. In other words, both equal-weighted and value-weighted negatively screened portfolios do not systematically underperform or overperform S&P 500 portfolio or all US stocks portfolio. Hypothesis testing of Sharpe ratio gives slightly different results: while the equal-weighted portfolio obtained from negative screening all public US stocks underperform the unscreened
portfolio at 10% significance level, the rests remain insignificant.

The statistically insignificant results might stem from the fact that the dataset used here does not have any severity index for each ESG-related news event. This makes it impossible for us to distinguish the major problems from the minor ones, and thus, ESG screening becomes less effective.
Chapter 8

Conclusion and Future Work

There is a long-running debate on the financial implications of ESG investing. It indicates whether investors have to sacrifice financial performance by investing socially responsible. ESG investors can apply different screening strategies to construct a socially responsible portfolio. In this thesis, we analyze two strategies, namely negative and positive screening. After implementing these strategies using controversial firms obtained from Dow Jones Adverse Media Entity dataset that covers 2012 to 2019, we get the following main results:

i) ESG risks generate negative excess stocks returns in the short-run. Specifically, it generates a cumulative abnormal return of -0.32% for a 21-day event window. It demonstrates that the market fails to fully incorporate ESG information in the short-run.

ii) In the long-run, a monthly excess return of -4.37% for portfolios of controversial firms is reported. The result is robust to the removal of outliers and different model specifications. However, it is not robust to industry-matched returns, which implies that the underperformance of controversial stocks can stem from underperforming industries. Within industry information, management, and human, the negative alphas remain statistically significant. Moreover, the negative alpha is not robust to several combinations of firm characteristics, including size, book-to-market ratio, and momentum.

iii) ESG portfolios with negative screening strategy do not significantly overperform or underperform passive investing portfolios, including S&P stocks and all US stocks publicly traded on NASDAQ, NYSE, and AMEX. This implies that for socially responsible investors, negative screening ESG stocks does not harm the financial performance of their portfolios, but neither does it brings extra financial profits.

iv) An equal-weighted portfolio with relative screening strategy, which selects worst ESG-performing stocks, underperform the best-performing ones by -0.54%, while a value-weighted portfolio outperforms S&P 500 portfolio by 0.29%. This suggests shorting or
longing controversial firms with extreme numbers of ESG incidents could bring excess returns.

One major concern of our study is the short sample period, which only covers seven years from 2012 to 2019. The sample size for some event subcategories, such as sanctions and association, is especially small during this sample period. This forms an obstacle for precise inference for two main reasons. Firstly, the asset pricing tests have relatively low statistical power, which implies a lower probability of detecting abnormal returns. Secondly, the results are sensitive to particular developments during the period. For instance, the economy in the US has been in an expansionary stage for most of the sample period and statistical techniques may not be able to fully explain such realizations of underlying factors. To provide evidence that the findings from the portfolio regressions in these seven years are not spurious, we could further run out-of-sample tests with, for instance, European stock data. A second concern is the lack of severity index in our dataset, which is mentioned in chapter 7. This impedes us from distinguishing major events from minor ones, and thus potentially introduce noises with minor events. Taking the significance of each news events by mining the raw text data of each news into account might overcome this issue, and this is left for future work.

A further improvement we could make is to apply buy-and-hold abnormal return (BHAR) method in addition to Jensen’s alpha approach. This involves calculating a stock’s un-adjusted return by geometrically compounding its monthly returns and subtracting the benchmark returns. Conrad and Kaul argue that BHAR is a better methodology than Jensen’s alpha for some statistical reasons.

One interesting subject for further study is the earnings surprises. Our hypothesis is that the intangible consequences of ESG incidents, such as reputational loss, are not immediately capitalized by the market. Instead, it only affects the stock price when it subsequently manifests in tangible outcomes, thus generating lower profits overtime. To confirm the hypothesis, we could firstly examine whether controversial firms have lower operating performance, such as return on equity (ROE) and return on assets (ROA). As weaker operating performance only results in abnormal returns if it is unexpected by the stock market, we could proceed with investigating the earning surprise, which is calculated as the actual earnings per share less the analysts’ forecast. Another interesting topic would be studying the values of positive ESG events and the financial implication of the positive screening strategy. This entails a dataset of positive news and is left for future study.
Bibliography


BIBLIOGRAPHY


MSCI. Esg-rating agency.


RepRisk. Esg-rating agency.


Sustainalytics. Esg-rating agency.


Appendix A

Heteroskedasticity and Autocorrelation Consistent Standard Errors

A.1 Ordinary Least Square (OLS)

Linear Regression assumes the following form:

\[ Y = \beta X + \epsilon \]  \hspace{1cm} (A.1.0.1)

It is normally assumed that \( \epsilon_i, ..., \epsilon_j \) are i.i.d. with \( E[\epsilon_i] = 0, \text{Var}(\epsilon_i) = \sigma^2 \)

The ordinary least square (OLS) estimator \( \hat{\beta} \) is defined as:

\[ \hat{\beta} = \arg\min ||Y - X\hat{\beta}||^2 \]  \hspace{1cm} (A.1.0.2)

This gives the solution:

\[ \hat{\beta} = (X^T X)^{-1} X^T Y \]
\[ = (X^T X)^{-1} X^T (X\beta + \epsilon) \]  \hspace{1cm} (A.1.0.3)
\[ = \beta + (X^T X)^{-1} X^T \epsilon \]

For the OLS estimators to be unbiased and efficient, the following assumptions need to
The errors have mean 0. \( \mathbb{E}[\epsilon] = 0 \).

- The variance of error is constant (homoskedasticity). \( \text{VAR}(\epsilon_i) = \sigma^2 \) for all \( i \).
- The errors are uncorrelated. \( \text{COV}(\epsilon_i, \epsilon_j) = 0 \) for all \( i \neq j \).
- The regression model is linear in parameters.
- \( X \) is independent of \( e \).
- No multi-collinearity.
- The errors are jointly normally distributed (optional).

If assumptions 1-3 hold, Gauss-Markov theorem suggests that OLS is the best linear unbiased estimator (BLUE). Put differently, OLS estimator will have the lowest variance, compared with other unbiased linear estimators.

However, the error terms of time series data are oftentimes serially correlated (autocorrelated) and heteroskedastic. In such cases, the point estimators are still unbiased, but the estimation for standard errors become inaccurate, and thus the statistical inference and tests provide unreliable results. Plus, OLS is no longer BLUE anymore.

\[
\mathbb{E}[\hat{\beta}] = \mathbb{E}[\beta + (X^T X)^{-1} X^T \epsilon] \\
= \beta + (X^T X)^{-1} X^T \mathbb{E}[\epsilon] \\
= \beta
\]  
(A.1.0.4)

\[
\text{COV}[\hat{\beta}] = \mathbb{E}[(\beta - \hat{\beta})(\beta - \hat{\beta})^T] \\
= \mathbb{E}[(X^T X)^{-1} X^T \epsilon(X^T X)^{-1} X^T \epsilon^T] \\
= (X^T X)^{-1} X^T \mathbb{E}[\epsilon \epsilon^T] X (X^T X)^{-1}
\]  
(A.1.0.5)

If the assumptions are met,

\[
\mathbb{E}[\epsilon \epsilon^T] = \text{COV}[\epsilon] \\
= \Omega \\
= \sigma^2 I
\]  
(A.1.0.6)
A.2 Heteroskedasticity and Serial-Correlation

\[
COV[\hat{\beta}] = (X^T X)^{-1} X^T E[\epsilon \epsilon^T] X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1} \tag{A.1.0.7}
\]

A.2 Heteroskedasticity and Serial-Correlation

If the homoskedasticity and non-serial-correlation properties are violated, \( \Omega \neq \sigma^2 I \) and thus the above formula for estimating the covariance matrix will be inaccurate. If successive errors are positively correlated, we are underestimating the standard error of parameters and t-statistic will be much too high, and vice versa.

A.2.1 Correlogram and Residual Plots

Correlogram and residual plot can provide a first inspection on these issues.

Correlogram

Correlogram is the standard means of visualization for autocorrelation function, which is calculated by:

\[
\rho(k) = \text{Cor} (X_{t+k}, X_t) = \frac{\text{Cov} (X_{t+k}, X_t)}{\sqrt{\text{Var} (X_{t+k})} \sqrt{\text{Var} (X_t)}} \tag{A.2.1.1}
\]

A plug-in estimation is then:

\[
\hat{\rho}(k) = \frac{\sum_{t=1}^{n-k} (x_{s+k} - \bar{x}) (x_s - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2}, \text{ for } k = 1, \ldots, n - 1 \tag{A.2.1.2}
\]

Confidence band is also shown in the correlogram. It is computed based on the asymptotic assumption that \( \rho(k) \) follows a \( N(0, \frac{1}{n}) \) distribution for long i.i.d. time series. Sample autocorrelation coefficients \( \hat{\rho}(k) \) that fall within the confidence band of \( \pm \frac{1.96}{\sqrt{n}} \) are considered to be different from 0 only by chance, whereas those outside are considered to be truly different from 0.

I plotted the correlogram for the time-series of excess returns of the portfolio that is consisted of all controversial stocks over the risk-free-rate. This series is used in chapter 4. A rough inspection of the correlogram shows that the autocorrelation exceeds the confidence bound for lags up to 7.
Heteroskedasticity and Autocorrelation Consistent Standard Errors

Figure A.1: Correlogram of time-series of portfolio return consisted of controversial firms over the risk-free-rate

Residual Plot

The residual plot is used to identify heteroskedasticity. In this case, the variance appears to increase with increased value of fitted excess return. Therefore, there is some slight deviation from the constant variance assumption.

