DEVELOPING MARKET SENTIMENT INDICATORS FOR COMMODITY PRICE FORECASTING USING MACHINE LEARNING

by

Tariq Sohail

A Thesis Submitted to the Faculty of Graduate Studies

The University of Manitoba

In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Department of Agribusiness and Agricultural Economics

University of Manitoba

Winnipeg, Manitoba, Canada

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

****

COPYRIGHT PERMISSION

Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree of

Master of Science

ABSTRACT

The objective of this study is to develop a market sentiment model for financial markets using machine learning, and to illustrate these methods using commodity price data. A market sentiment model may capture the fundamental and crowd psychology of the market, through a variable that uses positive and negative words and phrases. The commodity price used is the daily price of the spot crude oil exchange-traded fund (ETF), United States Oil Fund (USO). The forecasting power of the market sentiment model is compared with a traditional autoregressive model. The results showed that the autoregressive models did not have significant forecasting power for the oil data over the time period examined and the addition of the sentiment model did not improve the forecasting power. Machine learning is a relatively new forecasting method. Therefore, further research on this topic is needed before any firm conclusions can be drawn regarding the effectiveness of this approach.

Keywords: artificial intelligence, market sentiment indicator, machine learning, autoregressive model, crude oil, forecasting, commodity prices, Python, Amazon Mechanical Turks, Beautiful Soup, Natural Language Toolkit, Mechanize.

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Milton Boyd for the encouragement, support, inspiration and words of wisdom that he has so generously given me throughout the entire program and the last years of my Bachelor's program as well. I am also grateful for the assistance and guidance from committee members Dr. Lysa Porth and Dr. Tulsi Thulasiram. I would like to give my thanks and gratitude to Brock Porth whose support throughout this project was extremely helpful, I want to also express my gratitude to Alex Sahu at Goldman and Sachs, New York, for his invaluable input. I am also grateful of the support of Winnipeg Commodity Exchange Fellowship and the Department of Agribusiness and Agricultural Economics at the University of Manitoba.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1. BACKGROUND OF THE RESEARCH PROBLEM

The widely accepted view of the efficient market hypothesis was first introduced by Fama (1965). It holds that markets incorporate new market information into the stock price immediately which makes the market efficient. This hypothesis implies that it is difficult for researchers and financial professionals to build a successful model for forecasting stock and commodity prices, as the current market reflects all past and current information (Fortuny, Smedt, Martens, & Daelemans, 2014). However, with continued advances in computer technology software, there is renewed interest in using these advances to build a model to forecast stock and commodity prices. Therefore, the objective of this study is to develop a market sentiment model for financial markets using machine learning, and to illustrate these methods using commodity price data.

Market sentiment in this study refers to a variable using positive and negative news (words and phrases) for a stock in order to create a sentiment indicator variable. The sentiment indicator variable will be positive if the words and phrases are positive and negative if the words and phrases are negative. For example, theoretically, if the sentiment indicator variable increases, then the stock price would be expected to increase. Market sentiment may capture the fundamentals of a market or its tone and psychology.

Previous literature has shown that the problem of financial price forecasting is complex (Hajek, Oleg, & Myskova, 2013) and there is an extensive amount of literature on the relationship of news market sentiment analysis and returns in the equity market

(Kearney & Liu, 2014). The study here will be one of the first studies to examine the relationship between market sentiment generated from machine learning and commodity prices. In this case the data used, USO, the exchange-traded fund (ETF) tracks West Texas Intermediate (WTI) crude oil futures returns.

1.2. METHODS

There are various methods that market sentiment indexes are created from and they are derived from different sources of information. For this study, the market sentiment picture of the market is created by using full news articles from a news site that caters primarily to financial and commodity news, Seeking Alpha.

For this study, three different kinds of information sources that market sentiment is generated from are defined broadly as: 1) corporation-expressed market sentiment, 2) media-expressed market sentiment and 3) internet-expressed market sentiment are used (Liu & Kearney, 2014). The first, corporation expressed market sentiment is based on corporate public disclosures which are issued by companies entailing current and forward looking textual statements which can be used by researchers.

The second, media-expressed market sentiment, is news information emanating from news websites such as Bloomberg News, Financial Times, The Wall Street Journal, and Reuters. The third, internet-expressed market sentiment is based on textual information obtained from chatrooms and comments on financial news sites such as Seeking Alpha. Although, some researchers consider this noise in market sentiment analysis, it may still hold influence in constructing daily market sentiment (Liu &

Kearney, 2014). However, this study uses only Seeking Alpha because they allow their websites to be repetitively pinged by a robot.

In the case of USO, the ETF used here, corporation, media, and internet-expressed market sentiment sources are used. Once targeted financial news sources are identified, a web-scraping software, Python's library called *Mechanize* is used. It emulates a browser, and downloads articles which are parsed in order to extract relevant details which includes title, description, keywords and then the actual content of the article. Once articles are downloaded and parsed, they are stored in a manner that is efficient and easy to access for use. After storing the data, the articles are cleaned and classified in this case as either positive or negative sentiment. The next step is training the model using the bag of words model approach where the model is trained manually through many articles by using *Amazon Mechanical Turks*, a crowdsourcing platform used for large human intelligence tasks.

1.3. DATA: BACKGROUND ON UNITED STATES OIL FUND (USO)

Daily oil price data used is United States Oil Fund, LP (USO), "is a commodity pool that continuously issues common shares that can be purchased and sold on the NYSE Arca". The objective of the fund is for the "daily changes in percentage terms of its shares' per share net asset value to reflect the daily changes in percentage terms of the spot price of light, sweet crude oil delivered to Cushing, Oklahoma." (USO Prospectus, 2016). As of April 30, 2016, it had $3.63 billion Assets Under Management (AUM) and 52.73% of its Total Assets are the nearest WTI crude oil futures (June 2016). The rest of the Total Assets are held in Treasuries and other cash equivalents. In order to engage in meaningful research, USO was chosen because it tracks the nearest month WTI crude oil futures contract listed on NYMEX which allows for the short-term fluctuations in the crude oil market. Data on USO prices were retrieved from Yahoo Finance by downloading daily closing price data starting April 4, 2016 to July 8, 2016. The prices are then converted into natural logarithmic returns (e.g. percentage price changes) and used in the analysis.

