

Allen, David E.; McAleer, Michael; Singh, Abhay K.

Working Paper

Daily Market News Sentiment and Stock Prices

Tinbergen Institute Discussion Paper, No. 15-090/III

Provided in Cooperation with:

Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Allen, David E.; McAleer, Michael; Singh, Abhay K. (2015) : Daily Market News Sentiment and Stock Prices, Tinbergen Institute Discussion Paper, No. 15-090/III, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/125094>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

TI 2015-090/III
Tinbergen Institute Discussion Paper



Daily Market News Sentiment and Stock Prices

David E. Allen^a

Michael McAleer^b

Abhay K Singh^c

^a *Centre for Applied Financial Studies, Adelaide, University of South Australia, and University of Sydney, Australia;*

^b *National Tsing Hua University, Taiwan, Erasmus University Rotterdam, Tinbergen Institute, the Netherlands, and Complutense University of Madrid, Spain;*

^c *Edith Cowan University, Perth Australia.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Daily Market News Sentiment and Stock Prices

David E. Allen^a, Michael McAleer^b, Abhay K Singh^{c,*}

^a*Centre for Applied Financial Studies, Adelaide, University of South Australia, and School of Mathematics and Statistics, Sydney, University of Sydney*

^b*Department of Quantitative Finance National Tsing Hua University, Taiwan and Econometric Institute Erasmus School of Economics Erasmus University Rotterdam and Tinbergen Institute, The Netherlands and Department of Quantitative Economics Complutense University of Madrid, Spain*

^c*School of Business, Edith Cowan University, Perth Australia*

Abstract

In recent years there has been a tremendous growth in the influx of news related to traded assets in international financial markets. This financial news is now available via print media but also through real-time online sources such as internet news and social media sources. The increase in the availability of financial news and investor's ease of access to it has a potentially significant impact on market price formation as these news items are swiftly transformed into investors sentiment which in turn drives prices. Various commercial agencies have started developing their own financial news data sets which are used by investors and traders to support their algorithmic trading strategies. Thomson Reuters News Analytics (TRNA)¹ is one such data set. In this study we use the TRNA data set to construct a series of daily sentiment scores for Dow Jones Industrial Average (DJIA) stock index component companies. We use these daily DJIA market sentiment scores to study the influence of financial news sentiment scores on the stock prices of these companies using a multi-factor model. We use an augmented Fama French Three Factor Model to evaluate the additional effects of financial news sentiment on stock prices in the context of this model. Our results suggest that even when market factors are taken into account, sentiment scores have a significant effect on Dow Jones constituent company returns and that lagged daily sentiment scores are

*Corresponding author. Tel.: +61863045605

Email address: a.singh@ecu.edu.au (Abhay K Singh)

¹http://thomsonreuters.com/products/financial-risk/01_255/News_Analytics_-_Product_Brochure-_Oct_2010_1_.pdf

often significant, suggesting that information compounded in these scores is not immediately reflected in security prices and related return series.

JEL classification: G12 G140 C31

Keywords: Sentiment Analysis, Financial News, Factor Models, Asset Pricing

1. Introduction

Investors utilise the daily company news releases which are obtained via different sources including both traditional newspapers and on-line internet news and social media channels. The information in these news items become the basis of investor opinions, which could be termed news sentiment, and could be viewed as taking on either positive, negative or neutral values per news item, for each individual investor. The continuous release of financial news helps to update general investor's information sets in relation to financial markets and influences general investor sentiment. Investors' investment strategies which influence the market and the evolution of stock prices are potentially influenced by changes in these sentiments which are generated by the continuous flow of news items. Academic researchers and investment practitioners are always looking for new investment tools or factors which may help to predict moves in asset prices and improve on existing models. Recently, the role of market news sentiment, in particular machine-driven sentiment signals, and their implication for financial market processes, has captured the attention of both investment practitioners and academics. There is a growing body of research that argues that news items from different sources influence investor sentiment, and hence asset prices, asset price volatility and risk (Tetlock, 2007; Tetlock et al., 2008; Da et al., 2011; Barber and Odean, 2008; diBartolomeo and Warrick, 2005a; Mitra et al., 2008; Dzielinski et al., 2011).

Bhattacharya et al. (2009) explore news items that came out between 1996 and 2000 on 458 Internet initial public offerings (IPOs) and a matching sample of 458 non-Internet IPOs (a total of 171,488 news items). They classify each news item on a simple basis as being good news, neutral news, or bad news. They document that the media was more positive

for Internet IPOs in the period of the dramatic rise in share prices and more negative for Internet IPOs in the period of the dramatic fall in share prices. They then document that media hype was unable to explain the Internet bubble.

Dougal et al. (2012) report that a small set of financial columnists has a causal effect on short-term aggregate stock market prices. For 'bullish' journalists, the market reaction is consistently positive, whereas for others, 'bearish journalists', it is negative. They report that journalist effects are much stronger after extreme returns, suggesting that amplification or attenuation of existing sentiment is a mechanism underlying the financial media's influence. Engelberg and Parsons (2011) attempt to disentangle the causal impact of media reporting from the impact of the events being reported. They approach the problem by comparing the behaviors of investors with access to different media coverage of the same information event. They use zip codes to identify 19 mutually exclusive trading regions, corresponding to 19 large U.S. cities and local newspapers (e.g., the Houston Chronicle). For all earnings announcements of S&P 500 Index firms, they report that local media coverage strongly predicts local trading, after controlling for characteristics of the earnings surprise, firm, local investors, and reporting newspaper(s).

According to the Efficient Market Hypothesis (EMH), the prices of traded stocks rapidly reflect all the relevant information sets available. It is widely accepted that the degree and speed to which this applies, is not necessarily uniform across all the various markets or for the every asset class; the main reason being that all the relevant information is not necessarily available to everyone at the same time, and also different investors face different cost structures. This means there is a scope for insider trading. Investors are always seeking new and more timely sources of information relevant for asset pricing which can then be used to more swiftly predict price changes. Investor sentiment has been proven to be a determinant of stock returns (Baker and Wurgler, 2006). Recent work by Hafez and Xie (2012b) examines the effect of investor's sentiment using news based sentiment, generated from the RavenPack Sentiment Index as a proxy for market sentiment in a multi-factor

model. They report a strong impact of market sentiment on stock price predictability over 6 and 12 month time horizons.

In asset pricing, the Capital Asset Pricing model (CAPM) (Sharpe, 1964; Lintner, 1965) is the most commonly used model for pricing stocks. The CAPM assumes that an asset's returns are dependent on the return on the market portfolio. An alternative related approach is provided by Arbitrage Pricing Theory which was developed to potentially subsume multiple linear factors (APT) (Ross, 1976). The APT model thus facilitates the inclusion of additional independent risk factors which may affect asset returns. A popular multi-factor model is the Fama and French (1992, 1993) three factor model which was created by extending the basic CAPM to include size and book-to-market terms as additional explanatory factors in explaining the cross-section of stock returns. These so called multi-factor models are widely used by practitioners and fund managers to account for the risk factors in asset pricing. Cahan et al. (2009) and Hafez and Xie (2012b) found that news sentiment as calculated by the RavenPack sentiment scores could add diversification benefits to the traditional factor models like CAPM or the other multi-factor models.

In this paper we examine the sentiment scores provided by TRNA as a single factor in a simple regression model and then in a multi-factor model to evaluate their effect on the stock prices of the DJIA component companies. We use daily DJIA market sentiment scores constructed from high frequency sentiment scores for the various stocks in DJIA as provided by TRNA and augment the Fama French 3-factor regression model. The empirical analysis includes data from the time periods of the Global Financial Crisis and other periods of market turbulence to assess the effect of financial news sentiment on stock prices in both normal and in extreme market conditions. We use Ordinary Least Square (OLS) regression and Quantile Regression (QR) to evaluate the model around the mean and the tails of the stock return distributions. The choice of Dow Jones constituent companies as our sample means that this is a tough test of the influence of sentiment scores; given that these are likely to be some of the most closely analysed and highly traded companies in the US stock

market.

The paper is organized as follows: Section 1 provides an introduction, Section-2 features an introduction to sentiment analysis and an overview of the TRNA data set. Section-3 discusses the data and research methods used in the empirical exercise undertaken in this paper. The next section-4, discusses the major results and section-5 draws some conclusions.