Figure A.2: Residual plot of portfolio return consisted of controversial firms over the risk-free-rate
A.2.2 Tests for Heteroskedasticity and Serial-Correlation

White Test for Heteroskedasticity

White Test is a formal test for constant variance. The null hypothesis is $H_0 : \sigma_i^2 = \sigma^2$ and the alternative hypothesis is $H_1 : \sigma_i^2 \neq \sigma^2$. The squared residual from the original OLS is regressed against the original regressors along with their squares and cross-products. The lagrange multiplier (LM) is then calculated as the product of sample size and the $R^2$ value:

$$LM = nR^2$$ (A.2.2.1)

The LM statistics should follow a $\chi^2$ distribution, with degree of freedom $P-1$, where $P$ is the number of regressors in the auxiliary regression.

Durbin-Watson Test for Serial-Correlation

Assume that the regressors are strictly exogenous, which rule out the lagged dependent variables. Also assume the AR(1) case, where the error follows a first order Markov error process as follows:

$$u_t = \rho u_{t-1} + e_t, |\rho| < 1, e_{i.i.d} \sim N(0,c)$$ (A.2.2.2)

The null hypothesis is $H_0 : \rho = 0$ and the alternative hypothesis is $H_1 : \rho \neq 0$. The Durbin-Watson test statistics is computed as:

$$DW = \frac{\sum_{t=2}^{n} (v_t - v_{t-1})^2}{\sum_{t=1}^{n} v_t^2}$$ (A.2.2.3)

, where $v$ is the residual.

Simple algebra shows that DW statistic is closely related to the estimated $\rho$ from the large sample test:

$$DW \approx 2(1 - \hat{\rho})$$ (A.2.2.4)

$d = 2$ indicates no autocorrelation.
Heteroskedasticity and Autocorrelation Consistent Standard Errors

Breusch-Godfrey Test for Serial-Correlation

In presence of more lagged variables, the Breusch-Godfrey test is typically used. It considers the null hypothesis of non-autocorrelated errors against an alternative that they are AR(P). The equation for AR(P) is as follows:

$$ u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \varepsilon_t $$  \hspace{1cm} (A.2.2.5)

The simple regression model is first fitted by OLS to obtain a residuals $\hat{u}_t$. Then an auxiliary regression model is fitted as follows:

$$ \hat{u}_t = \alpha_0 + \alpha_1 X_{t,1} + \alpha_2 X_{t,2} + \rho_1 \hat{u}_{t-1} + \rho_2 \hat{u}_{t-2} + \cdots + \rho_p \hat{u}_{t-p} + \varepsilon_t $$ \hspace{1cm} (A.2.2.6)

The test statistics is computed as follows and it should follow a chi-square distribution:

$$ nR^2 \sim \chi_p^2 $$  \hspace{1cm} (A.2.2.7)

, where $R^2$ is obtained from the auxiliary regression.

A.2.3 Estimators Addressing Heteroskedasticity and Autocorrelation

White Estimator

In case of heteroskedasticity, the covariance matrix of the OLS takes the form:

$$ COV[\hat{\beta}] = \sigma^2 (X^TX)^{-1}X^T\Omega X(X^TX)^{-1} $$  \hspace{1cm} (A.2.3.1)

, which can be used for inference once the estimator is available. Based on this consideration, White (1980) proposes the following unbiased estimator:
\[
\begin{bmatrix}
\hat{u}_1^2 & 0 & 0 & \cdots & 0 \\
0 & \hat{u}_2^2 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & 0 \\
0 & \cdots & \cdots & \cdots & \hat{u}_{T-1}^2 \\
0 & \cdots & \cdots & \cdots & 0 \\
\end{bmatrix}
\]  
(A.2.3.2)

The choice for \( \hat{\Omega} \) leads to the following covariance variance estimator:

\[
l_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + \epsilon_t \Sigma^W_{\beta} = \frac{T}{T-k} (X'X)^{-1} \left( \sum_{t=1}^{T} \hat{u}_t^2 X_t X'_t \right) (X'X)^{-1}
\]  
(A.2.3.3)

**Newey-West Estimator**

Newey and West (1987) proposed a more general covariance estimator that is robust to heteroskedasticity and autocorrelation of the residuals of unknown form. This heteroskedasticity-and-autocorrelation consistent (HAC) is given by:

\[
\Sigma^{NW}_{\beta} = (X'X)^{-1} T \hat{\Omega} (X'X)^{-1}
\]  
(A.2.3.4)

where \( \hat{\Omega} \) is the long-run covariance estimators.

\[
\hat{\Omega} = \Gamma(0) + \sum_{j=1}^{p} \left[ 1 - \frac{j}{p+1} \right] [\hat{\Gamma}(j) + \hat{\Gamma}(-j)]
\]  
(A.2.3.5)

\[
\hat{\Gamma}(j) = \left( \sum_{t=1}^{T} \hat{u}_t \hat{u}_{t-j} X_t X'_{t-j} \right) \frac{1}{T}
\]  
(A.2.3.6)

In absence of serial-corrilation, \( \hat{\Omega} = \hat{\Gamma}(0) \) and it goes back to the White estimator. Note that Newey-West estimator must be calculated conditional on a maximum lag. Normally a lag exceeding the periodicity of the data is sufficient. For example, for monthly data a lag of 12 is often chosen.
A.3 Cluster-robust Inference

In this section, we present the fundamentals of cluster-robust inference. For these basic results, we first assume that the following: i) the model does not include cluster-specific fixed effects, ii) it is clear how to form the clusters, and iii) there are many clusters.

Clustered errors have two main consequences: they reduce the precision of parameter estimation and the standard estimator for the variance of parameters is biased downward from the true variance.

\[ y_g = X_g \beta + u_g, \quad g = 1, \ldots, G \]  
(A.3.0.1)

A.3.1 One-way Clustering

The key assumption for the clustered-covariance is that errors are uncorrelated across clusters, while they can be correlated within clusters.

\[ \mathbb{E} [u_{ig} u_{jg}' | x_{ig}, x_{jg}] = 0 \text{ for } g \neq g' \]  
(A.3.1.1)

The OLS estimator is:

\[ \hat{\beta} = (X'X)^{-1} X'y = \left( \sum_{g=1}^{G} X'_g X_g \right)^{-1} \sum_{g=1}^{G} X'_g y_g \]  
(A.3.1.2)

The estimated covariance matrix is:

\[ V[\hat{\beta}] = (X'X)^{-1} B (X'X)^{-1} \]  
(A.3.1.3)

with

\[ B = X'V[u|X]X \]  
(A.3.1.4)

Given that the error is independent across clusters, \( V[u|X] \) has a block-diagonal structure, and thus the previous equation simplifies to:

\[ B_{\text{clu}} = \sum_{g=1}^{G} X'_g \mathbb{E} [u_g u'_g | X_g] X_g \]  
(A.3.1.5)
The cluster-robust estimate of covariance matrix can be written as:

$$
\hat{\nabla}_{\text{clu}}[\hat{\beta}] = (X'X)^{-1} \hat{B}_{\text{clu}} (X'X)^{-1}
$$

(A.3.1.6)

\[ \hat{B}_{\text{clu}} = \sum_{g=1}^{G} X'_g \hat{u}_g \hat{u}_g' X_g \]  

(A.3.1.7)

### A.3.2 Multi-way Clustering

For multi-way clustering, the key assumption becomes:

$$
\mathbb{E}[u_i u_j|x_i, x_j] = 0 \text{ unless } i \text{ and } j \text{ share any cluster dimension}
$$

(A.3.2.1)

Then the formula for covariance estimation becomes:

$$
\hat{B} = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i x_j' \hat{u}_i \hat{u}_j 1[i, j \text{ share any cluster}]
$$

(A.3.2.2)

It can be easily shown that the two-way cluster-robust covariance estimation can be rewritten as:

$$
\hat{V}_{2\text{way}}[\hat{\beta}] = \hat{V}_{1}[\hat{\beta}] + \hat{V}_{2}[\hat{\beta}] - \hat{V}_{1\cap2}[\hat{\beta}]
$$

(A.3.2.3)
Appendix B

Hypothesis Testing with Sharpe Ratio Test

Assume that we have two portfolios: portfolio $i$ with a series of excess returns $r_{i1},\ldots,r_{it}$ and portfolio $n$ with excess returns $r_{n1},\ldots,r_{nt}$. Their mean vector $\mu$ and covariance matrix $\Sigma$ is defined as follows:

$$
\mu = \begin{bmatrix} \mu_i \\ \mu_n \end{bmatrix},
\Sigma = \begin{bmatrix} \sigma_i^2 & \sigma_{in} \\ \sigma_{ni} & \sigma_n^2 \end{bmatrix}
$$

The difference between two Sharpe ratios is given by:

$$
\delta = s_i - s_n = \frac{\mu_i}{\sigma_i} - \frac{\mu_n}{\sigma_n} = \frac{\mu_i}{\sqrt{\gamma_i - \mu_i^2}} - \frac{\mu_n}{\sqrt{\gamma_n - \mu_n^2}}
$$

(B.0.0.1)

where $\gamma_i = E[r_i^2]$ and $v = [\mu_i, \gamma_i, \mu_n, \gamma_n]$.