1.4. LITERATURE REVIEW

A popular study on the explanation of movements in stock returns was undertaken by Roll (1988). It used public firm-specific news events, economic news and news in the same industry and found that there was a modest relationship between prices and news. The study found that the average adjusted R^2 was 0.20 with daily data and 0.35 with monthly data. An interesting finding of the study also showed that explanatory power is not industry related but the coefficients in the regressions were different for different industries. Since this study was published, there has been limited success in the finance literature to show strong explanatory power between prices and news. However, further research on this subject of forecasting equity returns has shown that using artificial intelligence, soft computing and machine learning may allow for more accurate for forecasting stock returns (Zhang Patuwo & Hu, 1998).

Early research by Boyd and Kaastra (1996) has provided an introductory guide into the designing of a neural network forecasting model for financial and economic time series data. Neural networks are often used in pattern recognition, classification and forecasting, all of which has practical use in finance. In their study, they describe and outline an eight-step approach to designing a neural network forecasting model for financial time series data.

Another research study was done on examining the relationship between news by using the advancements in textual analysis and prices (Boudoukh, Feldman, Kogan & Richardson, 2012). In this study they were able to identify relevant news by type and tone and were able to find that there is a strong relationship between stock price changes and publicly available information. They found that the variance ratio of returns on news versus no news are 120% higher versus only 20% for unidentified news versus no news while stock price reversals occur on no news days (Ibid.). They also found that tone, content and relevance of news are important for creating a market sentiment indicator. However, there is difficulty in classifying the right news, which is important and which is irrelevant, hence the construction of stock market sentiment is close to accurate but not exact yet.

Further, a study done on stock price forecasting used corporate annual reports by using neural network and support vector regression. In this study (Hajek, Oleg, Myskova, 2013) they tried to show that the long-run behavior of stock price can be predicted more accurately when using qualitative textual information hidden in annual reports. They developed a qualitative market sentiment indicator from annual reports, though is a limitation in itself as it does not account for other financial news information that could be affecting the stock market sentiment. However, they found that large variance exists in the next year’s stock price return using the change in market sentiment found in the annual report (Ibid.).

Different proxies for market sentiment have also been used in studies to examine stock returns by using consumer confidence as a proxy for individual investor market sentiment. A study in 18 different industrialized countries, used market sentiment to forecast aggregate stock market returns. (Schmeling, 2009). Results showed for example, when the market sentiment is high, the stock returns tend to be low and vice versa. The study also found that this relationship holds for value stocks, growth stocks, small stocks and different forecasting horizons (Ibid.). Despite the study using a proxy for investor market sentiment, it nonetheless demonstrated that there might be forecasting power through using market sentiment analysis.

Additionally, a host of studies have been conducted in the past several years that have given way to further research. For example, Antweiler and Frank (2004) applied machine learning to examine internet stock message boards and used linear regression and volatility models to predict equity returns. They found that a positive shock to message board postings predicts negative returns the next day and an uptick in message postings induces volatility the next day. In the study, they studied the effect of more than 1.5 million messages posted on Yahoo! Finance and Raging Bull. In summary, the effect on stock returns was found to be statistically significant but economically small. Further, it was found that message board postings on equities reflect public information very rapidly.

Tetlock (2007) studied the period of 1984-1999 using a dictionary-based GI/Harvard method for analyzing daily news articles from Wall Street Journal, a popular financial news provider. The study used Value-at-Risk (VAR) model and found that high values of media pessimism induce downward pressure on market prices and low values of media pessimism lead to temporarily high market trading volume.

Further, a study (Huang et al., 2013) used a Naïve Bayes machine learning approach which is a similar text classification approach as the market sentiment used in this study. Results showed that using analyst reports from 1995-2008, investors place much more weight on negative than positive statements. They also found in this study that negative statements are more informative than positive ones about a firm's future performance as well. The models that provided these results were a linear regression model and an event-study method.

Chen et al. (2013) also used internet messages from 2005-2011 used dictionary-based L&M for content analysis from Seeking Alpha website and linear regression models which showed that the fraction of negative words contained in the articles and comments on the website negatively predict subsequent stock returns. They also showed that articles and comments predict future stock returns much more strongly than news articles alone.

Finally, Thulasriam et al. (2015) found that Twitter feed may increase predictive power of forecasting returns on the Dow Jones Industrial Average. However, there is another study that indicates that Twitter feed is possibly driving market returns and influencing the market (Liew and Budavari, 2016). Both conclusions show that maybe the market sentiment approach is useful for forecasting market prices while shedding some insight into whether the markets rapidly digest new information.

Therefore, the results of this literature lends support to the argument for using machine learning and artificial intelligence for creating market sentiment indicators as an alternative approach to traditional forecasting. Although, some of the studies cited have used a variation of methods that are similar to the method developed in this study, they illustrate that there is potential in approaches other than the traditional forecasting models.

The remainder of this study is organized as follows. Chapter 2 discusses the data used in the study, and Chapter 3 discusses the methods used in the study with a particular focus on the measures of market news sentiment and how it was created. Chapter 4 presents the empirical results of the basic relationship between market news sentiment and USO (crude oil) returns. Chapter 5 summarizes the research study and explores further research related to the study.