2. Background

It is widely observed that present financial markets are influenced or rather driven by the influx of critical information in the form of real time unanticipated news along with anticipated company announcements. Recently there has been a surge in studies exploring the relationship between stock price movements and news sentiment (Tetlock, 2007; Barber and Odean, 2008; Mitra et al., 2008; Leinweber and Sisk, 2011; Sinha, 2010; Huynh and Smith, 2013).

With the growth in the sources of real-time news the speed and breadth of the influx of news into the financial markets has increased tremendously. The volume and rate of incoming news makes it rather difficult for market participants to process all asset-specific news to make prompt investment decisions. To resolve this problem there are other sources of pre-processed news available from vendors like TRNA and Ravenpack, which provide direct indicators to the traders and other financial practitioners of changes in news sentiment. These sources use text mining tools to electronically analyse available textual news items. The analytics engines of these sources use pattern recognition and identification methods to analyse, words and their patterns, the novelty and relevance of the news items for a particular industry or sector. The type and characteristics of these news items are converted into quantifiable sentiment scores.

We use sentiment indicators provided by TRNA for our empirical analysis. Thomson Reuters was one of the first to implement a sophisticated text mining algorithm as an addition to its company and industry specific news database starting from January 2003 which

resulted in the present TRNA data set. As per the official TRNA data guide, “Powered by a unique processing system the Thomson Reuters News Analytics system provides real-time numerical insight into the events in the news, in a format that can be directly consumed by algorithmic trading systems”. Currently the data set is available for various stocks and commodities until October 2012. The TRNA sentiment scores are produced from text mining news items at a sentence level, which takes into account the context of a particular news item. This kind of news analytics makes the resulting scores more usable as they are mostly relevant to the particular company or sector. Every news item in the TRNA engine is assigned an exact time stamp and a list of companies and topics it mentions. A total of 89 broad fields are reported in the TRNA data set which are broadly divided into following 5 main categories².

1. Relevance: A numerical measure of how relevant the news item is to the asset.
2. Sentiment: A measure of the inherit sentiment of the news item quantifying it as either negative (-1), positive (1) or neutral (0).
3. Novelty: A measure defining how new the news item is; in other words whether it reports a news item that is related to some previous news stories.
4. Volume: Counts of news items related to the particular asset.
5. Headline Classification: Specific analysis of the headline.

Figure-1 provides some background information about the basic data structure. It features an illustration of the time and dating sequence, which is a 24 hour cycle centred on Greenwich meantime, plus an indication of some features of the basic data capture. The example features an announcement on the Hong Kong Stock Exchange about a bid by Hutcheson Whampoa for Hutch Telecom shares, at a price of HKD 2.2 per share. The date stamp is 08 January 2010 and the time stamp is 01:01:31.394 based on GMT.

²See the Handbook of News Analytics in Finance (Mitra & Mitra, 2011) for further details on TRNA data set

Figure 1: Date and Time features and data-capture.

Column Name	Description
DATE	Date (in GMT) when this news 'event' was transmitted. Example: "2003-01-31"
TIME	Time (in 24 hr GMT) when this news 'event' was transmitted. Example: "13:45:59"
UNIQUE_STORY_INDEX	A unique index which will always be the same for all news 'events' about the same story. This is a simple concatenation of STORY_DATE_TIME and PNAC but with the colons, dashes and single blank character removed from the STORY_DATE_TIME field.

```

TIMESTAMP  BCAST_REF  STOCK_RIC  ITEM_ID  RELEVANCE
SENTIMENT  SENT_POS  SENT_NEUT  SENT_NEG  LNKD_CNT1
LNKD_CNT2  LNKD_CNT3  LNKD_CNT4  LNKD_CNT5  LNKD_ID1
LNKD_ID2   LNKD_ID3   LNKD_ID4   LNKD_ID5   LNKD_IDPV1
LNKD_IDPV2 LNKD_IDPV3 LNKD_IDPV4 LNKD_IDPV5 ITEM_TYPE
ITEM_GENRE BCAST_TEXT DSELY_NAME PNAC STORY_TYPE
CROSS_REF  PROC_DATE  TAKE_TIME  STORY_DATE STORY_TIME
NAMED_ITEM TAKE_SEQNO ATTRIBTN  PROD_CODE  TOPIC_CODE CO
IDS  LANG_IND
08 JAN 2010 01:01:31.394  0001.HK  0001.HK  2010-01-08_
01.01.31.nHKF000044.A1.eedb8019  1  1  0.542338  0.449667
0.00799559 0 0 0 0 0
ALERT NOT DEFINED
HUTCHISON WHAMPOA <0013.HK> SAYS TO PAY HKD2.2 PER SHARE
IN GENERAL OFFER FOR HUTCH TELECOM SHARES 1 nHKF000044 S
08 JAN 2010 01:01:31 08 JAN 2010 01:01:31
1 RTRS Z E HK ASIA CN IN RET REA SHP MJL TEL
REAM RRL DBT ENR CONG LEN RTRS 0013.HK 0001.HK 2332.HK EN
08 JAN 2010 01:17:18.858 0001.HK 0001.HK 2010-01-08_
01.17.19.nHKF000044.T1.385e36a7 0.258199 1 0.536821
0.454702 0.00847696 1 1 1 1 1
2010-01-08_01.01.31.nHKF000044.A1.eedb8019
2010-01-08_
01.01.31.nHKF000044.A1.eedb8019 ARTICLE NOT DEFINED

```

Figure 2: TRNA-Snapshot of News Headlines Generated for BHP Billiton in Year 2011

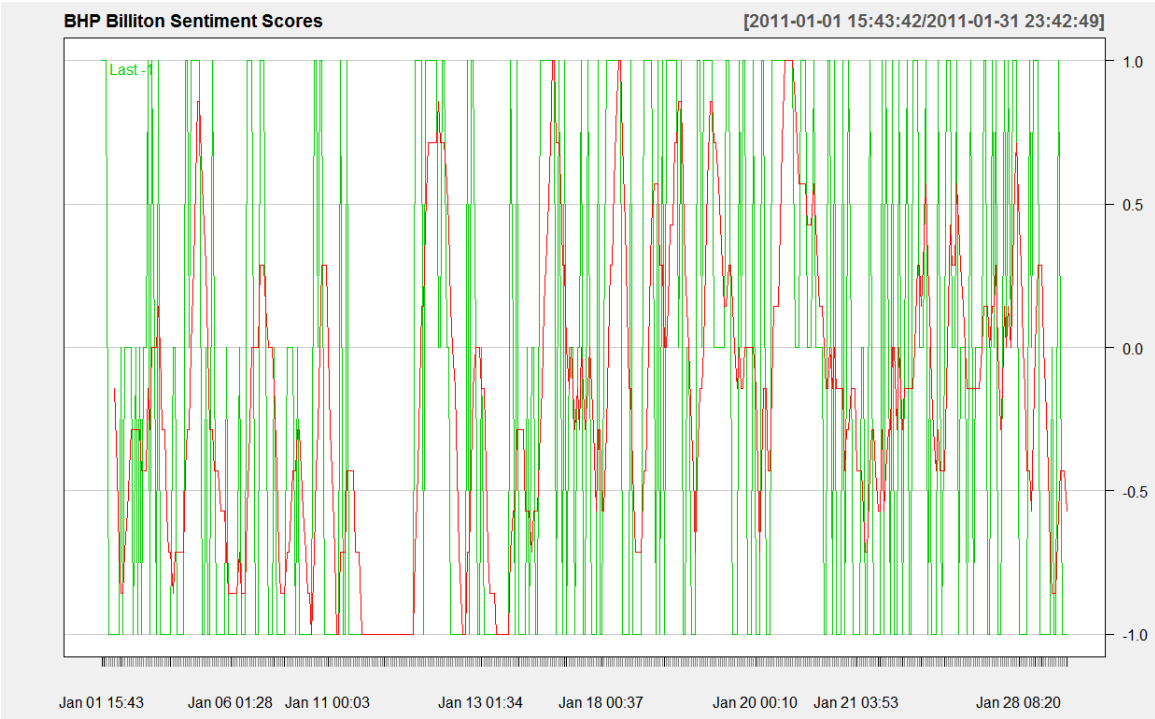
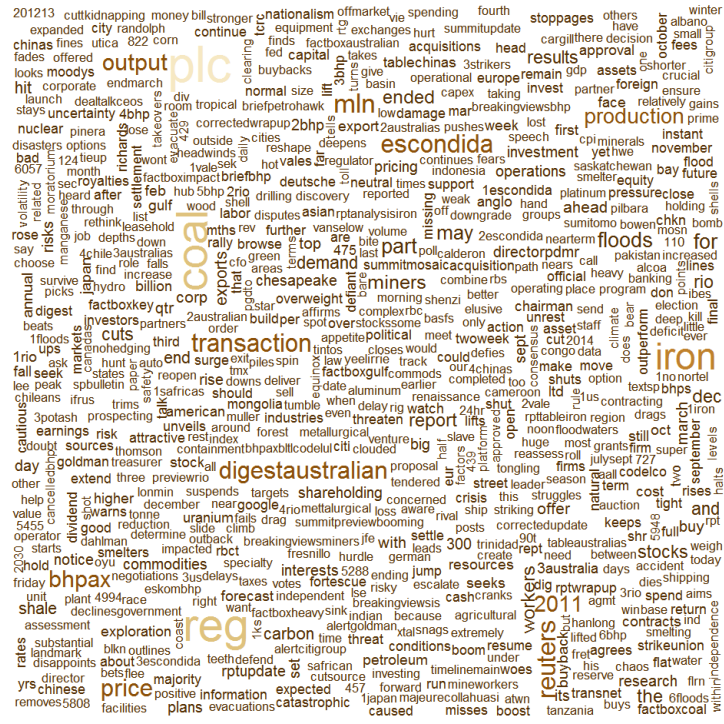


