Environmental sentiment and stock performance in the U.S. energy sector

By Xue Wang

Marcelo Bianconi, Advisor

A thesis submitted in partial fulfillment

of the requirements for the degree of

Master of Science

in

Economics

Tufts University

August 2017

Abstract

This thesis studies how energy sector stocks perform under environment-related risks. I constructed the environmental sentiment factor and 7 subgroup environment-related risk factors using Dynamic Factor Model and took the abnormal returns as financial performance measure. The regression results show that during January 2004 to October 2016, all energy company stocks had negative abnormal returns; renewable energy firms had even lower abnormal returns than non-renewables; the abnormal returns of nonrenewable energy firms could be decreased by the increase of environmental concerns and they are more sensitive about public's environmental sentiment and weather conditions than that of renewable energy firms.

In conclusion, based on the findings of this thesis, in a world with worse weather conditions and higher public's environmental concerns, investors would sell nonrenewable energy stocks due to the decreasing abnormal returns, causing the cost of equity for nonrenewable energy firms to increase. So fewer nonrenewable energy firms would afford the cost of raising fund in stock market.

Acknowledgements

I would first like to thank my thesis advisor Professor Marcelo Bianconi of Economics Department at Tufts University. I can always get timely and warmly help from Prof. Bianconi, he always motivates me when I was stuck in troubles with my research. He is a very learned scholar, providing innovative thoughts and guaranteeing that my thesis is going in the right direction.

I also thank Prof. Bianconi for his encourages when I was diffident and in deep anxiety during my thesis writing process. He acts both friend and mentor, I really appreciate all his help.

Finally, I would like to acknowledge my readers on the thesis committee, Professor Daniel Richards and Professor Jeffrey Zabel, for their insightful comments during the writing and defense process.

Table of Contents

1. Introduction

Climate change has been a heated issue around the world. Although there are still people who doubt the existence of global warming, the general consensus is that climate change has caused the high frequency of natural disasters. As a result, climate conferences are held, agreements are signed, green energy and green consumption are promoted. As the public's environmental concerns increase, risks and uncertainties for energy firms would change, which could possibly affect energy firm's stock performance. Thus, it would be interesting to study the relationship between public's environmental sentiment and stock returns of energy firms. In addition, I am also curious about whether the mentioned relations would differ between renewable and nonrenewable energy firms.

It is not new for economists to relate environment with capital market. On the one hand, researchers are interested in the relationship between external weather conditions and stock market performance. To study this topic, the aggregative stock market index and the observed exogenous climate data are used. People find that weather conditions such as pollution, temperature, cloud cover, the amount of sunshine and the length of daylight can influence capital market's performance by altering traders' or investors' sentiment and further their investment behaviors (Lepori 2016; Levy and Yagil 2011; Chang et al. 2008). In these studies, only weather data around stock exchange are used and the time scope of these studies is usually very short, the main focus is how the stock market's performance changes with traders'/investors' changed emotions due to weather conditions.

On the other hand, economists are curious about how corporations' inner environmental performance (CEP) relates to the corporations' financial performance (CFP). This is a controversial sector, no consenus has been reached so far. To study this topic, researchers compare the financial performance of corporations with good and bad environmental performance.

Studies show that different measures of financial performance and environmental performance give different results. The most traditional view is that corporations with superior environmental performance would get less financial returns, since the additional costs would make the economic returns as a sacrifice for social/environmental goals. Elshahat, Freedman, and Elshahat (2015) used reduced factors to measure corporates' environmental performance and found that the relationship between corporates' environmental conerns and stock market returns is negative, while its relationship with accounting returns is positive. Chava (2014) got similar results using the implied cost of capital (ICC) as a measure of financial performance. Some studies get results of "no relations". Hamilton, Jo, and Statman (1993) studied the abnormal returns of mutual funds that are either socially responsible or not. They finally found that social responsibility factors have no effect on mutual funds' excess returns and thus these factors are not efficiently priced by the stock market. Ng and Zheng (2016) also found that Green and non-green firms have comparable firm values, investors of non-green portfolios wouldn't get punishment or penalty by achieving lower returns.

Many empirical studies also get “positive relations”. People believe that superior corporate environmental performance (CEP) can reduce the total corporate risks (Herremans, Akathaporn, and McInnes 1993), improve companies’ reputation (Saeidi et al. 2015) and therefore give firms significantly positive returns; while the significantly negative returns would act as “supplement punishments” for inferior corporate environmental performance (Laplante, Dasgupta, and Mamingi 1998; Klassen and McLaughlin 1996; Heinkel, Kraus, and Zechner 2001; Khairollahi et al. 2016).

In summary, the studies of weather conditions and stock market performance focused on external variance—the climate condition. The external variance influences stock market performance via traders’/ investors’ emotional changes; while the studies of relations between corporate environmental performance and corporate financial performance focused on inner variance --- the corporate characters. In this paper, I went a step further to combine these two topics by focusing on how public’s environmental sentiment would influence stock performance in energy sector; or in other words, how would the capital market price the environmental sentiment factor. In addition, would the effects differentiate between renewable and nonrenewable energy firms?

My thesis differs from the former studies in the following aspects: First, this paper focuses on the environmental sentiment of the public, not only that of investors or traders; Second, the climate/weather data are not limited to the stock exchange sites but the whole U.S.; Third, the study’s subject is not the performance of the whole capital market but the individual securities in energy sector. This paper contributes a better understanding of the relations between public's environmental sentiment and stock returns in energy sector, and a proper construction of the abstract concept "environmental sentiment".

The rest of the paper is laid out as follows: Section 2 is a detailed introduction on the construction of environmental sentiment factor, including literature, construction methods, data and the construction results; Section 3 is the calculation of the financial performance, including the model, data and regression results; Section 4 is results and analysis for the relations between environmental sentiment and stock returns; Section 5 is the conclusion.

2.The Construction of Environmental Sentiment Factor

2.1 Possible Sentiment Measures

The main issue for this paper is to find a proper environmental sentiment measure. Based on a wide range of literature, there are three possible methods to use: survey method, social media method, and econometric method.

2.1.1 Survey Method and Environmental Sentiment

Survey method is the most traditional one, whereby researchers construct sentiment factors using data collected by surveys. New Environmental Paradigm (NEP) Scale (Dunlap, Van Liere 1978; Dunlap et al. 2000) that measures public's pro-environmental orientation is one of this type. The NEP scale is constructed based on a survey asking respondents' strength of their agreement or disagreement with fifteen statements. Although the NEP Scale has been widely used as a measure of environmental world view or paradigm, its shortcomings and validity have been extensively discussed by Amburgey and Thoman (2012). The annually released environment survey by Gallup is another application of the survey method. The bad timeliness and the difficulties to access survey data are main drawbacks of this method.

2.1.2 Social Media Method and Environmental Sentiment

Social media method is a newly developed one. With the explosion of information, the concept of big data is widely used. Many researchers start to take advantage of the new media to measure public's attitudes and to predict public's behaviors. Searching engine data and new media data like Twitter, Facebook have been used to predict the spread of influenza-like illness, the job searching activities, the unemployment rate and employment policy effects (Ginsberg et al. 2009; Askitas and Zimmermann 2009; D'Amuri and Marcucci 2010; Baker and Fradkin 2011). The social media method is also very popular in capital market when predicting investment sentiment. The practices include analyzing words in main financial newspapers or analyzing related internet search queries to indicate investment sentiment (Tetlock 2007; Garcia 2013; Da et al 2011; Dimpfl and Jank 2016).

However, very few studies addressed the application of social media data in measuring environmental sentiment. Kahn and Kotchen (2010) used the query data of term "global warming" in *Google Insights for Search* to indicate public's environmental concerns and studied its relations with business cycle. Connor et al (2010) verified the potentials for social media data to act as a substitute and supplement of traditional polling. Compared with survey method, the social media method can provide us more real-time and easy-to-access data. Therefore, social media could be a possible method for the construction of environmental sentiment factor.

2.1.3 Econometric Method and Environmental Sentiment

Econometric method is based on computing, econometric theory, and availability of rich datasets. In this paper, I used Dynamic Factor Model (DFM) to construct a proper environmental sentiment factor.

Dynamic Factor Model (DFM) was initially proposed by Geweke in 1997, with intentions to study fluctuations in economic activities (Stock and Watson 1989; 1998). The model is a good implementation when the number of series exceeds the number of time series observations. The general econometric expression of Dynamic Factor Model is as follows:

$$X_t = Pf_t + e_t \quad (1)$$

$$f_t = Rw_t + A_1 f_{t-1} + A_2 f_{t-2} + \cdots + A_{t-p} f_{t-p} + \eta_t \quad (2)$$

$$e_t = C_1 e_{t-1} + C_2 e_{t-2} + \cdots + C_{t-q} e_{t-q} + \varepsilon_t \quad (3)$$

In the above equations, X_t is the observed informational dependent variable, f_t is the unobserved coincident variable, e_t is the exogenous variable. Suppose there are N observable time series, X_t and e_t would be N×1 vectors. Basically, it is assumed that the vector of time series (X_t) has two parts: a common component driven by some common factors (f_t), and an "idiosyncratic component" driven by "idiosyncratic factors" (e_t). It is allowed to include an autoregressive process for the unobserved factor (as equation (2)) and the error term (as equation (3)) to capture the dynamic effects.

Although DFM has been widely used in studying macroeconomic issues, such as obtaining real-time estimates of HongKong's real economy state (Gerlach and Yiu 2005), detecting monetary policy's effects on economy (Bernanke, Boivin, and Eliasz 2005) and so on, few applications have been implemented in constructing a proper sentiment index. So in this paper, I did this implementation by constructing public's environmental sentiment factor via Dynamic Factor Model (DFM).