CHAPTER 2: DATA

2.1 UNITED STATES OIL (USO) FUND PRICE DATA

USO is an exchange-traded fund (ETF) and was introduced on April 10, 2006. In general, ETFs have been some of the fastest growing instruments in the investment management business. ETFs provide true accessibility, lower costs, transparency, liquidity, and tax efficiency, and even the average investor now can gain exposure to commodities. In cases of ETFs that track commodity futures, they also require no margin, no futures account, no approval from a broker required and are traded exactly like stocks on the stock exchange (Hill et al., 2015)

USO daily closing prices were obtained from Yahoo Finance website ranging from April 4, 2016 to July 8, 2016. There are 62 observations in the data and it excludes previous years as market sentiment data was not readily available. The logarithmic returns for every day were calculated for the USO using four decimal places and were used in all the econometric models with and without the market sentiment values. The USO daily returns skewness is 0.0052 which indicates the distribution is close to normal (Table 2.1). The minimum and maximum returns are -5.228% and 10.286%, respectively. Results showed limited correlation between the five lagged returns (Table 2.2). The USO daily returns when graphed implied the returns were normal (Figure 2.1). The scatter plot also implied that the USO returns had no significant outliers (Figure 2.2). A histogram of the frequency of USO returns shows that the returns were leaning towards the negative area (Figure 2.3).

2.2 MARKET SENTIMENT DATA INDICATORS

The market sentiment data indicators were constructed by using machine learning methods and applications. These included Python's Natural Language Toolkit, Python's Beautiful Soup library and Amazon's Mechanical Turks. The website used was Seeking Alpha where the data content was obtained from using these above tools and market sentiment indicators were created. This is further explained in chapter 3, methodology. In summary, market sentiment indicators were created for only 62 consecutive days.

TABLE 2.1 DESCRIPTIVE STATISTICS: DAILY USO RETURNS AND MARKET SENTIMENT

Variables	# Obs.	Mean	Std-dev	Min	Max	Skewness	Kurtosis
USO (Daily Return)	62	0.2842419	2.55631	-5.228 (a)	10.286 (b)	0.0052	0.0027
Market Sentiment	62	4.896403	3.24287	-2.275	16.125	0.0435	0.0446

- Notes: Obs = observations; n = 62
  Yahoo Finance website http://finance.yahoo.com/
- (a): for example, -5.228 is -5.228%
- (b): for example, 10.286 is 10.29%
- Data: USO daily logarithmic return (x 100) is a approximate percentage price change. USO is a crude oil ETF (exchange traded fund). Prices are obtained from Yahoo Finance ranging from April 4, 2016 to July 8, 2016.
- The data was found to be stationary using the Dickey-Fuller ADF test in USO daily returns and market sentiment data.

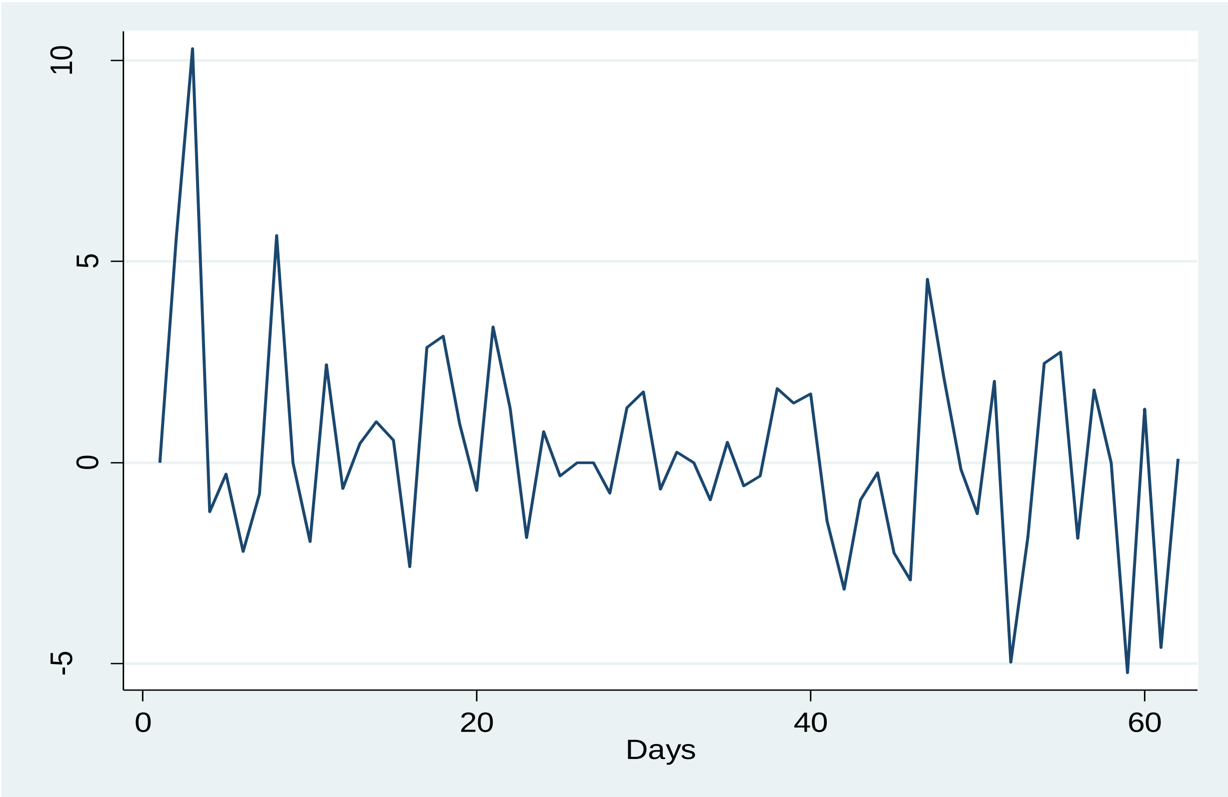
TABLE 2.2 CORRELATION MATRIX OF USO DAILY RETURNS, CRUDE OIL ETF (APRIL 4, 2016 TO JULY 8, 2016)