Figure 3: TRNA-Sentiment Scores Generated for BHP Billiton in Jan-2011

Figure-2, shows a snapshot of the headline text as reported in BCAST_REF field of the TRNA database for BHP Billiton during the year 2011. These are not the sentences which are analysed by TRNA to produce sentiment scores but are the headlines for the news item used to generate the TRNA sentiment and other relevant scores. As reported in TRNA, BHP Billiton generated more than 3000 news items in the year 2011. Figure-3 shows the sentiment scores (-1 to +1) for BHP Billiton during the month of January 2011, the red line is the moving average of the scores.

Similar to BHP, there are various news stories reported per day for the various DJIA traded stocks. These news stories result in sentiment scores which are either positive, negative or neutral for that particular stock. Figure-4 gives a snapshot of the sentiment scores for the DJIA's traded stocks during the year 2008. The bar chart of figure-4 shows that the most sentiment scores generated during the year 2008, which was also the period of Global Financial Crisis, were for the Citi Bank group (C.N) , General Motors (GM.N) and J. P Morgan (JPM.N). This is a reflection of the market sentiment during the GFC period, as these financial institutions were among the most effected during the GFC.

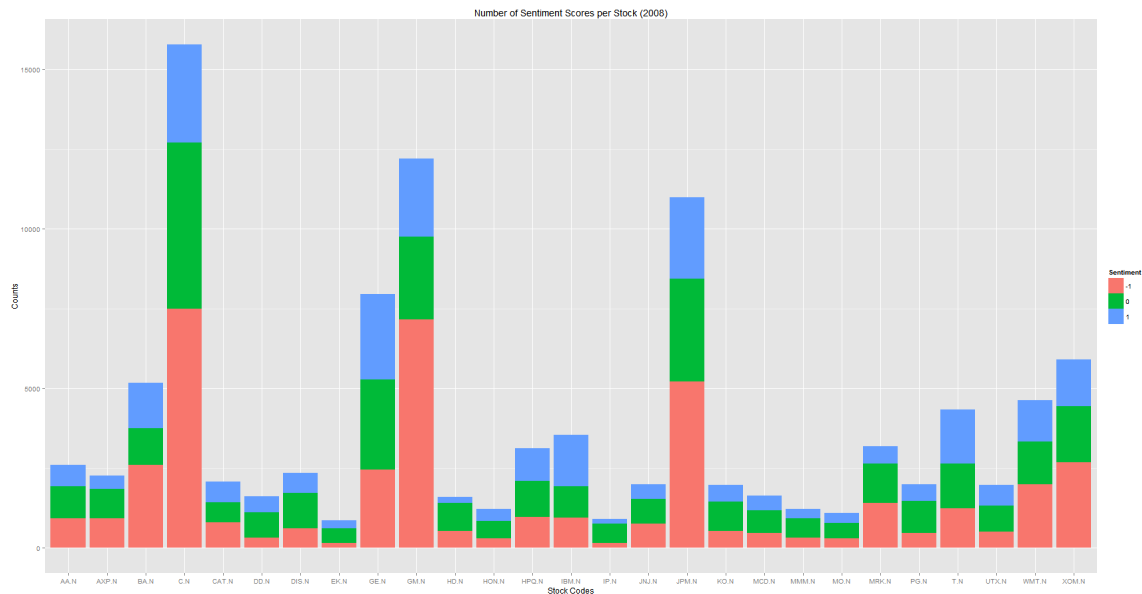
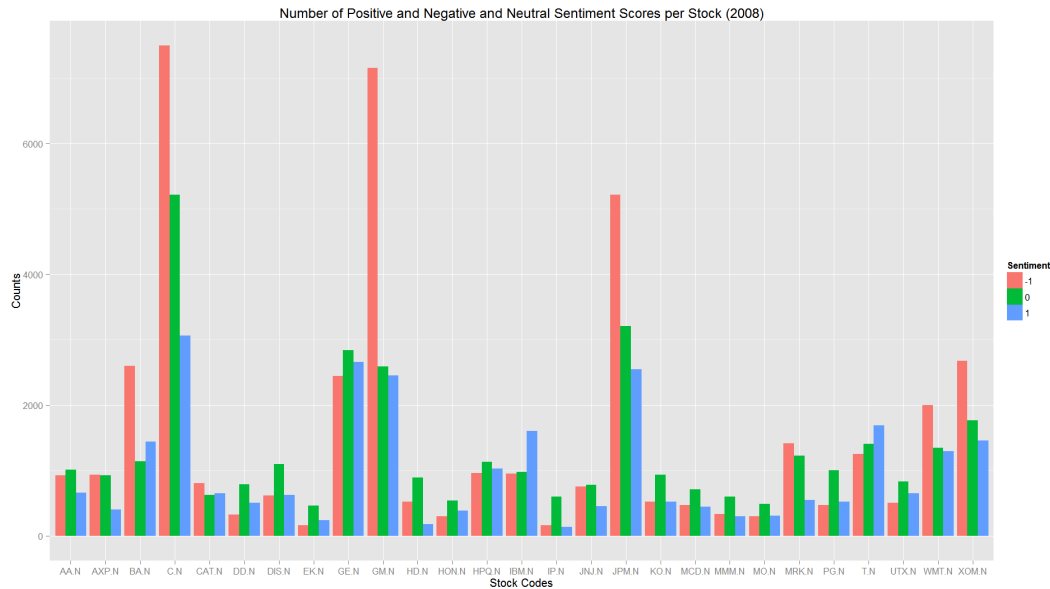


Figure 4: Sentiment Score Distribution for DJIA Stocks in 2008

Figure-5, shows the number of positive, negative or neutral sentiment scores stacked

against each other. Its evident from this figure that the number of negative and neutral sentiment news exceeded the number of positive sentiments for the majority of stocks. Again its in agreement with the context of the GFC period when the DJIA stock market index took a big plunge downwards.

Figure 5: Positive, Negative and Neutral Sentiment Score Distribution for DJIA Stocks in 2008



The applications of TRNA news data sets to financial research has recently gained interest. Dzielinski (2012), Groß-Klußmann and Hautsch (2011), Smales (2014), Huynh and Smith (2013), Borovkova and Mahakena (2015). Storckenmaier et al. (2012) and Sinha (2010), have shown the usefulness of the TRNA dataset in stock markets and in commodity markets for both high frequency and multi-day frequency. In this study we utilize the TRNA data set to analyse the effect of news sentiment on DJIA stocks at a daily frequency. We construct daily sentiment index score time series for the empirical exercise from the high frequency scores reported by TRNA. The specific data and research methods used are discussed in the next section.

3. Data & Methodology

The empirical analysis in this study analyses the effect of news sentiment on stock prices of the DJIA constituents by considering the daily DJIA market sentiment as an additional risk factor to explain stock returns. We construct daily sentiment scores for DJIA market by accumulating high frequency sentiment scores of the DJIA's constituents obtained from TRNA dataset. We use data from January 2006 to October 2012 to study the sensitivity of the daily stock returns to the daily market sentiment scores. The daily stock prices for all the DJIA traded stocks are obtained from Thomson Tick History database for the same time period.