It is assumed that $\sqrt{T}(\hat{v} - v) \rightarrow N(0, \Psi)$, where $\Psi$ is an unknown symmetric positive semi-definite matrix. Then this implies that:

$$
\sqrt{T}(\hat{\Delta} - \Delta) \overset{d}{\rightarrow} N(0; \nabla'f(v)\Psi \nabla f(v))
$$

(B.0.0.2)
Now given a consistent estimator \( \hat{\Psi} \), the standard error of the estimated difference can be computed as:

\[
s(\hat{\Delta}) = \sqrt{\frac{\nabla' f(\hat{v}) \hat{\Psi} \nabla f(\hat{v})}{T}}
\]

(G.0.0.3)

Given the standard error, the p-value can be estimated as:

\[
\hat{p} = 2\Phi\left(-\frac{|\hat{\Delta}|}{s(\hat{\Delta})}\right)
\]

(G.0.0.4)

Bootstrap leads to higher inference accuracy than standard inference procedures under the assumption of asymptotic normal distribution. Therefore, Ledoit and Wolf (2008) propose to test the null hypothesis by constructing a time-series bootstrap confidence interval with nominal level \( \alpha \). If zero is not included in the confidence interval, the null hypothesis of equal Sharpe ratios is rejected. The distribution function of the studentized difference is approximated by a distribution from the bootstrap samples.

\[
L\left(\frac{|\hat{\Delta} - \Delta|}{s(\Delta)}\right) \approx L\left(\frac{|\hat{\Delta}^* - \hat{\Delta}|}{s(\hat{\Delta}^*)}\right)
\]

(G.0.0.5)

The bootstrap data is generated using the circular block bootstrap method of Politis and Romano (1992), resampling blocks of pairs of excess returns from the observed values with replacement. The estimator \( \hat{\Psi} \) is obtained from HAC kernel inference, specifically, a prewhitened quadratic spectral kernel estimation of Andrews and Monahan (1992). The standard error and p-value are calculated using the same formulas above.

To conduct this bootstrap, we can either follow Ledoit and Wolf (2008) and use the calibration procedure for selecting the required fixed block size, or follow Behr, Guettler, and Truebenbach (2012) that selects the size manually to preserve the autocorrelation structure in the data. We follow the latter one and set the block size to 5. For the number of bookstraps, we set the number to 5000, as this is sufficient for reliable inference suggested by Ledoit2008. As to the standard error, we calculated it based on HAC kernel estimation.
Appendix C

Code

C.1 Event Study

```python
import pandas as pd
import numpy as np
from datetime import datetime
import matplotlib.pyplot as plt
import warnings
import datetime as dt
from scipy import stats
import math
import operator

''' read data '''
data_ret = pd.read_pickle('data_ret_new.pkl')
df_negative = pd.read_pickle('df_negative_new_1.pkl')
df_negative = df_negative[df_negative.date > dt.datetime(2012,1,3)]
data_ret = data_ret[data_ret.date > dt.datetime(2012,1,3)]
data_ret.index= data_ret['date']
data_ret = data_ret.drop(columns = ['date'])
data_ret= data_ret.iloc[:,(data_ret.values < 1).all(axis=0) & (data_ret.values > -1).all(axis=0)]

stocks = data_ret.columns.unique().tolist()
stocks.remove('^GSPC')

''' parameter initialization '''
window = 10
estimation_period = 200
windows_indexes = range(-window, window + 1)
estimation_indexes = range(-estimation_period-window, -window)
dr_equities_window = pd.DataFrame(index=windows_indexes)
dr_equities_estimation = pd.DataFrame(index=estimation_indexes)
dr_market_window = pd.DataFrame(index=windows_indexes)
dr_market_estimation = pd.DataFrame(index=estimation_indexes)
```
for each stock, compute daily stock and market return during the window period

get the date on which negative event happens

def get_event_dates(df_negative, ticker):
    neg_event_dates = df_negative[df_negative.ticker == ticker].date.tolist()
    neg_event_dates

for neg_event in neg_event_dates:
    start_idx = evt_idx - window - estimation_period
    end_idx = evt_idx - window
    if start_idx < 0 or evt_idx + window >= data_ret.shape[0]:
        continue
    new_dr_equities_estimation = data_ret[symbol][start_idx:end_idx]
    new_dr_equities_estimation.index = estimation_indexes
    dr_equities_estimation[col_name] = new_dr_equities_estimation

new_dr_market_estimation = data_ret['^GSPC'][start_idx:end_idx]
new_dr_market_estimation.index = estimation_indexes
dr_market_estimation[col_name] = new_dr_market_estimation

start_idx = evt_idx - window
end_idx = evt_idx + window + 1
new_dr_equities_window = data_ret[symbol][start_idx:end_idx]
new_dr_equities_window.index = windows_indexes
dr_equities_window[col_name] = new_dr_equities_window

new_dr_market_window = data_ret['^GSPC'][start_idx:end_idx]
new_dr_market_window.index = windows_indexes
dr_market_window[col_name] = new_dr_market_window

initialize regression_estimation and expected_return datasets

reg_estimation = pd.DataFrame(index=dr_market_estimation.columns,
                             columns=['intercept', 'beta', 'rse'])
er = pd.DataFrame(index=dr_market_window.index, columns=dr_market_window.columns)
compute the expected return of each date using the regression

```python
for col in dr_market_estimation.columns:
    x = dr_market_estimation[col]
    y = dr_equities_estimation[col]
    slope, intercept, _, _, _ = stats.linregress(x, y)
    reg_estimation['beta'][col] = slope
    reg_estimation['intercept'][col] = intercept
    reg_estimation['rse'][col] = sum((y - slope * x - intercept)**2)/(len(x)-2)
    er[col] = intercept + dr_market_window[col] * slope
```

```python
er.columns.name = 'Expected return'
```

calculate abnormal return

```python
ar = dr_equities_window - er
ar.columns.name = 'Abnormal return'
```

remove AR with NA or all values as 0

```python
ar = ar.dropna(axis = 1, how = 'any')
ar = ar.loc[:, (ar != 0).any(axis=0)]
```

calculate cumulative abnormal return

```python
car = ar.apply(np.cumsum)
car.columns.name = 'Cum Abnormal Return'
```

calculate mean for CAR and AR

```python
mean_car = car.mean(axis=1)
mean_car.name = 'CAR'
mean_ar = ar.mean(axis=1)
mean_ar.name = 'AR'
```

compute mean and variance for CAAR

```python
cols = car.columns.tolist()
var_car = (window*2+1) * rse
mean_val = mean_car[window]
print('mean = %.5f%%(mean_val))

N = len(cols)
var_caar = sum(var_car)/(N*N)
t_car = mean_val/math.sqrt(var_caar)
pval = stats.t.sf(np.abs(t_car), N-1)*2
print('t-statistic = %6.3f, pvalue = %6.4f%%(t_car, pval))
```

plot CAR

```python
x = mean_car.index.values
y = mean_car.values
z = mean_ar.values
fig, ax = plt.subplots(figsize=(8,5))
ax.plot(x, y, label='label')
ax.set_xlabel('window')
```
ax.set_ylabel('CAR')
plt.title('cumulative abnormal return(CAR) within the event window')
plt.legend()

''' compute CAR by label '''
def car_by_label(label):
    ''' select ticker and date for negative event in this label category '''
    negative_by_label = df_negative[df_negative.label2 == label][['ticker', 'date']]  # negative_by_label['date_str'] = negative_by_label['date'].apply(lambda x: x.strftime('%Y-%m-%d'))
    negative_by_label['col_names'] = negative_by_label['ticker'] + ' ' + negative_by_label['date_str']
    ''' select these events from CAR '''
    col_names = negative_by_label['col_names'].values.tolist()
    col_names = [col_name for col_name in col_names if col_name in car.columns.tolist()]
    car_sub = car[col_names]
    ''' calculate mean CAR for these events '''
    mean_car = car_sub.mean(axis=1)
    ''' calculate variance CAR for all events in this category '''
    cols = car_sub.columns.tolist()
    N = len(cols)
    rse = reg_estimation.loc[reg_estimation['rse'].index.isin(cols), 'rse']
    var_car_i = (window*2+1) * rse
    std_car = math.sqrt(sum(var_car_i)/N**2)
    t_car = mean_car[window]/std_car
    pval = stats.t.sf(np.abs(t_car), N-1)*2
    return mean_car, std_car, pval

mean_dic = {}
pval_dic = {}
std_dic = {}
mean_cum_val = {}
label2 = df_negative['label2'].unique().tolist()
for label in label2:
    print(label)
    mean_car, std_car, pval = car_by_label(label)
    print(mean_car)
    mean_dic[label] = mean_car
    mean_cum_val[label] = mean_car[window]
    pval_dic[label] = round(pval,5)
    std_dic[label] = std_car

''' plot CAR for each event subcategory during the event window '''
fig, axs = plt.subplots(6,3, figsize=(10, 8), constrained_layout=True)
axs = axs.ravel()
for i, var in enumerate(mean_dic):
C.2 Jensen's Alpha

```python
x = mean_dic[var].index.values
y = mean_dic[var].values
yerr = std_dic[var]
ymin = y - 1.65*yerr
ymax = y + 1.65*yerr
axs[i].plot(x, y)
axs[i].set_title(var)
plt.show()

''' plot CAR and confidence interval at the end of event window '''
yerror=[std_dic[i]*1.65 for i in std_dic]
mean_window_end = {k: mean_dic[k][window] for k in mean_dic}
plt.figure(figsize=(9,5))
plt.xticks(rotation=90)
x = range(len(mean_window_end))
y = list(mean_window_end.values())
plt.axhline(color='black')
plt.errorbar(x, y, yerr=yerror, fmt='o', lolims=True, uplims=True)
plt.xticks(x, list(mean_window_end.keys()))
plt.title('CARs at the end of window period (day 20) by event category')
plt.xlabel('event category')
plt.ylabel('CAR')
plt.show()
```