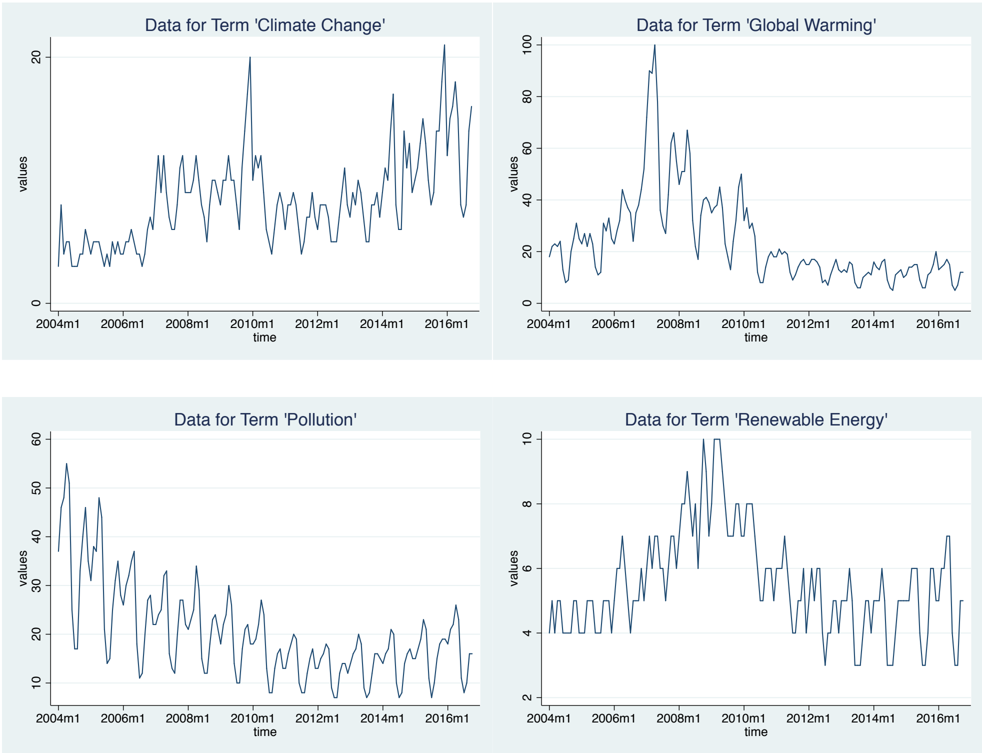
2.2 Data Used in DFM

According to equation (1), the first step to get the coincident factor f_t is to figure out the N×1 vector X_t. So firstly I got a group of time series that could have relations with, or could reflect public's environmental sentiment; then I applied the Dynamic Factor Model to these series to get the reduced single factor. Following this procedure, I divided these time series into five categories according to literature: Google Search queries, macroeconomic status, environment conditions, energy consumption, and environment regulation stringency. All data are limited to the United States.

2.2.1 Google Search Queries

To apply the social media method, I used the search frequency data of query terms "pollution" "climate change" "global warming" and "renewable energy" in the Google Trend. All data are monthly from January 2004 to October 2016, throughout the U.S.. As the official website of Google Trend explains, each data point is divided by the total searches of the geography and time range it represents. So the resulting numbers are then scaled on a range from 0 to 100 based on a topic's proportion to all searches on all topics.

Figure 2.2.1 Data for Google Search Queries



Data Source: Google Trend

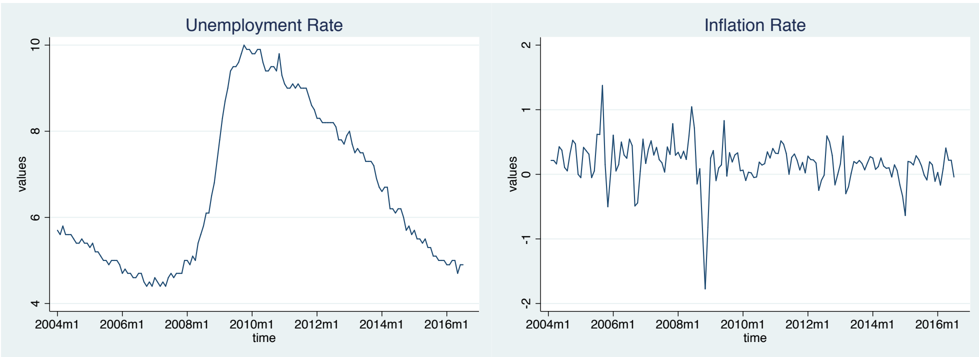
Table 2.2.1. Summary Statistics for Google Search Queries

variable	Observation	Mean	Standard Deviation	Min	Max
Climate Change	154	8.39	3.72	3	21
Global Warming	154	24.23	17.65	5	100
Pollution	154	20.17	9.74	7	55
Renewable Energy	154	5.56	1.58	3	10

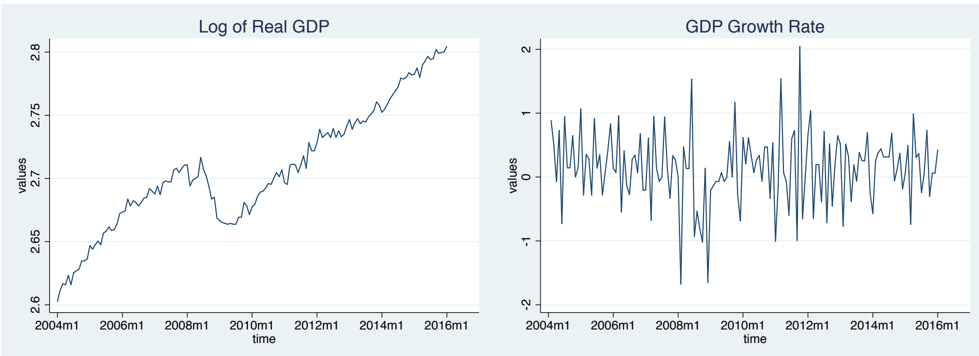
2.2.2 Macroeconomic Status

It is traditionally believed that social class or income level could affect individuals' environmental concerns (Liere and Dunlap 1980). When family's wealth is low or when the economic condition is bad (e.g. high unemployment rate), the public would care more about their financial problems and thus pay less attention to the environment (Soretz and Ott 2015; Kahn and Kotchen 2010). So I hypothesized that environmental concerns would be lower during recessions, this hypothesis was studied by Kahn and Kotchen in 2010. In this sector, I chose the monthly US data of inflation (from January 2004 to October 2016), the unemployment rate (from January 2004 to October 2016), and the log of real GDP (from January 2004 to January 2016) as part of X_t.

Figure 2.2.2 Macroeconomic Status



Data source: U.S. Bureau of Labor Statistics Data source: U.S. Bureau of Labor Statistics



Data source: Y Charts

Table 2.2.2. Statistic Summary Table for Economic Status Data

Variable	observation	Mean	Standard Deviation	Min	Max
Unemployment Rate	154	6.61	1.82	4.40	10.00
lnGDP	154	2.71	0.05	2.60	2.80
Inflation	154	0.17	0.34	-1.77	1.38

2.2.3 Environment Conditions

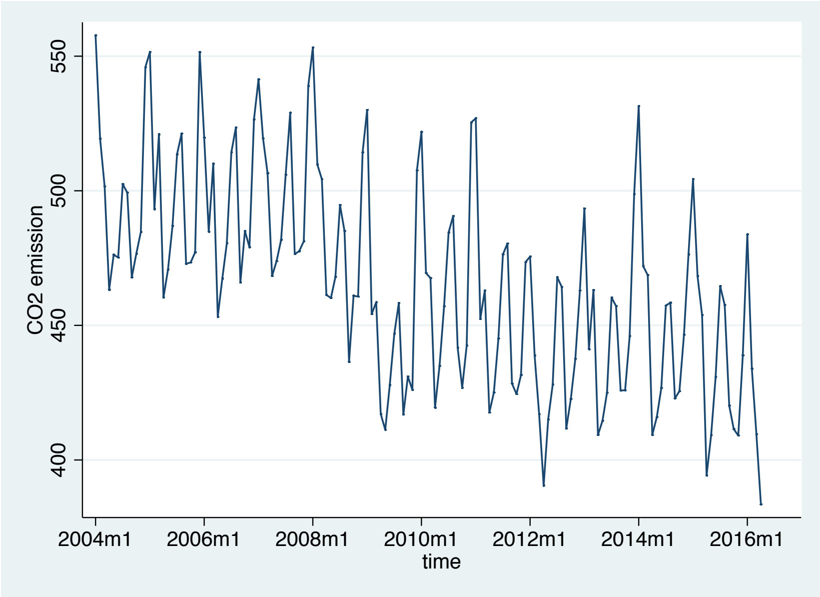
Together with economic status, the current environment conditions are also recognized as a factor that affects public's environmental sentiment. Liere and Dunlap (1980) mentioned "residence hypothesis" when studying the influence factors of environmental concerns. They proposed that people in urban areas who are exposed to higher levels of pollution and other types of environmental deteriorations care more about the environment. Soretz and Ott (2015) used the same hypothesis to study the relations between green attitude and economic growth. To capture the environment conditions, I chose the U.S. monthly data of CO2 emission and yearly data on Climate Extremes Index (CEI).

CEI is an index used as a framework for quantifying observed changes in climate within the U.S.. It was first introduced in early 1996 with the goal of describing and presenting a complex set of multivariate and multidimensional climate changes in the United States[1]. According to the official website of National Oceanic and Atmospheric Administration (NOAA), the U.S. CEI is the arithmetic average of several components including extreme temperatures

[1] https://www.ncdc.noaa.gov/extremes/cei/introduction

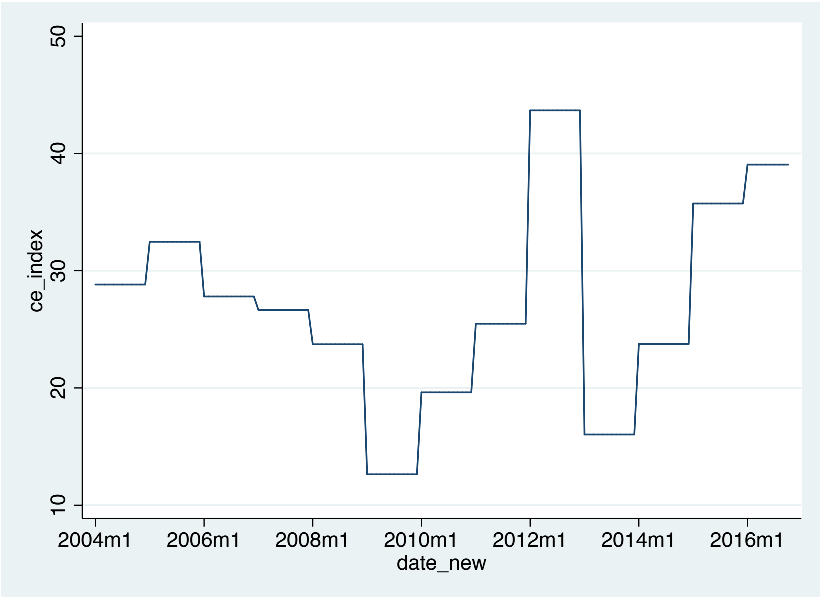
(maximum/minimum temperature much above/below normal), severe drought, degree and frequency of precipitation, and tropical storm/hurricane wind. Higher CEI indicates worse environment conditions.