The TRNA provides high frequency sentiment scores calculated for each news item reported for various stocks and commodities. These TRNA scores for the stocks traded in DJIA can be aggregated to obtain a daily market sentiment score series for the DJIA stock index components. A news item s_{it} received at time t for a stock i is classified as a positive (+1), negative (-1) or neutral (0). $I_{s_{it}}^+$ is a positive classifier (1) for a news item s_{it} and $I_{s_{it}}^-$ is the negative (-1) classifier for a news item s_{it} . TRNA reported sentiment scores have a probability level associated with them, $prob_{s_{it}}^+$, $prob_{s_{it}}^-$, $prob_{s_{it}}^0$ for positive, negative and neutral sentiments, which is reported by TRNA in the Sentiment field. Based on the probability of level of the sentiment, denoted by $P_{s_{it}}$ for an individual news item s_{it} , for a particular company, all the daily sentiments can be combined across all companies in the DJIA to obtain a daily average sentiment indicator S_t . We use the following formula to obtain the average sentiment indicator $S_t = \sum S_{i,t-1}$, which is obtained by summing all the individual news items for each company that makes up the DJIA, on that particular day. The expression inside the brackets in equation (1) is the daily sentiment score for each of the individual companies that make up the DJIA. This is then summed across all constituents to get the score for that day S_t .

$$S_t = \sum S_{i,t-1} = \sum_{i=1}^{i=30} \left[\frac{\sum_{q=t-1}^{t-Q} I_{s_{iq}}^+ P_{s_{iq}} + \sum_{q=t-1}^{t-Q} I_{s_{iq}}^- P_{s_{iq}}}{n_+ + n_- + n_0} \right] \quad (1)$$

The time period considered are $t - Q, \dots, t - 1$ which covers all the news stories (and respective scores) for a 24 hour (Q) period prior to t when relevant news for the companies included in DJIA were received. In our subsequent analysis, the previous days score is used in the regressions, plus various lags. The neutral news sentiments are not included in the numerator in equation-1 but it is included in the denominator which is the sum of number of positive, negative and neutral sentiment scores (n_+, n_-, n_0) for the day, for all companies in the DJIA, thereby lowering the score in case the day had a lot of neutral scores.

Table-1 lists the various stocks traded in DJIA along with their RIC (Reuters Instrument Code) and time period. We use the TRNA sentiment scores related to these stocks to obtain the aggregate daily sentiment for the market. The aggregated daily sentiment score S represents the combined score of the sentiment scores reported for the stocks on a particular date.

Table 1: DJIA Stocks with Thomson Tick History RIC Codes

RIC Code	Stocks	First Date	Last Date
.DJI	Dow Jones INDU AVERAGE	1-Jan-96	17-Mar-13
AA.N	ALCOA INC	2-Jan-96	18-Mar-13
GE.N	GENERAL ELEC CO	2-Jan-96	18-Mar-13
JNJ.N	JOHNSON&JOHNSON	2-Jan-96	18-Mar-13
MSFT.OQ	MICROSOFT CP	20-Jul-02	18-Mar-13
AXP.N	AMER EXPRESS CO	2-Jan-96	18-Mar-13
GM.N	GENERAL MOTORS	3-Jan-96	18-Mar-13
GMGMQ.PK	GENERAL MOTORS	2-Jun-09	15-Aug-09
JPM.N	JPMORGAN CHASE	1-Jan-96	18-Mar-13
PG.N	PROCTER & GAMBLE	2-Jan-96	18-Mar-13
BA.N	BOEING CO	2-Jan-96	18-Mar-13
HD.N	HOME DEPOT INC	2-Jan-96	18-Mar-13
KO.N	COCA-COLA CO	2-Jan-96	18-Mar-13
SBC.N	SBC COMMS	2-Jan-96	31-Dec-05
T.N	AT&T	3-Jan-96	18-Mar-13
C.N	CITIGROUP	2-Jan-96	18-Mar-13
HON.N	HONEYWELL INTL	2-Jan-96	18-Mar-13
XOM.N	EXXON MOBIL	1-Dec-99	18-Mar-13
MCDw.N	MCDONLDS CORP	6-Oct-06	4-Nov-06
MCD.N	MCDONALD'S CORP	1-Jan-96	18-Mar-13
EK.N	EASTMAN KODAK	1-Jan-96	18-Feb-12
EKDKQ.PK	EASTMAN KODAK	19-Jan-12	18-Mar-13
IP.N	INTNL PAPER CO	2-Jan-96	18-Mar-13
CAT.N	CATERPILLAR INC	2-Jan-96	18-Mar-13
HPQ.N	HEWLETT-PACKARD	4-May-02	18-Mar-13
MMM_w.N	3M COMPANY WI	18-Sep-03	27-Oct-03
MMM.N	MINNESOTA MINING	1-Jan-96	18-Mar-13
UTX.N	UNITED TECH CP	2-Jan-96	18-Mar-13
DD.N	DU PONT CO	2-Jan-96	18-Mar-13
IBM.N	INTL BUS MACHINE	2-Jan-96	18-Mar-13
MO.N	ALTRIA GROUP	2-Jan-96	18-Mar-13
WMT.N	WAL-MART STORES	2-Jan-96	18-Mar-13
DIS.N	WALT DISNEY CO	2-Jan-96	18-Mar-13
INTC.OQ	INTEL CORP	20-Jul-02	18-Mar-13
MRK.N	MERCK & CO	2-Jan-96	18-Mar-13

Note: There are more than 30 stocks in the table because at different points in time the DJIA constituents change

The stocks with insufficient data are removed from the analysis and the stocks prices for EK.N and EKDKQ.PK are combined together to get a uniform timeseries. We employ regression analysis as given in equation-2 to quantify the effect of the daily sentiment index

on stock prices. The equation includes five lagged values of sentiment plus the square of the sentiment score (S_t^2) generated from previous day's news (the result of undertaking functional form tests in the regression analysis).

$$r_{it} = \alpha_i + \sum_{q=1}^5 \beta_{i,s_{t-q}} S_{t-q} + \beta_{i,s} S_t + \beta_{i,s^2} S_t^2 + \varepsilon_{it} \quad (2)$$

Where, r_i are the logarithmic returns for a stock traded in DJIA, $\beta_{i,s_{t-q}}$ is the coefficient on lagged sentiment score S_{t-q} for five lags, β_{i,s^2} is the coefficient on the square of the sentiment score S_t^2 .

We also augment the Fama French Factor model by introducing the sentiment score as an additional risk factor to evaluate the affect of market news sentiment on stock prices. The basic Fama French factor model is given by

$$r_{it} = r_{f_t} + \beta_i(r_{M_t} - r_{F_t}) + s_i SMB_t + h_i HML_t + \alpha_i + e_{it} \quad (3)$$

$$r_{it} = r_{f_t} + \sum_{q=1}^5 \beta_{i,s_{t-q}} S_{t-q} + \beta_{i,s} S_t + \beta_{i,s^2} S_t^2 + \beta_{i,m}(r_{M_t} - r_{F_t}) + s_i SMB_t + h_i HML_t + \alpha_i + e_{it} \quad (4)$$

In equation-3, SMB stands for Small Minus Big and represents the size premium and HML, which is short for High Minus Low, represents value premium (Fama and French, 1992, 1993). We introduce the sentiment factor in the above model (equation-4) to evaluate the joint effect of these four risk factors on the stock prices of the DJIA. The Fama French factors are obtained from Kenneth R French data library³. This empirical exercise is first evaluated using OLS followed by Quantile Regression (Koenker and Bassett Jr, 1978). Evaluation of the above two models (3 & 4) using QR quantifies the effect of market sentiment for extremal stock returns (Allen et al., 2011).

³mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Quantile Regression is modelled as an extension of classical OLS (Koenker and Bassett Jr, 1978), in quantile regression the estimation of conditional mean as estimated by OLS is extended to similar estimation of an ensemble of models of various conditional quantile functions for a data distribution. In this fashion quantile regression can better quantify the conditional distribution of $(Y|X)$. The central special case is the median regression estimator that minimizes a sum of absolute errors. We get the estimates of remaining conditional quantile functions by minimizing an asymmetrically weighted sum of absolute errors, here weights are the function of the quantile of interest. This makes quantile regression a robust technique even in presence of outliers. Taken together the ensemble of estimated conditional quantile functions of $(Y|X)$ offers a much more complete view of the effect of covariates on the location, scale and shape of the distribution of the response variable.