Parameters	**Return**	**Lag Return 1**	**Lag Return 2**	**Lag Return 3**	**Lag Return 4**	**Lag Return 5**
Return	1.0000					
Lag Return 1	-0.0678	1.0000				
	(0.671)					
Lag Return 2	-0.2063	-0.0681	1.0000			
	(0.172)	*(0.678)*				
Lag Return 3	0.0546	-0.1852	-0.0867	1.0000		
	(0.717)	(0.176)	*(0.601)*			
Lag Return 4	-0.1003	0.0497	-0.3080	0.0.983	1.0000	
	(0.456)	*(0.701)*	*(0.039)*	*(0.512)*		
Lag Return 5	-0.0929	-0.0997	0.0464	-0.2712*	0.0844	1.0000
	(0.492)	*(0.460)*	*(0.732)*	*(0.041)*	*(0.532)*	

*** Correlation is significant at the 0.05 level (2-tailed).**

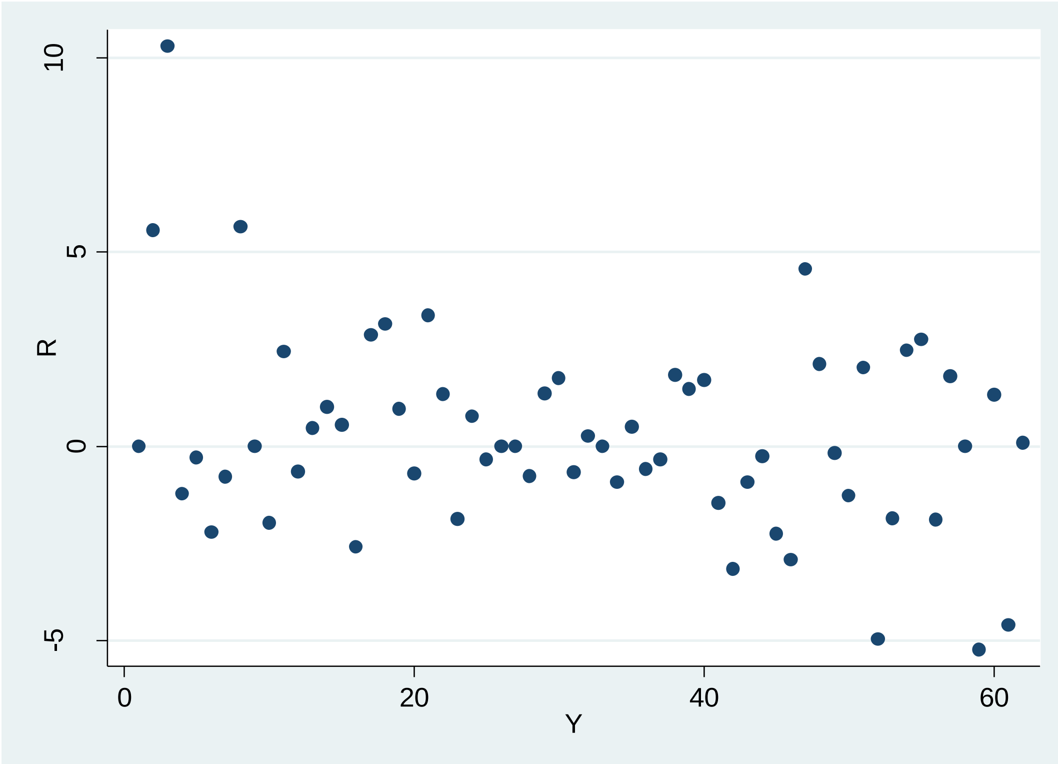
N = 62.

Figure 2.1 USO DAILY RETURNS, CRUDE OIL ETF (APRIL 4, 2016 TO JULY 8, 2016)



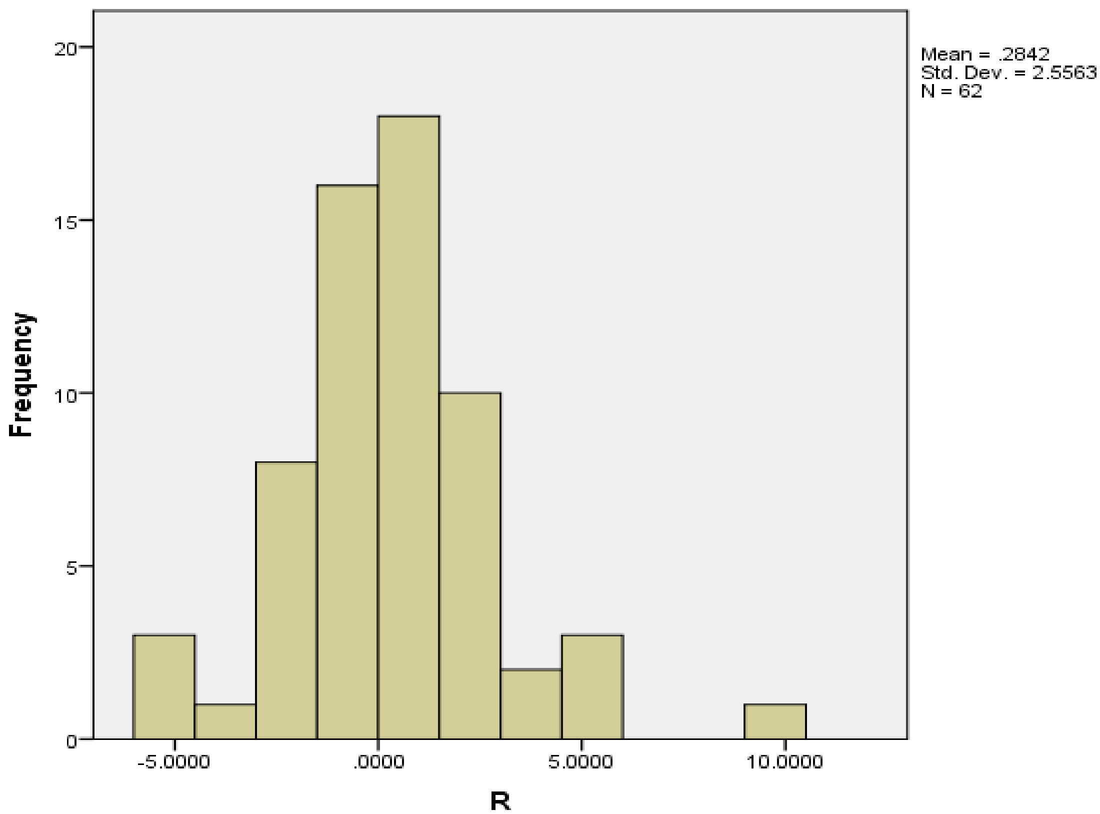
Notes: D = # of days starting April 4, 2016 to July 8, 2016
R = USO daily logarithmic returns (x 100) which is a approximate percentage price change
N = 62