C.2 Jensen's Alpha

```python
import pandas as pd
import numpy as np
from datetime import datetime
import matplotlib.pyplot as plt
import matplotlib.dates as mdates
import warnings
import datetime as dt
from scipy import stats
import math
import operator
from statsmodels.graphics.tsaplots import plot_acf

''' read and preprocess data '''
factors = pd.read_pickle('factors_monthly.pkl')
factors = factors[(factors.index >= '2013-1-1') & (factors.index < '2019-2-1')]

factors_5 = pd.read_csv('fama_french_5_factor.csv')
factors_5 = factors_5.iloc[12:-5]
factors_5 = factors_5[['Mkt-RF', 'SMB', 'HML', 'RMW', 'CMA', 'RF']]
factors_5['const'] = 1

data_ret = pd.read_csv('monthly_return.csv', header=1)
data_ret = data_ret.iloc[:, :]
data_ret.index = data_ret.ticker
data_ret = data_ret.iloc[:, 1:]```
data_ret /= 100
data_ret['date'] = data_ret.index
data_ret.index = data_ret.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
data_ret = data_ret.drop('date', axis=1)
data_ret = data_ret.dropna(axis=1)
data_ret_1 = data_ret.dropna(axis=1)
data_ret = data_ret.iloc[:, (data_ret.values < np.percentile(data_ret_1, 99)).all(axis=0) & (data_ret.values > np.percentile(data_ret_1, 1)).all(axis=0)]

df_negative = pd.read_pickle('df_negative_new_1.pkl')
stock_tickers = data_ret.columns.tolist()

''' equal-weighted portfolio '''
event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
monthly_rets = []
num_firms = []

while event_starting_date < pd.Timestamp(year=2018, month=2, day=1):
    event_ending_date = relativedelta(months=12) + event_starting_date
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)

    ''' find stocks with negative news within the last 12 months '''
    df = df_negative[(df_negative.date < event_ending_date) & (df_negative.date >= event_starting_date)]
tickers = df.ticker.unique().tolist()
new_tickers = list(set(tickers) & set(stock_tickers))
num_firms.append(len(new_tickers))

    ''' compute monthly return for every stock '''
    rets = data_ret[new_tickers][portfolio_starting_date:portfolio_ending_date]
    monthly_ret = np.mean(rets, axis=1).values[0]
    monthly_rets.append(monthly_ret)
    event_starting_date += relativedelta(months=1)

    ''' Fama-French five-factor model '''
factors_5['R'] = monthly_rets
factors_5['R-RF'] = factors_5['R'] - factors_5['RF']
mod = sm.OLS(factors_5['R-RF'], factors_5[['SMB', 'HML', 'Mkt-RF', 'RMW', 'CMA', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

    ''' Carhart four-factor model '''
factors['R'] = monthly_rets
factors['R-RF'] = factors['R'] - factors['RF']
factors['const'] = 1
factors.columns = ['Mkt-RF', 'SMB', 'HML', 'RF', 'MKT', 'Mom', 'R', 'R-RF', 'const']
C.2 Jensen’s Alpha

```python
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'Mom', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

'''Fama-French three-factor model'''
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'Mkt-RF', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

'''residual and ACF plot'''
resids = res.resid.values.tolist()
y_fitted = res.fittedvalues
plot_acf(resids, lags=20)
fig, ax = plt.subplots(figsize=(6, 3))
_ = ax.scatter(y_fitted, resids)
plt.xlabel('fitted values')
plt.ylabel('residual')
plt.title('residual vs. fitted value')

'''run the regression by event category'''
labels = df_negative.label2.unique().tolist()
alphas = []
pvals = []
num_firms_list = []

for label in labels:
    df_sub = df_negative[df_negative.label2 == label]
    print(f'=== currently running regression for label: {label}')
    event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
    monthly_rets = []
    num_firms = []
    while event_starting_date < pd.Timestamp(year=2018, month=2, day=1):
        event_ending_date = relativedelta(months=12) + event_starting_date
        portfolio_starting_date = event_ending_date
        portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
        df = df_sub[(df_sub.date < event_ending_date) & (df_sub.date >= event_starting_date)]
tickers = df.ticker.unique().tolist()
new_tickers = list(set(tickers) & set(stock_tickers))
num_firms.append(len(new_tickers))
rets = data_ret[new_tickers][portfolio_starting_date: portfolio_ending_date]
monthly_ret = np.mean(rets, axis=1).values[0]
monthly_rets.append(monthly_ret)
```
event_starting_date += relativedelta(months=1)

factors['R'] = monthly_rets
factors['R-RF'] = factors['R'] - factors['RF']
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'Mom', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})

alphas.append(res.params[-1])
pvals.append(res.pvalues[-1])
num_firms_list.append(np.mean(num_firms))

d = {'label': labels, 'alpha': alphas, 'p-value': pvals, 'num_firms': num_firms_list}
res = pd.DataFrame(d)
res.index = res.label
res = res.drop(columns='label')
res['num_firms'] = res['num_firms'].astype(int)
res = res.sort_values('num_firms', ascending=False)
res = res[['alpha', 'p-value', 'num_firms']]
res

'value-weighted portfolios'
market_cap = pd.read_csv('market_cap.csv', header=1)
market_cap = market_cap.iloc[1:, :]
market_cap.index = market_cap.ticker.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
market_cap = market_cap.drop('ticker', axis=1)
market_cap_1 = market_cap.dropna(axis=1)
mc_max_cut = np.percentile(market_cap_1, 99)
mc_min_cut = np.percentile(market_cap_1, 1)
market_cap_1 = market_cap_1.iloc[:, (market_cap_1.values < mc_max_cut).all(axis=0) & (market_cap_1.values > mc_min_cut).all(axis=0)]
market_cap = market_cap_1

event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
monthly_rets = []
num_firms = []
while event_starting_date < pd.Timestamp(year=2018, month=2, day=1):
    event_ending_date = relativedelta(months=12) + event_starting_date
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)

    df = df_negative[(df_negative.date < event_ending_date)&(df_negative.date >= event_starting_date)]
tickers = df.ticker.unique().tolist()
new_tickers = list(set(tickers) & set(stock_tickers) & set(market_cap.columns))
num_firms.append(len(new_tickers))
rets = data_ret[new_tickers][portfolio_starting_date:portfolio_ending_date]
m = market_cap[new_tickers][portfolio_starting_date:portfolio_ending_date]
rets = rets[sorted(rets.columns)]
mc = rets[sorted(mc.columns)]
if (sorted(mc.columns) != sorted(rets.columns)) or (len(mc.columns) != len(rets.columns)):
    print('False!')
value_weighted_rets = mc*rets/np.sum(mc, axis=1).values[0]
monthly_ret = np.mean(value_weighted_rets, axis=1).values[0]
monthly_rets.append(monthly_ret)
event_starting_date += relativedelta(months=1)

'''value-weighted four-factor model'''
factors['R'] = monthly_rets
factors['R-RF'] = factors['R'] -factors['RF']
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'Mom', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

'''winsorize stock returns'''
q95 = np.percentile(data_ret.values,95)
q5 = np.percentile(data_ret.values,5)
data_ret_1 = data_ret.iloc[:,(data_ret.values < q95).all(axis=0) & (data_ret.values > q5).all(axis=0)]
stock_tickers = data_ret_1.columns.tolist()
event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
monthly_rets = []
num_firms = []
while event_starting_date < pd.Timestamp(year=2018, month=2, day=1):
    event_ending_date = relativedelta(months=12) + event_starting_date
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
    df = df_negative[(df_negative.date < event_ending_date)&(df_negative.date >= event_starting_date)]
tickers = df.ticker.unique().tolist()
new_tickers = list(set(tickers) & set(stock_tickers))
num_firms.append(len(new_tickers))
rets = data_ret_1[new_tickers][portfolio_starting_date:portfolio_ending_date]
monthly_ret = np.mean(rets, axis=1).values[0]
monthly_rets.append(monthly_ret)
event_starting_date += relativedelta(months=1)

factors['R'] = monthly_rets
factors['R-RF'] = factors['R'] -factors['RF']
factors['const'] = 1
factors.columns = ['Mkt-RF', 'SMB', 'HML', 'RF', 'MKT', 'Mom', 'R', 'R-RF', 'const']
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'Mom'],
```python
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