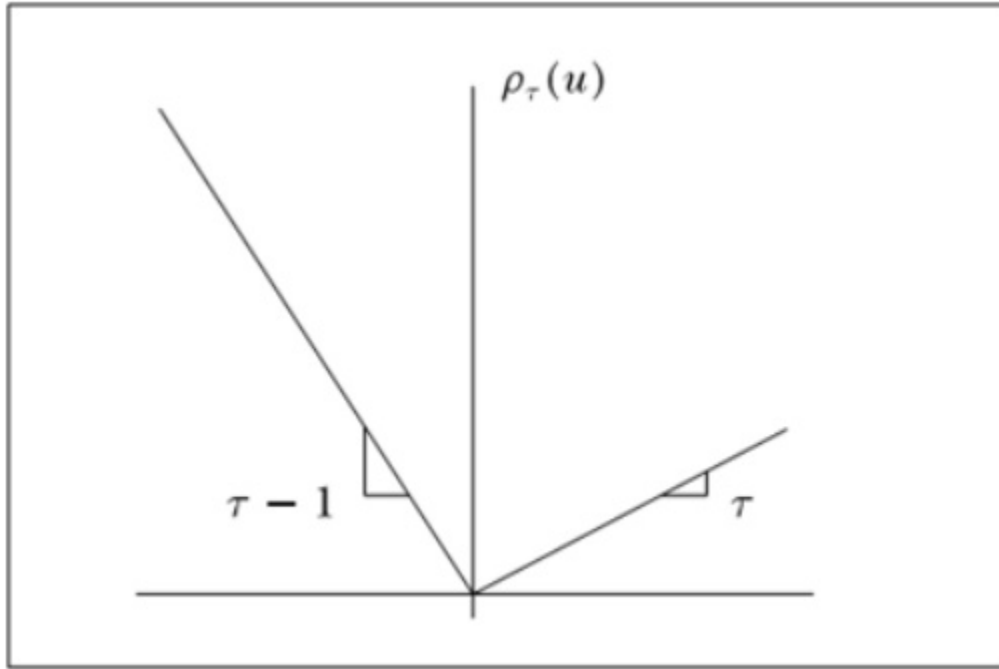
For parameter estimation in quantile regression, quantiles as proposed by Koenker and Bassett Jr (1978) can be defined through an optimization problem. To solve an OLS regression problem a sample mean is defined as the solution of the problem of minimising the sum of squared residuals, in the same way the median quantile (0.5%) in quantile regression is defined through the problem of minimising the sum of absolute residuals. The symmetrical piecewise linear absolute value function assures the same number of observations above and below the median of the distribution.

The other quantile values can be obtained by minimizing a sum of asymmetrically weighted absolute residuals, (giving different weights to positive and negative residuals). Solving

$$\min_{\xi \in R} \sum \rho_{\tau}(y_i - \xi) \tag{5}$$

Where $\rho_{\tau}(\cdot)$ is the tilted absolute value function as shown in figure-6, which gives the τ th sample quantile with its solution. Taking the directional derivatives of the objective function with respect to ξ (from left to right) shows that this problem yields the sample quantile as its solution.

Figure 6: Tilted absolute value function



After defining the unconditional quantiles as an optimization problem, it is easy to define conditional quantiles similarly. Taking the least squares regression model as a base to proceed, for a random sample, y_1, y_2, \dots, y_n , we solve

$$\min_{\mu \in R} \sum_{i=1}^n (y_i - \mu)^2 \quad (6)$$

which gives the sample mean, an estimate of the unconditional population mean, EY . Replacing the scalar μ by a parametric function $\mu(x, \beta)$ and then solving

$$\min_{\mu \in R^p} \sum_{i=1}^n (y_i - \mu(x_i, \beta))^2 \quad (7)$$

gives an estimate of the conditional expectation function $E(Y|x)$.

Proceeding the same way for quantile regression, to obtain an estimate of the conditional median function, the scalar ξ in the first equation is replaced by the parametric function $\xi(x_t, \beta)$ and τ is set to $1/2$. The estimates of the other conditional quantile functions are obtained by replacing absolute values by $\rho_\tau(\cdot)$ and solving

$$\min_{\mu \in \mathbb{R}^p} \sum \rho_{\tau}(y_i - \xi(x_i, \beta)) \quad (8)$$

The resulting minimization problem, when $\xi(x, \beta)$ is formulated as a linear function of parameters, and can be solved very efficiently by linear programming methods. Further insight into this robust regression technique can be obtained from Koenker (2005).

We run these comprehensive tests as represented by equations (2) and (4) above and report the results in Table (2). In some of the subsequent regression analysis lagged and squared sentiment terms are omitted from the analysis to reduce the size of tables and volume of reported results. The intention was to undertake an omnibus analysis of the influence of sentiment scores over the whole time period, and then to focus on subsets of the sample as captured by market downturns, in the case of the GFC, or the extremes of the return distributions, as captured by the quantile regression analysis.

The TRNA dataset is different from the daily stock price dataset as it reports news sentiments also for the days when there is no trading in the DJIA. This becomes an issue to consider in the empirical analysis. To mitigate this issue we test two different models; one with the sentiment for the day and the other with a 7 day simple moving average (SMA) of the sentiment score. Taking the SMA of the sentiment accounts for the news reported on the weekends for the stocks. We will now discuss the major results whilst highlighting the main steps of the methodology.

4. Discussion of the Results

The empirical analysis is first conducted for the whole time period of 2007 to 2012 and then for sub-periods of two years in a moving window; 2007-2008, 2008-2009, 2009-2010, 2010-2011 and 2011-2012. We will discuss the results obtained from the complete time period and two moving window periods of 2007-2008 and 2010-2011. These results are representative of the factor behaviour across the whole sample and in two rather different market conditions. The period for 2007-2008 represents the time of GFC and 2010-2011 can

be considered as being representative of more normal market conditions. Figure-7 shows the plot of DJIA constituent's daily returns and daily sentiment scores in percentage (dotted line), the figure shows that the daily sentiment scores move really closely with stock returns. The market sentiment generated from the individual sentiment scores represents the stock return movement in the DJIA stock market. The figure also shows that the DJIA was quite volatile during late 2007 to 2009 due to the GFC but was relatively normal during 2010-2011. We will first discuss the results obtained from daily sentiment scores followed by the results obtained using seven day SMA of the daily sentiment scores.

4.1. Results from the Daily Sentiment Scores

4.1.1. Linear Regression (OLS)

Table-2 and table-3, give the OLS results for Model-1 (equation-2) and Model-2 (equation-4) as obtained from the complete data set (2007 to 2012). The results in table-2 and table-3 quantify the sensitivity of DJIA traded stock returns to the daily DJIA market news sentiment scores. The Model-1 results, reported in the table-2, clearly demonstrate the significant effect of market news sentiment when taking aggregate market sentiment scores as an independent risk factor for stock price returns, in the case of Model-1 virtually all the stocks including the DJIA show a statistically significant contemporaneous sentiment beta coefficient ($\beta_{i,s}$). The only exception being Coca-Cola (KO.N). However, the influence of the squared sentiment term is equally pervasive and is highly significant in all but two of the companies, Altria Group and Merck (MO.N, MRK.N). This term was added when Ramsey reset tests of functional form suggested a squared term was important for the sentiment variable. This makes intuitive sense in that squaring this variable, which is bounded between +1 for highly positive sentiment and -1, for highly negative sentiment, would serve to emphasize the extreme positive and negative sentiment scores, which are likely to have more impact on trading and prices. The inclusion of lagged sentiment scores also proved important in that in only four of the cases did the inclusion of lagged scores up to five lags not have at least one of these lags proving significant. This has potential implications for the informational

DJIA Stocks Return Timeseries Vs Sentiment Plot (2007-2012)