Figure 2.2.3a CO_2 Emissions (monthly): from 2004m1 to 2016m4



Data Source: U.S. Energy Information Administration (EIA)

Figure 2.2.3b The U.S. Climate Extreme Index (monthly): from 2004m1 to 2016m10



Data Source: National Oceanic and Atmospheric Administration (NOAA)

2.2.4 Energy Consumption

Shifting from fossil fuels to renewable energy can be a main method to cope with climate change. So I assumed that higher public's environmental concerns would come with higher renewable energy consumption and lower nonrenewable energy consumption. Therefore, changes in renewable and nonrenewable energy consumption could reflect public's environmental sentiment.

Table 2.2.4 Summary Statistics for Log Form of Energy Consumption: 2004m1-2016m6

Variable	Observations	Mean	Standard deviation	Min	Max
Coal	150	0.49	0.18	-0.17	0.76
Natural Gas	150	0.71	0.20	0.38	1.20
Petroleum	150	1.11	0.07	0.96	1.27
Nuclear energy	150	-0.37	0.08	-0.56	-0.25
Solar Energy	150	-4.67	0.77	-5.84	-2.88
Wind Energy	150	-2.87	0.90	-4.67	-1.58
Biomass Energy	150	-1.11	0.17	-1.43	-0.86

Data Source: U.S. Energy Information Administration (EIA)

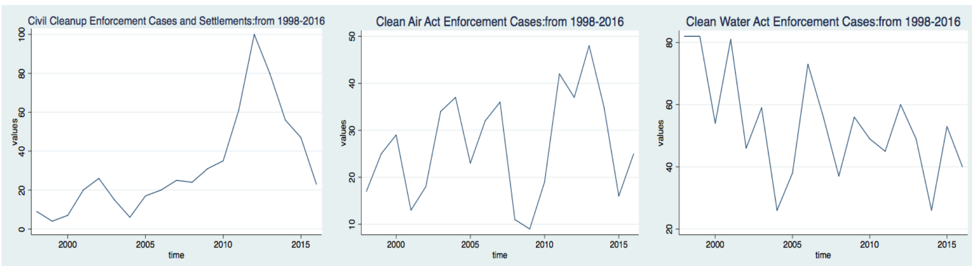
2.2.5 Environment Regulation Stringency

Many researchers have analyzed the reason why the links between public's environmental concerns and their pro-environmental activities are so weak. They found that people usually attribute the responsibility of protecting environment to government (Chyong et al. 2006). Public opinions were also found influential on public policies (Burnstein 2003; Shove 2010). So I hypothesized that public's environmental concerns can affect the environment regulation stringency. When the public is more concerned about the environment, they put more pressure on government's regulation policy, tightening the related environment regulation.

Brunel and Levinson (2016) summarized 5 traditional measures of regulation stringency by reviewing former studies. The first measure is "addressing simultaneity through natural experiments". One widely used natural experiment relies on the U.S. Clean Air Act; the second one is to use a specific regulation as an indicator of the overall environmental regulatory stringency; the third one is to use composite indexes that come from surveys of government officials, business managers or counts of regulation; the fourth one is to use emissions, pollution or energy use data to indicate the regulation; the last measure is to use the public sector expenditures or enforcement cases to measure the regulation stringency.

In this thesis, I applied the fifth regulation stringency measure by counting the number of EPA's yearly enforcement cases on civil cleanup, Clean Air Act (CAA) and Clean Water Act (CWA).

Figure 2.2.5 Environment Regulation Stringency Indicators (yearly): from 1998 to 2016



Data source: United States Environmental Protection Agency (EPA)

Table 2.2.5 Summary Statistics for Environment Regulation Stringency: from 1998 to 2016

Variable	Obs	Mean	Std. Dev.	Min	Max
Civil cleanup	19	31.89474	26.06167	4	100
Clean air act	19	26.63158	11.18635	9	48
Clean water act	19	53.26316	17.03213	26	82

2.3 Dynamic Factor Model Results

All series mentioned in 2.2 are first-difference stationary and a Johansen test indicates no cointegration at the 5% significance level. Due to computational problems, I failed to directly apply DFM to all series, so I implemented the DFM twice in separate steps to get my interested coincident factor. Firstly, I applied DFM to the first-differenced series in all five categories to get sub-group factors; secondly, I obtained the single common factor by applying DFM to these sub-group factors. Also because of computational difficulties, I calculated 7 sub-group factors from the 5 series categories, they are "*Google search factor*", *"macroeconomic factor"*, *"weather factor", "pollution factor (CO2 emission)*", "*nonrenewable energy consumption factor", "renewable energy consumption factor", and "regulation factor".*

Table 2.3a Components of Constructed Factors

Sub-group factors	Included in environmental sentiment factor?	Component series	Categories
Google Search factor	Yes	"pollution" "climate change" "global warming" "renewable energy"	google search queries
macroeconomic factor[2]	No	inflation CPI log of GDP unemployment rate	macroeconomic status
weather factor	Yes	CEI	environment condition
pollution factor	Yes	CO2 emission	
nonrenewable energy consumption factor	Yes	coal consumption natural gas consumption petroleum consumption	energy consumption
renewable energy consumption factor	Yes	nuclear energy consumption solar energy consumption wind energy consumption biomass energy consumption	
regulation factor	Yes	civil cleanup enforcement clean air act enforcement clean water act enforcement	environment regulation stringency

Maintaining the stationarity of the obtained sub-group factors, I implemented DFM for the second time. Since both steps have the same logic, I would only explain details of the second step.

With sub-group factors, I rewrote the DFM equations (1) (2) (3) as the following format:

$$X_{it} = p_i f_t + e_{it} \quad , \text{i=1,2, 3…6} \qquad (4)$$

$$f_t = a f_{t-1} + \eta_t \qquad (5)$$

[2] Macroeconomic factor acts as a control variable when study the relations between stock performance and environmental sentiment factor.

$$e_{it} = c_i e_{i,t-1} + \varepsilon_{it} \qquad (6)$$

As equation (5) and (6), I allowed AR(1) process for the unobserved factor f_t and the idiosyncratic component e_{it}. Note that although I got 7 sub-group factors through the first step, I only used 6 of them to construct the single composite index f_t. The factor I excluded is the macroeconomic factor, since macroeconomic condition is an important control variable in the following sections. Then, the maximum likelihood estimator is implemented by writing the model in state-space form and using the Kalman filter to execute the log likelihood.

Table 2.3b DFM Results: Maximum Likelihood Estimation

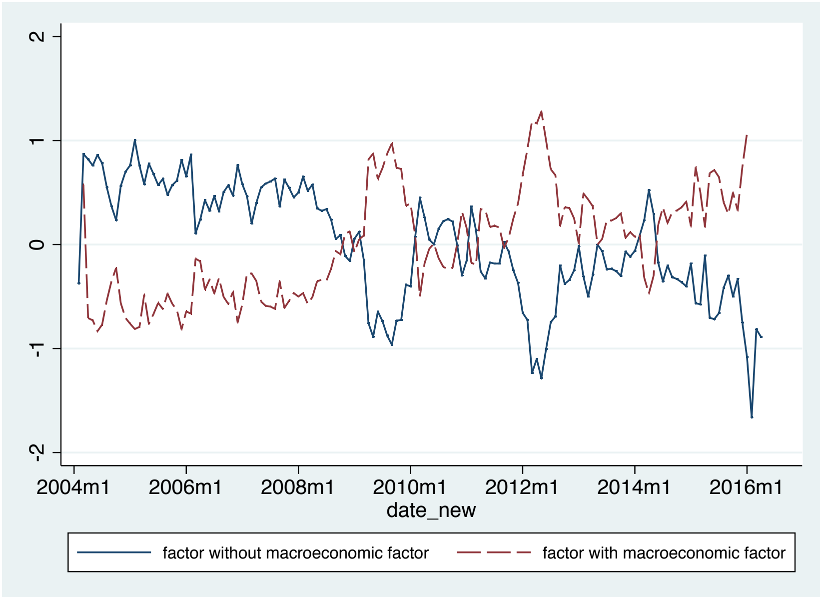
Parameters	(1)without macroeconomic factor[3]		(2)with macroeconomic factor[4]	
	estimates	z-values	estimates	z-values
a	0.6712***	7.97	0.6563***	6.89
c_1	0.9538***	32.08	0.9511***	28.85
c_2	0.1777	0.37	0.1358	0.16
c_3	-0.5002***	-6.03	-0.5084	-5.98
c_4	-0.0026	-0.5	-0.0026	-0.49
c_5	0.9066***	40.66	0.9174***	39.76
c_6	0.4792***	6.27	0.4874***	6.95
c_7			-0.1374	-1.42
p_1	-0.6521***	-7.72	0.6395***	6.09
p_2	0.4875***	7.31	-0.4821***	-5.18
p_3	-0.0055	-0.28	0.0054	0.27
p_4	-0.0915	-0.91	0.0961	0.81
p_5	-0.0312	-0.62	0.0401	0.91
p_6	1.0963***	3.63	-0.9928***	-3.29
p_7			0.0404	0.87
var(e_1)	0.2052**	2.38	0.2156*	1.93
var(e_2)	0.0458	0.97	0.0414	0.6
var(e_3)	0.2636***	6.45	0.2613***	6.14
var(e_4)	10.594*	1.83	10.8893*	1.83
var(e_5)	0.5742**	2.78	0.5284*	2.54
var(e_6)	7.9063***	7.4	7.8299**	7.21
var(e_7)			0.6569***	8.06
log likelihood	-1263.4987		-1395.1001	

Note: (i) Significance levels are + 0.1; *0.05; **0.01; ***0.001.

(ii) In theory, there could be more than one coincident factor f_t ; however, due to computational problems, I failed to get results when including more than one f_t.

[3] (1) stands for the case when f_t is constructed without macroeconomic factor.
[4] (2) stands for the case when f_t is constructed with macroeconomic factor.

Figure 2.3a Constructed Environmental Sentiment Factor: With and Without Macroeconomic Factor



Note: The two constructed environmental sentiment factors are all deseasonalized.