"""industry-matched portfolio"""

industry = pd.read_csv('industry.csv')
industry = industry[['ticker', 'industry', 'industry_sub']]  # industry columns
industry.columns = ['ticker', 'industry1', 'industry2']
industry.loc[(industry.industry1 == 'Food, Beverage & Tobacco') |
             (industry.industry1 == 'Household & Personal Products') |
             (industry.industry2 == 'Textiles, Apparel & Luxury Goods')]
    , 'industry'] = 'NoDur'

industry.loc[(industry.industry2 == 'Household Durables') |
              (industry.industry2 == 'Leisure Products')]
    , 'industry'] = 'Durbl'

industry.loc[(industry.industry1 == 'Automobiles & Components') |
              (industry.industry1 == 'Capital Goods') |
              (industry.industry2 == 'Paper & Forest Products')]
    , 'industry'] = 'Manuf'

industry.loc[(industry.industry1 == 'Software & Services') |
              (industry.industry1 == 'Technology Hardware & Equipment') |
              (industry.industry1 == 'Semiconductors & Semiconductor Equipment')]
    , 'industry'] = 'BusEq'

industry.loc[industry.industry1 == 'Energy', 'industry'] = 'Enrgy'
industry.loc[industry.industry1 == 'Telecommunication Services', 'industry'] = 'Telcm'
industry.loc[industry.industry1 == 'Utilities', 'industry'] = 'Utils'
industry.loc[industry.industry2 == 'Chemicals', 'industry'] = 'Chems'

industry.loc[(industry.industry1 == 'Retailing') |
             (industry.industry1 == 'Food & Staples Retailing')]
    , 'industry'] = 'Shops'

industry.loc[(industry.industry1 == 'Health Care Equipment & Services')]
    , 'industry'] = 'Hlth'

industry.loc[(industry.industry1 == 'Diversified Financials')]
    , 'industry'] = 'Money'
```
industry.loc[(industry.industry1 == 'Media & Entertainment') | 
(industry.industry1 == 'Consumer Services') | 
(industry.industry1 == 'Commercial & Professional Services') | 
(industry.industry1 == 'Transportation') | 
(industry.industry2 == 'Metals & Mining') | 
(industry.industry2 == 'Construction Materials') | 
(industry.industry2 == 'Containers & Packaging') 
, 'industry'] = 'Other'

industry = industry.dropna(subset = ['industry'])

''substract the industry return from the company monthly return '',

industry = industry.drop(columns = ['Unnamed: 0'])

industry_portfolios = pd.read_csv('12_portfolio_equal_weighted.csv')

industry_portfolios['Unnamed: 0'] = industry_portfolios['Unnamed: 0'].apply(lambda x: datetime.strptime(str(x),'%Y%m'))

industry_portfolios.columns = ['date', 'NoDur', 'Durbl', 'Manuf', 'Enrgy', 'Chems', 'BusEq', 'Telcm', 'Utils', 'Shops', 'Hlth', 'Money', 'Other']

industry_portfolios.index = industry_portfolios.date

industry_portfolios = industry_portfolios.iloc[:,1:]

industry_portfolios /=100

industry_portfolios['date'] = industry_portfolios.index

industry_portfolios = industry_portfolios[(industry_portfolios.date >= datetime.strptime('2011-11-01','%Y-%m-%d')) & (industry_portfolios.date <= datetime.strptime('2019-01-01','%Y-%m-%d'))]


for t in data_ret_1.columns:
  indstr = industry[industry.ticker == t].industry.values.tolist()
  if len(indstr) == 0:
    data_ret_1.drop(columns = t,inplace=True)
  else:
    indstr = indstr[0]
    data_ret_1[t] -= industry_portfolios[indstr]

'' construct the industry-matched portfolio''

stock_tickers = data_ret_1.columns.tolist()

event_starting_date = pd.Timestamp(year=2012, month=1, day=1)

monthly_rets = []
```python
num_firms = []
while event_starting_date < pd.Timestamp(year=2018, month=2, day=1):
    event_ending_date = relativedelta(months=12) + event_starting_date
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
    df = df_negative[(df_negative.date < event_ending_date) & (df_negative.date >= event_starting_date)]
    tickers = df.ticker.unique().tolist()
    new_tickers = list(set(tickers) & set(stock_tickers))
    num_firms.append(len(new_tickers))
    rets = data_ret_1[new_tickers][portfolio_starting_date:portfolio_ending_date]
    monthly_ret = np.mean(rets, axis=1).values[0]
    monthly_rets.append(monthly_ret)
    event_starting_date += relativedelta(months=1)

factors['R'] = monthly_rets
factors['R-RF'] = factors['R'] - factors['RF']
factors['const'] = 1
factors.columns = ['Mkt-RF', 'SMB', 'HML', 'RF', 'MKT', 'Mom', 'R', 'R-RF', 'const']
mod = sm.OLS(factors['R'], factors[['SMB', 'HML', 'MKT', 'Mom', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

'''characteristics matched portfolio'''
mc_all_stocks = pd.read_csv('market_cap_all_stocks_1.csv')
mc_all_stocks.date = mc_all_stocks.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
mc_all_stocks.index = mc_all_stocks.date
mc_all_stocks = mc_all_stocks.drop(columns=['date'])

bv_all_stocks = pd.read_csv('bv_all_stocks_1.csv')
bv_all_stocks.date = bv_all_stocks.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
bv_all_stocks.index = bv_all_stocks.date
bv_all_stocks = bv_all_stocks.drop(columns=['date'])

yearly_ret_all_stocks = pd.read_csv('yearly_return_all_stocks_1.csv')
yearly_ret_all_stocks.date = yearly_ret_all_stocks.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
yearly_ret_all_stocks.index = yearly_ret_all_stocks.date
yearly_ret_all_stocks = yearly_ret_all_stocks.drop(columns=['date'])

mc_all_stocks = mc_all_stocks.dropna(axis=1)
bv_all_stocks = bv_all_stocks.dropna(axis=1)
yearly_ret_all_stocks = yearly_ret_all_stocks.dropna(axis=1)
mc_esg_stocks = pd.read_csv('yearly_market_cap_esg_stocks.csv',
```
header=1)
mc_esg_stocks = mc_esg_stocks.iloc[1:, :]
mc_esg_stocks.index = mc_esg_stocks.ticker
mc_esg_stocks = mc_esg_stocks.iloc[:, 1:]

bv_esg_stocks = pd.read_csv('bv_esg_stocks.csv', header=1)
bv_esg_stocks = bv_esg_stocks.iloc[1:, :]
bv_esg_stocks.index = bv_esg_stocks.ticker
bv_esg_stocks = bv_esg_stocks.iloc[:, 1:]

yearly_ret_esg_stocks = pd.read_csv('yearly_return_esg_stocks.csv', header=1)
yearly_ret_esg_stocks = yearly_ret_esg_stocks.iloc[1:, :]
yearly_ret_esg_stocks.index = yearly_ret_esg_stocks.ticker
yearly_ret_esg_stocks = yearly_ret_esg_stocks.iloc[:, 1:]

mc_esg_stocks.columns = [t.split('.', 1)[0].lower() for t in mc_esg_stocks.columns.tolist()]
bv_esg_stocks.columns = [t.split('.', 1)[0].lower() for t in bv_esg_stocks.columns.tolist()]
yearly_ret_esg_stocks.columns = [t.split('.', 1)[0].lower() for t in yearly_ret_esg_stocks.columns.tolist()]
mc_all_stocks.columns = [t.split('.', 1)[0].lower() for t in mc_all_stocks.columns.tolist()]
bv_all_stocks.columns = [t.split('.', 1)[0].lower() for t in bv_all_stocks.columns.tolist()]
yearly_ret_all_stocks.columns = [t.split('.', 1)[0].lower() for t in yearly_ret_all_stocks.columns.tolist()]
monthly_ret_all_stocks = pd.read_csv('monthly_ret_all_stocks_1.csv')
monthly_ret_all_stocks.date = monthly_ret_all_stocks.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
monthly_ret_all_stocks.index = monthly_ret_all_stocks['date']
monthly_ret_all_stocks = monthly_ret_all_stocks.iloc[:, 1:]
monthly_ret_all_stocks /= 100
monthly_ret_all_stocks.columns = [t.split('.', 1)[0].lower() for t in monthly_ret_all_stocks.columns.tolist()]
d = [mc_all_stocks.columns, bv_all_stocks.columns,
yearly_ret_all_stocks.columns, monthly_ret_all_stocks.columns]
tickers = list(set.intersection(*map(set,d)))
tickers = [t for t in tickers if t not in df_negative.ticker.
unique().tolist()]
mc_all_stocks = mc_all_stocks[tickers]
bv_all_stocks = bv_all_stocks[tickers]
yearly_ret_all_stocks = yearly_ret_all_stocks[tickers]
monthly_ret_all_stocks = monthly_ret_all_stocks[tickers]
data_ret_1 = data_ret
data_ret_1 = data_ret_1[(data_ret_1.index >= '2012-01-01') & (data_ret_1.index <= '2018-12-31')]
i, m = 0, 0
event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
```python
monthly_rets = []
while event_starting_date < pd.Timestamp(year=2018, month=1, day=1):
    event_ending_date = relativedelta(months=12) + event_starting_date
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)

    ''' divide the stocks into 125 portfolios, based on market
    capitalization, price to market ratio, and past year
    return '''
    
    if m % 12 == 0:
        q1, cutoff1 = pd.qcut(mc_all_stocks.iloc[i], 5, labels=False, retbins=True)
        q1 = pd.DataFrame(q1)
        q1.columns = ['group']
        cutoff1_1 = cutoff1.tolist() 

        ''' first level '''
        q1_1 = q1.groupby('group').groups
        q1_1_values = list(q1_1.values())

        q1_2 = []
        cutoff1_2 = []

        for t in q1_1_values:
            bv_stocks = bv_all_stocks[t]
            ''' second level '''
            q1_2, cutoff_2 = pd.qcut(bv_stocks.iloc[i], 5, labels=False, retbins=True)
            cutoff1_2.append(cutoff_2.tolist())
            q1_2 = pd.DataFrame(q1_2)
            q1_2.columns = ['group']
            q1_2 = q1_2.groupby('group').groups
            q1_2_values.append(list(q1_2.values()))

        q1_2_values = sum(q1_2_values, [])
        cutoff1_3 = []
        q1_3_values = []

        for t in q1_2_values:
            yr_stocks = yearly_ret_all_stocks[t]
            ''' third level '''
            q1_3, cutoff_3 = pd.qcut(yr_stocks.iloc[i], 5, labels=False, retbins=True)
            cutoff1_3.append(cutoff_3.tolist())
            q1_3 = pd.DataFrame(q1_3)
            q1_3.columns = ['group']
```