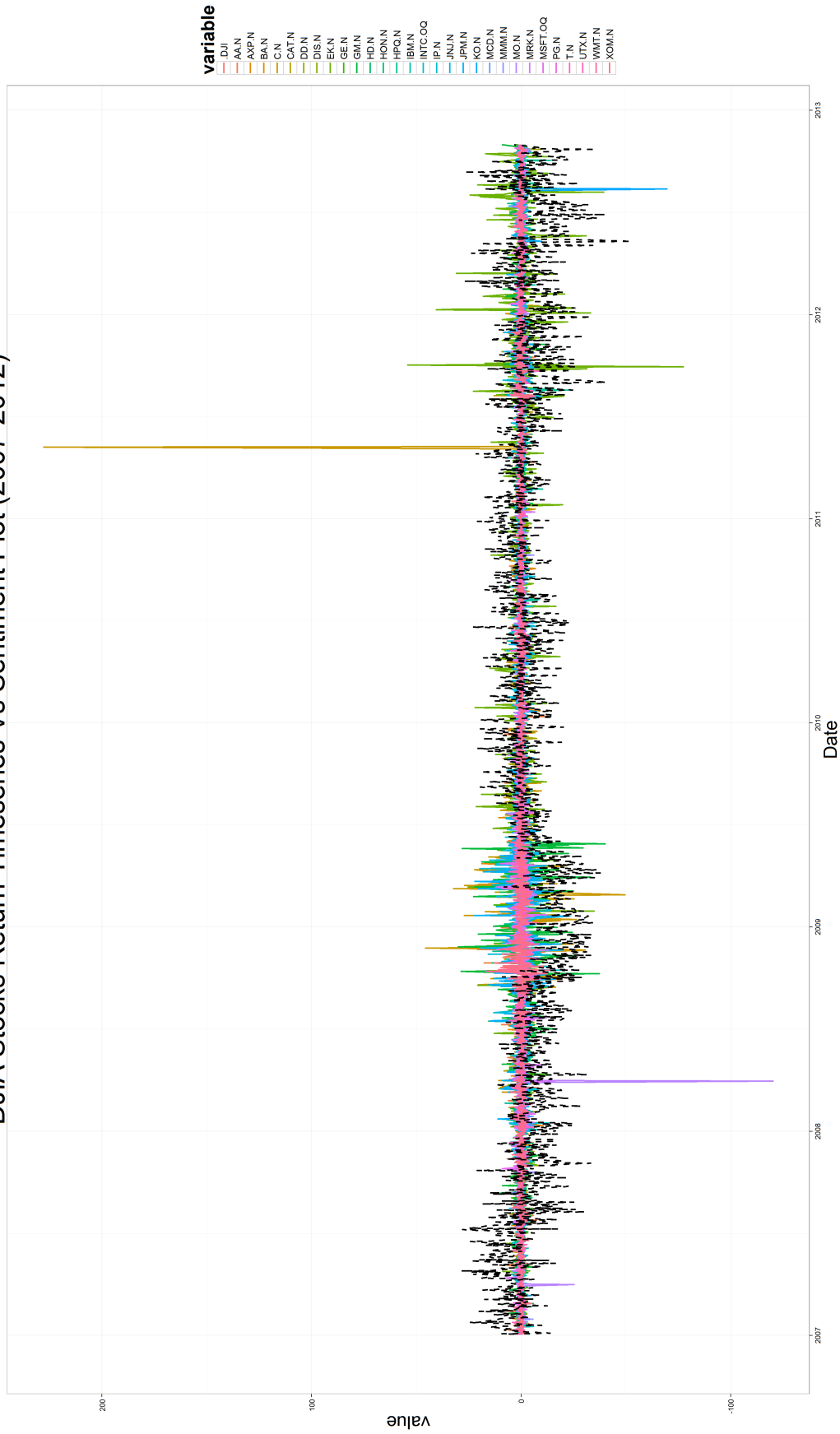


Figure 7: DJIA constituent's returns vs sentiment scores (2007-2012)

efficiency of the market in relation to these sentiment scores in that information captured lagged up to 5 days was proving to be significant in some of these regressions. The acid test is provided by model two in that this second regression analysis, also run over the whole period, includes the market portfolio and the two other Fama French factors as well. If variants of the sentiment score and its 5 lags remain significant this will be fairly strong evidence for a lack of market efficiency, at least in the context of this commonly adopted model of the return setting process.

The results from Model-2 , shown in Table 3, give a similar picture, although the DJIA (the traded asset) returns are dependent on the daily sentiment factor even when the other three factors are added as controls (Market Return, SMB, HML), and more than half the other stocks still show a significant relationship with at least one of the sentiment terms, usually a lagged sentiment term. Indeed, the results show that 17 out of 30 assets, or 57% of the sample, still have at least one significant sentiment even after controlling for Fama French factors. These 17 companies are, in terms of Table 1, the following: AA_N, BA_N, DD_N, EK_N, GE_N, GM_N, HPQ_N, IBM_N, INTC_OQ, IP_N, JPM_N, KO_N, MCD_N, MRK_N, PG_N, T_N, and WMT_N. This means that even though the significance of the sentiment terms revealed in the regression application of Model 1 is partly absorbed by the other three independent risk factors there is still a frequent and significant news sentiment factor. These results demonstrate that the market news sentiment has a significant effect on asset returns (Model-1), which can also remains in contention after the market return and the returns of the other two factors SMB and HML, as incorporated in Model-2, are included. This suggests that even these highly liquid and closely scrutinised Dow Jones Company returns are not informationally efficient with respect to some of the news captured in these sentiment scores. The adjusted R-Squares, not reported in the tables, suggest that the sentiment scores capture about 5% of the variation in the return series of the sample over the full sample period whilst the market factors typically capture 50-60%. However, the important point is that the influence of the sentiment score frequently remains significant,

even after the the market factors are taken into account.

As the market conditions are never the same for the whole data time period, we also conduct the analysis in a two year moving window (shifting by one year). The choice of the two year moving period accounts for any changing dynamics in the stock market. We will now discuss the linear regression results for the year 2007-2008 which can be considered a representative of the turbulent times of the GFC, and the years 2010-2011 which was relatively a normal period for DJIA.

Table-4 presents the OLS results for Model-1 and Model-2 without the lagged and squared sentiment terms obtained using aggregated daily sentiment scores during 2007-2008. The results for Model-1 in this sub period are identical to the results from whole data period, they all show the sentiment measure has a significant effect on stock prices. The results from Model-2 (the augmented Fama French model) for this sub period which includes the GFC show a significant effect of the daily sentiment scores on a large number of stock returns, 15 in total numbers, or in excess of 50%, which is similar to the analysis for the whole data period, which was 17 in number, but the former analysis included lagged and squared sentiment terms. (These were dropped from the subsequent analysis to keep the scale of the results and tables to a manageable size). The influence in extremes and the squared sentiment term echoes previous findings by Engelberg and Parsons (2011) and Dougal et al. (2012). There are clearly many stocks displaying significant market news sentiment effects along with the effects of Fama French factors in times of financial distress, which are comparable to the results for the total data time period, but the results in our previous analysis included both lagged and squared sentiment terms. The results from just including a contemporaneous sentiment score in the regression analysis suggest that this coefficient is significant more frequently in our sample in times of financial distress. This suggests that during times of market turbulence stock price return series are still effected by market news sentiment scores which may or may not get absorbed in the effects of other asset pricing factors. These results indicate a significant effect of the daily news sentiment scores during the times of financial

crisis.

We also consider the results for the period of year 2010-2011 to contrast them with the results in table-4. Table-5⁴ gives the results obtained for the two year period of 2010 to 2011, this can be considered a relatively less turbulent time period in the US than 2007-2008. The results from the time period of 2010-2011 show a significant $\beta_{i,s}$ effects for 8 out of 30 assets.

The results shown in table-4 and table-5 indicate that the market news sentiment scores affect a greater number of asset price return series during times of market turbulence and this effect is not completely absorbed by other market factors. We will now further check this analysis using QR to evaluate if the $\beta_{i,s}$ is more significant for the lower tail returns, which has been suggested by the results from OLS for periods of market downturns. We also evaluate if the relationship between the stock returns and news sentiment changes across the quantiles.

⁴The analysis is conducted for all the 5 moving 2 year periods, the rest of the results can be obtained from the authors on request.