Table 2.3b presents the estimates of a (coefficient of AR(1) process of unknown factor f_t), c_i (coefficient of AR(1) process of exogenous part e_{it}), p_i and the variance of the disturbances e_i in the DFM for the 6 or 7 (with macroeconomic factor) sub-group factors. The subscript numbers 1,2...7 indicate "*Google Search factor*", "*nonrenewable energy consumption factor", "renewable energy consumption factor", "weather factor", "regulation factor", "pollution factor (CO2 emission)*" and *"macroeconomic factor"*, respectively.

For column (1), the estimated autoregressive coefficient of f_t was 0.6712 (a=0.6712), which indicates a high persistence in the coincident factor. The coefficients for "*renewable energy consumption factor"* (p_3), "*weather factor"* (p_4) and "*regulation factor*" (p_5) were insignificant, which means the unobserved factor (f_t) is not a significant predictor for them; however, the coefficients for "*Google Search factor*" (p_1), "*nonrenewable energy* *consumption factor"* (p_2) and *"pollution factor"* (p_6) were significant and signs for p_2 and p_6 were positive, while the sign for p_1 was negative. So I concluded that the unobserved factor (f_t) is a significant predictor for "*nonrenewable energy consumption factor"* and "*pollution factor"* (CO2 emission), it is also a significant predictor for the reverse of "*Google Search factor*".

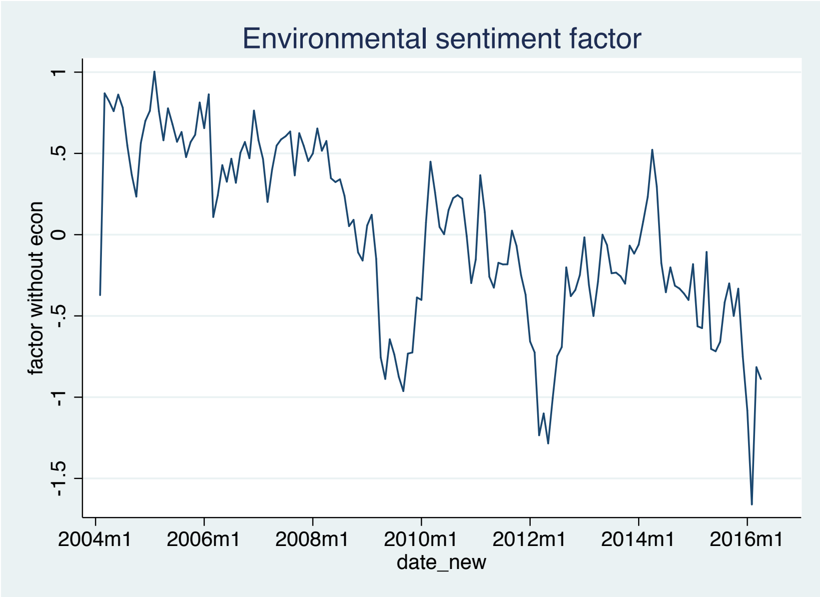
For column (2), I included "*macroeconomic factor*" when constructing the coincident factor, it showed high persistence just as the case with column (1); the significance for coefficients $p_1 - p_6$ remained consistent with column (1). The only difference is that the significant coefficients (p_1, p_2 and p_6) flipped their signs compared with the results without "*macroeconomic factor*". So I expected that the two obtained common factors (f_t constructed with and without *macroeconomic factor*) would behave inversely just as Figure 2.3a showed.

After getting the estimated parameters, I calculated the environmental sentiment factor by running the Kalman smoother with the case that didn't include *macroeconomic factor*. The obtained *environmental sentiment factor* suffered from severe seasonality problems, so I deseasonalized it using the following method:

$$f_t = \gamma_t + \theta_t Month + e_t \quad (7)$$

Where $Month$ is time dummies; f_t is a factor suffering from seasonality; t indicates different months and $t \in [1,11]$; e_t is the deseasonalized factors. The procedures to obtain deseasonalized sub-group factors are the same as above.

Figure 2.3b Deseasonalized Environmental Sentiment Factor (F)



Note: The deseasonalized environmental sentiment factor above is constructed without macroeconomic factor.

Figure 2.3c Deseasonalized Control Variable: Macroeconomic Factor

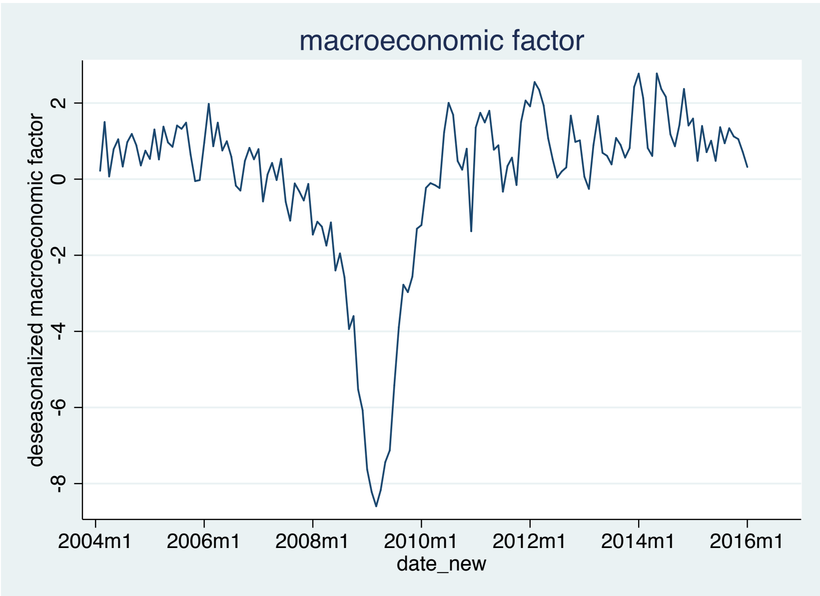
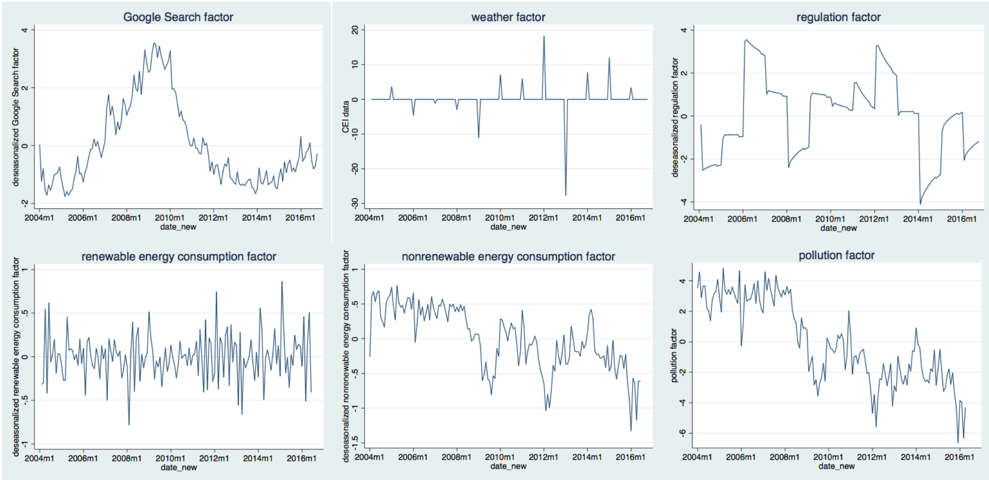


Figure 2.3d Deseasonalized Component Subgroup Factors: DFM



2.4 Economic Implication of Obtained Environmental Sentiment Factor

In the following paragraphs, I took the coincident factor that constructed without *macroeconomic factor* as the obtained *environmental sentiment factor*. So far, I am not sure whether the *environmental sentiment factor* indicates "environmental concerns" or "environmental relief", since for Dynamic Factor Model, the coefficients' signs of the unobserved factor is very flexible (the coefficients' signs flipped compared with the case that includes *macroeconomic factor*). So further analysis is needed before figuring out the economic implication of the obtained *environmental sentiment factor F*.

Table 2.4 gives some clues. Following the assumptions in Section 2.2, I expected the "*environmental concerns*" to be positively related with *weather factor*, *regulation factor*, *Google Search factor*, *renewable energy consumption factor* and *macroeconomic factor*; while it should be negatively related with *nonrenewable energy consumption factor*. According to Table 2.4, the obtained *environmental sentiment factor* showed opposite relations with the above subgroup factors, acting reversely with "*environmental* *concerns*". So I concluded that the economic implication of the *environmental sentiment factor F* is "environmental relief". That is, the decrease of F indicates the increase of environmental concerns. This economic implication would be confirmed further by the regression results in next section.

Table 2.4 Correlation Matrix for Environmental Sentiment Factor F and Subgroup Factors

	Environmental sentiment factor (F)	Weather Factor	Google search factor	Nonrenewable energy consumption factor	Renewable energy consumption	Regulation Factor	Pollution factor (CO_2 emission)	Macroeconomic factor
Environmental sentiment factor (F)	1							
Weather Factor	-0.0937	1						
Google search factor	-0.0901	-0.016	1					
Nonrenewable energy consumption factor	0.8477*	-0.0441	-0.053	1				
Renewable energy consumption	-0.0682	-0.135	-0.0802	-0.0209	1			
Regulation Factor	-0.1853*	-0.0716	0.1725*	-0.1502	-0.0255	1		
Pollution factor (CO_2 emission)	0.7961*	-0.0292	0.0553	0.8309*	-0.0099	-0.0502	1	
Macroeconomic factor	0.126	0.1499	-0.7874*	0.1083	0.0295	-0.0231	-0.0102	1

Note: (i) * 0.05 significance level. (ii) The pollution factor and nonrenewable energy consumption factor show high correlation which exceeds 0.8, so there could be multicollinearity problem when taking both of them as independents.

3. Stock Performance Measure: Abnormal Returns

3.1 Models for Stock Returns

According to the Efficient Market Hypothesis (EMH), it is impossible to beat the market, since capital market efficiency has made stock prices to reflect all relevant information in the market. So in an efficient market, no additional information could be used to get excess returns.

Abnormal returns or excess returns which denoted by α_{it} in equation (8) describe stock returns that asset pricing models failed to capture. In other words, the abnormal returns are the difference between stocks' actual returns and stocks' expected returns calculated by asset pricing models, they are payoffs for investors to endure risks that aren't studied in models. "Abnormal returns" is a good measure of stock performance when study the effects of risk factors that are not included in asset pricing models (Herremans, Akathaporn, and McInnes 1993; Lorraine, Collison, and Power 2004). It is also used to implement event study, analyzing how would stock returns react when some information is released(Flammer 2015). In this paper, I used companies' abnormal returns as a measure of stock performance and studied its relationship with *environmental sentiment factor*.