```python
c.2 jensen's alpha

q1_3 = q1_3.groupby('group').groups
q1_3_values.append(list(q1_3.values()))

q1_3_values = sum(q1_3_values, [])
q1_3_values = [q1_3_values[i].tolist() for i in range(125)]

'''compute the equal-weighted monthly return for each of the 125 portfolios'''
monthly_returns_all_groups = []
for t in q1_3_values:
    df = monthly_ret_all_stocks[t]
    df = df[(df.index < portfolio_starting_date + pd.offsets.DateOffset(years=1))
            & (df.index > portfolio_starting_date)]
    ret = np.mean(df, axis=1).values.tolist()
    # print(len(ret))
    monthly_returns_all_groups.append(ret)

i += 1

'''find the stocks with negative ESG news within the previous 12 months'''
df = df_negative[(df_negative.date < event_ending_date) &
                 (df_negative.date >= event_starting_date)]
tickers = df.ticker.unique().tolist()
new_tickers = list(set(tickers) & set(mc_esg_stocks.columns) &
                  set(bv_esg_stocks.columns) & set(yearly_ret_esg_stocks.columns) & set(data_ret_1.columns))

'''assign them to each of the 125 portfolios based on the precomputed cutoff'''
mc_esg = pd.DataFrame(mc_esg_stocks.iloc[0][new_tickers])
mc_esg.columns = ['mc']
bv_esg = pd.DataFrame(bv_esg_stocks.iloc[0][new_tickers])
bv_esg.columns = ['bv']

yr_esg = pd.DataFrame(yearly_ret_esg_stocks.iloc[0][
                      new_tickers])

yr_esg.columns = ['yr']
c_esg = mc_esg.join(bv_esg).join(yr_esg)
c_esg['mc_group'] = pd.cut(c_esg['mc'], bins=cutoff1_1, labels=False, include_lowest=True)
for i, c in enumerate(cutoff1_2):
    c_esg.loc[c_esg.mc_group == i,'bv_group'] = \
    pd.cut(c_esg[c_esg.mc_group == i]['bv'], bins=c, labels=False, include_lowest=True)
```

for i in range(5):
    for j in range(5):
        c_esg.loc[(c_esg.mc_group == i) & (c_esg.bv_group == j), 'yr_group'] = 
        pd.cut(c_esg[(c_esg.mc_group == i) & (c_esg.bv_group == j)]['yr'], bins = cutoff1_3[i*5+j],
        labels=False, include_lowest=True)

c_esg['group'] = c_esg['mc_group'] * 25 + c_esg['bv_group'] * 5 + c_esg['yr_group']

''' compute the adjusted return for ESG-related stocks '''
''' compute the equal-weighted monthly return '''
rets = data_ret_1[new_tickers][portfolio_starting_date:
portfolio_ending_date]
for t in rets.columns:
    group = c_esg.loc[c_esg.index == t, 'group'].values[0]
    if np.isnan(group):
        rets.drop(columns=t, inplace=True)
    else:
        group_monthly_ret = monthly_returns_all_groups[int(group)]
        rets[t] -= group_monthly_ret[(m % 12)]

monthly_ret = np.mean(rets, axis=1).values[0]
monthly_rets.append(monthly_ret)

event_starting_date += relativedelta(months=1)
m += 1
factors['R'] = monthly_rets
factors['const'] = 1
factors.columns = ['Mkt-RF', 'SMB', 'HML', 'RF', 'MKT', 'Mom', 'R','const']
mod = sm.OLS(factors['R'], factors[[ 'SMB', 'HML', 'MKT', 'Mom', 'R','const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

C.3 Panel Regression

import pandas as pd
import numpy as np
from datetime import datetime, timedelta
from dateutil.relativedelta import relativedelta
import matplotlib.pyplot as plt
import warnings
import datetime as dt
from scipy import stats
import pandas_datareader.data as web
import statsmodels.api as sm
import statsmodels.formula.api as smf
from datetime import datetime
from linearmodels import PooledOLS

warnings.filterwarnings("ignore")
pd.set_option('display.max_columns', None)

''' preprocess '''

''' dividend yield '''
div = pd.read_csv('div_yield.csv', header = 1)
div = div.iloc[1:, :]
div.index = div.ticker
div = div.iloc[:, 1:]

''' log of market capitalization, in terms of billion of dollar '''
mc = pd.read_csv('market_cap.csv', header = 1)
mc = mc.iloc[1:, :]
mc.index = mc.ticker
mc = mc.iloc[:, 1:]
mc /= 1000000000
mc = np.log(mc)

''' log of book-to-market ratio '''
bm = pd.read_csv('book_to_market_ratio.csv', header=1)
bm = bm.iloc[1:, :]
bm.index = bm.ticker
bm = bm.iloc[:, 1:]
bm = np.log(bm)

''' log of reciprocal of stock price '''
price = pd.read_csv('monthly_price.csv', header=1)
price = price.iloc[1:, :]
price.index = price.ticker
price = price.iloc[:, 1:]
price = np.log(price)
prc = 1/price
prc = np.log(prc)

''' log of cumulative returns '''
ret23 = np.log((price.shift(2)/price.shift(3)))
ret46 = np.log((price.shift(4)/price.shift(6)))
ret712 = np.log((price.shift(7)/price.shift(12)))

div = div.shift(1)
mc = mc.shift(1)
bm = bm.shift(1)
prc = prc.shift(1)

df_negative = pd.read_pickle('df_negative_new_1.pkl')
df_negative = df_negative[['ticker', 'date', 'label1', 'label2']]
df_negative.date = df_negative.date + pd.offsets.MonthEnd(0)
df_negative.index = df_negative.ticker
df_negative = df_negative[['date', 'label1', 'label2']]
df_negative_1 = df_negative
'' create a dummy, which is the count of negative events for the last month''

```
def_negative = pd.get_dummies(df_negative['date']).groupby(
    df_negative.index).sum()
df_negative.columns = [i + relativedelta(months=1) + pd.offsets.
    MonthEnd(0) for i in df_negative.columns]
df_negative['tickers'] = df_negative.index
df_negative = pd.melt(df_negative,
    id_vars='tickers',
    value_vars=list(df_negative.columns)[:-1],
    var_name='date',
    value_name='dummy')
df_negative.date = df_negative.date.apply(lambda x: x.strftime('%Y-%m-%d'))
```

'' create a set of dummies, which count the negative events for the last month per event subcategory''

```
labels = df_negative_1.label2.unique().tolist()
l = labels[0]
df_negative = df_negative_1[df_negative_1.label2 == l]
df_negative = pd.get_dummies(df_negative['date']).groupby(
    df_negative.index).sum()
df_negative.columns = [i + relativedelta(months=1) + pd.offsets.
    MonthEnd(0) for i in df_negative.columns]
df_negative['tickers'] = df_negative.index
df_negative = pd.melt(df_negative,
    id_vars='tickers',
    value_vars=list(df_negative.columns)[:-1],
    var_name='date',
    value_name='dummy')
df_negative.date = df_negative.date.apply(lambda x: x.strftime('%Y-%m-%d'))
df_negative.columns = ['tickers', 'date', l]
```

```
for l in labels[1:]:
    df_new = df_negative_1[df_negative_1.label2 == l]
    df_new = df_new[['date']].groupby(df_new.index).sum()
    df_new.columns = [i + relativedelta(months=1) + pd.offsets.
        MonthEnd(0) for i in df_new.columns]
    df_new['tickers'] = df_new.index
df_new = pd.melt(df_new,
    id_vars='tickers',
    value_vars=list(df_new.columns)[:-1],
    var_name='date',
    value_name='dummy')
    df_new.date = df_new.date.apply(lambda x: x.strftime('%Y-%m-%d'))
    df_new.columns = df_new.columns[:-1].tolist() + [l]
    df_negative = df_new.merge(df_negative, on=['tickers', 'date'], how='outer')
```

df_negative = df_negative.fillna(0)
mc = mc.iloc[2:, :]
mc = mc.iloc[:, (mc.values < np.percentile(mc, 99)).all(axis=0) & (mc.values > np.percentile(mc, 1)).all(axis=0)]

prc = prc.iloc[2:, :]
prc = prc.dropna(axis=1)
prc = prc.iloc[:, (prc.values < np.percentile(prc, 99)).all(axis=0) & (prc.values > np.percentile(prc, 1)).all(axis=0)]

div = div.iloc[2:, :]
div = div.dropna(axis=1)
div = div.iloc[:, (div.values < np.percentile(div, 99)).all(axis=0) & (div.values > np.percentile(div, 1)).all(axis=0)]

bm = bm.iloc[2:, :]
bm = bm.dropna(axis=1)
bm = bm.iloc[:, (bm.values < np.percentile(bm, 99)).all(axis=0) & (bm.values > np.percentile(bm, 1)).all(axis=0)]

ret23 = ret23.iloc[3:, :]
ret23 = ret23.dropna(axis=1)
ret23 = ret23.iloc[:, (ret23.values < np.percentile(ret23, 99)).all(axis=0) & (ret23.values > np.percentile(ret23, 1)).all(axis=0)]

ret46 = ret46.iloc[6:, :]
ret46 = ret46.dropna(axis=1)
ret46 = ret46.iloc[:, (ret46.values < np.percentile(ret46, 99)).all(axis=0) & (ret46.values > np.percentile(ret46, 1)).all(axis=0)]

ret712 = ret712.iloc[12:, :]
ret712 = ret712.dropna(axis=1)
ret712 = ret712.iloc[:, (ret712.values < np.percentile(ret712, 99)).all(axis=0) & (ret712.values > np.percentile(ret712, 1)).all(axis=0)]

monthly_return = pd.read_csv('monthly_return.csv', header=1)
monthly_return = monthly_return.iloc[1:, :]
monthly_return.index = monthly_return['ticker']
monthly_return = monthly_return.drop(columns=['ticker'])
monthly_return = monthly_return.dropna(axis=1)
mc_max_cut = np.percentile(monthly_return, 99)
mc_min_cut = np.percentile(monthly_return, 1)
monthly_return = monthly_return.iloc[:, (monthly_return.values < mc_max_cut).all(axis=0) & (monthly_return.values > mc_min_cut).all(axis=0)]
monthly_return['ticker'] = monthly_return.index
monthly_return = pd.melt(monthly_return, id_vars='ticker', value_vars=list(monthly_return.columns)[:-1],
```python
var_name = 'tickers',
value_name = 'return'
monthly_return.columns = ['date', 'tickers', 'return']
monthly_return['return'] = monthly_return['return'] / 100