FIGURE 2.2 SCATTER PLOT OF USO DAILY RETURNS (APRIL 4, 2016 TO JULY 8, 2016)



Notes: Y = # of days starting April 4, 2016 till July 8, 2016
N = 62.

FIGURE 2.3 HISTOGRAM OF USO DAILY PERCENTAGE RETURNS, CRUDE OIL ETF (APRIL 4, 2016 TO JULY 8, 2016)



Notes: R = daily returns of USO

CHAPTER 3: METHODOLOGY

3.1 AR(5) MODEL DESCRIPTION

This study develops a market sentiment indicator for USO and examines whether it has power in predicting crude oil prices (USO ETF). First, a base model of order-five, AR(5), of 5 lagged daily returns of USO is constructed (see equation 3.1) in order to forecast returns. The rationale for using five lags are that there are five trading days in a week and it is reasonable to assume that in a given week, prices from the day before have tremendous impact on the next day's prices. This model was to examine the explanatory power of five lagged returns for USO returns using an F-test.

The second AR(5) model is then estimated which incorporates a one period lagged market sentiment variable to see whether it has added more predictive power to the model by considering the F-test with the previous AR(5) model's F-test statistic (see equation 3.2). The third AR(5) model estimated used also included lagged market sentiment, for a total of 10 variables, 5 AR plus 5 sentiment lagged (see equation 3.3). This is to identify whether the forecasting power of the model has been further enhanced. The F-test is an indicator of the explanatory power of the model.

The first AR(5) model without market sentiment can be expressed as:

$$R_t = C_0 + R_1 x_{t-1} + R_2 x_{t-2} + R_3 x_{t-3} + R_4 x_{t-4} + R_5 x_{t-5} \tag{3.1}$$

Where: R_t *is the current logged return and* $R_p x_{t-p}$ *is the independent lagged price variable*

The second AR(5) model with market sentiment can be expressed as:

$$R_t = C_0 + R_1x_{t-1} + R_2x_{t-2} + Rx_{t-3} + R_4x_{t-4} + R_5x_{t-5} + sent_{1ty} \qquad (3.2)$$

Where: $sent_{1ty}$ *is the daily market sentiment variable and will be compared with the model without sentiment using the F-test*

The third AR(5) model with lagged market sentiment can be expressed as:

$$R_t = C_0 + R_1x_{t-1} + R_2x_{t-2} + R_3x_{t-3} + R_4x_{t-4} + R_5x_{t-5} + \mathrm{sent}_{1yt-1} + \mathrm{sent}_{2yt-2} + \mathrm{sent}_{3yt-3} + \mathrm{sent}_{4yt-4} + \mathrm{sent}_{5yt-5} \qquad (3.3)$$

Where: $sent_{pyt-p}$ *is the lagged daily market sentiment variable* and will be *compared with the model with sentiment using the F-test*

3.2 MARKET SENTIMENT INDICATORS

The market sentiment was constructed in the following steps:

1. In order to test the hypothesis, a large amount of data set was required to create market sentiment indicator, on which a classifier was trained and classified.

2. Full news articles were used for this purpose from financial news site as this has been shown to be more credible and powerful in creating a market sentiment indicator (Zhang and Skiena, 2010).

3. To acquire the data set from targeted news sites, a web-scraping software was developed to access Seeking Alpha.

4. Not all sites allowed users to automatically download their articles. To overcome this problem, an existing python library called *Mechanize* was used which emulates a browser, so every website accessed through this library believes they are being accessed by a human. All modern web servers have a “Robots.txt protocol” or “Robots exclusion standard” which is a set of rules which the server uses to verify the “User-Agent” and forbid access to automated programs. The issue was resolved only for Seeking Alpha.

5. Seeking Alpha provided RSS news feed which could be listened to programmatically to download articles. Since Seeking Alpha organizes articles by stock/ETF, all articles in the USO category were assumed to be relevant. Further, a scraper was designed to read in its raw format and extract the relevant details from the article.

6. In order to parse through large scale news text on Seeking Alpha, a Python HTML parsing library known as *BeautifulSoup* was used which re-creates the tree structure of HTML code from the raw source code without data loss.

7. The text also showed how many comments were for each single article. The hypothesis was that more comments on an article means more weight it will have on USO.

8. Once downloaded and parsed, it was stored in an efficient manner which was user-friendly. The data sets were stored locally and cached the classification which makes the analysis steps more efficient.

9. After the data was cleaned, the text was classified into categories of positive and negative using Bayesian classification technique in conjunction with the bag of words learning model. Text classification essentially classifies articles by their content and each article is represented by a bag of words (a set which has repeating elements). This method identifies which words occur and counts the frequency of how many times it is repeated in the article. For example, consider an article A whose class is given by R. In the case of R, there are two classes, positive and negative. This can be expressed mathematically using Bayes Theorem:

$$p(a|r) = p(r|a)\ p(a) / p(r) \tag{3.4}$$

10. Phrases were classified using the SentiWordNet database, a publicly available resource for opinion mining. SentiWordNet is used by Hamouda (2015) in scoring movie review with 68% accuracy which is higher than classical classification methods. Classification tasks were all done using Python's open source *Natural Language Toolkit (NLTK).*

11. Before the bag of words model can be applied to the data set, it has to be trained by creating positive clusters and negative clusters of words. This

is done by taking a subset of the data, which are many articles and have them manually classified as positive or negative.