To get the abnormal returns, I used three well recognized asset pricing models: CAPM model, Fama-French three-factor model (1993) and Fama-French five-factor model (2015). The general form of asset pricing models could be written as:

$$R_{it} - R_{Ft} = \alpha_{it} + \beta_1(R_{mt} - R_{Ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t \quad (8)$$

Where R_{it} is the stock returns for a public firm, R_{Ft} is the risk-free rate, $R_{mt}$is the market returns. All these returns only vary with time. We call "$R_{it} - R_{Ft}$" company premium, which measures investors' payoffs for their assumed risks; following the same logic, "$R_{mt} - R_{Ft}$" is called market premium; SMB_t[5], HML_t[6], RMW_t[7] and CMA_t[8] are size factor, value factor, profitability factor and investment factor separately, they are all returns and only vary with time; α_{it} is the abnormal returns or excess returns that I used as a measure of stock performance.

For traditional CAPM model (one-factor model), $\beta_2 = \beta_3 = \beta_4 = \beta_5 = \alpha_{it} = 0$; For Fama-French three-factor model (1993), $\beta_4 = \beta_5 = \alpha_{it} = 0$; For Fama-French five-factor model (2015), only $\alpha_{it} = 0$.

3.2 Understand and Obtain Abnormal Returns

I rewrote equation (8) to form the econometric model, η_{it} is the error term:

$$R_{it} - R_{Ft} = \alpha + \beta_1(R_{mt} - R_{Ft}) + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \eta_{it} \quad (9)$$

In equation (9), α is the average of all constant terms in regression results for individual equities, which constitutes a part of the abnormal returns α_{it};

5 SMB is size factor that describes the difference between the returns on a portfolio of small market capitalization (the bottom 50%) stocks and the returns on a portfolio of large market capitalization (the top 50%) stocks.
6 HML is value factor that describes the difference between the return on portfolio of high (the top 30%) book-to-market stocks and the return on a portfolio of low (the bottom 30%) book-to-market.
7 RMW is Profitability factor that describes the difference between the returns on diversified portfolios of stocks with robust and weak profitability.
8 CMA is Investment factor that measures the difference between the returns on diversified portfolios of stocks with low and high investment firms, which we call conservative and aggressive.

the error term η_{it} indicates random returns for specific risk factors of each individual equities, it is another part of the abnormal returns and $\mathrm{E}(\eta_{it}) = 0$. So the following relations should hold:

$$\alpha_{it} = \alpha + \eta_{it} \qquad (10)$$

$$\mathrm{E}(\alpha_{it}) = E(\alpha + \eta_{it}) = E(\alpha) \quad (11)$$

In the above equations, α has a non-zero expected value, so we can understand α as the additional returns for investors to assume risks that were not included in asset pricing models; While η_{it} is random returns related with stock's characters and it has a zero expected value, so we can understand η_{it} as returns that influenced by unsystematic risks. According to basic investment theory, investors won't get paid for taking unsystematic risks since these risks could be cleared if the portfolio is well diversified. Deep understanding of the abnormal returns α_{it} would help us in explaining the regression results in the following sections.

To get the the abnormal returns for each stock (stock performance measure), I firstly applied the three asset pricing models to equity dataset and got the expected company premium $(\widehat{R_{\iota t} - R_{Ft}})$; secondly, as equation (12), I got the individual stock's abnormal returns by excluding the expected company premium from the actual company premium. In the first step, the regression results of asset pricing models give a constant term, this constant term is part of the abnormal returns, that is α; In the second step, the predicted abnormal returns for each stock are the whole part of abnormal returns, that is α_{it}.

As equation (13), I regressed the predicted abnormal returns (α_{it}) on the *environmental sentiment factor* and its component subgroup factors separately to study the effects of *environmental sentiment factor* on equity abnormal returns:

$$\hat{\alpha}_{it} = real(company\ premium) - expected(company\ premium) \quad (12)$$

$$\hat{\alpha}_{it} = \gamma_0 + \gamma_1 D_i + \gamma_2 F_t + \gamma_3 F_t * D_i + \gamma_4 C_t + \gamma_5 C_t * D_i + u'_i + v'_t + \eta'_{it} \quad (13)$$

In equation (13), D_i is a dummy variable "*renewable*", $D_i = 1$ for renewable energy firms, $D_i = 0$ for nonrenewable energy firms; F_t is the obtained *environmental sentiment factor* or its component subgroup factors; $F_t * D_i$ is the interact terms to study whether environment-related factors influence renewable and nonrenewable energy firms differently; C_t is the control term, in this paper, it indicates *macroeconomic factor;* $C_t * D_i$ is the interact term of *renewable* and *macroeconomic factor*; u'_i is firm fixed effect that controlled by including firm dummies in regression; v'_t is time fixed effect that controlled by adding month and year dummies in regression; η'_{it} is error term.

3.3 Equity Data Information

The equity dataset includes 448 American energy equities, among which 53 belong to renewable energy firms and 395 belong to nonrenewable energy firms. All stock price data are monthly, from January 2004 to October 2016.

Table 3.3 Equity Information

Sector	Sub-sector	Number of equity	Total number
Renewable energy industry	Biofuels	20	53
	Renewable Energy Equipment	24	
	Renewable Energy Project Development	9	
Nonrenewable energy industry	Coal Operations	1	395
	Exploration & Production	223	
	Integrated Oils	2	
	Integrated Utilities	1	
	Midstream - Oil & Gas	69	
	Oil & Gas Services & Equip	80	
	Refining & Marketing	19	

Data source: Bloomberg

The realized stock market returns were calculated using $R_{it} = \ln(p_{i,t}) - \ln(p_{i,t-1})$, where $p_{i,t}$ is the close price each month for each stock. I used the monthly data of S&P500 index return as market return "R_{mt}" and the monthly yield of 3 month T-bill as the risk-free rate "$\mathrm{R_{Ft}}$". The monthly data of $\mathrm{R_{mt}}$ and $\mathrm{R_{Ft}}$ are both from FRED, running from January 2004 to October 2016. The monthly data for $\mathrm{SMB_t}$, $\mathrm{HML_t}$, $\mathrm{RMW_t}$ and $\mathrm{CMA_t}$ are from the website[9] of Kenneth R. French. All returns data are in decimal format.

3.4 Regression Results of Asset Pricing Models

Following the first step to get abnormal returns in section 3.2, I got the regression results of asset pricing models shown in Table 3.4a. The results include applying three asset pricing models using the full dataset, the renewable energy firms' dataset and the nonrenewable energy firms' dataset.

[9] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

The coefficient of market premium, which is also called "market beta", indicates the sensitivity of asset's returns to market returns. Market beta is a measure of systematic risks or market risks, higher market beta means higher systematic risks. From Table 3.4a, I found that renewable energy firms tend to have a higher market beta than nonrenewable energy firms, except for the case under Fama-French five-factor model. I further analyzed the significance of the discrepancy and formed Table 3.4b, which shows that market beta is not significantly different for renewable and nonrenewable energy firms, so renewable and nonrenewable energy firms have similar market risks.

The constant term α in Table 3.4a remains significantly negative under all asset pricing models with all different datasets. The negative signs for constant term α indicate that energy sector investors get less returns than expected. Besides, I found that the abnormal returns for renewable energy firms are about 3.5% lower than that for nonrenewable energy firms. The difference is statistically significant at 5% significance level as shown in Table 3.4b. So we can conclude that the actual returns of renewable energy firms' investors are even lower than that of nonrenewable energy firms' investors. According to the second hypothesis proposed in Hamilton, Jo, and Statman (1993), the above results can be explained as: the market priced the green characters of renewable energy firms, since they increased the value of renewable energy firms by driving down the expected returns.

Figure 3.4a Abnormal Returns for U.S. Energy Firms: 2004m1 to 2016m10

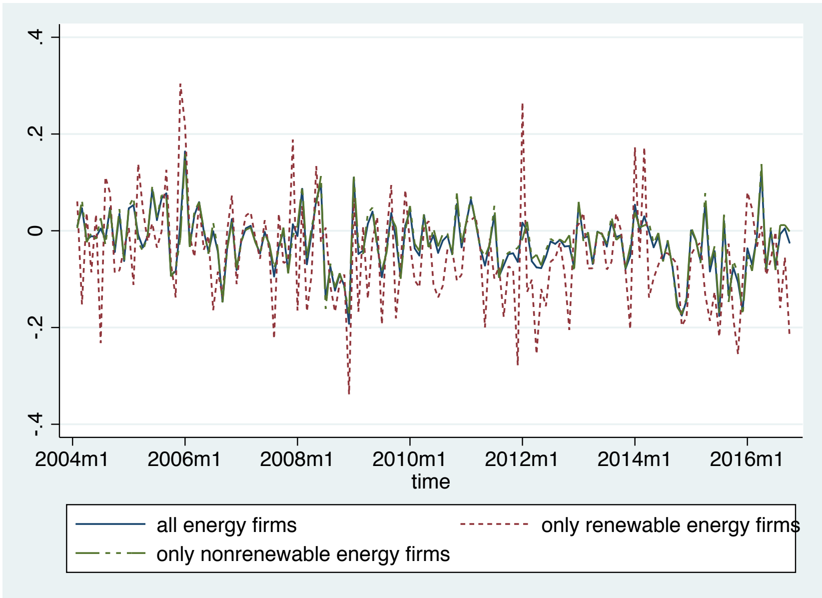


Table 3.4a Regression Results for Asset Pricing Model

Dependent variable: company premium									
	one-factor model			three-factor model			five-factor model		
	all energy firms	renewables	Non-renewables	all energy firms	renewables	Non-renewables	all energy firms	renewables	Non-renewables
Market premium	1.13599	1.19268	1.12777	1.01546	1.04754	1.01125	0.97387	0.96183	0.97577
	(0.04451)***	(0.14291)***	(0.04672)***	(0.04707)***	(0.16075)***	(0.04901)***	(0.05000)***	(0.17785)***	(0.05165)***
Smb				0.51287	0.85664	0.46986	0.49393	0.78519	0.4578
				(0.07594)***	(0.22215)***	(0.08067)***	(0.08168)***	(0.23952)***	(0.08678)***
Hml				0.1201	-0.11475	0.14772	0.28602	0.06447	0.30826
				(0.07060)*	(0.28)	(0.07090)**	(0.08556)***	(0.31219)	(0.08790)***
Rmw							-0.07497	-0.36521	-0.03928
							(0.11893)	(0.4052)	(0.12348)
Cma							-0.61353	-0.69345	-0.59144
							(0.13810)***	(0.59841)	(0.13593)***
Constant	-0.02771	-0.05877	-0.02378	-0.02772	-0.05902	-0.02377	-0.02721	-0.0573	-0.02341
	(0.00191)***	(0.00748)***	(0.00182)***	(0.00191)***	(0.00753)***	(0.00181)***	(0.00187)***	(0.00762)***	(0.00177)***
R-Squared	0.01707	0.01155	0.01843	0.0182	0.01313	0.01955	0.01863	0.01358	0.01998
N	47,607	5,344	42,263	47,607	5,344	42,263	47,607	5,344	42,263