Table 2: OLS results for equation (2) $r_{it} = r_{ft} + \sum_{q=1}^5 \beta_{i,s_{t-q}} S_{t-q} + \beta_{i,s} S_t + \beta_{i,s^2} S_t^2 + \beta_{i,m} (r_{M_t} - r_{F_t}) + s_i SMB_t + h_i HML_t + \alpha_i + e_{it}$ (2007-2012)

Assets	Model-1							
	α_i	$\beta_{i,s}$	β_{i,s^2}	$\beta_{i,s_{t-1}}$	$\beta_{i,s_{t-2}}$	$\beta_{i,s_{t-3}}$	$\beta_{i,s_{t-4}}$	$\beta_{i,s_{t-5}}$
.DJI	0.137986	2.73029	-7.85521	-1.01924	-0.804547	-0.0203932	0.0043049	-0.637285
p-value	0.00006***	<0.00001***	0.00065***	0.00222***	0.03213**	0.95396	0.0126**	0.03103**
AA.N	0.20508	5.31857	-15.2339	-0.607751	-1.597	-0.0823751	0.765416	-2.16863
p-value	0.01782**	0<0.00001***	0.00642***	0.42588	0.06497*	0.90758	0.33264	0.00274***
AXP.N	0.194612	4.08222	-12.2917	-1.37204	-1.97586	0.294143	0.214318	-0.947163
p-value	0.00829***	0.00001***	0.02673**	0.06681*	0.03148**	0.65350	0.75749	0.15586
BA.N	-0.0231	1.021	-7.61565	-1.18537	-0.735104	0.181495	-0.0667433	-0.890306
p-value	0.5551	0<0.00001***	0.01132**	0.01635**	0.15187	0.70251	0.88907	0.06578*
C.N	0.529841	4.65043	-29.7868	-1.75731	-1.79745	0.0216818	0.206178	1.65431
p-value	0.05756*	0.00008***	0.00365***	0.26896	0.11217	0.98766	0.90128	0.39589
CAT.N	0.24049	3.95832	-12.3972	-0.792343	-1.18373	-0.572356	0.305597	-0.812152
p-value	0.00064***	<0.00001***	0.00005***	0.17689	0.04341**	0.32755	0.60067	0.16035
DD.N	0.214019	3.21144	-13.0084	-1.11488	-1.03066	-0.0664237	0.206562	-0.887012
p-value	0.00036***	<0.00001***	<0.00001***	0.02561**	0.03877**	0.89375	0.67754	0.07160*
DIS.N	0.193687	3.07448	-10.7058	-1.03192	-1.17203	-0.108417	0.143491	-0.971893
p-value	0.00008***	<0.00001***	0.00084***	0.03732**	0.02549**	0.82818	0.76640	0.02592**
EK.N	0.128772	3.97302	-24.326	0.163378	-3.75778	0.209942	-2.45072	1.91409
p-value	0.37901	0.00672***	0.00980***	0.90770	0.00240***	0.86984	0.15859	0.34409
GE.N	0.174458	3.59796	-11.2941	-1.05448	-1.23363	-0.00767782	0.368185	-0.593438
p-value	0.00533***	<0.00001***	0.00349***	0.06445*	0.04263**	0.98999	0.52162	0.31006
GM.N	0.141322	7.24532	-19.7329	-0.782115	-1.57659	0.330005	-0.232982	-2.41238
p-value	0.35320	<0.00001***	0.02657**	0.47263	0.17495	0.76275	0.85950	0.02691**
HD.N	0.125165	2.4002	-7.05276	-0.871644	-0.607102	-0.461726	-0.307233	-0.521621
p-value	0.01382**	<0.00001***	0.00697***	0.04424	0.22989	0.35714	0.53234	0.23377
HON.N	0.177117	3.21358	-7.75937	-1.02377	-0.791729	0.127923	-0.0586206	-0.6914
p-value	0.00084***	<0.00001***	0.01322**	0.04265**	0.14521	0.79939	0.90916	0.13411
HPQ.N	0.122568	3.46494	-10.6119	-0.458124	-1.66109	0.0693741	0.12303	-1.25208
p-value	0.02041**	<0.00001***	0.00047***	0.38642	0.00552***	0.88991	0.82339	0.01040**
IBM.N	0.185276	2.31471	-8.75753	-0.949011	-0.459865	-0.349807	0.0864106	-0.685205
p-value	<0.00001***	<0.00001***	0.00018***	0.00638***	0.20894	0.35474	0.81334	0.04332**
INTC.OQ	0.114102	3.76848	-9.55474	-1.28688	-1.22036	-0.310351	0.0492284	-1.5132
p-value	0.04496**	<0.00001***	0.00340***	0.00503***	0.02414**	0.55810	0.91724	0.00100***
IP.N	0.259391	3.48988	-15.2358	-1.11935	-2.56763	0.355795	1.37082	-0.985386
p-value	0.00239***	<0.00001***	0.00507***	0.09467*	0.00806***	0.64556	0.06161*	0.13078
JNJ.N	0.0660116	1.46996	-4.15905	-0.621795	-0.513731	0.080363	-0.127797	-0.362566
p-value	0.01758**	<0.00001***	0.00862***	0.01488**	0.05624*	0.76759	0.63377	0.16198
JPM.N	0.299876	4.11058	-21.6266	-1.37303	-1.78123	-0.28903	-0.752693	-0.753218
p-value	0.00016***	<0.00001***	0.00003***	0.10180	0.04085**	0.72198	0.33084	0.31375
KO.N	0.11786	0.457806	-9.69104	-0.501395	-0.424206	-0.385138	0.616034	-0.257325
p-value	0.00026***	0.71666	0.02954**	0.33792	0.17914	0.17914	0.11360	0.48105
MCD.N	0.129126	1.66271	-4.45897	-1.12844	-0.45838	0.358509	0.0297395	-0.148185
p-value	0.00009***	<0.00001***	0.03265**	0.00101***	0.15899	0.26481	0.93093	0.65031
MMM.N	0.111438	2.39718	-7.00857	0.82842	-0.92369	-0.15667	0.110522	-0.605769
p-value	0.00888***	<0.00001***	0.00586***	0.02684**	0.04006**	0.67105	0.81358	0.09532*
MO.N	-0.0300664	1.85193	-0.639913	-0.874459	0.141911	1.18394	0.168842	-2.03656
p-value	0.76808	0.00040***	0.85203	0.01618**	0.84449	0.20802	0.63862	0.15746
MRK.N	0.0678925	2.87238	-1.70766	-1.03392	-0.246556	0.205999	-0.02164	-1.10735
p-value	0.17347	<0.00001***	0.53483	0.02407**	0.59231	0.64154	0.95823	0.01906**
MSFT.OQ	0.117699	2.89658	-6.79389	-1.19668	-0.707106	0.339589	-0.0728021	-1.1623
p-value	0.01598**	<0.00001***	0.01534**	0.00598***	0.10882	0.46994	0.86698	0.00507***
PG.N	0.0949701	1.32634	-5.78187	-1.01027	-0.403711	0.00681	-0.174579	0.205577
p-value	0.00184***	<0.00001***	0.00237***	0.00027***	0.21751	0.98207	0.54154	0.44449
T.N	0.136047	2.85145	-7.24379	-0.836519	-0.833862	-0.35389	-0.133207	-0.393793
p-value	0.00295***	<0.00001***	0.01998**	0.03027**	0.08482	0.37867	0.72052	0.24390
UTX.N	0.136047	3.07784	-6.58689	-1.13717	-0.866721	-0.0291935	0.100285	-0.666738
p-value	0.00347***	<0.00001***	0.01497**	0.01588**	0.04023**	0.94879	0.81717	0.09672*
WMT.N	0.095745	1.94242	-5.34269	-0.998112	-0.550654	-0.238649	-0.349501	-0.261618
p-value	0.00565***	<0.00001***	0.01610**	0.00516***	0.09645*	0.52736	0.31334	0.40425
XOM.N	0.128929	2.71869	-6.68583	-1.37415	-0.640345	-0.10375	0.216478	-0.834602
p-value	0.00660***	<0.00001***	0.02662**	0.00388***	0.17329	0.81617	0.62502	0.02704**

P-values < * show statistical significance at 90% confidence level, ** =95%, *** =1%. HAC standard errors