12. In training this model, *Amazon Mechanical Turks*, a crowdsourcing platform where large human intelligence tasks like this can be done, was used. Once manually classified, the model can be then trained using the bag of words approach. The next step is to split the words of each article in a set of articles into words and count the frequency of each word to create a distribution. After this step, there are two "bags" of words, one positive and another for negative.

13. The positive and negative market sentiment indicator was then summed up for each and a total market sentiment indicator was calculated. This net sentiment indicator was then used in the three AR models.

In this model, it captures the frequency of the words and not just their presence or absence. Stop words and conjunctions are removed from these articles when stored which have no meaning to estimating the market sentiment of a stock. Once a probability distribution has been calculated for a particular word, a market sentiment score is decided upon to give to that word. The market sentiment rating methodology was developed by Sahu et al. (2015) and was provided for the purpose of this paper. The sentiment rating methodology utilizes Python's Natural Language Toolkit (Python, 2015) and Princeton University's open source WordNet database (Miller, 1995).

CHAPTER 4: RESULTS

4.1 AR MODEL WITHOUT MARKET SENTIMENT

The autoregressive AR(5) model used crude oil USO daily returns (but without the market sentiment variable) in order to identify a base model for forecasting power so it can be later compared to the other two sentiment models where market sentiment variable is included. The regression estimates without the market sentiment variable (Table 4.1) shows mostly statistically insignificant coefficients, indicating very limited forecasting power. Table 4.2 shows that the AR(5) model without the market sentiment variable is not statistically significant at the 5 percent level (F value 0.93), again indicating very limited forecasting power.

The residuals were found to have some skewness, with a skewness value of 0.1710 (Table 4.3). The kurtosis of the residuals was found to be 0.1734 which indicates there is positive kurtosis (Table 4.3). The model was also found to have basically no autocorrelation in the residuals as shown in the Durbin-Watson statistic, in Table 4.2 and a plot of residuals shown in Figure 4.1

4.2 AR MODEL WITH LAGGED MARKET SENTIMENT VARIABLE

A lagged market sentiment variable may help explain today's return, for forecasting purposes. However, the results showed that the AR(5) model with a lagged sentiment variable was statistically insignificant (F value 0.79) at the five percent level, indicating very limited forecasting power (Table 4.2). The model also showed coefficients to be not significant at the 5 percent level, again indicating very limited forecasting power (Table 4.4). The model's residuals were found to have skewness of

0.1845 and kurtosis of 0.1874 indicating some positive skewness and positive kurtosis as shown in Table 4.3.

4.3 AR MODEL WITH FIVE LAGGED MARKET SENTIMENT VARIABLES

When the third AR(5) model was used along with five lagged market sentiment variables. The F-test statistic was 0.79 and not significant at the 5 percent level, indicating very limited forecasting power (Table 4.2). The coefficients were also found to be statistically insignificant at the 5 percent level as shown in Table 4.5, again indicating very limited forecasting power. Further, the model's residuals had a skewness value of 0.2282 and a kurtosis of 0.4916 as shown in Table 4.3, indicating some skewness and kurtosis.

4.4 SUMMARY OF RESULTS

The results showed that all three models were not statistically significant at the 5 percent level, as indicated by the F-test. AR(5) models were used to test the hypothesis regarding forecasting power of the daily market sentiment indicators. The addition of the market sentiment indicators also did not increase the forecasting power of the autoregressive models. The purpose of the study was to develop market sentiment indicator model. Models and software included, Python Natural Language Toolkit, Python Beautiful Soup Library, and Amazon Mechanical Turks. These methods showed relatively limited forecasting power, and are insightful for future study of machine learning with applications to commodities and stocks.

The results were not surprising as this study was a first step in developing market sentiment indicators and applying them to forecasting of commodity prices. Overall, there could be potential in using these market sentiment indicators for price forecasting. Also, while it may be that the market may have been difficult to predict using an AR model for the period of the data used, other econometric models may show otherwise.

TABLE 4.1 ESTIMATES FOR AR(5) WITHOUT MARKET SENTIMENT VARIABLE, CRUDE OIL ETF

Parameters	Estimate	Std. Error	t value	p value
Constant	0.1856342	0.3034895	0.61	0.543
Lag Return 1	-0.0812494	0.1389047	-0.58	0.561
Lag Return 2	-0.2708005	0.1473583	-1.84	0.072
Lag Return 3	0.0129712	0.1265373	0.10	0.919
Lag Return 4	-0.1515417	0.1258219	-1.20	0.234
Lag Return 5	-0.0619866	0.1254743	-0.49	0.623

N = 62.

TABLE 4.2 OVERALL MODEL EQUATION TEST STATISTICS, CRUDE OIL ETF

Model	prob. > F	R^2	Adj. R^2	Durbin –Watson
AR(5) without market sentiment	0.93 (0.4674)	0.0838	-0.0060	1.939844
AR(5) with one period lagged market sentiment	0.77 (0.5952)	0.0848	-0.0250	1.948989
AR(5) with five period lagged market sentiment	0.79 (0.6373)	0.1467	-0.0387	1.914082

N = 62.

TABLE 4.3 SKEWNESS AND KURTOSIS TESTS FOR NORMALITY IN RESIDUALS, CRUDE OIL ETF

Model	Skewness	Kurtosis
AR(5) without market sentiment	0.1710	0.1734
AR(5) with one lagged market sentiment	0.1845	0.1874
AR(5) with five lagged market sentiment	0.2282	0.4916

N = 62.

