Norwegian School of Economics
Bergen, fall 2018

Sustainability Disclosure and Performance

An Empirical Study of U.S. Petroleum Corporations

Sigurd Ketilson Rykke Berstad & Theodor Lange Dokkan

Supervisor: Maximilian Rohrer

Master Thesis, MSc in Economics and Business Administration, Finance

NORWEGIAN SCHOOL OF ECONOMICS

Preface

This master's thesis marks the end of our Master of Science in Economics and Business Administration at the Norwegian School of Economics. We have studied the nature of eight U.S. petroleum companies' attitude towards sustainability by using textual data analysis.

During our major in finance, both of us had *Applied Textual Data Analysis for Business and Finance*. The subject yielded a thorough introduction to the programming language R along with the quantification of textual data to conduct analysis. Moreover, both of us possess a genuine interest for the field of sustainability, gained from subjects at NHH and previous internships. Thus, to apply textual data analysis to investigate petroleum companies' attitude towards sustainability was a fitting issue for our thesis.

Moreover, we would like to offer our gratitude to our advisor Maximillian Rohrer. Thank you for offering valuable guidance and feedback throughout the semester.

Finally, we would like to thank our families. Without their support, the road towards submission would not have been the same. They motivated us to proceed and not forget to have fun at the same time. Thank you.

Abstract

This thesis investigates if some of the largest U.S. petroleum companies' written disclosure regarding sustainability is correlated with their sustainability performance. We measure the companies' written disclosure with textual data analysis, conducted on annual reports (10-K), event reports (8-K) and sustainability reports. Furthermore, we represent the sustainability performance with the indicators greenhouse gas emissions, total recordable incident rate and environmental expenditures, which represent the environmental, social and economic dimension of sustainability, respectively. We find that the companies' combined written disclosure is not correlated with the companies' sustainability performance. However, when analysing the disclosure document measures individually, we observe that how much the companies write about sustainability correlates less with their performance, compared to when they write positive or negative.

Contents

1. Introduction

In 2011, a court in Ecuador fined Chevron 8.6 billion USD for damages on crops and farm animals, as well as increased local cancer rates due to of reckless oil drilling (The New York Times, 2011). Yet, in Chevron's sustainability report for 2011 (Chevron, 2011), the company cherishes their land-conservation program in Indonesia, whereas the incident in Ecuador is not mentioned. Moreover, Global Witness (2018) claims that in 2013, ExxonMobil signed a 120 million USD deal with the Liberian government, even though they knew the government was tainted by corruption. However, they state nothing about this accusation in their sustainability report for 2013, while claiming that *"ExxonMobil is committed to the highest standards of business conduct and anti-corruption wherever we operate"* (ExxonMobil, 2013). These examples illustrate that U.S. petroleum companies' sustainability disclosure is potentially not correlated with their sustainability performance. Hence, we question their integrity by investigating the following thesis:

"Is the written disclosure regarding sustainability correlated with the sustainability performance, among eight of the largest U.S. petroleum companies?"

We define sustainability in accordance with Global Reporting Initiative's (GRI) (2011) definition, being *"the development that meets the needs of the present without compromising the ability of future generations to meet their own needs"*. In the same report, GRI divide sustainability into three dimensions: environmental, social and economic. To represent the companies' sustainability performance, we extract three indicators, one for each of the three dimensions. The performance measures are the companies' greenhouse gas emissions (GHG), total recordable incident rate (TRIR) and environmental expenditures. These are retrieved from the Carbon Majors database (Paul, Heede, & Flugt, 2017), sustainability reports and annual reports (10-K), respectively.

We represent the written disclosure by quantitative measures which describes how the petroleum companies write about sustainability. These measures are computed with word frequencies and sentiment analysis on the companies' annual reports, event report (8-K) and

sustainability reports. From these documents, we construct six independent variables, which henceforth are referred to as the textual variables.

In order to investigate the correlation between the companies' written disclosure and their performance, we construct three regression tables. The table of the environmental, social and economic dimension has GHG, TRIR and environmental expenditures as its dependent variable, respectively. Moreover, the independent variables are a combination of textual and control variables. First, we analyse the correlation between the textual variables and the sustainability performance when all textual variables are included in the regression. Second, we assess the correlation by reviewing the textual variables individually. Our hypothesis, being that the companies' written disclosure regarding sustainability is not correlated with their performance, is rejected if the textual variables' coefficients are significant.

While there are several limitations to our thesis, our main result supports that the eight U.S. petroleum companies' combined written disclosure regarding sustainability is not correlated with their sustainability performance. However, when we interpret the textual variables individually, the results suggest that how much they write about sustainability correlates less with their performance compared to when they write positive or negative.

2. Background

United Nations states that *"after more than a century and a half of industrialization, deforestation, and large-scale agriculture, quantities of greenhouse gases in the atmosphere have risen to record levels not seen in three million years"* (United Nations, u.d.). These emissions contribute to global warming, which in turn leads to for example higher sea levels and more extreme weather (National Ocean Service, 2018). Moreover, Carbon Majors (2017) states that 50 % of GHG emissions origins from production and use of oil and gas products. Therefore, we argue that emissions from petroleum companies have a considerable global impact on the environment.

The petroleum industry is also accountable for several incidents which had an impact on local environments. For instance, the Exxon Valdez oil tanker released about 350 000 barrels of oil into the Prince William Sound in Alaska in 1989 (Shin, 2017). The incident resulted in the death of 2000 sea otters, about 250 000 thousand birds and pollution of 1990 square kilometers of shoreline (Bright Hub Engineering, u.d.). Another example is the Deepwater Horizon incident, in which four million barrels of oil were spilled and 11 persons were killed (Milman, 2018). Thus, consequences from petroleum incidents are potentially extensive. This also means that petroleum companies' incident risk management may have an impact on stakeholders like the government, local communities and natural habitats.

By improving their local and global impact, we argue the petroleum companies can also increase their profitability. Alshehhi, Nobanee & Khare (2018) finds that 78 % of publications regarding the relationship between corporate sustainability and financial performance, finds a positive relation. For instance, increased focus on sustainability can save costs and improve a company's reputation. Efficiency-increasing efforts, like new drilling technology, can reduce energy consumption, which in turn reduces costs and emissions. Further, companies which address sustainability have become increasingly popular for investors (Forbes, 2017). Thus, by gaining a reputation of being socially responsible, firms can expect a positive development in their stock prices and their ability to raise capital. We do observe that several petroleum companies, like BP (BP, 2018) and Equinor (Equinor, 2018), address sustainability by redesignating themselves to broad energy

companies, with an increased focus on renewable energy. Nevertheless, renewable energy is still only a small portion of their operations (Aras & Crowther, 2008).

Even though some petroleum companies have initiated sustainability efforts, we question their true incentives. An analysis of 49 Fortune-500 firms by Orniston & Wong (2013), found that corporate social responsibility efforts predicted subsequent corporate social irresponsibility. The study raises the discussion of moral licensing, meaning that good responsible behaviour justifies social irresponsibility. We observe examples of potential moral licensing in the petroleum industry. E.g., according to The U.S. Fish and Wildlife Service (Fears, 2017), "*the single most important achievement for polar bear conservation is decisive action to address Arctic warming".* Hence, the emissions from petroleum companies are contributing to destroying polar bear habitats. However, we observe that ExxonMobil (ExxonMobil, u.d.) and ConocoPhillips (Environmental Studies Program, 2005) ironically have polar bear protection programs. Thus, the petroleum companies may also address sustainability to justify their negative local and global impacts.

Furthermore, if 78 % of publications suggests there is a business case for sustainability, why do not all companies implement sustainability? Crowther (2002) emphasize that financial and sustainability performance are by some regarded as conflicting with one another. This perception of sustainability might exist because there is no universal definition of what sustainability encompasses (Johnston, Everard, Santillo, & Robèrt, 2007). Due to this vague definition, top corporate executives may place less faith in sustainability research, and thus trust their own beliefs and experiences instead. For instance, U.S. top executives' political affiliation can reveal deviating beliefs in terms of sustainability. Compared to democrats, republicans tend to be more sceptical regarding sustainability subjects, like human contribution to global warming (Popovich & Albeck-Ripka, 2017). The deviating beliefs on sustainability may clarify why Chin, Hambrick & Trevino (2013) find that companies with republican CEOs tend to invest less in corporate social responsibility than democratic CEOs. Therefore, beliefs on sustainability among top executives can affect companies' will to address the subject.

To summarize, the petroleum industry's impact on the world is considerable, which emphasizes the need to address sustainability in this sector. By addressing sustainability, the petroleum companies can benefit by saving costs and improve their reputation. We do

observe that some petroleum companies are starting to change their impact on the world. Yet, their true incentives for addressing sustainability may be questionable. For example, sustainability may be accommodated to justify irresponsible behaviour. In addition, different beliefs among U.S. top executives may also contribute to explain why sustainability is not implemented in a greater extent.

3. Literature Review

In this section we review two articles which can be related to textual data analysis on sustainability-related issues. Barkemeyer, Comyns, Figge & Napolitano (2014) conduct sentiment analysis on CEO statements from corporate sustainability reports to see if the rhetoric used in sustainability reporting accurately reflects sustainability performance. Their results indicate that sustainability reporting has not matured over the period they choose, and that the rhetoric used in the CEO statements is superficial rather than accountable. Wen (2014) develops two main algorithms, the first regress word frequencies from 10-K annual reports against a published sustainability score. The second conduct sentiment analysis on the same annual reports to analyse the positive and negative sentiment on sentences which contain sustainability-related words. The result of the first main algorithm indicates there is no significant correlation between the frequency of sustainability-related words and sustainability score. As regards the second main algorithm, no conclusions are made.

First of all, we contribute to the existing literature by using several document types. The other articles apply one type of document on which they perform textual analysis. We conduct analyses on the 10-K annual reports, 8-K event reports and sustainability reports, yielding six textual variables. Thus, the analysis scope of existing literature, both in terms of document types and number of textual variables, is less extensive. Second, the terms we use to describe sustainability is chosen with a quantitative and qualitative analysis, in contrast to Wen's (2014) article in which similar terms are chosen solely with algorithms.

4. Companies and Years

According to the Climate Accountability Institute (2017), the world's 100 largest fossil fuel producers are linked to 71 % of industrial greenhouse gas emissions since 1988. Of those 100 fossil fuel producers, we review the following eight U.S. corporations:

Anadarko Petroleum Corporation
Apache Corporation
Chevron Corporation
ConocoPhillips
ExxonMobil Corporation
Hess Corporation
Marathon Oil Corporation
Occidental Petroleum Corporation

We select U.S. companies to utilize the data disclosed on the U.S. Securities and Exchange Commission's (SEC) web pages (8-Ks and 10-Ks). In addition, we do not include other countries due to deviating reporting standards and rules, and thus incomparable reports. Moreover, we consider the years 2004 - 2015, since we observe that these years involves relatively consistent sustainability reporting compared to the years prior to 2004. We limit the years to 2015, since the Carbon Majors dataset does not include GHG emissions for the years 2016 - 2018.

5. Data Foundation

5.1 10-K Annual Reports

U.S. public companies are required to produce an annual report report each year and file it with the SEC. The commission requires that annual reports follow the form of 10-K. The 10-K form contain a detailed picture of a company's business, the risk it faces, and the operating and financial results for the fiscal year. (U.S. Securities and Exchange Commision, 2011)

	Total	Average	Standard Deviation	Min	Max
Words in Annual Reports	2898065	30188	6078	17860	42595

Figure 1 - Descriptive Statistics of Annual Reports

We gather one annual report per firm over the 12 years, yielding a total of 96 reports. A 10-K report from our sample has 30188 words on average. The standard deviation of 6078 indicates that relatively to the mean, there is small variation in the annual reports, compared to the other document types (Figure 2 and 3). The relatively low variation can be explained by the strict set of rules entailed with the 10-K annual reports.

5.2 Sustainability Reports

The sustainability reports follow the reporting standards established by the Global Reporting Initiative. These standards include information about a company's economic, environmental and social impact. Furthermore, these reports present the corporation's values and governance model, and demonstrate the link between the company's strategy and its efforts to address sustainability. (Global Reporting Initiative, u.d.)

We mainly retrieve the sustainability reports from GRI's web pages, however, some are collected from the companies' websites. We collect a total of 57 reports, most of which were published during the later years of those we regard. The reports from recent years contain

tables with key performance indicators, which can be used to assess a company's sustainability performance.

	Total	Average	Standard Deviation	Min	Max
Words in Sustainability Reports	841024	14755	7775	546	32541

Figure 2 - Descriptive Statistics of Sustainability Reports

As can be seen from figure 2, the average number of words for each report is 14755, meaning the sustainability reports tend to be shorter than the annual reports. However, with a standard deviation of 7775 words, these reports have relatively greater variation in length compared to the annual reports.