data = df_negative.merge(mc, how = 'inner', on = ['tickers', 'date'])
    .merge(prc, how = 'inner', on = ['tickers', 'date'])
    .merge(bm, how = 'inner', on = ['tickers', 'date']).merge(div, how = 'inner', on = ['tickers', 'date'])
    .merge(ret23, how = 'inner', on = ['tickers', 'date']).merge(ret46, how = 'inner', on = ['tickers', 'date'])
    .merge(ret712, how = 'inner', on = ['tickers', 'date']).merge(monthly_return, how='inner', on=['tickers', 'date'])
data.columns = df_negative.columns.tolist() + ['mc', 'prc', 'bm', 'div', 'ret23', 'ret46', 'ret712', 'return']
data = data.replace([np.inf, -np.inf], np.nan)
data1 = data.dropna()

# data1['dummy'] = data1['dummy'].clip(0, 1)

'''Fama-Macbeth regression'''
def ols_coef(x, formula):
    return smf.ols(formula, data=x).fit().params

def fm_summary(p):
    s = p.describe().T
    s['std_error'] = s['std']/np.sqrt(s['count'])
    s['tstat'] = s['mean']/s['std_error']
    s['pval'] = stats.t.sf(np.abs(s['tstat']), s['count']-1)*2
    return s[['mean', 'std_error', 'tstat', 'pval']]
gamma = data1.groupby('date').apply(ols_coef, 'ret ~ 1 + dummy + mc + prc + bm + div + ret23 + ret46 + ret712')
res = fm_summary(gamma)
res.pval = [round(x,3) for x in res.pval.values.tolist()]
gamma = data1.groupby('date').apply(ols_coef, 'ret ~ 1 + Association +Sanctions+Financial+Corruption+Information+Human+Workplace+Production_Supply+Environmental+Management+Workforce+Regulatory+Fraud+Anti_Competitive+Ownership+Product_Service+Discrimination_Workforce+mc + prc + bm + div + ret23 + ret46 + ret712')
res = fm_summary(gamma)
res.pval = [round(x,3) for x in res.pval.values.tolist()]
res

''pooled OLS with double-clustered standard error''
data1.date = data1.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
data1 = data1.set_index(['tickers', 'date'])
data1['const'] = 1
mod = PooledOLS(data1['ret'], data1[['dummy', 'mc', 'prc', 'bm', 'div', 'ret23', 'ret46', 'ret712', 'const']])
```
C.4 Relative Screening

```python
res = mod.fit(cov_type='clustered', cluster_entity=True,
              cluster_time=True)
```

---

```python
import pandas as pd
import numpy as np
import random
from dateutil.relativedelta import relativedelta
from datetime import date, datetime
import matplotlib.pyplot as plt
from scipy import stats
import statsmodels.api as sm

''' read data '''
df_negative = pd.read_pickle('df_negative_new_1.pkl')
df_negative = df_negative[['ticker', 'date', 'label2']]
df_negative['news_count'] = df_negative.groupby('ticker')['date'].rank(ascending=True)
df_negative = df_negative.sort_values(['ticker', 'date'])

data_ret = pd.read_csv('monthly_return.csv', header=1)
data_ret = data_ret.iloc[:,]
data_ret['ticker'] = data_ret['ticker'].apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
data_ret.index = data_ret['ticker']
data_ret = data_ret.iloc[:,1:]
data_ret /= 100
data_ret = data_ret.dropna(axis='columns')

stock_tickers = list(set(data_ret.columns) & set(df_negative.ticker.unique().tolist()))
df_negative = df_negative[df_negative.ticker.isin(stock_tickers)]
data_ret = data_ret[stock_tickers]

monthly_ret_all = pd.read_csv('monthly_ret_all_stocks_1.csv')
monthly_ret_all.index = monthly_ret_all['date']
monthly_ret_all = monthly_ret_all.iloc[:,1:]
monthly_ret_all /= 100
monthly_rets_all = monthly_ret_all.mean(axis=1).values.tolist()

monthly_ret_sp500 = pd.read_csv('monthly_return_sp500.csv', header=1)
monthly_ret_sp500 = monthly_ret_sp500.iloc[:,]
monthly_ret_sp500['ticker'] = monthly_ret_sp500.ticker.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
monthly_ret_sp500.index = monthly_ret_sp500.ticker
monthly_ret_sp500 = monthly_ret_sp500.iloc[:,1:]
monthly_ret_sp500 /=100

all_monthly_rets = monthly_ret_all.mean(axis=1).values.tolist()
sp500_monthly_rets = monthly_ret_sp500.mean(axis=1).values.tolist()
```
event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
event_ending_date = relativedelta(months=1) + event_starting_date
portfolio_starting_date = event_ending_date
portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
df = df_negative[(df_negative.date < event_ending_date)&(df_negative.date >= event_starting_date)]

'''value-weighted portfolio returns for worst 30 stocks'''

event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
event_ending_date = pd.Timestamp(year=2012, month=4, day=1)
df = df_negative[(df_negative.date < event_ending_date)&(df_negative.date >= event_starting_date)]
tickers = ['a' for i in range(30)]
monthly_rets = []
turn_overs = []
while event_ending_date < pd.Timestamp(year=2019, month=1, day=1):
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
    df = df_negative[(df_negative.date < event_ending_date)]
    news_count = pd.DataFrame(df.groupby('ticker')['news_count'].max())
    prev_tickers = tickers
tickers = news_count.sort_values('news_count', ascending=False)[:30].index.tolist()
turnover_rate = 1-len(list(set(prev_tickers) & set(tickers)))/30
    turnover_rate = turnover_rate
    rets = data_ret[tickers][portfolio_starting_date:portfolio_ending_date].values[0]
    mc = mc_all[tickers][portfolio_starting_date:portfolio_ending_date]
    mc_sum = np.sum(mc, axis=1).values[0]
    monthly_ret = np.sum((rets * mc)/ mc_sum , axis=1).values[0]
    monthly_rets.append(monthly_ret)
    event_ending_date += relativedelta(months=1)
monthly_rets_worst = monthly_rets

'''value-weighted portfolio returns for best 30 stocks'''

event_starting_date = pd.Timestamp(year=2012, month=1, day=1)
event_ending_date = pd.Timestamp(year=2012, month=4, day=1)
df = df_negative[(df_negative.date < event_ending_date)&(df_negative.date >= event_starting_date)]
tickers = ['a' for i in range(30)]
monthly_rets = []
turn_overs = []
while event_ending_date < pd.Timestamp(year=2019, month=1, day=1):
```python
portfolio_starting_date = event_ending_date
portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)

df = df_negative[(df_negative.date < event_ending_date)]
news_count = pd.DataFrame(df.groupby('ticker')['news_count'].max())

prev_tickers = tickers
tickers = news_count.sort_values('news_count', ascending=True)[:30].index.tolist()

turnover_rate = 1-len(list(set(prev_tickers) & set(tickers)))/30

turn_overs.append(turnover_rate)

rets = data_ret[tickers][portfolio_starting_date:portfolio_ending_date].values[0]
mc = mc_all[tickers][portfolio_starting_date:portfolio_ending_date]
mc_sum = np.sum(mc,axis=1).values[0]
monthly_ret = np.sum((rets * mc)/mc_sum ,axis=1).values[0]
monthly_rets.append(monthly_ret)

event_ending_date += relativedelta(months=1)