Table 3.4b Analysis of Coefficients' Discrepancy

	CAPM model				Three-Factor Model				Five-Factor Model			
	renewable	Non-renewable	difference	z-score	renewable	Non-renewable	difference	z-score	renewable	non-renewable	difference	z-score
market beta	1.19268	1.12777	0.0649	0.4317	1.04754	1.01125	0.0363	0.2159	0.96183	0.97577	-0.0139	-0.0753
	(0.1429)	(0.04672)			(0.16075)	(0.04901)			(0.17785)	(0.05165)		
abnormal returns	-0.05877	-0.02378	-0.0350	-4.5452	-0.05902	-0.02377	-0.0353	-4.5516	-0.0573	-0.02341	-0.0339	-4.3322
	(0.00748)	(0.00182)			(0.00753)	(0.00181)			(0.00187)	(0.00177)		

Note: We get the z-score by z=$\frac{\beta_1-\beta_2}{\sqrt{se_1^2+se_2^2}}$

4. Environmental Sentiment and Abnormal Returns

4.1 With 3 Component Subgroup Factors

After getting the regression results of asset pricing models, I executed equation (12) to calculate the predicted abnormal returns $\hat{\alpha}_{it}$. Then I merged the panel data of security information with time series data of various factors and regressed the obtained abnormal returns on these factors just as shown in equation (13). To get a more precise estimation, I applied Newey-West estimation while controlling company and time fixed effects. The regression results are shown in Table 4.1a. To better understand how renewable and nonrenewable energy firms differ, I did same regressions using separate datasets. See Table 4.1b-c.

For Table 4.1a-c, I only used 3 out of 6 component subgroup factors. The excluded factors are *renewable energy consumption factor*, *nonrenewable energy consumption factor* and *pollution factor (CO2 emission),* since the energy consumption factors and *pollution factor* (CO2 emission) seem to be relatively endogenous compared with the other three. I added them back in section 4.2.

Table 4.1a Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F)

Dependent variable	abnormal returns for one-factor model			abnormal returns for three-factor model			abnormal returns for five-factor model		
	Benchmark	(1)	(2)	Benchmark	(3)	(4)	Benchmark	(5)	(6)
Renewable		-0.348	-0.352		-0.348	-0.351		-0.348	-0.351
Environmental sentiment factor (F)		0.0470***			0.0460***			0.0443***	
F*renewable		0.0146			0.0146			0.0146	
Google Search factor	-0.0148***		-0.003	-0.0137***		-0.00278	-0.0137***		-0.00301
Weather factor			-0.00235***			-0.00249***			-0.00288***
Macroeconomic factor		0.00693***	0.00815***		0.00700***	0.00832***		0.00714***	0.00858***
Regulation factor			-0.0155***			-0.0156***			-0.0155***
Google Search factor*renewable			-0.00186			-0.00192			-0.00189
Weather factor*renewable			0.00527*			0.00524*			0.00522*
Macroeconomic factor*renewable		0.00167	-0.000277		0.00169	-0.000286		0.00168	-0.000274
Regulation factor*renewable			-0.00329			-0.00332			-0.00334
constant	0.161***	0.0593*	0.0249	0.167***	0.0684**	0.0352	0.157***	0.0799**	0.0494+
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	47607	44059	44059	47607	44059	44059	47607	44059	44059
F	3.523	3.835	5.358	3.473	3.745	5.057	3.423	3.766	4.689

+ 0.1; *0.05; **0.01;***0.001

Note: (i) The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 447 company dummies (448 equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table.

(ii) Theoretically, if properly constructed, the results in Table 4.1a should be the consolidation of Table 4.1b and Table 4.1c together with dummy variable "*renewable*"; However, comparing the results of Table 4.1a-c, the theoretical relations seemed to failed. The reason for this failure is that I didn't interact "renewable" with my firm dummies and time dummies when using the full dataset. So in this paper, I emphasize the results of separate datasets.

(iii) The benchmark factor "Google Search factor" is a relatively direct measure of environmental concerns. Compare the benchmark factor and the environmental sentiment factor, we can conclude that the economic implication of environmental sentiment factor is "environmental relief", not "environmental concerns". So the decrease of environmental sentiment factor F means an increase of environmental concerns. The economic implication of environmental sentiment factor always holds for all regressions.

Table 4.1b Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F): Renewables

Dependent variable	abnormal returns for one-factor model		abnormal returns for three-factor model		abnormal returns for five-factor model	
	(1)	(2)	(3)	(4)	(5)	(6)
Environmental sentiment factor (F)	0.0390+		0.0379		0.0357	
Google Search factor		0.0123		0.0123		0.0123
Weather factor		0.00265		0.00246		0.00204
Macroeconomic factor	0.0108*	0.0122*	0.0109*	0.0124*	0.0110*	0.0127*
Regulation factor		-0.0206*		-0.0212*		-0.0212*
constant	-0.208*	-0.171	-0.207*	-0.175	-0.210*	-0.174
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes
N	4967	4967	4967	4967	4967	4967
F	3.953	2.739	4.124	2.741	4.118	2.658

+ 0.1; *0.05; **0.01; ***0.001

Note: The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 52 renewable energy company dummies (53 renewable energy equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table.

Table 4.1c Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F): Non-renewables

Dependent variable	abnormal returns for one-factor model		abnormal returns for three-factor model		abnormal returns for five-factor model	
	(1)	(2)	(3)	(4)	(5)	(6)
Environmental sentiment factor (F)	0.0497***		0.0488***		0.0471***	
Google Search factor		-0.00514		-0.0049		-0.00516
Weather factor		-0.00233***		-0.00247***		-0.00286***
Macroeconomic factor	0.00663***	0.00760***	0.00670***	0.00776***	0.00683***	0.00802***
Regulation factor		-0.0153***		-0.0154***		-0.0152***
constant	-0.0353	-0.0791**	-0.0311	-0.0735*	-0.0176	-0.0570+
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes
N	39092	39092	39092	39092	39092	39092
F	3.682	3.8	3.540	3.637	3.575	3.594

+ 0.1; *0.05; **0.01; ***0.001

Note: The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 394 nonrenewable energy company dummies (395 nonrenewable energy equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table.

I divided each table (Table 4.1a-c and Table 4.2a-c) into three sections based on the different asset pricing models used to predict abnormal returns. In Table 4.1a-c, column (1) (3) (5) are regression results for *environmental sentiment factor* (F) while controlling *macroeconomic factor*, together with the dummy variable "*renewable*" and interact terms. Column (2) (4) (6) are regression results for 3 component subgroup factors: *Google Search factor, weather factor and regulation factor*, "*renewable*" and interact terms while also controlling *macroeconomic factor.*

In Table 4.1a, I included a benchmark column in each section. This column is the results of regressing abnormal returns on a single factor: *Google Search factor*. The benchmark column is used to verify the economic implication of the *environmental sentiment factor* F, since the *Google Search factor* is a relatively direct social media method of public's environmental concerns. The increase of *Google Search factor* can be understood as the increase of public's environmental concerns.

In Table 4.1a, the coefficient of benchmark factor is negative while the coefficient of *environmental sentiment factor* is positive. Together with the factor's correlation matrix (see section 2.4) and the DFM estimation results (see section 2.3), I further verified that the obtained *environmental sentiment factor F* means "environmental relief". Therefore, a decrease in *environmental sentiment factor* F implies a decrease of "environmental relief" and an increase of public's "environmental concerns". Regression results should be explained more cautiously.

Another important thing to remember when explaining the regression results is: the dependent variable in this section is $\hat{\alpha}_{it}$, which includes two parts α and η_{it}. The "risk return tradeoff" is not applicable when explaining $\hat{\alpha}_{it}$. So I won't expect to see positive relations between risks and abnormal returns.

For column (1) (3) (5) of Table 4.1a, when using full dataset, *environmental sentiment factor F* and *macroeconomic factor* appear to have significantly positive effects on abnormal returns under each section, while the "renewable" and interact terms are insignificant. So I concluded that better *economic conditions* could improve market's prospects of the whole stock market including the energy industry, which lead to higher abnormal returns for energy firms; while higher *environmental concerns* (lower F) would make investors less confident about energy industry's future performance, so the environmental concerns show a negative relation with the whole energy industry's abnormal returns. With full dataset, I failed to see differences in these effects between renewable and nonrenewable energy firms.

In column (2) (4) (6) of Table 4.1a, when studying the component subgroup factors of *environmental sentiment factor F*, I found that *weather factor* and *regulation factor* have significantly negative relations with abnormal returns, while the relations for *macroeconomic factor* continue to be significantly positive. No interact terms are significant except for "*weather factor*renewable*" and its coefficient is positive.

These results are reasonable, since when the *environment regulation* becomes stricter, there would be more requirements for energy firms and there would be higher possibility for energy firms to violate the regulations, which may cause additional costs and impair energy companies' financial performance. So the market or the investors would lower their expectations of stock returns and buy less shares of energy stocks, further lower energy firms' stock price. This effect doesn't differ between renewables and non-renewables which indicates that the market believes renewables won't do better under the same regulation status.

The analysis of *macroeconomic factor* remains the same, so I won't repeat here. For *weather factor*, it has different effects on abnormal returns for renewable and nonrenewable energy firms. For non-renewables, when the climate change becomes severer (higher *weather factor*), there would be less abnormal returns since the market expects a worse future for nonrenewable energy firms; For renewables, the whole effect of *weather factor* on abnormal returns would be positive, which means worse *weather conditions* would give renewable energy firms more future profit opportunities thus increasing its abnormal returns. So under this weather risk factor, investors believe renewable and nonrenewable firms would perform differently, and data show that renewables performed better during January 2004 to October 2016. But note that the significance level of the coefficient of "weather factor *renewable" is 0.05, relatively low compared with other coefficients.