5.3 8-K Event Reports

In addition to filing annual reports on form 10-K, public companies must report certain corporate events on a current basis. Examples on such events are termination of a material definitive agreement, bankruptcy or acquisition of assets. These major events require the 8-K report form. (U.S. Securities and Exchange Commision, u.d.)

	Total	Firm Average per Year	Firm Standard Deviation per Year	Min	Max
Words in 8-K Event Reports	358963	3739	2870	53	16960

Figure 3 - Descriptive Statistics of Event Reports

We observe that the total number of words is lower than both the annual reports and the sustainability reports. Moreover, the firm standard deviation per year is relatively high, compared to the other document types. The standard deviation indicates a large variation in published reports.

	Total	Firm Average per Year	Firm Standard Deviation per Year	Min	Max
Number of 8-K Event Reports	1190	12.4	4.4	1	26

Figure 4 - Descriptive Statistics of Event Reports

For the years 2004 - 2015, our selected companies published a total of 1190 event reports, which yields an average of 148.8 reports per company. On average, the companies file 12.4

reports annually, which is approximately one report per month. However, due to the standard deviation of 4.4, and a range of 25, the reporting tends to be inconsistent. To investigate which extent the 8-Ks emphasize sustainability, we conduct an analysis to determine if the terms in the event reports resembles the annual or the sustainability reports (Appendix 12.1). Our result suggests that the event reports are more similar to the 10-Ks, and thus contain financial content rather than sustainability-oriented issues.

5.4 Vector to Analyse Written Disclosure on Sustainability

In order to analyse the written disclosure on sustainability, we construct the following vector, which consists of sustainability-related terms, and their inflected forms.

environment	*greenhouse*
safety	*citizenship*
water	*conservation*
emission	*integrity*
health	*flare*
human	*biodiversity*
social	*community*
incident	*diversity*
climate	*transparency*
sustainable	*renewable*
spill	*corruption*
stakeholder	*wildlife*
waste	*footprint*

Figure 5 - The Relevant Words Vector

The inflected forms are omitted in figure 5 due to readability purposes. To construct this vector, we use the following quantitative assessment. First, terms from all of the sustainability reports are assembled in a vector, and terms with a frequency of less than 200

are excluded. Then, inflected forms are collapsed into their shortest form, and their frequencies added together. Subsequently, the 200 most frequent words are assembled into a vector (Appendix 12.2). This approach may reveal the most common and relevant words for sustainability in the petroleum industry. For example, an industry specific word is “flare” or “flaring”, which is a common cause of direct greenhouse gas emissions (Emam, 2015).

Finally, we use a qualitative assessment to determine which terms from the frequency vector fit into the dimensions of sustainability. We regard this assessment as necessary because many words in the initial frequency vector are unrelated to sustainability, like “company” and “operation”. Furthermore, some of the terms, like “responsible”, are used in different contexts in the text documents. E.g. “Responsible” is mainly used in annual reports and event reports to describe a responsible attitude towards financial risk, whereas in the sustainability reports, the term is used to address responsibility towards externalities. Thus, the written disclosure on sustainability can be biased if those words are included.

6. Independent Textual Variables

In this section, we construct six independent textual variables, with the use of two textual approaches on 10-Ks, 8-Ks and sustainability reports. The first is the rate, for each firm, of sustainability-related words relative to the total amount of words for a document type over one year. The second is the sentiment of sentences containing sustainability-related words on the various document types. The first method examines the relative amount of which the companies write about sustainability-oriented issues, and the second investigates to which extent they write positive or negative about sustainability.

6.1 Relevant Words' Fraction of 10-Ks and 8-Ks

We calculate the *Relevant Words' Fraction 10-K* by taking the sum of the frequency of the words from the *Relevant Words Vector* for each 10-K, and dividing it by the total number of words in the respective 10-K. Furthermore, in the 8-Ks, the sum of the sustainability-related words is divided by the total amount of words for all 8-Ks for a given year and firm, resulting in 96 observations. However, many of the 8-K reports contain no sustainability-related words, which yields several zero values. The purpose of both variables is to measure how much the companies write about sustainability-oriented matters. Therefore, an increase in the rate represents that the companies' written disclosure regarding sustainability increases.

	Mean	Standard Deviation	Missing Observations
Relevant Words' Fraction 10-K	0.49 %	0.15 %	0
Relevant Words' Fraction 8-K	0.05 %	0.06 %	0

Figure 6 - Descriptive Statistics of Relevant Words' Fraction

As can be seen from figure 6, the average implies that the words related to sustainability in the 10-Ks and 8-Ks constitutes 0.49 % and 0.05 % of the report, respectively. Thus, sustainability is addressed in a greater extent in the annual reports compared to the event reports. Additionally, we observe that the standard deviation is greater than the mean for the

event reports, which implies a relatively large variation between firms and years compared to the 10-Ks.

6.2 Sustainability Report Dummy

Since we only have 57 sustainability reports, a potential relevant words' fraction of sustainability reports will include 39 missing observations. The relatively low number of observations increases the possibility for biased coefficients in the regressions. Therefore, we create a dummy for whether they file a sustainability report as a proxy for the amount of companies' written disclosure. Thus, if the companies disclose a report, it represents an increase in the written disclosure regarding sustainability.

6.3 Sentiment on Sustainability-Related Sentences in 10-Ks, 8-Ks and Sustainability Reports

For the annual, event and sustainability reports, our script locates all sentences that contains at least one sustainability word from the *Relevant Words Vector.* Similar to when the *Relevant Words' Fraction 8-K* is conducted, all 8-K reports for a given year and firm are collapsed to a single document. We retrieve the sentences by using regular expressions, which is a "search language" used to locate words, sentences or other textual patterns. Furthermore, in our sentiment analysis, the sentences are scanned for words present in the Harvard IV dictionary. We select this general-purpose dictionary due to lack of sustainability-related dictionaries. The dictionary assigns positive and negative words in the sentences a sentiment score. Moreover, the sentiment score of all collapsed sentences for a given year and firm are summarized, yielding a netto sentiment score. Finally, the netto sentiment is divided by the total number of words of all collapsed sentences for a given year and firm. This last computation in our selected sentiment analysis ensures that the scores are comparable for a document type with differing lengths. We select the three sentiment variables as proxies for whether the companies' written disclosure is positive or negative in terms of sustainability.

	Mean	Standard Deviation	Missing Observations
Sentiment of Sustainability-Related Sentences in 10-K	0.015	0.012	0
Sentiment of Sustainability-Related Sentences in 8-K	0.027	0.043	49
Sentiment of Sustainability-Related Sentences in Sustainability Reports	0.042	0.006	39

Figure 7 - Descriptive Statistics of Sentiment of Sustainability-Related Sentences

We observe that the sentiment is more positive in the sustainability reports than in the other documents. In addition, the sentiment variation in the sustainability reports is also relatively less than the 10-Ks and the 8-Ks. Furthermore, the event reports have a high standard deviation compared to its mean. This can be explained by the high variation in the length of the 8-K reports, which affects the denominator in the sentiment score. The missing values in *Sentiment of Sustainability-Related Sentences in 8-K* can be explained by the low frequency of sustainability words in these documents. As regards the missing values in the sentiment on sustainability reports, these are explained by the lack of reports for some of the early years. Thus, the inference credibility related to *Sentiment of Sustainability-Related Sentences in Sustainability Reports* is reduced.

7. The Environmental Dimension

Global Reporting Initiative (2011) states the environmental dimension concerns *"an organisation's impact on living and non-living natural systems, including ecosystems, land, air and water."* Considering that greenhouse gas emissions have an impact on all examples in GRI's description, we regard emissions as representative for the environmental dimension. Therefore, we consider the following dependent variable:

Emission Rate = Greenhouse Gas Emissions / Production

Greenhouse gas emissions are constituted by Scope 1 and Scope 3 emissions of metric tonnes of CO2 equivalent, whereas production is measured in thousand barrels of oil equivalent. Scope 1 is defined by direct emissions from operations that are owned or controlled by the organisation (Global Reporting Initiative, 2011). For instance, CO2 emissions from fuel powering the oil pumps or chemical release from processing petroleum products at refineries. Scope 3, on the other hand, is indirect emissions from the entire company supply chain. By using a combination of direct and indirect emissions, differences due to various positions in the supply chain are accounted for. For instance, Marathon Oil might have relatively more oil pumps and less refineries than Anadarko. Thus, if only direct emissions are analyzed, Anadarko may have a different emission intensity dependent on their positioning. Hence, by including all indirect emissions, GHG pollution from the whole supply chain is included for each barrel produced. To make GHG emissions comparable across firms, we adjust for production. This computation also ensures that reduced emissions are not a reflection of decreased production. The data used to compute the *Emission Rate* is gathered from the Carbon Majors database and the 10-Ks.

	Mean(Emissions/Production)	Standard Deviation	Missing Observations
Anadarko Petroleum Corp	0.40	0.09	0
Apache Corp	0.51	0.09	0
Chevron Corp	0.41	0.03	0
ConocoPhillips	0.48	0.08	0
ExxonMobil Corp	0.40	0.02	0
Hess Corp	0.40	0.01	0
Marathon Oil Corp	0.41	0.05	0
Occidental Petroleum Corp	0.44	0.07	0

Figure 8 - Descriptive Statistics of the Emission Rate for the Years 2004 - 2015

Figure 8 shows the descriptive statistics for the *Emission Rate* from 2004 to 2015. On average, the companies pollute 0.43 metric tonnes of CO2 equivalent per thousand barrels of oil equivalent. The range of the mean is 0.11, and the range of the standard deviation is 0.08. There is no missing data for any of the companies.

We conduct a logarithmic transformation of the *Emission Rate* to account for a potential non-linear trend over time. It is likely that reducing emissions are more feasible from a higher level, but it turns increasingly difficult when emissions are lower. E.g., diminishing technological developments, related to energy use in production or emission filters at refineries. Moreover, the emissions can be reduced, but never turn negative, making a linear relationship unfeasible.

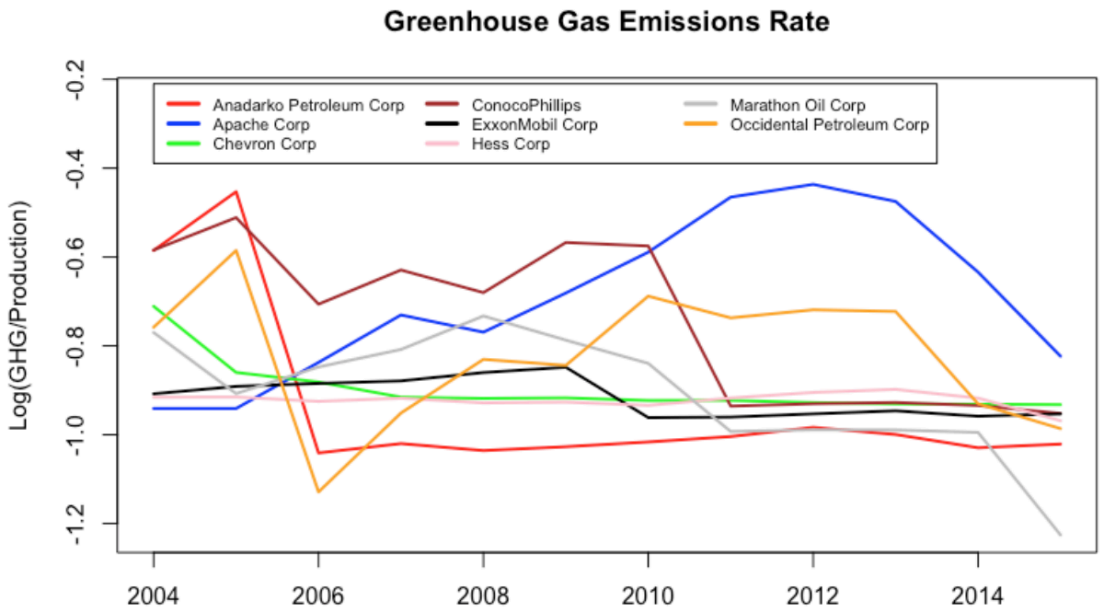


Figure 9 - The Logarithmic Emission Rate vs. Years

Visually, figure 9 does not indicate any time-specific variation. On the one hand, government regulations can yield a time-fixed effect. E.g., higher fuel standards which was imposed by the Obama administration in 2012 (The White House - Office of the Press Secretary, 2012) can lead to lower *Emission Rate*. On the other hand, the technological developments in shale oil production in North-America, has increased the activity for this segment (Cai, et al., 2015). Since shale oil production has a 18-21 % higher emissions intensity than conventional crude oil production (Cai, et al., 2015), the advancements in this segment can lead to higher *Emission Rate* for the later years of those we regard. Additionally, volatility in the oil price can also result in time-specific effects. For instance, higher oil price opens production on fields with a higher break-even price. Those fields often have greater energy use (Gavenas, Rosendahl, & Skjerpen, 2015), which in turn leads to more GHG emissions per barrel produced. If time-specific effects are not adjusted for, we can get correlation between our textual variables and performance variables, even though the correlation is potentially driven by time-specific variation. We account for this by including year dummies in the regression.