TABLE 4.4 ESTIMATES FOR AR(5) WITH ONE LAGGED MARKET SENTIMENT VARIABLE, OIL ETF

Parameters	Estimate	Std. Error	t value	p value
Constant	0.2939327	0.5527884	0.53	0.597
Lag Return 1	-0.0857173	0.1414885	-0.61	0.547
Lag Return 2	-0.2605473	0.1549906	-1.68	0.099
Lag Return 3	0.0129855	0.1277257	0.10	0.919
Lag Return 4	-0.1474271	0.1282011	-1.15	0.256
Lag Return 5	-0.0681631	0.129343	-0.53	0.601
Lag Sentiment 1	-0.0220693	0.093769	-0.24	0.815

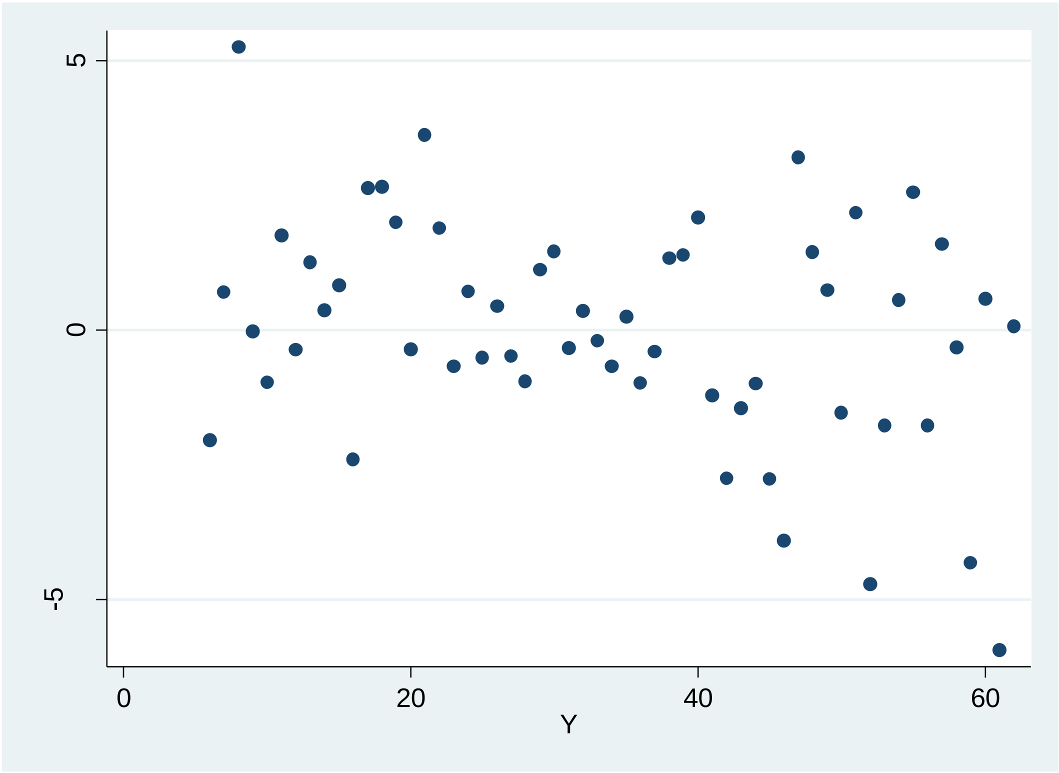
N = 62.

TABLE 4.5 ESTIMATES FOR AR(5) WITH FIVE LAGGED MARKET SENTIMENT VARIABLES, OIL ETF

Parameters	Estimate	Std. Error	t value	p value
Constant	1.032551	1.234716	0.84	0.407
Lag Return 1	-0.0936714	0.1483131	-0.63	0.531
Lag Return 2	-0.290503	0.1625732	-1.79	0.081
Lag Return 3	0.0075521	0.1315804	0.06	0.954
Lag Return 4	-0.1147581	0.1320238	-0.87	0.389
Lag Return 5	-0.0722987	0.1313163	-0.55	0.585
Lag Sentiment 1	-0.0041886	0.0980609	-0.04	0.966
Lag Sentiment 2	-0.0809549	0.0948882	-0.85	0.398
Lag Sentiment 3	-0.1397645	0.0916733	-1.52	0.134
Lag Sentiment 4	0.0502045	0.0975064	0.51	0.609
Lag Sentiment 5	0.0009366	0.0969937	0.01	0.992

N = 62.

FIGURE 4.1 RESIDUALS SCATTER PLOT FOR AR(5) MODEL WITHOUT MARKET SENTIMENT, CRUDE OIL ETF (APRIL 4, 2016 TO JULY 8, 2016)



N = 62.

CHAPTER 5: SUMMARY

5.1 THESIS SUMMARY

The objective of this study was to develop a market sentiment model for financial markets using machine learning, and to illustrate these methods using commodity price data. Market sentiment in this study refers to a variable using positive and negative news (words and phrases) for a stock in order to create a sentiment indicator variable. The sentiment indicator variable will be positive if the words and phrases are positive and negative if the words and phrases are negative. For example, theoretically, if the sentiment indicator variable increases, then the stock price would be expected to increase. Market sentiment may capture the fundamentals of a market or its tone and psychology.

The commodity used was the daily oil price, using the spot crude oil USO ETF, from April 4, 2016 to July 8, 2016. A market sentiment indicator was created using machine learning methods, with the software Python Natural Language Toolkit and Beautiful Soup. The results showed that the base autoregressive model did not have significant statistical forecasting power and that the addition of sentiment variables did not add to the forecasting power of the model.

This may indicate that markets were relatively difficult to forecast during the period of the data. Machine learning and sentiment indicators are a relatively new forecasting method. Therefore, further research on this topic is needed before any firm conclusions can be drawn regarding the effectiveness of this approach. Typically, financial markets are often believed to be relatively efficient, and so it is not surprising that it was difficult to forecast commodity prices using an autoregressive model and a model with market sentiment variables.