Table 4.1b and Table 4.1c would be helpful in better understanding the former analysis. The regression results for nonrenewable energy firms shown in Table 4.1c are comparable with Table 4.1a, so I won't bother in explaining it. The results for renewable energy firms (Table 4.1b) are different from that for all energy firms and nonrenewable energy firms, the differences are summarized as follows.

Firstly, the coefficient of *environmental sentiment factor* is only significant in column (1) at 0.05 significance level; secondly, the coefficients of *weather* *factor* are insignificant under each column of (2) (4) (6). Thirdly, for significant coefficients in Table 4.1b, although the signs are consistent with Table 4.1a, the significance level is much lower (significant at 0.05).

So according to Table 4.1b, I can say that neither *environmental sentiment factor* nor the *weather factor* could significantly influence renewable energy firms' abnormal returns; their abnormal returns are significantly affected by the *environment regulation factor* and *macroeconomic factor*, and the effects are indifferent compared with non-renewables. To understand these results, we should notice that the *environmental sentiment factor F* is not a simple average of component sub-group factors. It is possible for *environmental sentiment factor F*'s coefficient to be insignificant while the coefficients of some sub-group factors are significant due to the existence of the "idiosyncratic factors" mentioned in section 2.

4.2 With 5 and 6 Component Subgroup Factors

I redid Table 4.1a-c by including *renewable energy consumption factor*, *nonrenewable energy consumption factor* and *pollution factor*.

Table 4.2a Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F)

Dependent variable	abnormal returns for one-factor model				abnormal returns for three-factor model				abnormal returns for five-factor model			
	Benchmark	(1)	(2)	(3)	Benchmark	(4)	(5)	(6)	Benchmark	(7)	(8)	(9)
Renewable		-0.348	-0.347***	-0.354		-0.348	-0.347***	-0.356		-0.348	-0.347***	-0.356
Environmental sentiment factor (F)		0.0470***				0.0460***				0.0443***		
F*renewable		0.0146				0.0146				0.0146		
Google Search factor	-0.0148***		-0.0044	-0.00162	-0.0137***		-0.00432	-0.000933	-0.0137***		-0.00448	-0.00105
Weather factor			-0.00259***	-0.00282***			-0.00272***	-0.00300***			-0.00317***	-0.00346***
Macroeconomic factor		0.00693***	0.00800***	0.00824***		0.00700***	0.00815***	0.00845***		0.00714***	0.00840***	0.00870***
Regulation factor			-0.0141***	-0.0156***			-0.0140***	-0.0158***			-0.0142***	-0.0161***
Renewable energy consumption factor			-0.0186**	-0.0212**			-0.0182**	-0.0213**			-0.0214**	-0.0247***
Nonrenewable energy consumption factor			0.0268***	0.0156+			0.0301***	0.0160+			0.0269***	0.0126
pollution factor				0.00562**				0.00694***				0.00705***
Google Search factor*renewable			-0.00383	-0.00419			-0.00389	-0.00424			-0.00386	-0.0042
Weather factor*renewable			0.00591**	0.00587**			0.00589**	0.00586**			0.00587**	0.00584**
Macroeconomic factor*renewable		0.00167	-0.00185	-0.00169		0.00169	-0.00187	-0.00172		0.00168	-0.00186	-0.00171
Regulation factor*renewable			-0.00161	-0.00192			-0.00162	-0.00193			-0.00165	-0.00195
Renewable energy consumption factor* renewable			0.0215	0.0205			0.022	0.021			0.0221	0.0211
Nonrenewable energy consumption factor*renewable			0.0281+	0.0143			0.0283+	0.0147			0.0283+	0.0149
pollution factor*renewable				0.00276				0.00272				0.00268
constant	0.161***	0.0593*	0.0496+	0.0609	0.167***	0.0684**	0.0591*	0.0729	0.157***	0.0799**	0.0695*	0.0835
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	47607	44059	44059	44059	47607	44059	44059	44059	47607	44059	44059	44059
F	3.523	3.835	4.134	3.82	3.473	3.745	4.03	3.673	3.423	3.766	4.008	3.622

+ 0.1; *0.05; **0.01; ***0.001

Note: (i) The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 447 company dummies (448 equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table. (ii) Theoretically, if properly constructed, the results in Table 4.1a should be the consolidation of Table 4.1b and Table 4.1c together with dummy variable "*renewable*"; However, comparing the results of Table 4.1a-c, the theoretical relations seemed to failed. The reason for this failure is that I didn't interact "renewable" with my firm dummies and time dummies when using the full dataset. So in this paper, I emphasize the results of separate datasets.

Table 4.2b Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F): Renewables

Dependent variable	abnormal returns for one-factor model			abnormal returns for three-factor model			abnormal returns for five-factor model		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Environmental sentiment factor (F)	0.0390+			0.0379			0.0357		
Google Search factor		0.0105	0.0117		0.0104	0.0122		0.0105	0.0123
Weather factor		0.00292	0.00283		0.00275	0.00262		0.00227	0.00214
Macroeconomic factor	0.0108*	0.0119*	0.0120*	0.0109*	0.0121*	0.0122*	0.0110*	0.0124*	0.0125*
Regulation factor		-0.0170+	-0.0176+		-0.0173+	-0.0182+		-0.0176+	-0.0185+
Renewable energy consumption factor		0.00511	0.00390		0.00575	0.00396		0.00261	0.000798
nonrenewable energy consumption factor		0.0485+	0.0431		0.0520+	0.0439		0.0487+	0.0406
pollution factor			0.00250			0.00374			0.00377
constant	-0.208*	-0.275**	-0.281**	-0.207*	-0.279**	-0.288**	-0.210*	-0.285**	-0.294**
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	4967	4967	4967	4967	4967	4967	4967	4967	4967
F	3.953	3.836	3.86	4.124	4.066	4.156	4.118	4.087	4.199

+ 0.1; *0.05; **0.01; ***0.001

Note: The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 52 renewable energy company dummies (53 renewable energy equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table.

Table 4.2c Regression Results for Abnormal Returns ($\hat{\alpha}_{it}$) and Environmental Sentiment Factor (F): Non-renewables

Dependent variable	abnormal returns for one-factor model			abnormal returns for three-factor model			abnormal returns for five-factor model		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Environmental sentiment factor (F)	0.0497***			0.0488***			0.0471***		
Google Search factor		-0.00672	-0.00381		-0.00662	-0.0031		-0.00680+	-0.00324
Weather factor		-0.00256***	-0.00282***		-0.00268***	-0.00300***		-0.00313***	-0.00345***
Macroeconomic factor	0.00663***	0.00725***	0.00753***	0.00670***	0.00739***	0.00773***	0.00683***	0.00764***	0.00798***
Regulation factor		-0.0140***	-0.0157***		-0.0138***	-0.0159***		-0.0140***	-0.0161***
Renewable energy consumption factor		-0.0190**	-0.0218**		-0.0186**	-0.0220***		-0.0218***	-0.0253***
nonrenewable energy consumption factor		0.0274***	0.0137		0.0307***	0.0141		0.0307***	0.0106
pollution factor			0.00639***			0.00771***			0.00782***
constant	-0.0353	-0.0394	-0.0352	-0.0311	-0.0293	-0.0241	-0.0176	-0.0172	-0.0119
Company fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect controlled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	39092	39092	39092	39092	39092	39092	39092	39092	39092
F	3.682	5.107	3.973	3.540	4.742	3.995	3.575	4.487	4.074

+ 0.1; *0.05; **0.01; ***0.0

Note: The regressions in all columns controlled both company fixed effects and time fixed effects. To control company fixed effects, I included 394 nonrenewable energy company dummies (395 nonrenewable energy equities in total); to control time fixed effects, I included 11 month dummies and 12 year dummies (13 years in total). Their estimates are omitted in the above table.

In Table 4.2a-c, column (1)(4)(7) are regression results for *environmental sentiment factor F*, *macroeconomic factor*, "*renewable*" and interact terms; column (2)(5)(8) are regression results for 5 component subgroup factors (*Google Search factor, weather factor, regulation factor, renewable energy consumption factor* and *nonrenewable energy consumption factor), "renewable"* and interact terms while controlling *macroeconomic factor;* column (3)(6)(9) are regression results for 6 component subgroup factors including *pollution factor* (CO2 emission).

From Table 4.2a we can see that the regression results for subgroup factors are quite robust, since adding more subgroup factors in the regression model didn't change the coefficients for most of the subgroup factors dramatically. The exceptions are *pollution factor* and *nonrenewable energy consumption factor*. When *pollution factor* was included, *nonrenewable energy consumption factor* became less significant (insignificant in column (9)). This reveals the existence of multicollinearity problem, since we also detected the correlation between the two factors exceeds 0.8. F-test rejected the null hypothesis that the coefficients of *nonrenewable energy consumption factor* and *pollution factor* are jointly insignificant. To deal with multicollinearity problem, I kept one of the two factors in model. In the following explanation, I kept the *nonrenewable energy consumption factor*.

The *renewable energy consumption factor* had significantly negative relations with abnormal returns for nonrenewable energy firms (as shown in Table 4.2c), while the relations for renewable energy firms were insignificant; the *nonrenewable energy consumption factor* had significantly positive relations with abnormal returns for both nonrenewable and renewable energy firms. In sum, we saw positive relations between nonrenewable energy consumption and energy firms' abnormal returns while negative relations between renewable energy consumption and energy firms' abnormal returns. These relations are consistent with my hypotheses in section 2, because I assumed that when *environmental concerns* increase, nonrenewable energy consumption would decrease and renewable energy consumption would increase.

From Table 4.2a-c, I can conclude that after including more factors, the abnormal returns for renewable energy firms are still only significantly influenced by *macroeconomic factor*, *environment regulation factor* and other risk factors that failed to be included in the model; while *weather factor*, *energy consumption factor* and *environmental sentiment factor* don't seem to be as influential as for nonrenewable energy firms.