Deviating technology between firms can lead to differences in the GHG per barrel produced. If one firm holds more energy-efficient technology than another, GHG emissions per barrel

produced can differ. Furthermore, while Scope 3 covers indirect emissions in the supply chain, it does not cover differences in the GHG rate due to deviating horizontal positions. Different horizontal positions refer to e.g. that one firm is focused on conventional crude, whereas another is more focused on deepwater production, which can yield differences in GHG intencity. These arguments suggest the presence of firm-specific effects, and thus we include firms dummies in the regression. These ensure, on the same foundation as the year dummies, that the correlations we observe on the textual variables are adjusted for differences across firms.

7.1 Results

In this section, we investigate whether the regression coefficients are significant, in the sense that the companies' written sustainability disclosure is correlated with their *Emission Rate*. Our hypothesis is that there is no correlation, meaning the coefficients are statistically insignificant. In the following regressions, we include eleven year dummies and seven firm dummies as control variables. The consequence of adding the dummies is that variation due to firm- or time-specific effects are accounted for. However, the dummies may consume variation in the textual variables. If for instance one textual variable and a control variable are collinear, a potential correlation between the textual variable and the dependent variable can be hidden in the control variable. To account for potential collinearity, all coefficients' significance is compared to a regression without the control variables. If a textual variable is significant in the regressions without control variables, but turns insignificant with the inclusion of controls, we cannot neglect the possibility that the dummies have consumed the variation. Moreover, if a variable is not significant in the regression with and without firm and year dummies, our conclusion of no correlation is more robust.

	Dependent variable:							
	ln(Greenhouse Gas Emissions/Production)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relevant Words' Fraction 10-K	-8.752						-9.803	
	(7.774)						(8.318)	
Relevant Words' Fraction 8-K		3.421					4.380	
		(13.621)					(13.907)	
Sustainability Report Dummy			0.025				0.022	
			(0.019)				(0.020)	
Sentiment of Sustainability-Related Sentences in 10-K				0.617			0.843	
				(1.257)			(1.376)	
Sentiment of Sustainability-Related Sentences in 8-K					-0.499			
					(0.305)			
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-2.354**		
						(0.995)		
Observations	96	96	96	96	47	57	96	96
Adjusted R^2	0.300	0.289	0.304	0.290	0.575	0.668	0.291	0.297
Note:								*p<0.1; **p<0.05; ***p<0.01

Figure 10 - Regression Table of the Environmental Dimension

Control variables (Appendix 12.11, 12.12 and 12.13) are not shown in figure 10 above for readability purposes. The table consists of eight regressions in which the coefficients estimates are presented, significance level indicated by (*) and standard deviations in parentheses. We use a 10 % level of significance throughout the analyses, since the level of significance should be set as a decreasing function of sample size (Gill, 1999), and our sample size is relatively small. Since missing observations reduce the credibility of coefficient estimates (Kang, 2013), we omit the variables *Sentiment of Sustainability-Related Sentences in 8-K* and *Sentiment of Sustainability-Related Sentences in Sustainability Reports* in the multiple regression (7).

None of the coefficients in the multiple regression (7) are significantly different from zero. If *Relevant Word's Fraction 10-K* increases with 0.1 percentage points, the *Emission Rate* decreases with 0.98 %, all else equal. If *Relevant Word's Fraction 8-K* increases with 0.1 percentage points, the *Emission Rate* increases with 0.44 %. Furthermore, if a company file a sustainability report, the *Emission Rate* increases with 2.2 %. Finally, if *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *Emission Rate* increases with 0.84 %. The additional adjusted R^2, gained by going from the control variables regression (8) to the multiple regression (7), is -0.6 percentage points. The negative development suggests that the additional information the textual variables give the model, does not outweigh the loss in efficiency. Hence, we conclude that we cannot reject the

hypothesis that the textual variables are not correlated with the *Emission Rate* after adjusting for firm- and time-fixed effects.

When we compute regression (7) without time- and firm-fixed effects (Appendix 12.7), only *Sentiment of Sustainability-Related Sentences in 10-K* is significantly different from zero. This variable loses significance after the inclusion of controls. Therefore, we cannot neglect the possibility that its insignificance can be a result of consumed variation by the dummies. Moreover, the insignificance of *Relevant Word's Fraction 10-K, Relevant Word's Fraction 8-K* and *Sustainability Report Dummy,* in both the regressions with and without control variables, gives a more solid conclusion for these three variables.

The negative coefficient of *Relevant Words' Fraction 10-K* in regression (1) indicates that the companies write more about sustainability-related issues in the annual reports when the *Emissions Rate* decreases, and vice versa. If *Relevant Words' Fraction 10-K* increases with 0.1 percentage points, the *Emission Rate* decreases with 0.88 %, all else equal. Moreover, the coefficient is not significantly different from zero. Hence, we have a statistical basis to conclude there is no correlation between the frequency of sustainability-related words in annual reports and the *Emission Rate* when time- and firm-fixed effects are adjusted for. However, when we conduct regression (1) without time- and firm-fixed effects (Appendix 12.7), the coefficient is significant. Thus, we cannot neglect that time- and firm specific effects consume some of the variation, and hence make *Relevant Words' Fraction 10-K* insignificant.

If *Relevant Words' Fraction 8-K* in regression (2) increases with 0.1 percentage points, the *Emission Rate* increases with 0.34 %. The coefficient of *Relevant Words' Fraction 8-K in* regression (2) is not significant, implying we have the statistical foundation to conclude there is no correlation between the frequency of sustainability-related words in the 8-Ks and the *Emission Rate*. In addition, when we calculate regression (2) without time- and firm-fixed effects (Appendix 12.7), the coefficient is also insignificant. Hence, our conclusion is more robust.

The coefficient of the variable *Sustainability Report Dummy* from regression (3) supports a positive correlation, which implies that if they publish a sustainability report, they have a greater *Emission Rate*. If a company file a sustainability report, the GHG *Emission Rate*

increases with 2.5 %, all else equal. However, the coefficient is not significantly different from zero. Hence, we have the statistical basis to conclude there is no correlation between publishing a sustainability report and their *Emission Rate*. In addition, when we construct regression (3) without time- and firm-fixed effects (Appendix 12.7), the coefficient is also insignificant, which strengthens the conclusion.

Furthermore, the coefficient sign of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) indicates that if the companies write more positively about sustainability-oriented issues in their annual reports, the *Emission Rate* increases. If *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *Emission Rate* increases with 0.62 %. Moreover, the coefficient of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) is not significantly different from zero, which suggests that we can not reject the null hypothesis stating there is no correlation between the variable and the *Emission Rate*. However, in regression (4) without time- and firm-fixed effects (Appendix 12.7), the coefficient is significant. Thus, the coefficient's insignificance can be explained by consumed variation in the control variables.

The coefficient of *Sentiment of Sustainability-Related Sentences in 8-K* in regression (5) indicates a negative correlation, meaning that if the companies write more positively about sustainability-related issues in their event reports, the *Emission Rate* decreases. If *Sentiment of Sustainability-Related Sentences in 8-K* increases with 0.01, the *Emission Rate* decreases with 0.5 %. However, we can not reject the null hypothesis stating there is no correlation between the *Sentiment of Sustainability-Related Sentences in 8-K* and the *Emission Rate.* Moreover, in regression (5) without time- and firm-fixed effects (Appendix 12.7), the coefficient is not significant. Thus, the robustness of the results increases. However, since the 8-K reports seldom address sustainability, this variable has many missing observations. Thus, we cannot neglect the possibility that the coefficient is biased.

Regarding regression (6), its coefficient of *Sentiment of Sustainability-Related Sentences in Sustainability Reports* indicates a negative correlation. The negative correlation suggests that if the companies write more positively about sustainability-related issues in their sustainability reports, the *Emission Rate* decreases. If *Sentiment of Sustainability-Related Sentences in Sustainability Reports* increases with 0.01, the *Emission Rate* decreases with 2.4 %, all else equal. Finally, we can reject the null hypothesis, hence we conclude that the

Sentiment of Sustainability-Related Sentences in Sustainability Reports and the *Emission Rate* are negatively correlated. However, his regression also has many missing observations in the earlier years, which in turn indicates that we can not neglect the possibility of biased results. For instance, bias can be introduced if the sentiment on sustainability reports are significantly higher or lower in the early years we regard. The potential bias reduces the credibility of the significance.

8. The Social Dimension

According to GRI (2018), the social dimension of sustainability *"concerns an organization's impacts on the social systems within which it operates".* A social system is an organisation of individuals into groups or structures that have different functions, characteristics, origin or status (Business Dictionary, u.d.). Hence, we consider the employees as a social system within an organization. Thus, the safety of the employees can be used to represent an impact the organization has on an internal social system. Since the TRIR describes the safety of a company's employees, the rate fits into the social dimension of sustainability.

The TRIR is total employee incidents per 200 000 work hours. According to the Society of Petroleum Engineers (2017), the most common causes of incidents among employees are road accidents, falls, explosions or fires.

	Mean	Standard Deviation	Missing Observations
Anadarko Petroleum Corp	0.47	0.19	4
Apache Corp	1.25	0.40	4
Chevron Corp	0.27	0.12	0
ConocoPhillips	0.13	0.05	4
ExxonMobil Corp	0.30	0.07	0
Hess Corp	0.95	0.72	0
Marathon Oil Corp	0.64	0.17	3
Occidental Petroleum Corp	0.39	0.07	2

Figure 11 - Descriptive Statistics TRIR for the Years 2004 - 2015

The mean rate across firms and years is 0.54, meaning that there is about half an incident per 200 000 work hours. The range of the mean and standard deviation is 1.12 and 0.67, respectively. The *TRIR* consists of 79 observations, all of which are gathered from the sustainability reports. The missing observations are skewed to the early years, which can be explained by the inconsistency in sustainability reporting these years. This suggests that there might be greater uncertainty to the regression estimates in the early years compared to the later. Contrary to the other companies, Marathon Oil calculates the *TRIR* for employees and contractors combined. However, an F-test supports that deviating variation is not an issue (Appendix 12.3), and potential level differences can be solved with firm-fixed effects.

In the initial phase of addressing the *TRIR*, there is a lot of potential for risk reduction, yielding relatively large decreases in the rate. However, after the first adoptions of safety policies and efforts, it becomes increasingly hard to find new potential for risk reduction. In addition, some incidents will never be avoidable, like human failure. Therefore, there will always be some risk of incidents. Moreover, a company can never have a negative number of incidents, making it impossible for the *TRIR* to be negative. Hence, we choose to conduct a logarithmic transformation of *TRIR* to account for a percent-wise development. In addition, a test for normality rejects the null hypothesis of a normally distributed variable for the level variable, but normality can not be rejected for the logged *TRIR* (Appendix 12.5). This result supports a logarithmic variable.