''value-weighted benchmark returns''
mc_sp500 = pd.read_csv('mc_sp500.csv',header = 1)
mc_sp500 = mc_sp500.iloc[:,1:]
mc_sp500.ticker = mc_sp500.ticker.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
mc_sp500.index = mc_sp500.ticker
mc_sp500 = mc_sp500.iloc[:,1:]
mc_all = pd.read_csv('monthly_market_cap_all_stocks_1.csv')
mc_all.date = mc_all.date.apply(lambda x: datetime.strptime(x, '%Y-%m-%d'))
mc_all.index = mc_all.date
mc_all = mc_all.iloc[:,1:]
all_tickers = mc_all.columns
all_tickers = [t.split('.', 1)[0].lower() for t in all_tickers]
mc_all.columns = all_tickers
all_tickers = monthly_ret_all.columns
all_tickers = [t.split('.', 1)[0].lower() for t in all_tickers]
monthly_ret_all.columns = all_tickers
stocks_sp500 = list(set(mc_sp500.columns) & set(monthly_ret_sp500.columns))
mc_sp500 = mc_sp500[stocks_sp500]
monthly_ret_sp500 = monthly_ret_sp500[stocks_sp500]
stocks_all = list(set(mc_all.columns) & set(monthly_ret_all.columns))
mc_all = mc_all[stocks_all]
monthly_ret_all = monthly_ret_all[stocks_all]
sp500_weighted_returns = (monthly_ret_sp500 * mc_sp500).div(mc_sp500.sum(axis=1),axis=0).sum(axis=1).values.tolist()
all_weighted_return = (monthly_ret_all * mc_all).div(mc_all.sum(axis=1),axis=0).sum(axis=1).values.tolist()
tickers = list(set(df_negative.ticker.unique()) & set(mc_all.columns))
```
C.5 Negative Screening

```python
import pandas as pd
import numpy as np
import random
from dateutil.relativedelta import relativedelta
from datetime import date, datetime
import matplotlib.pyplot as plt
from scipy import stats
import statsmodels.api as sm

''' read data '''
df_negative = pd.read_pickle('df_negative_new_1.pkl')
df_negative = df_negative[['ticker', 'date', 'label2', 'label1']]
df_negative['news_count'] = df_negative.groupby('ticker')['date'].rank(ascending=True)
```


```python
def_negative = df_negative.sort_values([‘ticker’, ‘date’])

monthly_ret_all = pd.read_csv(‘monthly_ret_all_stocks_1.csv’)  
monthly_ret_all.date = monthly_ret_all.date.apply(lambda x:  
    datetime.strptime(x, ‘%Y-%m-%d’))
monthly_ret_all.index = monthly_ret_all[‘date’]  
monthly_ret_all = monthly_ret_all.iloc[:,1:]  
monthly_ret_all /= 100  

all_tickers = monthly_ret_all.columns  
all_tickers = [t.split(‘.’, 1)[0].lower() for t in all_tickers]  
monthly_ret_all.columns = all_tickers

monthly_ret_sp500 = pd.read_csv(‘monthly_return_sp500.csv’, header=1)  
monthly_ret_sp500 = monthly_ret_sp500.iloc[1:,:]  
monthly_ret_sp500.ticker = monthly_ret_sp500.ticker.apply(lambda x:  
    datetime.strptime(x, ‘%Y-%m-%d’))  
monthly_ret_sp500.index = monthly_ret_sp500.ticker  
monthly_ret_sp500 = monthly_ret_sp500.iloc[:,1:]  
monthly_ret_sp500 /=100

all_monthly_rets = monthly_ret_all.mean(axis=1).values.tolist()  
sp500_monthly_rets = monthly_ret_sp500.mean(axis=1).values.tolist()

mc_sp500 = pd.read_csv(‘mc_sp500.csv’, header=1)  
mc_sp500 = mc_sp500.iloc[:,1:]  
mc_sp500.ticker = mc_sp500.ticker.apply(lambda x:  
    datetime.strptime(x, ‘%Y-%m-%d’))  
mc_sp500.index = mc_sp500.ticker  
mc_sp500 = mc_sp500.iloc[:,1:]  
mc_all = pd.read_csv(‘monthly_market_cap_all_stocks_1.csv’)  
mc_all.date = mc_all.date.apply(lambda x:  
    datetime.strptime(x, ‘%Y-%m-%d’))  
mc_all.index = mc_all.date  
mc_all = mc_all.iloc[:,1:]  
all_tickers = mc_all.columns  
all_tickers = [t.split(‘.’, 1)[0].lower() for t in all_tickers]  
mc_all.columns = all_tickers  
stocks_sp500 = list(set(mc_sp500.columns) & set(monthly_ret_sp500.columns))  
mc_sp500 = mc_sp500[stocks_sp500]  
monthly_ret_sp500 = monthly_ret_sp500[stocks_sp500]  
stocks_all = list(set(mc_all.columns) & set(monthly_ret_all.columns))  
mc_all = mc_all[stocks_all]  
monthly_ret_all = monthly_ret_all[stocks_all]  
sp500_weighted_returns = (monthly_ret_sp500 * mc_sp500).div(  
    mc_sp500.sum(axis=1), axis=0).sum(axis=1).values.tolist()  
all_weighted_return = (monthly_ret_all * mc_all).div(mc_all.sum(  
    axis=1), axis=0).sum(axis=1).values.tolist()
```
'value-weighted negatively screening portfolios'

'perform negative screening on all public traded US stocks'

```python
threshold = 0
event_ending_date = pd.Timestamp(year=2012, month=2, day=1)
monthly_rets_all_public = []
num_stocks_list = []
stock_tickers = monthly_ret_all.columns.tolist()

while event_ending_date < pd.Timestamp(year=2019, month=1, day=1):
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
    df = df_negative[(df_negative.date < event_ending_date)]
    count = pd.DataFrame(df.groupby('ticker')['news_count'].max())
    negative_tickers = count[count.news_count > threshold].index.tolist()
    tickers = [x for x in stock_tickers if x not in negative_tickers]
    num_stocks = len(tickers)
    num_stocks_list.append(num_stocks)
    rets = monthly_ret_all[tickers][portfolio_starting_date:portfolio_ending_date]
    mc = mc_all[tickers][portfolio_starting_date:portfolio_ending_date]
    mc_sum = np.sum(mc, axis=1).values[0]
    monthly_ret = np.sum((rets * mc) / mc_sum, axis=1).values[0]
    monthly_rets_all_public.append(monthly_ret)
    event_ending_date += relativedelta(months=1)

factors['R'] = monthly_rets_all_public
factors['RF'] = all_weighted_return[3:-1]
factors['R-RF'] = factors['R'] - factors['RF']
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'MOM', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

'perform negative screening on SP500 stocks'

threshold = 0
event_ending_date = pd.Timestamp(year=2012, month=2, day=1)
monthly_rets_sp500 = []
num_stocks_list = []
stock_tickers = monthly_ret_sp500.columns.tolist()

while event_ending_date < pd.Timestamp(year=2019, month=1, day=1):
    portfolio_starting_date = event_ending_date
    portfolio_ending_date = portfolio_starting_date + relativedelta(months=1)
    df = df_negative[(df_negative.date < event_ending_date)]
    count = pd.DataFrame(df.groupby('ticker')['news_count'].max())
    negative_tickers = count[count.news_count > threshold].index.tolist()
```

tickers = [x for x in stock_tickers if x not in negative_tickers]
num_stocks = len(tickers)
num_stocks_list.append(num_stocks)
rets = monthly_ret_sp500[tickers][portfolio_starting_date: portfolio_ending_date]
mc = mc_sp500[tickers][portfolio_starting_date: portfolio_ending_date]
mc_sum = np.sum(mc, axis=1).values[0]
monthly_ret = np.sum((rets * mc)/ mc_sum ,axis=1).values[0]
monthly_rets_sp500.append(monthly_ret)
event_ending_date += relativedelta(months=1)

factors['R'] = monthly_rets_sp500
factors['RF'] = sp500_weighted_returns[3:-1]
factors['R-RF'] = factors['R'] - factors['RF']
mod = sm.OLS(factors['R-RF'], factors[['SMB', 'HML', 'MKT', 'MOM', 'const']])
res = mod.fit(cov_type='HAC', cov_kwds={'maxlags': 12})
res.summary()

''' plot cumulative returns '''
date = pd.date_range(start='2012-02-01', end='2018-12-31', freq='MS')
cumsum_all_stocks = np.cumsum(all_weighted_return[3:-1])
cumsum_sp500 = np.cumsum(sp500_weighted_returns[3:-1])
cumsum_all_screened = np.cumsum(monthly_rets_all_public)
cumsum_sp500_screened = np.cumsum(monthly_rets_sp500)
data = pd.DataFrame(data = {'date': date, 'sp500': cumsum_sp500, 'all public': cumsum_all_stocks, 'all public negative screened': cumsum_all_screened, 'sp500 negative screened': cumsum_sp500_screened})
plt.plot('date', 'sp500', data = data)
plt.plot('date', 'sp500 negative screened', data = data)
plt.xlabel('year')
plt.ylabel('cumulative return')
plt.legend()}
Declaration of Originality

The signed declaration of originality is a component of every semester paper, Bachelor’s thesis, Master’s thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

**Title of work (in block letters):**

**AN IN-DEPTH ANALYSIS OF THE IMPACT OF ESG INVESTING ON RETURNS USING LARGE-SCALE NEWS DATA**

**Authored by (in block letters):**

For papers written by groups the names of all authors are required.

**Name(s):**

TIAN QI

WANG

**First name(s):**

**With my signature I confirm that**

- I have committed none of the forms of plagiarism described in the Citation etiquette information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.
- I am aware that the work may be screened electronically for plagiarism.
- I have understood and followed the guidelines in the document *Scientific Works in Mathematics*.

**Place, date:**

Zürich, 15.08.2019

**Signature(s):**

TIAN QI

WANG

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.