4.3 Summary of Results

From all of the former analysis and tables, we can get a brief summary of the results:

a. All energy firms have negative abnormal returns from January 2004 to October 2016; renewable energy firms get even lower abnormal returns (α) than nonrenewable energy firms, indicating higher company value for renewable energy firms.
b. *Environmental sentiment factor* has positive relations with abnormal returns for nonrenewable energy firms, while the relations for renewable energy firms are only marginally significant under CAPM model. The decrease of *environmental sentiment factor* indicates an increase of public's environmental concerns.

c. *Macroeconomic factor* always has significantly positive relations with abnormal returns for both renewable and nonrenewable energy firms, and the relations never differentiate between renewables and non-renewables, no matter which dataset I used.

d. For renewable energy firms, only *environment regulation factor* continues to have significantly negative effects on abnormal returns. The total negative abnormal returns $\hat{\alpha}_{it}$ of renewables are mainly from the effects of environment regulation, economic conditions (the control variable) and other risk factors that we failed to include in model.

e. For nonrenewable energy firms, except for *Google Search factor*, all other component subgroup factors have significant relations with abnormal returns. The *environment regulation factor* (negative coefficient), *weather factor* (negative coefficient), *renewable energy consumption factor* (negative coefficient), *nonrenewable energy consumption* (positive coefficient) are all found to be statistically significant.

In summary, although renewable and nonrenewable energy firms have comparable market beta, renewable energy firms provide lower stock returns and gain higher company value than non-renewables. For renewable energy firms, the total predicted abnormal returns $\hat{\alpha}_{it}$ are mainly from *environment regulation factor* and other risk factors that omitted in the model (as equation(13)); for nonrenewable energy energy firms, this total predicted abnormal returns $\hat{\alpha}_{it}$ mainly come from *environmental sentiment factor*, *weather factor, environment regulation factor* and *energy consumption factors*.

So when holding all other risk factors constant, facing with same level of environment-related risks, renewable energy firms would have higher total abnormal returns than nonrenewable energy firms.

5. Conclusions

To study the relationship between environmental sentiment and stock performance in energy sector, I constructed an *environmental sentiment factor* and 7 subgroup factors using the Dynamic Factor Model and calculated the predicted abnormal returns $\hat{\alpha}_{it}$ using asset pricing models. Then I implemented Newey-West regressions to equation (13). With the panel dataset for 448 U.S. energy equities, I found that all energy firms got lower actual returns than expected; renewable energy firms had higher company values due to the lower abnormal returns from January 2004 to October 2016; the predicted abnormal returns for nonrenewable energy firms are more sensitive about environmental sentiment and weather conditions, while the predicted abnormal returns for renewable energy firms are more closely related with environment regulation and other risks that omitted in the regression.

Therefore, I conclude that when the weather conditions become worse or when the public's environmental concerns become higher, investors would favor renewable energy equities and be upset about the nonrenewable energy ones. Because under same environment risk factors, renewable energy equities give higher abnormal returns, which adds more attractiveness for renewable energy firms' stocks.

So pessimistically, more investors would sell nonrenewable energy stocks, driving up the cost of equity; therefore, as environment conditions getting worse or public's environmental concerns getting higher, there would be less nonrenewable energy firms to raise funds in stock markets.

Bibliography

Askitas, Nikolaos, and Klaus F Zimmermann. 2009. "Google Econometrics and Unemployment Forecasting." Applied Economics Quarterly 55 (2): 107–20.

Amburgey, Jonathan W., and Dustin B. Thoman. 2012. "Dimensionality of the new ecological paradigm: Issues of factor structure and measurement." Environment and Behavior 44, no. 2: 235-256.

Baker, Scott, and Andry Fradkin. 2011. "What drives job search? Evidence from Google search data." Discussion Papers: 10-020.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz. 2005. "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." The Quarterly Journal of Economics 120 (1): 387–422.

Brunel, Claire, and Arik Levinson. 2016. "Measuring the Stringency of Environmental Regulations." Review of Environmental Economics and Policy 10 (1): 47–67.

Burstein, Paul. 2003. "The Impact of Public Opinion on Public Policy: A Review and an Agenda." Political Research Quarterly 56 (1): 29–40.

Chang, Shao-Chi, Sheng-Syan Chen, Robin K. Chou, and Yueh-Hsiang Lin. 2008. "Weather and Intraday Patterns in Stock Returns and Trading Activity." Journal of Banking & Finance 32 (9): 1754–66.

Chava, Sudheer. 2014."Environmental Externalities and Cost of Capital." *Management Science* 60 (9): 2223–47.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. "In Search of Attention." The Journal of Finance 66 (5): 1461–99.

D'Amuri, Francesco, and Juri Marcucci. 2010. "'Google It!' Forecasting the US Unemployment Rate with A Google Job Search Index." SSRN Scholarly Paper ID 1594132. Rochester, NY: Social Science Research Network.

Dimpfl, Thomas, and Stephan Jank. 2016. "Can Internet Search Queries Help to Predict Stock Market Volatility?" SSRN Scholarly Paper ID 2740779. Rochester, NY: Social Science Research Network.

Dunlap, Riley E., Kent D. Van Liere, Angela G. Mertig, and Robert Emmet Jones. 2000. "New Trends in Measuring Environmental Attitudes: Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale." Journal of Social Issues 56 (3): 425–42.

Dunlap, Riley E., and Kent D. Van Liere. 1978. "The 'New Environmental Paradigm.'" *The Journal of Environmental Education* 9 (4): 10–19.

Elshahat, Islam, Martin Freedman, and Ahmed Elshahat. 2015. "Are Index Measures Better than Individual Measures in Assessing the Association between Environmental and Financial Performance? A

Multivariate Analysis Study.” *International Journal of Critical Accounting* 7 (2): 191–212.

Flammer, Caroline. 2015. “Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach.” *Management Science* 61 (11): 2549–68.

Fama, Eugene F., and Kenneth R. French. 1993. “Common Risk Factors in the Returns on Stocks and Bonds.” Journal of Financial Economics 33 (1): 3–56.

———. 2007. “Disagreement, Tastes, and Asset Prices.” Journal of Financial Economics 83 (3): 667–89.

———. 2015. “A Five-Factor Asset Pricing Model.” Journal of Financial Economics 116 (1): 1–22.

García, Diego. 2013. “Sentiment during Recessions.” The Journal of Finance 68 (3): 1267–1300.

Gerlach, Stefan, and Matthew S. Yiu. 2005. “A Dynamic Factor Model of Economic Activity in Hong Kong.” SSRN Scholarly Paper ID 735830. Rochester, NY: Social Science Research Network.

Geweke, John.1978. The dynamic factor analysis of economic time series models. Social Systems Research Institute, University of Wisconsin-Madison.

Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, and Larry Brilliant. 2009. “Detecting Influenza Epidemics Using Search Engine Query Data.” *Nature* 457 (7232): 1012–14.

Hamilton, Sally, Hoje Jo, and Meir Statman. 1993. “Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds.” *Financial Analysts Journal* 49 (6): 62–66.

Herremans, Irene M., Parporn Akathaporn, and Morris McInnes. 1993. “An Investigation of Corporate Social Responsibility Reputation and Economic Performance.” *Accounting, Organizations and Society* 18 (7): 587–604.

Heinkel, Robert, Alan Kraus, and Josef Zechner. 2001. “The Effect of Green Investment on Corporate Behavior.” Journal of Financial and Quantitative Analysis 36 (4): 431–49.

Kahn, Matthew E., and Matthew J. Kotchen. 2010. “Environmental Concern and the Business Cycle: The Chilling Effect of Recession.” Working Paper 16241. National Bureau of Economic Research.

Khairollahi, Farshid, Farhad Shahveisi, Alireza Vafaei, and Mohammad Alipour. 2016. “From Iran: Does Improvement in Corporate Environmental Performance Affect Corporate Risk Taking?” *Environmental Quality Management* 25 (4): 17–33.

Klassen, Robert D., and Curtis P. McLaughlin. 1996. “The Impact of Environmental Management on Firm Performance.” *Management Science* 42 (8): 1199–1214.

Laplante, Benoît, Susmita Dasgupta, and Nlandu Mamingi. 1998. “Capital Market Responses to Environmental Performance in Developing Countries.” SSRN Scholarly Paper ID 604915. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=604915.

Lepori, Gabriele M. 2016. “Air Pollution and Stock Returns: Evidence from a Natural Experiment.” *Journal of Empirical Finance* 35 (January): 25–42.
Levy, Tamir, and Joseph Yagil. 2011. “Air Pollution and Stock Returns in the US.” *Journal of Economic Psychology* 32 (3): 374–83.
Liere, Kent D. Van, and Riley E. Dunlap. 1980. “The Social Bases of Environmental Concern: A Review of Hypotheses, Explanations and Empirical Evidence.” *Public Opinion Quarterly* 44 (2): 181–97.
Lorraine, N. H. J., D. J. Collison, and D. M. Power. 2004. “An Analysis of the Stock Market Impact of Environmental Performance Information.” *Accounting Forum* 28 (1): 7–26.
Ng, Alex and Zheng, Di. 2016. “The Investment Performance of Green Energy Firms.” Working paper, Department of Accounting and Finance, School of Business and Economics, Thompson Rivers University, Canada.
O'Connor, Brendan, Ramnath Balasubramanyan, Bryan R. Routledge, and Noah A. Smith. 2010. "From tweets to polls: Linking text sentiment to public opinion time series." ICWSM 11, no. 122-129: 1-2.
Saeidi, Sayedeh Parastoo, Saudah Sofian, Parvaneh Saeidi, Sayyedeh Parisa Saeidi, and Seyyed Alireza Saaeidi. 2015. “How Does Corporate Social Responsibility Contribute to Firm Financial Performance? The Mediating Role of Competitive Advantage, Reputation, and Customer Satisfaction.” *Journal of Business Research* 68 (2): 341–50.
Soretz, Susanne, and Ingrid Ott. 2015. “Green Attitude and Economic Growth.” In . ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft. https://www.econstor.eu/handle/10419/113208.
Stock, James H., and Mark W. Watson. 1989. “New Indexes of Coincident and Leading Economic Indicators.” NBER Macroeconomics Annual 4 (January): 351–94.
———. 1998. “A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series.” Working Paper 6607. National Bureau of Economic Research.
Shove, Elizabeth. 2010. "Beyond the ABC: climate change policy and theories of social change." Environment and planning A 42, no. 6: 1273-1285.
Tetlock, Paul C. 2007. “Giving Content to Investor Sentiment: The Role of Media in the Stock Market.” The Journal of Finance 62 (3): 1139–68.
Tsen, Chyong-Huey, Grace Phang, Haslinda Hasan, and Merlyn Rita Buncha. 2006. “Going Green: A Study of Consumers’ Willingness to Pay for Green Products in Kota Kinabalu.” International Journal of Business and Society; Sarawak 7 (2): 40–54.