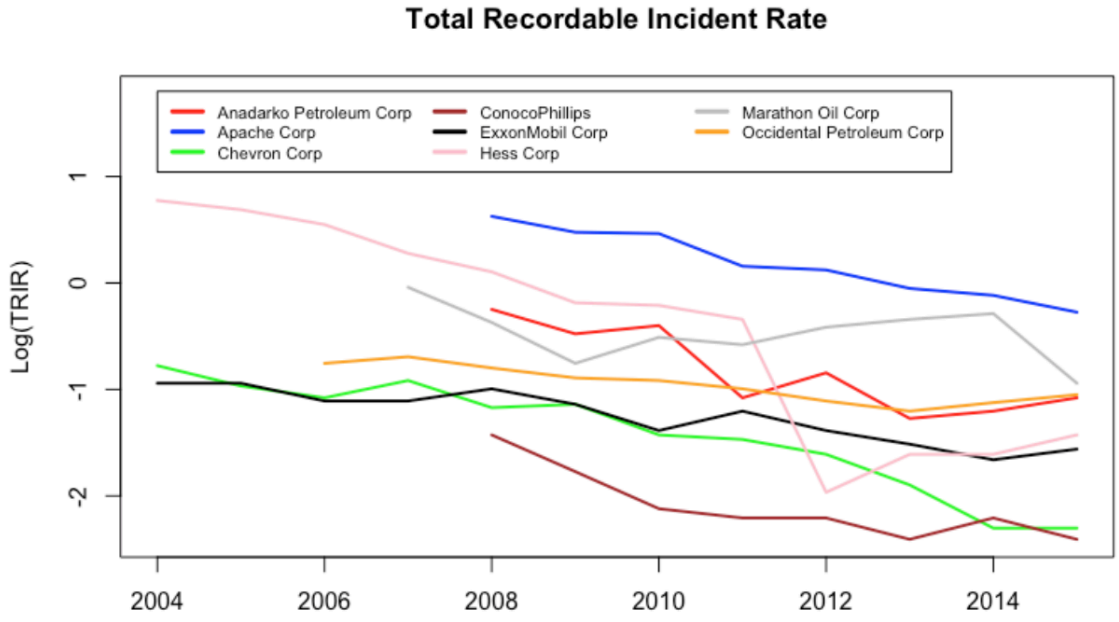


Figure 12 - The Logarithmic TRIR vs. Years

Visually, one can argue for a linear time-trend in figure 12, which implies a percent-wise development for the *TRIR*. However, year-specific shocks in the *TRIR* can for instance be explained by safety measures imposed by government regulation and/or automatization of hazardous processes across the industry. Moreover, a decrease in the oil price can increase the incident rate, in the sense that petroleum companies tend to cut costs in health and safety measures (Marsh, 2016) in such an event. Due to a sharp oil price decrease, both in 2009 and 2015, shocks in the *TRIR* might be present in our data. Therefore, in order to accommodate

potential time-specific variation in our data, eleven year-specific dummies are added to our regression.

Furthermore, the visual level differences in figure 12 indicate firm-specific variation. Differing focus on technology might also yield differences across firms. For example, installation of newer and safer machines among some companies are likely to reduce the *TRIR*. Individual firms can also have different internal safety policies and training of their employees in order to avoid incidents. The firm's position, both horizontally and vertically, can affect their risk of incidents. For example, onshore production leads to greater risk exposure compared to offshore production (Oil & Gas Journal, 2014). Moreover, firm size matter if there is a non-linear relationship between incidents and number of employees. For instance, if the relative number of employees working with risky drilling operations decreases as the firm grows. Hence, we adjust for firm-specific effects by adding seven firm dummies.

8.1 Results

In the following regressions, we include eleven year dummies and seven firm dummies as control variables. The consequence of adding the dummies is that variation due to firm- or time-specific effects are accounted for. However, the dummies may consume variation in the textual variables. Thus, as in the environmental dimension, the regressions with time- and firm-fixed effects below are also compared to the regressions without the control variables (Appendix 12.8).

	Dependent variable:							
	ln(Total Recordable Incident Rate)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relevant Words' Fraction 10-K	50.006						26.547	
	(52.165)						(52.763)	
Relevant Words' Fraction 8-K		56.768					76.491	
		(80.323)					(78.305)	
Sustainability Report Dummy			0.208*				0.157	
			(0.111)				(0.117)	
Sentiment of Sustainability-Related Sentences in 10-K				15.629**			11.742	
				(7.453)			(8.180)	
Sentiment of Sustainability-Related Sentences in 8-K					-0.192			
					(1.182)			
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-1.452		
						(9.637)		
Observations	79	79	79	79	39	55	79	79
Adjusted R^2	0.824	0.823	0.832	0.834	0.939	0.797	0.833	0.825
Note:							*p<0.1; **p<0.05; ***p<0.01	

Figure 13 - Regression Table for the Social Dimension

Control variables are not shown in the regression table for readability purposes. The table consists of eight regressions in which the coefficients estimates are presented, significance level indicated by (*) and standard deviations in parentheses. We still use a 10 % significance level to determine significance. We omit the variables *Sentiment of Sustainability-Related Sentences in 8-K* and *Sentiment of Sustainability-Related Sentences in Sustainability Reports* in the multiple regression (7).

None of the coefficients in the multiple regression (7) are significantly different from zero. If *Relevant Word's Fraction 10-K* increases with 0.1 percentage points, the *TRIR* increases with 2.65 %, all else equal. If *Relevant Word's Fraction 8-K* increases with 0.1 percentage

points, the *TRIR* increases with 7.65 %. Furthermore, if a company file a sustainability report, the *TRIR* increases with 15.7 %. Finally, if *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *TRIR* increases with 11.7 %. The additional adjusted R^2, gained by going from the control variables regression (8) to the multiple regression (7), is 0.8 percentage points. This positive development suggests that the additional information the textual variables give the model, outweighs the loss in efficiency. However, we conclude that we cannot reject the null hypothesis that the textual variables are not correlated with the *TRIR* after adjusting for firm- and time-fixed effects.

For the multiple regression (7) without time- and firm-fixed effects (Appendix 12.8), *Relevant Word's Fraction 10-K* and *Relevant Word's Fraction 8-K* are significantly different from zero. These variables lose significance after the inclusion of controls. Thus, we cannot neglect the possibility that their insignificance in the regression with control variables can be a result of consumed variation by the dummies. Regarding the variables *Sustainability Report Dummy* and *Sentiment of Sustainability-Related Sentences in 10-K,* these are insignificant in both the regressions with and without control variables. Therefore, our conclusion for these two is more solid.

The positive coefficient of *Relevant Words' Fraction 10-K* in regression (1) indicates that they write more about sustainability-related issues in the annual reports when the *TRIR* increases, and vice versa. If *Relevant Words' Fraction 10-K* increases with 0.1 percentage points, the *TRIR* increases with 5.0 %, all else equal. Moreover, the coefficient is not significantly different from zero. Hence, we have statistical basis to conclude there is no correlation between the frequency of sustainability-related words in annual reports and the *TRIR* when time- and firm-fixed effects are adjusted for. However, when we compute regression (1) without time- and firm-fixed effects (Appendix 12.8), the coefficient is significantly different from zero. Thus, we cannot neglect that time- and firm specific effects consume the variation, and thus make *Relevant Words' Fraction 10-K* insignificant.

If *Relevant Words' Fraction 8-K* in regression (2) increases with 0.1 percentage points, the *TRIR* increases with 5.68 %. The coefficient of *Relevant Words' Fraction 8-K in* regression (2) is not significant, implying we have the statistical foundation to conclude there is no correlation between the frequency of sustainability-related words in the 8-Ks and the *TRIR*. However, when we compute regression (2) without time- and firm-fixed effects (Appendix

12.8), the coefficient is significant. Hence, the firm and year dummies might consume variation in the variable, which can disguise whether the textual variable and *TRIR* actually are correlated.

The coefficient of *Sustainability Report Dummy* from regression (3) supports a positive correlation, which implies that if the companies publish a sustainability report, the *TRIR* increases. If a company file a sustainability report, the *TRIR* increases with 20.8 %, all else equal. In addition, the coefficient is significantly different from zero. Hence, we conclude there is correlation between publishing a sustainability report and their *TRIR*. However, when regression (3) is conducted without time- and firm-fixed effects (Appendix 12.8), the coefficient is not significant. This may indicate that initially, some variation in the excluded dummies is reflected in the *Sustainability Report Dummy.* When the time- and firm specific effects are included, the dummies might isolate a more valid correlation between the *Sustainability Report Dummy* and *TRIR*.

Furthermore, the coefficient of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) indicates that if they write more positively about sustainability-related issues in their annual reports, the *TRIR* grows. If *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *TRIR* increases with 15.6 %. Moreover, the coefficient of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) is significantly different from zero, which suggests that we can conclude there is correlation between the variable and the *TRIR*. However, for regression (4) without time- and firm-fixed effects (Appendix 12.8), the coefficient is not significant. Thus, we probably have a case of confounding, which means that the fixed effects are correlated with both *TRIR* and the *Sustainability Report Dummy.* When the dummies are included, they may reveal a more legitimate correlation between the *Sustainability Report Dummy* and *TRIR*.

The coefficient of *Sentiment of Sustainability-Related Sentences in 8-K* in regression (5) indicates a negative correlation, meaning that if they write more positively about sustainability-related issues in their event reports, the *TRIR* decreases. If *Sentiment of Sustainability-Related Sentences in 8-K* increases with 0.01, the *TRIR* decreases with 0.19 %. However, we can not reject the null hypothesis stating there is no correlation between the *Sentiment of Sustainability-Related Sentences in 8-K* and the *TRIR*. However, when we calculate regression (5) without time- and firm-fixed effects (Appendix 12.8), the coefficient

is significantly different from zero. Therefore, we cannot neglect that the dummies consume some of the variation. In addition, this regression also has relatively few observations. Thus, collinearity and missing observations might be an issue when determining whether the variable is correlated.

Regarding regression (6), the coefficient of *Sentiment of Sustainability-Related Sentences in Sustainability Reports* indicates a negative correlation. The negative correlation suggests that if they write more positively about sustainability-related issues in their sustainability reports, the *TRIR* decreases. If *Sentiment of Sustainability-Related Sentences in Sustainability Reports* increases with 0.01, the *TRIR* decreases with 1.45 %, all else equal. However, we can not reject the null hypothesis stating there is no correlation between the *Sentiment of Sustainability-Related Sentences in Sustainability Reports* and the *TRIR*. Moreover, in regression (6) without time- and firm-fixed effects (Appendix 12.8), the coefficient is not significantly different from zero. This suggests that collinearity is not an issue for whether the coefficient is correlated. Nevertheless, many missing observations reduces the credibility of our conclusion.

9. The Economic Dimension

GRI (2011) defines the economic dimension of sustainability as *"the organization's impacts on the economic conditions of its stakeholders, and on economic systems at local, national, and global levels"*. In the same report, they describe the economic indicators as the *"flow of capital among different stakeholders",* among others. By regarding government as a stakeholder, the environmental expenditures can be considered as a flow of capital between the firms and the government. Therefore, to represent the economic dimension of sustainability, we regard the environmental expenditures.

	Mean	Standard Deviation	Missing Observations
Anadarko Petroleum Corp	99.08	32.39	0
Apache Corp	57.73	44.80	0
Chevron Corp	2525.00	691.67	0
ConocoPhillips	785.58	341.29	0
ExxonMobil Corp	4689.67	1139.97	0
Hess Corp	13.92	2.91	0
Marathon Oil Corp	677.57	129.71	5
Occidental Petroleum Corp	59.75	29.46	0

Figure 14 - Descriptive Statistics of the Enironmental Expenditures for the Years 2004 - 2015

The range of the mean is 4675.75 million USD, which indicates large level differences between the firms. The only company with missing observations is Marathon Oil, with five values missing.

When we collect the environmental expenditures from the annual reports, we observe that the companies have deviating formulations. For instance, "*accrued liabilities for remediation*" (Marathon petroleum's 10-K), "*accrued environmental costs*" (ConocoPhillips' 10-K) and "*reserve for estimated remediation liabilities*" (Hess' 10-K). These formulations are a concern if they include different measures when calculating the environmental expenditures. In addition, Polk (2009) states that: *"environmental costs and liabilities can take various forms, the key facts are often difficult to ascertain and the underlying environmental laws (and their enforcement) are constantly changing"*. This

implies that environmental expenditures are not clearly defined and can be subject to different interpretations, which in turn can explain the different wording used among the companies. To account for potential deviating interpretations, environmental expenditures is standardized by deducting the mean and dividing by the standard deviation, for each firm over time. Hence, the resulting variable will have a mean of zero and a standard deviation of one. Assuming that each company's interpretation of the environmental expenditures remains constant, we have accounted for the different descriptions.