One reason for the challenge in forecasting using market sentiment could be due to the fact that the market sentiment variable was created using only the Seeking Alpha database. It did not include news from Bloomberg, Wall Street Journal, and other financial news sites. Further, it did not incorporate Twitter and other social media sites that may have added forecast power to the market sentiment model. In addition, the return data used was limited to only 62 observations which may not have been a sufficient time period to test as it may have been a period where the market was difficult to forecast. In conclusion, the market sentiment method developed in this study is a relatively new area, and further studies are required to better understand whether market sentiment indicators serve as useful forecasting tools.

5.2 LIMITATIONS AND FUTURE RESEARCH

Some of the limitations to this study were that the data period was relatively short and the market may have been relatively difficult to forecast in this time period. The second limitation was that only one commodity, crude oil, was used. An additional commodity or other commodities may show different results, and more commodities should be examined in the future. Future research could also consider alternative functional forms, such as an exponential weighting method for the lagged market sentiment variables in order to put more weight on the recent periods and less weight on earlier periods, and this may enhance the model's forecasting power.

REFERENCES

Antweiler, W., & Frank, Z. M. 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. *The Journal of Finance 59(3),* 1259-1294. DOI: 10.1111/j.1540-6261.2004.00662.x. Accessed on May 25, 2016.

Boudoukh, J., Feldman, R. Kogan S., and Richardson, M. 2012. Which News Moves Stock Prices? A Textual Analysis. NBER Working Paper No. 18725. Accessed on May 25, 2016.

Chen, H., De, P., Hu, Y., & Hwang, B. (2013). Customers as Advisors: The Role of Social Media in Financial Markets. Available at SSRN. http://ssrn.com/abstract=1807265. Accessed on May 27, 2016.

Fama, F. E., 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance 25,* 383-417. Accessed on May 14, 2016.

Fortuny, J. E., Smedt, T., Martens, D., & Daelemans, W. 2014. Evaluating and Understanding Text-Based Stock Price Prediction Models. *Journal of Information Processing and Management 50,* 426-441. Accessed on May 24, 2016.

Gorton, G., Rouwenhorst, K.G., 2006. Facts and Fantasies about Commodity Futures.

Financial Analysts Journal 62 (2), 69–97. Accessed on June 1, 2016.

Hajek P., Oleg, V., & Myskova, R. 2013. Forecasting Stock Prices using sentiment Information in Annual Reports – A Neural Network and Support Vector Regression Approach. *Wseas Transactions on Business and Economics, 10,* 293-305. Accessed on May 29, 2016.

Huang, A., Zang, A. Y., & Zheng, R. 2013. Large Sample Evidence on the Informativeness of Text in Analyst Reports. Available from http://ssrn.com/abstract=1888724. Accessed on May 28, 2016.

Hill M. J., Nadig D., & Hougan, M. 2015. A Comprehensive Guide to Exchange-Traded Funds (ETFs). CFA Research Foundation Publications, 1–181. Available from: http://www.cfapubs.org/toc/rf/2015/2015/3. Accessed on May 12, 2016.

Kearney, C., & Liu, S. 2014. Textual sentiment in finance: A Survey of Methods and Models. *International Review of Financial Analysis 33,* 171-185. Accessed on May 15, 2016.

Kaastra, I. and Boyd, M. 1996. Designing a Neural Network for Forecasting Financial and Economic Time Series. *Neurocomputing 10,* 215-236. Accessed on September 30, 2016.

Liew, J. and Budavári, T. Do Tweet Sentiments Still Predict the Stock Market? 2016. Available at SSRN: http://ssrn.com/abstract=2820269. Accessed on September 15, 2016

Python's Natural Language Toolkit. 2015. https://www.python.org & http://www.nltk.org. Accessed on April 15, 2016.

Miller, G. (1995). WordNet: A Lexical Database for English. Communications of the ACM Vol. 38, No. 11: 39-41. https://wordnet.princeton.edu. Accessed on April 15, 2016.

Roll, R. 1988. R^2. *Journal of Finance, 43,* 541–566. Accessed on May 13, 2016.

Schmeling, M. 2009. Investor sentiment and Stock Returns: Some International Evidence. *Journal of Empirical Finance 16,* 394-408. Accessed on June 2, 2016.

Sahu, A., Porth, B., Porth, L., and Boyd, M. 2015. The Contribution of sentiment to Predicting Stock Price Movement Using Statistical Machine Learning. Working Paper, International Agricultural Risk, Insurance, and Finance Conference, Washington, D.C., USA, June 2015. Accessed on January 28, 2016.

Tetlock, C. P. 2007. Giving Content to Investor sentiment: The Role of Media in the Stock Market. *The Journal of Finance 62(3)*, 1139-1168. DOI: 10.1111/j.1540-6261.2007.01232.x. Accessed on May 12, 2016.

Thulasiraman, R., Thulasiraman, P. and Qasem, M. 2015. Twitter Sentiment Classification using Machine Learning Techniques for Stock Markets. *International Conference on Advances in Computing, Communications and Informatics.* DOI: 10.1109/ICACCI.2015.7275714. Accessed on September 16, 2016.

Thomson Reuters. Ipsos Primary Consumer Sentiment Index. 2002. http://thomsonreuters.com/en/products-services/financial/investment-management/ipsos-primary-consumer-sentiment-index.html. Accessed on January 14, 2016.

United States Oil Fund Prospectus. 2016. Securities Exchange Commission.

http://www.unitedstatescommodityfunds.com/documents/pdfs/unitedstatesoilfund-pro-20160301.pdf. Accessed on February 25, 2016.

Zhang, G., Patuwo, B.E., and Hu, M.Y. 1998. Forecasting with Artificial Neural Networks: The State of the Art. *International Journal of Forecasting 14*, 35–62. Accessed on May 5, 2016.