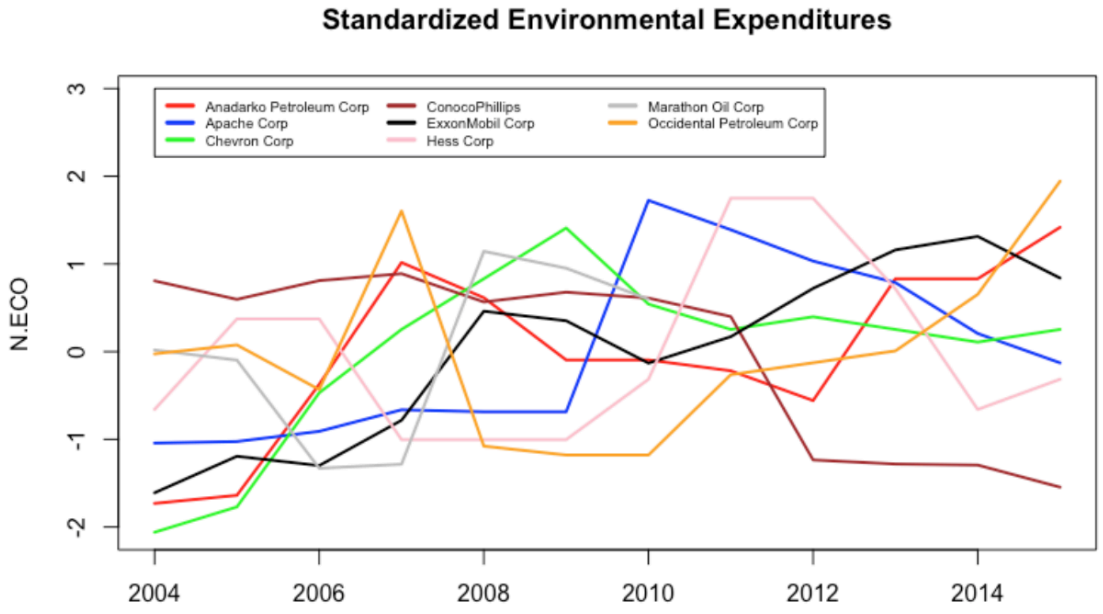


Figure 15 - Standardized Environmental Expenditures vs. Years

The figure above shows the standardized *Environmental Expenditures* versus the years 2004 - 2015. Visually, it is hard to comment on time- and firm-specific variation. Nevertheless, the funding of U.S. Environmental Protection Agency (EPA) over various years can yield year-specific variation. For instance, if EPA's funding increases, more environmental incidents are addressed and hence the increase in the companies' *Environmental Expenditures* (Savage, 2017). To remedy the issue with time-specific variation, we include year dummies in the regressions.

The standardization of *Environmental Expenditures* adjusts for deviations across firms in the dependent variable. However, it does not account for firm-specific differences in the independent variables. The written disclosure on sustainability may vary across firms due to

cultural differences, author's writing style or deviating focus. Moreover, the plot in Appendix 12.10 indicates level differences between firms in the independent variables. Therefore, we cannot neglect that the differing focus across firms is reflected in the textual variables. Thus, we include firm dummies, since these also account for differences in the independent variables.

9.1 Results

Eleven year-specific dummy variables and seven firm-specific dummies are included in the following regressions. The potential consequences of adding dummies are addressed in the environmental dimension.

	Dependent variable:							
	Standardized Environmental Expenditures							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Relevant Words' Fraction 10-K	55.077						10.779	
	(125.728)						(137.309)	
Relevant Words' Fraction 8-K		-51.451					-21.853	
		(223.767)					(231.450)	
Sustainability Report Dummy			-0.006				-0.073	
			(0.306)				(0.320)	
Sentiment of Sustainability-Related Sentences in 10-K				20.270			20.463	
				(20.571)			(23.029)	
Sentiment of Sustainability-Related Sentences in 8-K					-7.230			
					(5.328)			
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-35.481		
						(24.485)		
Observations	91	91	91	91	47	52	91	91
Adjusted R^2	-0.045	-0.047	-0.047	-0.033	-0.148	0.191	-0.078	-0.033
Note:								*p<0.1; **p<0.05; ***p<0.01

Figure 16 - Regression Table for the Economic Dimension

Control variables are not shown in the regression table for readability purposes. The table consists of eight regressions in which the coefficients estimates of the textual variables are presented, significance level indicated by (*) and standard deviations in parentheses. We use a significance level of 10 %. The variables *Sustainability-Related Sentences in 8-K* and *Sentiment of Sustainability-Related Sentences in Sustainability Reports* are omitted from regression (7) on the same argumentation as in the other dimensions.

None of the coefficients in the multiple regression (7) are significantly different from zero. If *Relevant Word's Fraction 10-K* increases with 0.1 percentage points, the *Environmental Expenditures* increases with 0.01 standard deviations, all else equal. If *Relevant Word's*

Fraction 8-K increases with 0.1 percentage points, the *Environmental Expenditures* decreases with 0.02 standard deviations. Furthermore, if a company file a sustainability report, the *Environmental Expenditures* decreases with 0.07 standard deviations. Finally, if *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *Environmental Expenditures* increases with 0.2 standard deviations. The additional adjusted R^2, calculated by going from the control variables regression (8) to the multiple regression (7), is -4.5 percentage points. This negative development suggests that the additional information the textual variables give the model, does not outweigh the loss in efficiency. To conclude, we cannot reject the null hypothesis stating the textual variables are not correlated with the *Environmental Expenditures* after adjusting for firm- and time-fixed effects.

When we compute regression (7) without time- and firm-fixed effects (Appendix 12.9), only *Relevant Words' Fraction 10-K* is significantly different from zero. This variable loses significance after the inclusion of controls. Therefore, we cannot neglect the possibility that its insignificance can be a result of consumed variation by the dummies. The variables *Relevant Word's Fraction 10-K, Relevant Word's Fraction 8-K* and *Sustainability Report Dummy* were insignificant in both the regressions with and without control variables. Hence, our conclusion of no significance for these variables is strengthened.

The coefficient of *Relevant Words' Fraction 10-K* in regression (1) indicates that when the companies write more about sustainability-related issues in the annual reports, the *Environmental Expenditures* increases, and vice versa. If *Relevant Words' Fraction 10-K* increases with 0.1 percentage points, the *Environmental Expenditures* increases with 0.06 standard deviations, all else equal. Nevertheless, the coefficient is not significant. Hence, we have the statistical basis to conclude there is no correlation between the frequency of sustainability-related words in annual reports and the *Environmental Expenditures*. However, when we calculate regression (1) without time- and firm-fixed effects (Appendix 12.9), the coefficient is significant. Thus, we cannot neglect that time- and firm specific effects consume variation, and thus makes *Relevant Words' Fraction 10-K* insignificant.

If *Relevant Words' Fraction 8-K* in regression (2) increases with 0.1 percentage points, the *Environmental Expenditures* decreases 0.05 standard deviations, all else equal. The coefficient of *Relevant Words' Fraction 8-K in* regression (2) is not significant, implying we

have the statistical foundation to conclude there is no correlation between the frequency of sustainability-related words in the 8-Ks and the *Environmental Expenditures*. In addition, when we compute regression (2) without time- and firm-fixed effects (Appendix 12.9), the coefficient is also insignificant. Hence, our conclusion is more solid.

The coefficient of the variable *Sustainability Report Dummy* from regression (3) supports a negative correlation, implying that if they publish a sustainability report, their *Environmental Expenditures* decreases. If a company file a sustainability report, the *Environmental Expenditures* decreases with 0.006 standard deviations, all else equal. However, the coefficient is not significant. Hence, we have statistical foundation to conclude there is no correlation between publishing a sustainability report and their *Environmental Expenditures*. In addition, when regression (3) is conducted without time- and firm-fixed effects (Appendix 12.9), the coefficient is also insignificant, which strengthens the conclusion.

Furthermore, the coefficient of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) indicates that if they write more positively about sustainability-related issues in their annual reports, the *Environmental Expenditures* grows. If *Sentiment of Sustainability-Related Sentences in 10-K* increases with 0.01, the *Environmental Expenditures* increases with 0.2 standard deviations. Moreover, the coefficient of *Sentiment of Sustainability-Related Sentences in 10-K* in regression (4) is not significantly different from zero, which suggests we have the statistical foundation to state there is no correlation between this variable and the *Environmental Expenditures*. Furthermore, when regression (4) is conducted without time- and firm-fixed effects (Appendix 12.9), the coefficient is also insignificant. Thus, the statement of no significance is more solid.

The coefficient of *Sentiment of Sustainability-Related Sentences in 8-K* in regression (5) indicates a negative correlation, meaning that if they write more positively about sustainability-related issues in their event reports, the *Environmental Expenditures* decreases. If *Sentiment of Sustainability-Related Sentences in 8-K* increases with 0.01, the *Environmental Expenditures* decreases with 0.07 standard deviations. However, we can not reject a null hypothesis stating there is no correlation between the *Sentiment of Sustainability-Related Sentences in 8-K* and the *Environmental Expenditures*. In addition, when computing regression (5) without time- and firm-fixed effects (Appendix 12.9), the

coefficient is also insignificant, which makes our conclusion more robust. Yet, relatively many missing observations may yield biased results, and thus make inference uncertain.

Regarding regression (6), the coefficient of *Sentiment of Sustainability-Related Sentences in Sustainability Reports* indicates a negative correlation. The negative correlation suggests that if they write more positively about sustainability-related issues in their sustainability reports, the *Environmental Expenditures* decreases. If *Sentiment of Sustainability-Related Sentences in Sustainability Reports* increases with 0.01, the *Environmental Expenditures* decreases with 0.35 standard deviations, all else equal. Finally, we cannot reject the null hypothesis, hence we have the statistical foundation to conclude that *Sentiment of Sustainability-Related Sentences in Sustainability reports* and the *Environmental Expenditures* are not correlated. In addition, when regression (6) is calculated without firm-and year specific effects, its coefficient is also insignificant, which strengthens our conclusion. Nevertheless, the relatively many missing observations reduces the credibility of our conclusion for this variable.

10. Final Thoughts

10.1 Conclusion

The goal of this thesis is to investigate how much the companies' written disclosure on sustainability correlates with their performance regarding a sustainable development. All the three regressions with four textual variables have no correlation with the textual variables and the sustainability performance. Furthermore, in the environmental, social and economic dimension, the additional adjusted R^2 is -0.6 %, 0.8 % and -4.5 %, respectively. These arguments suggest that the textual variables combined can barely describe anything of the companies' sustainability performance.

When analysing the textual variables individually in the environmental dimension, the only statistically significant coefficient is *Sentiment of Sustainability-Related Sentences in Sustainability Reports.* Regarding the social dimension, *Sustainability Report Dummy* and *Sentiment of Sustainability-Related Sentences in 10-K* are the only significant coefficients. In terms of the economic dimension, there are no significant coefficients. None of the frequency variables (*Relevant Words' Fraction in 10-K and Relevant Words' Fraction in 8-K)* are correlated with any of the sustainability performance indicators. Therefore, it appears that how much they write about sustainability correlates less with their performance compared to when they write positive or negative. Furthermore, considering that no variables derived from the 8-K reports are correlated with the performance indicators, we argue that these documents are not a suited reflection of the companies' sustainability performance.

Whether the petroleum companies perform accordingly to their written disclosure raises the question of integrity. The border between hypocrisy and integrity is not clear, but a greyscale. While there are many limitations to our thesis, most of the results speaks in favor of our hypothesis. Thus, our findings indicate that the eight U.S. petroleum companies are tilted towards hypocrisy at the grayscale.

10.2 Limitations

The way we choose to measure the sustainability dimensions does not necessarily correspond with the companies' efforts to address a sustainable development. For instance, if a given petroleum company choose to focus on the social dimension of sustainability by another performance indicator than TRIR, our results would not be accurate. If we had included more dependent variables, and conducted several regressions for each dimension, the sustainability performance could have been measured more precisely.

Moreover, there are several ways for a company to write about sustainability. The companies can communicate their written sustainability disclosure through other channels than the documents investigated in this thesis. Examples are the companies' websites, quarterly reports or other documents. Therefore, it is perhaps insufficient to analyse the 10-Ks, 8-Ks and sustainability reports alone.

Regarding the *Relevant Words Vector*, one can argue that it would have been more precise to create a relevant words vector for each dimension of sustainability. For instance, to have a relevant words vector containing words related to the social dimension exclusively, and used this for the regression with TRIR as dependent variable. Doing this for all dimensions, the textual variables of the three different regressions could be more specific for their respective dependent variable. Thus, the variables could possibly explain more of the sustainability performance. However, this would have been difficult. The potential three relevant words vectors would have been constituted by even fewer words, and thus yielded even more missing values. Moreover, some words are used in different contexts between the documents we investigate, and to create a relevant words vector for each document type can improve the results. However, to make relevant words vectors that are document- and dimension specific would have required 9 vectors, which will increase the complexity of our approach greatly. In addition, since we also use a qualitative approach when choosing the relevant words, our subjective opinion regarding sustainability may fail to reflect how the reports address sustainability.

Furthermore, our sentiment analysis is also subject to shortcomings. It is uncertain that the Harvard IV dictionary give a precise quantitative measure of their sustainability sentiment.

Furthermore, taking the netto sentiment and dividing the score by the amount of words in the respective document is not necessarily beneficial for small documents. For instance, this approach can yield unreasonably high positive or negative sentiment scores if there are few relevant sentences. In our case, the sentiment of the 8-Ks, which sometimes have a small number of sustainability-related sentences, can be subject to relatively high absolute values.

As regards measurement errors, the sustainability performance indicators TRIR and environmental expenditures are uncertain. The reports often disclose performance indicators for a 4-5 years period back. The result is that a performance indicator for a given year can be found in several reports. Hence, the ambiguity, since a performance indicator for a given year in a given report can be different in another report. To remedy the issue, we choose to regard the performance indicators from the newest reports. In so doing, potential hindsight changes by the petroleum companies in the indicators is accounted for. However, it does not eliminate the issue of measurement errors completely.

10.3 Future Work

The scope of our thesis can be adjusted to make it applicable for several industries, such as the car or coal industry. A possible modification in the scope of sustainability is to narrow down the analysis to solely investigate greenhouse gas emissions. In so doing, the data foundation for the dependent variable can be expanded significantly. Moreover, our thesis focuses on the companies' negative externalities. It can be interesting to include performance indicators related to positive externalities as well. How do companies make a positive impact on the world? And how do they write and promote that impact?

11. References

Alshehhi, A., Nobanee, H., & Khare, N. (2018). *The Impact of Sustainability Practices on Corporate Financial Performance: Literature Trends and Future Research Potential.*

Aras, G., & Crowther, D. (2008). *Corporate Sustainability Reporting: A Study in Disingenuity?* Springer.

Barkemeyer, R., Comyns, B., Figge, F., & Napolitano, G. (2014). *CEO statements in sustainability reports: Substantive information or background noise?*

BP. (2018). *bp*. Hentet fra https://www.bp.com/en/global/corporate/who-we-are.html

Bright Hub Engineering. (u.d.). Hentet fra Web area for Bright Hub Engineering: https://www.brighthubengineering.com/marine-history/66248-the-exxon-valdez-oil-spill-and-its-consequences-on-the-environment/

Business Dictionary. (u.d.). *BD - Business Dictionary*. Hentet fra http://www.businessdictionary.com/definition/social-system.html

Cai, H., Brandt, A. R., Yeh, S., Englander, J. G., Han, J., Elgowainy, A., & Wang, M. Q. (2015). *Environmental Science and Technology*. Hentet fra https://pubs.acs.org/doi/abs/10.1021/acs.est.5b01255?journalCode=esthag

Carbon Disclosure Project. (2017). *New report shows just 100 companies are source of over 70 % of emissions*. Hentet fra https://www.cdp.net/en/articles/media/new-report-shows-just-100-companies-are-source-of-over-70-of-emissions.

Chevron. (2011). *2011 Corporate Responsibility Report.* Chevron.

Chin, M., Hambrick, D. C., & Trevino, L. K. (2016). *Political Ideologies of CEOs: The Influence of Executives' Values on Corporate Social Responsibility.*

Crowther, D. (2002). *A Social Critique of Corporate Reporting.*

Emam, E. A. (2015). *GAS FLARING IN INDUSTRY: AN OVERVIEW.*

Environmental Studies Program. (2005). *Fish and Wildlife of Alaska's North Slope - Polar Bears.* ConocoPhillips.

Equinor. (2018). *Equinor*. Hentet fra Equinor - About us: https://www.equinor.com/en/about-us.html#statoil-at-a-glance

ExxonMobil. (u.d.). Hentet fra https://corporate.exxonmobil.com/en/current-issues/arctic/community/community-of-respect

ExxonMobil. (2013). *Corporate Citizenship Report 2013.* ExxonMobil.

Fears, D. (2017). *The Washington Post*. Hentet fra https://www.washingtonpost.com/news/energy-environment/wp/2017/01/09/without-action-on-climate-change-say-goodbye-to-polar-bears/?utm_term=.e5d9f9fb3fc1

Forbes. (2017, August 16). *Socially-Responsible Investing: Earn Better Returns From Good Companies*. Hentet fra https://www.forbes.com/sites/moneyshow/2017/08/16/socially-responsible-investing-earn-better-returns-from-good-companies/#38a060a0623d

Gavenas, E., Rosendahl, K. E., & Skjerpen, t. (2015). *CO2-emissions from Norwegian oil and gas extraction.*

Gill, J. (1999). The Insignificance of Null Hypothesis Significance Testing. I *Political Research Quarterly* (ss. 647-674).

Global Reporting Initiative. (2011). *Sustainability Reporting.*

Global Reporting Initiative. (u.d.). *GRI - Empowering Sustainable Decision*. Hentet fra https://www.globalreporting.org/information/about-gri/Pages/default.aspx

Global Reporting Initiative. (u.d.). *GRI - Empowering Sustainable Decisions*. Hentet fra https://www.globalreporting.org/information/sustainability-reporting/Pages/default.aspx

Global Witness. (2018). *Catch me if you can: Exxon complicit in corrupt liberian oil sector*. Hentet fra https://www.globalwitness.org/en/campaigns/oil-gas-and-mining/catch-me-if-you-can-exxon-complicit-corrupt-liberian-oil-sector/.

Griffin, D. P. (2017). *The Carbon Majors Database - CDP Carbon Majors Report 2017.*

Johnston, P., Everard, M., Santillo, D., & Robèrt, K.-H. (2007). *Reclaiming the Definition of Sustainability.*

Kang, H. (2013). *The prevention and handling of the missing data.*

Marsh. (2016). *Can Energy Firms Break the Historical Nexus Between Oil Price Falls and Large Losses?* Hentet fra https://www.marsh.com/content/dam/marsh/Documents/PDF/UK-en/Can%20Energy%20Firms%20Break%20Historical%20Nexus-3-2016.pdf

Milman, O. (2018, June 28). *The Guardian*. Hentet fra https://www.theguardian.com/environment/2018/jun/28/bp-deepwater-horizon-oil-spill-report

National Ocean Service. (2018, June 25). *Is sea level rising?* Hentet fra https://oceanservice.noaa.gov/facts/sealevel.html.

Oil & Gas Journal. (2014, May 5). *Oil & Gas Journal*. Hentet fra https://www.ogj.com/articles/print/volume-112/issue-5/special-report-offshore-petroleum-operations/offshore-drilling-risks-mdash-1-study-risk-indicators-have-varying-impact-on-mitigation.html

Ormiston, M. E., & Wong, E. M. (2013). *License to Ill: The Effects of Corporate Social Responsibility and CEO Moral Identity on Corporate Social Irresponsibility.*

Paul, G., Heede, R., & Flugt, I. v. (2017). *Climate Accountability Institute*. Hentet fra http://climateaccountability.org/publications.html

Polk, D. (2009). *Environmental Disclosure in SEC Filings* .

Popovich, N., & Albeck-Ripka, L. (2017, December 14). *The New York Times*. Hentet fra https://www.nytimes.com/interactive/2017/12/14/climate/republicans-global-warming-maps.html

Savage, C. (2017, September 27). *The New York Times*. Hentet fra https://www.nytimes.com/2017/09/27/us/politics/scott-pruitt-epa-justice-department-funding.html

Shin, A. (2017, March 16). *The Washington Post* . Hentet fra https://www.washingtonpost.com/lifestyle/magazine/the-exxon-valdez-oil-spill/2017/03/14/d131b630-f876-11e6-9845-576c69081518_story.html?noredirect=on&utm_term=.7be17f3b126b

Society of Petroleum Engineers. (2017, November 16). *Road Accidents Top List of Fatality Causes in Oil and Gas Industry for 2014*. Hentet fra https://www.spe.org/en/print-article/?art=3608

The New York Times. (2011, February 14). *Ecuador Judge Orders Chevron to Pay $9Billion*. Hentet fra https://www.nytimes.com/2011/02/15/world/americas/15ecuador.html.

The White House - Office of the Press Secretary. (2012, August 28). *The White House - President Barack Obama*. Hentet fra https://obamawhitehouse.archives.gov/the-press-office/2012/08/28/obama-administration-finalizes-historic-545-mpg-fuel-efficiency-standard

U.S. Securities and Exchange Commision. (u.d.). Hentet fra https://www.sec.gov/fast-answers/answersform8khtm.html

U.S. Securities and Exchange Commision. (2011). *U.S. Securities and Exchange Commision*. Hentet fra https://www.sec.gov/fast-answers/answersreada10khtm.html

United Nations. (u.d.). *United Nations*. Hentet fra http://www.un.org/en/sections/issues-depth/climate-change/

Wen, J. (2014). *A Business Analytics Approach to Corporate Sustainability Analysis.*

12. Appendices

12.1 Comparison of 10-K, 8-K and Sustainability Reports

```
######## SCRIPT TO COMPARE THE DIFFERENT SOURCES #######

# Acquiring the data
load("/Users/theodor_dokkan/Documents/Dokumenter/NHH/Master/Masteroppgave/Dataobjekter/Raw data/sus.pdf.Rdata")
load("/Users/theodor_dokkan/Documents/Dokumenter/NHH/Master/Masteroppgave/Dataobjekter/Raw data/K8.Rdata")
load("/Users/theodor_dokkan/Documents/Dokumenter/NHH/Master/Masteroppgave/Dataobjekter/Raw data/k10 (1).Rdata")

# Requiring packages
require(tm)
```

```
## Loading required package: tm
```

```
## Warning: package 'tm' was built under R version 3.4.3
```

```
## Loading required package: NLP
```

```
##
## FALSE  TRUE
##    75    25
```

```
# Making a dtm of the k8
corpus.sus <- Corpus(VectorSource(sus.pdf))
dtm.sus <- TermDocumentMatrix(corpus.sus, control = list(removePunctuation = T,
                                                         stopwords = T,
                                                         removeNumbers = T,
                                                         wordLengths = c(2, 20),
                                                         tolower = T))

# Converting SUS to a matrix
sus.mat <- as.matrix(dtm.sus)
sus.freq <- rowSums(sus.mat)
sus.freq <- sort(sus.freq, decreasing = TRUE)
sus.freq <- sus.freq[1:100]

# Comparing how equal sus is to 8K
table(names(sus.freq) %in% names(k.8.freq))
```

```
##
## FALSE  TRUE
##    87    13
```

```
# Making dtm of the k10

corpus.k10 <- Corpus(VectorSource(k10))
dtm.k10 <- TermDocumentMatrix(corpus.k10, control = list(removePunctuation = T,
                                                         stopwords = T,
                                                         removeNumbers = T,
                                                         wordLengths = c(2, 20),
                                                         tolower = T))
# Converting to a matrix
k10.mat <- as.matrix(dtm.k10)
k.10.freq <- rowSums(k10.mat)
k.10.freq <- sort(k.10.freq, decreasing = TRUE)
k.10.freq <- k.10.freq[1:100]

# Making a dtm of the k8
corpus.k8 <- Corpus(VectorSource(K8))
dtm.k8 <- TermDocumentMatrix(corpus.k8, control = list(removePunctuation = T,
                                                       stopwords = T,
                                                       removeNumbers = T,
                                                       wordLengths = c(2, 20),
                                                       tolower = T))

# Converting K8 to a matrix
k8.mat <- as.matrix(dtm.k8)
k.8.freq <- rowSums(k8.mat)
k.8.freq <- sort(k.8.freq, decreasing = TRUE)
k.8.freq <- k.8.freq[1:100]

# Comparing how equal 10Ks are to 8Ks
table(names(k.10.freq) %in% names(k.8.freq))
```

Initially, the raw data is loaded. The first, second and third vector contains sustainability reports, 8-K reports and 10-K reports as vector elements, respectively. Next, the required packages to conduct to the analysis is acquired. We subsequently create a term document matrix of the 10-K reports. A term document matrix has the different documents as one dimension, and the various terms in the other dimension. The inputs are naturally the term frequencies in the various documents. Moreover, the row sums of the term document matrix are calculated and made into a vector. Thus, the vector contains all the different terms' frequency in all the 10-K reports combined. This frequency vector is sorted in a decreasing order and trimmed so that only the 100 most frequent terms are shown.

The exact same approach is applied for the 8-K reports to create an equivalent frequency vector. In order to compare the similarities between the 8-K's and 10-K's, we investigate how many terms these two vectors have in common and create a table of the results. The vectors have 25 common words. A frequency vector for the sustainability reports is also created. When we compare the similarity to the 8-K's, we find that the two vectors have 13 common words. Thus, the data supports a higher degree of similarity between the 8-K's and the 10-K's.

12.2 Relevant Words Vector

The code first finds the most common words using the same procedure as explained in Appendix 12.1. An issue with this frequency vector is that inflected forms of the words often occur within the frequent words list. For example, “environment”, “environments” and “environmental”, or “community” and “communities”. These words mean essentially the same but can be displayed as different words in the frequent words vector. Hence, similar words are combined into one vector and the shortest form of the word is kept. The final result is the following vector of the most frequent words in the sustainability reports:

develop	report	operation	employee	environment	business	company	manage	safety	energy
6527	5966	5773	5659	5509	5419	5384	5145	4994	4983
community	program	water	emission	project	process	product	percent	performance	local
4644	4595	4221	4198	3927	3623	3589	3526	3452	3126
global	health	million	provide	corporate	improve	training	include	right	national
2934	2929	2877	2768	2738	2669	2569	2553	2501	2495
industry	human	social	source	support	supply	impact	stakeholder	state	system
2489	2417	2403	2401	2393	2299	2259	2178	2137	2135
issue	contract	policy	world	including	ability	natural	standard	operate	continue
2084	2083	2060	2011	1922	1920	1914	1882	1829	1829
activity	reduce	conduct	assess	technology	country	change	spill	implement	workforce
1797	1746	1743	1743	1706	1706	1688	1653	1651	1647
facility	incident	economic	partner	government	opportunity	total	asset	information	based
1646	1628	1600	1600	1574	1562	1561	1560	1559	1503
addition	initiative	climate	increase	security	direct	diversity	engagement	risks	focus
1498	1487	1459	1418	1416	1405	1398	1394	1394	1374
practices	effort	commitment	carbon	education	identify	level	united	investment	responsibility
1339	1333	1309	1305	1301	1299	1299	1294	1291	1288
value	action	result	potential	approach	sustainable	areas	metric	protect	group
1285	1256	1251	1225	1225	1211	1192	1189	1184	1183
group	leaders	revyw	years	approximately	address	build	people	public	response
1183	1180	1147	1144	1117	1116	1110	1108	1106	1096
significant	related	women	working	efficyncy	strategy	board	north	ensure	place
1096	1028	1015	1008	1002	998	996	993	984	979
petroleum	plans	annual	operating	services	waste	compliance	understand	market	share
977	951	949	934	932	929	928	927	926	923
reduction	number	principles	member	greenhouse	future	citizenship	challenge	design	drilling
916	909	898	889	888	869	861	860	857	845
fyld	organizations	since	effective	benefit	first	requirements	chemical	guinea	research
843	839	836	832	826	821	816	801	797	796
integrity	across	corporation	responsible	exploration	association	school	america	study	growth
792	789	788	773	765	759	757	746	745	743
voluntary	quality	governance	early	major	enhance	needs	committee	example	produced
740	736	734	731	718	716	710	706	696	690
flaring	contributions	conservation	meeting	assurance	three	committed	progress	available	ipyca
688	683	681	680	678	675	669	665	664	664
managing	expectations	current	learn	around	within	large	barrels	interest	material
656	654	653	653	652	650	650	649	644	642
internal	tonnes	equatorial	participate	organization	transparency	technical	emergency	billion	political
639	636	634	633	631	630	629	628	621	621
planning									
620									

Figure 17 - The Most Frequent Words in the Sustainability Reports

Note that a few words are still in their inflected forms in this vector due to very low frequency of the indefinite form of the word. This is due to the initial cleaning of the vector, which removes words with less than 200 in frequency. The script has its limitations by for example not being able to handle “ing”-endings, and two places within the 200 most frequent words it has captured a word without meaning, “ipyca” and “fyld”.

12.3 Marathon Oil

TRIR.Pop is the annual mean TRIR across all firms except Marathon Oil, while TRIR.Marathon is the TRIR for Marathon Oil.

```
        Shapiro-Wilk normality test

data:  TRIR.Pop
W = 0.90009, p-value = 0.159
```

```
        Shapiro-Wilk normality test

data:  TRIR.Marathon
W = 0.97687, p-value = 0.9463
```

Figure 18 - Normality Tests Related to Marathon Oil

The test yields that we can not reject a null hypothesis stating the two populations are normally distributed. Therefore, an F-test can be conducted to investigate whether there is a difference in variation between the two populations.

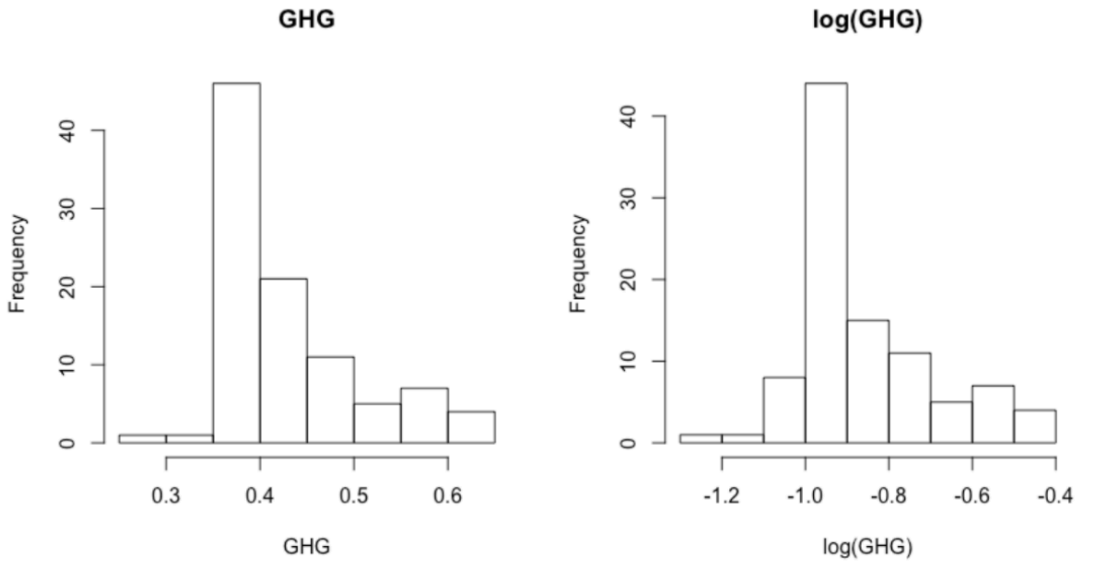
```
        F test to compare two variances

data:  TRIR.Pop and TRIR.Marathon
F = 2.5176, num df = 11, denom df = 8, p-value = 0.2002
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5933033 9.2241439
sample estimates:
ratio of variances
          2.517631
```

Figure 19 - F test to Compare two Variances

The result indicates that we can not reject the null hypothesis for no differences in variation between the two populations. Thus, the tests yield that there is no potential problem of deviating variances, and Marathon Oil can be included without further issues.

12.4 Greenhouse Gas Emissions




```
        Shapiro-Wilk normality test

data:  df$GHG
W = 0.86279, p-value = 5.875e-08
```

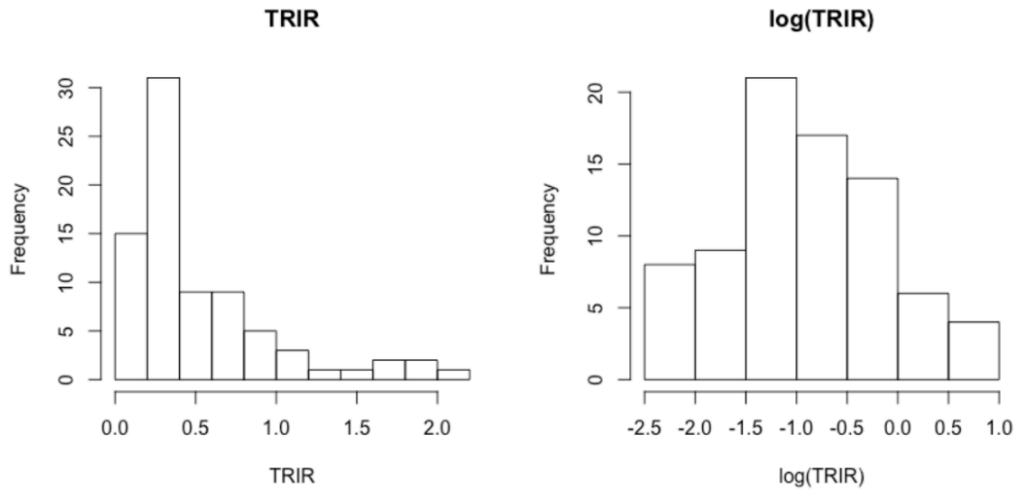
```
        Shapiro-Wilk normality test

data:  log(df$GHG)
W = 0.90662, p-value = 4.385e-06
```

Figure 20 - Normality Test for Greenhouse Gas Emissions

Both tests for normal distribution of GHG rejects null hypothesis that the variable is normally distributed.

12.5 Total Recordable Incident Rate




```
        Shapiro-Wilk normality test

data:  df$TRIR
W = 0.7905, p-value = 3.042e-09
```

```
        Shapiro-Wilk normality test

data:  log(df$TRIR)
W = 0.97869, p-value = 0.209
```

Figure 21 - Normality Test for TRIR

We see from the level plot and the Shapiro-Wilk normality test of TRIR that we can reject the null hypothesis of a normal distribution. After logging, the log(TRIR) plot indicates a normal distribution and the corresponding normality test states that we can not reject a null hypothesis of a normal distribution.

12.6 Standardized Environmental Expenditures

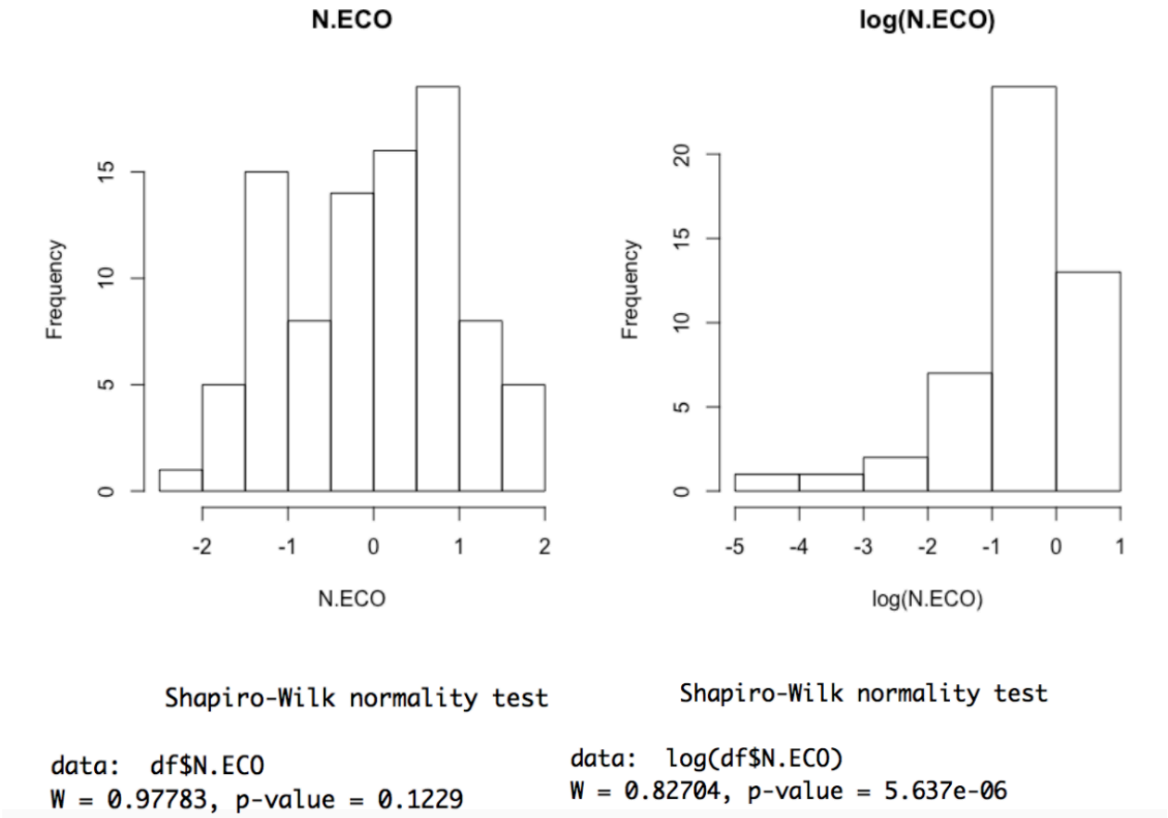


Figure 22 - Normality Test for the Environmental Expenditures

The results indicate that we cannot reject the null hypothesis of a normal distribution of the level standardized environmental expenditures, while we can reject a normal distribution when logging the variable.

12.7 Greenhouse Gas Emission Rate without Time- and Firm-Fixed Effects

	Dependent variable:						
	ln(Greenhouse Gas Emissions/Production)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relevant Words' Fraction 10-K	-10.452**						-8.347
	(4.819)						(5.073)
Relevant Words' Fraction 8-K		6.210					3.506
		(12.080)					(12.496)
Sustainability Report Dummy			-0.022				-0.008
			(0.015)				(0.016)
Sentiment of Sustainability-Related Sentences in 10-K				-1.520**			-1.218*
				(0.598)			(0.651)
Sentiment of Sustainability-Related Sentences in 8-K					0.403		
					(0.271)		
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-2.664*	
						(1.353)	
Observations	96	96	96	96	47	57	96
Adjusted R^2	0.038	-0.008	0.011	0.054	0.026	0.049	0.060
Note:							*p<0.1; **p<0.05; ***p<0.01

Figure 23 - Greenhouse Gas Emission Rate without Time- and Firm-Fixed Effects

12.8 Total Recordable Incident Rate without Time- and Firm-Fixed Effects

	Dependent variable:						
	ln(Total Recordable Incident Rate)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relevant Words' Fraction 10-K	-199.297***						-161.578***
	(55.223)						(51.814)
Relevant Words' Fraction 8-K		-496.986***					-503.783***
		(123.046)					(124.274)
Sustainability Report Dummy			-0.237				-0.109
			(0.191)				(0.170)
Sentiment of Sustainability-Related Sentences in 10-K				-2.136			-8.351
				(7.280)			(6.847)
Sentiment of Sustainability-Related Sentences in 8-K					8.398***		
					(2.353)		
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-6.559	
						(17.386)	
Observations	79	79	79	79	39	55	79
Adjusted R^2	0.134	0.164	0.007	-0.012	0.236	-0.016	0.271
Note:							*p<0.1; **p<0.05; ***p<0.01

Figure 24 - Total Recordable Incident Rate without Time- and Firm-Fixed Effects

12.9 Standardized Environmental Expenditures without Time- and Firm-Fixed Effects

	Dependent variable:						
	Standardized Environmental Expenditures						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relevant Words' Fraction 10-K	139.403**						125.607*
	(65.043)						(69.437)
Relevant Words' Fraction 8-K		192.399					66.727
		(160.788)					(171.352)
Sustainability Report Dummy			0.290				0.235
			(0.202)				(0.211)
Sentiment of Sustainability-Related Sentences in 10-K				-4.775			-8.182
				(8.222)			(8.738)
Sentiment of Sustainability-Related Sentences in 8-K					-2.385		
					(2.931)		
Sentiment of Sustainability-Related Sentences in Sustainability Reports						-21.007	
						(20.305)	
Observations	91	91	91	91	47	52	91
Adjusted R^2	0.038	0.005	0.012	-0.007	-0.007	0.001	0.031
Note:							*p<0.1; **p<0.05; ***p<0.01

Figure 25 - Standardized Environmental Expenditures without Time- and Firm-Fixed Effects

12.10 Level Differences in the Independent Variables

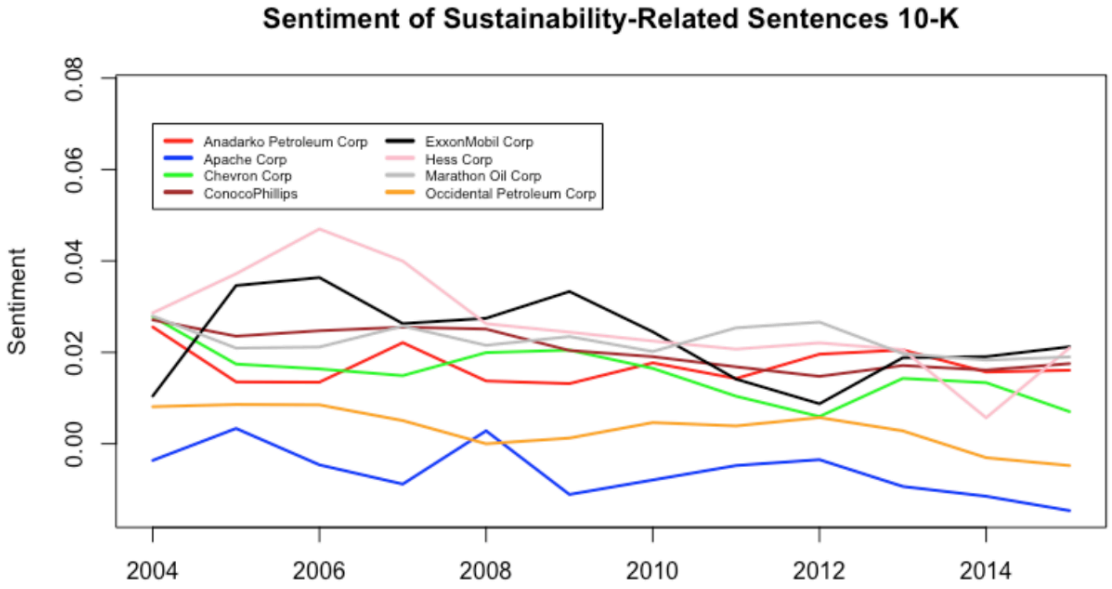


Figure 26 - Level Differences in Sentiment of Sustainability-Related Sentences 10-K

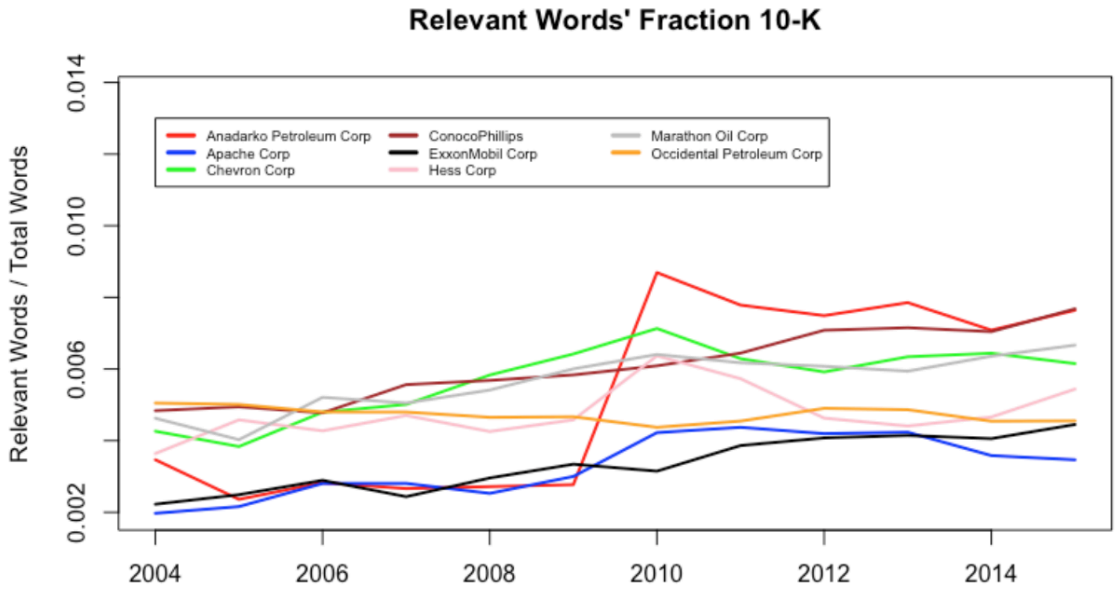


Figure 27 - Level Differences in Relevant Words' Fraction 10-K

12.11 Control Variables in the Environmental Dimension

```
Call:
lm(formula = df$GHG ~ df$Year + df$Firm)

Residuals:
      Min        1Q    Median        3Q       Max
-0.163588 -0.032463 -0.009753  0.036292  0.188346

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                     0.4362423  0.0271746  16.053  < 2e-16 ***
df$Year2005                     0.0110965  0.0305416   0.363  0.71736
df$Year2006                    -0.0595118  0.0305416  -1.949  0.05499 .
df$Year2007                    -0.0386176  0.0305416  -1.264  0.20989
df$Year2008                    -0.0340288  0.0305416  -1.114  0.26867
df$Year2009                    -0.0238637  0.0305416  -0.781  0.43699
df$Year2010                    -0.0180884  0.0305416  -0.592  0.55541
df$Year2011                    -0.0394549  0.0305416  -1.292  0.20028
df$Year2012                    -0.0339315  0.0305416  -1.111  0.27003
df$Year2013                    -0.0373037  0.0305416  -1.221  0.22566
df$Year2014                    -0.0636784  0.0305416  -2.085  0.04039 *
df$Year2015                    -0.0900776  0.0305416  -2.949  0.00422 **
df$FirmApache Corp              0.1064931  0.0249371   4.270 5.51e-05 ***
df$FirmChevron Corp             0.0077232  0.0249371   0.310  0.75762
df$FirmConocoPhillips           0.0811439  0.0249371   3.254  0.00169 **
df$FirmExxonMobil Corp         -0.0005278  0.0249371  -0.021  0.98317
df$FirmHess Corp               -0.0029911  0.0249371  -0.120  0.90484
df$FirmMarathon Oil Corp        0.0065671  0.0249371   0.263  0.79299
df$FirmOccidental Petroleum Corp  0.0429241  0.0249371   1.721  0.08921 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 28 - Control Variables in the Environmental Dimension

12.12 Control Variables in the Social Dimension

```
Call:
lm(formula = df$TRIR ~ df$Firm + df$Year)

Residuals:
     Min       1Q   Median       3Q      Max
-1.14589 -0.16975  0.03239  0.18464  0.58479

Coefficients:
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                       0.14270    0.23087   0.618  0.53885
df$FirmApache Corp                1.00105    0.16371   6.115 7.89e-08 ***
df$FirmChevron Corp              -0.85194    0.15338  -5.555 6.72e-07 ***
df$FirmConocoPhillips            -1.26894    0.16371  -7.751 1.30e-10 ***
df$FirmExxonMobil Corp           -0.67594    0.15338  -4.407 4.40e-05 ***
df$FirmHess Corp                  0.15636    0.15338   1.019  0.31207
df$FirmMarathon Oil Corp          0.28778    0.16003   1.798  0.07716 .
df$FirmOccidental Petroleum Corp -0.25073    0.15721  -1.595  0.11600
df$Year2005                      -0.09255    0.26733  -0.346  0.73041
df$Year2006                      -0.33573    0.25185  -1.333  0.18754
df$Year2007                      -0.37207    0.24220  -1.536  0.12975
df$Year2008                      -0.47744    0.22722  -2.101  0.03983 *
df$Year2009                      -0.67810    0.22722  -2.984  0.00411 **
df$Year2010                      -0.75594    0.22722  -3.327  0.00150 **
df$Year2011                      -0.90732    0.22722  -3.993  0.00018 ***
df$Year2012                      -1.11929    0.22722  -4.926 6.92e-06 ***
df$Year2013                      -1.22983    0.22722  -5.413 1.15e-06 ***
df$Year2014                      -1.25642    0.22722  -5.530 7.38e-07 ***
df$Year2015                      -1.32278    0.22722  -5.822 2.43e-07 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 29 - Control Variables in the Social Dimension

12.13 Control Variables in the Economic Dimension

```
Call:
lm(formula = df$N.ECO ~ df$Year + df$Firm)

Residuals:
     Min       1Q   Median       3Q      Max
-1.89778 -0.73645  0.04104  0.67802  1.62995

Coefficients:
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                     -8.175e-01  4.354e-01  -1.878   0.0645 .
df$Year2005                      2.030e-01  4.880e-01   0.416   0.6787
df$Year2006                      3.340e-01  4.880e-01   0.684   0.4959
df$Year2007                      7.915e-01  4.880e-01   1.622   0.1092
df$Year2008                      8.944e-01  4.880e-01   1.833   0.0710 .
df$Year2009                      8.414e-01  4.880e-01   1.724   0.0890 .
df$Year2010                      1.007e+00  4.880e-01   2.064   0.0426 *
df$Year2011                      1.315e+00  5.075e-01   2.590   0.0116 *
df$Year2012                      1.100e+00  5.075e-01   2.168   0.0335 *
df$Year2013                      1.171e+00  5.075e-01   2.307   0.0239 *
df$Year2014                      9.833e-01  5.075e-01   1.938   0.0566 .
df$Year2015                      1.169e+00  5.075e-01   2.304   0.0241 *
df$FirmApache Corp              -6.942e-16  3.984e-01   0.000   1.0000
df$FirmChevron Corp             -5.830e-16  3.984e-01   0.000   1.0000
df$FirmConocoPhillips           -6.774e-16  3.984e-01   0.000   1.0000
df$FirmExxonMobil Corp          -7.481e-16  3.984e-01   0.000   1.0000
df$FirmHess Corp                -4.821e-16  3.984e-01   0.000   1.0000
df$FirmMarathon Oil Corp         2.359e-01  4.728e-01   0.499   0.6194
df$FirmOccidental Petroleum Corp -6.345e-16  3.984e-01   0.000   1.0000
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Figure 30 - Control Variables in the Economic Dimension