Table 4: OLS Results for 2007 to 2008

Assets	Model-1		Model-2				
	α_i	$\beta_{i,s}$	α_i	$\beta_{i,m}$	s_i	h_i	$\beta_{i,s}$
.DJI	-0.0089	0.9143	-0.0089	0.9143	-0.1707	-0.0185	0.2339
p-value	0.5209	0***	0.5209	0***	0***	0.3223	0.0371
AA.N	-0.0365	1.6875	-0.0365	1.6875	-0.3923	-0.926	2.0683
p-value	0.7518	0***	0.7518	0***	0.0102	0***	0.0263**
AXP.N	-0.0393	1.338	-0.0393	1.338	-0.0891	0.885	2.0838
p-value	0.6632	0***	0.6632	0***	0.4543	0***	0.0042***
BA.N	-0.0954	0.8917	-0.0954	0.8917	-0.2978	-0.3157	0.3912
p-value	0.2115	0***	0.2115	0***	0.0032***	0.0022***	0.5236
C.N	-0.1675	1.3522	-0.1675	1.3522	-0.1185	3.2247	2.4466
p-value	0.2488	0***	0.2488	0***	0.5358	0***	0.0364**
CAT.N	0.0672	1.101	0.0672	1.101	0.258	-0.4182	1.4883
p-value	0.3323	0***	0.3323	0***	0.0049***	0***	0.0078***
DD.N	-0.0274	1.0564	-0.0274	1.0564	0.1949	-0.1217	0.8178
p-value	0.6444	0***	0.6444	0***	0.013**	0.1272	0.0869*
DIS.N	-0.0418	1.0765	-0.0418	1.0765	-0.2244	-0.1969	-0.3931
p-value	0.4779	0***	0.4779	0***	0.004***	0.013**	0.4061
EK.N	-0.0849	0.9225	-0.0849	0.9225	-0.0344	0.337	2.5054
p-value	0.4458	0***	0.4458	0***	0.815	0.0235**	0.0052***
GE.N	-0.0423	0.9427	-0.0423	0.9427	0.1315	0.6314	1.4338
p-value	0.5845	0***	0.5845	0***	0.1973	0***	0.0214**
GM.N	-0.1197	1.4012	-0.1197	1.4012	0.3318	1.3386	4.6245
p-value	0.5688	0***	0.5688	0***	0.2311	0***	0.0063***
HD.N	-0.0741	0.9958	-0.0741	0.9958	0.3771	0.3246	-0.6691
p-value	0.3485	0***	0.3485	0***	0.0003***	0.0023***	0.2921
HON.N	-0.0042	0.9745	-0.0042	0.9745	-0.105	-0.2424	0.2243
p-value	0.9488	0***	0.9488	0***	0.2238	0.006***	0.6698
HPQ.N	0.0043	0.9588	0.0043	0.9588	-0.192	-0.589	-0.3209
p-value	0.953	0***	0.953	0***	0.0466**	0***	0.5845
IBM.N	0.0375	0.7718	0.0375	0.7718	-0.0104	-0.1301	0.5531
p-value	0.4992	0***	0.4992	0***	0.8865	0.081*	0.215
INTC.OQ	0.0296	1.1848	0.0296	1.1848	0.0404	-0.677	0.5746
p-value	0.6987	0***	0.6987	0***	0.6896	0***	0.3488
IP.N	-0.0642	1.0933	-0.0642	1.0933	-0.0225	0.1846	1.8104
p-value	0.5021	0***	0.5021	0***	0.8586	0.151	0.0189**
JNJ.N	-0.0219	0.571	-0.0219	0.571	-0.3115	-0.1757	-0.574
p-value	0.6093	0***	0.6093	0***	0***	0.0024***	0.0966*
JPM.N	0.0613	1.1586	0.0613	1.1586	0.3861	2.4748	0.265
p-value	0.5149	0***	0.5149	0***	0.0019***	0***	0.726
KO.N	0.0103	0.6326	0.0103	0.6326	-0.092	-0.3477	-0.0045
p-value	0.8628	0***	0.8628	0***	0.2433	0***	0.9925
MCD.N	0.0843	0.6216	0.0843	0.6216	0.0077	-0.0947	0.057
p-value	0.1629	0***	0.1629	0***	0.9226	0.2432	0.9066
MMM.N	-0.024	0.7382	-0.024	0.7382	-0.0314	-0.0925	0.1957
p-value	0.6601	0***	0.6601	0***	0.6625	0.2073	0.6554
MO.N	-0.3152	0.646	-0.3152	0.646	-0.316	-0.4703	0.502
p-value	0.2559	0***	0.2559	0***	0.3876	0.207	0.8218
MRK.N	-0.0227	0.8926	-0.0227	0.8926	-0.2418	-0.3943	-0.0347
p-value	0.7952	0***	0.7952	0***	0.0363**	0.0008***	0.9607
MSFT.OQ	-0.0216	1.0514	-0.0216	1.0514	-0.3081	-0.6811	0.1704
p-value	0.7476	0***	0.7476	0***	0.0006***	0***	0.7518
PG.N	-0.0257	0.6087	-0.0257	0.6087	-0.2911	-0.1267	-0.8377
p-value	0.5823	0***	0.5823	0***	0***	0.0436**	0.0258**
T.N	0.0279	0.8917	0.0279	0.8917	-0.4973	0.0132	0.4437
p-value	0.6634	0***	0.6634	0***	0***	0.8786	0.39
UTX.N	0.0232	0.9565	0.0232	0.9565	-0.0013	-0.1418	0.0371
p-value	0.6611	0***	0.6611	0***	0.9848	0.0469**	0.9306
WMT.N	0.0869	0.6751	0.0869	0.6751	0.1541	-0.1315	-0.0444
p-value	0.1308	0***	0.1308	0***	0.0423**	0.0887*	0.9234
XOM.N	0.0212	1.1401	0.0212	1.1401	-0.7656	-0.7169	-0.8423
p-value	0.7347	0***	0.7347	0***	0***	0***	0.095*

P-values < * show statistical significance at 90% confidence level, ** = 95%, *** = 1%. HAC standard errors.

Table 5: OLS Results for 2010 to 2011

Assets	Model-1		Model-2				
	α_i	$\beta_{i,s}$	α_i	$\beta_{i,m}$	s_i	h_i	$\beta_{i,s}$
.DJI	-0.0015	0.9195	-0.0015	0.9195	-0.2397	-0.0596	0.2553
p-value	0.8733	0***	0.8733	0***	0***	0.0061***	0.0081***
AA.N	-0.1661	1.5462	-0.1661	1.5462	-0.0696	0.1764	1.5186
p-value	0.0089***	0***	0.0089***	0***	0.5884	0.2166	0.0167**
AXP.N	-0.0079	1.155	-0.0079	1.155	-0.2375	0.5855	-0.1274
p-value	0.878	0***	0.878	0***	0.0243**	0***	0.8055
BA.N	0.0277	1.1304	0.0277	1.1304	-0.0761	0.1275	0.9725
p-value	0.5827	0***	0.5827	0***	0.4573	0.2616	0.0539*
C.N	0.364	1.552	0.364	1.552	0.7332	0.5441	1.6796
p-value	0.4323	0.0007***	0.4323	0.0007***	0.4363	0.6025	0.7169
CAT.N	0.0334	1.3445	0.0334	1.3445	0.2135	-0.0461	0.1424
p-value	0.4988	0***	0.4988	0***	0.034**	0.6792	0.7733
DD.N	0.0175	1.1463	0.0175	1.1463	0.0063	-0.0443	0.4426
p-value	0.6455	0***	0.6455	0***	0.935	0.6044	0.2435
DIS.N	-0.0005	1.0902	-0.0005	1.0902	-0.2539	0.0044	0.95
p-value	0.9916	0***	0.9916	0***	0.0055***	0.9655	0.0344**
EK.N	-0.3606	1.0824	-0.3606	1.0824	-0.4943	1.4954	2.2086
p-value	0.2109	0.0002***	0.2109	0.0002***	0.3984	0.0216**	0.4431
GE.N	0.0127	1.1052	0.0127	1.1052	-0.2634	0.5272	1.1525
p-value	0.7635	0***	0.7635	0***	0.0022***	0***	0.0065***
GM.N	-0.2122	1.2128	-0.2122	1.2128	0.1241	0.3274	0.5294
p-value	0.0293**	0***	0.0293**	0***	0.5356	0.1425	0.5645
HD.N	0.0384	0.8534	0.0384	0.8534	-0.0728	0.0597	-0.1395
p-value	0.4163	0***	0.4163	0***	0.4482	0.5746	0.7675
HON.N	0.0212	1.2167	0.0212	1.2167	0.0517	-0.1394	0.8375
p-value	0.5695	0***	0.5695	0***	0.4942	0.0974*	0.0249**
HPQ.N	-0.1526	1.1533	-0.1526	1.1533	-0.376	-0.4501	2.8054
p-value	0.0386**	0***	0.0386**	0***	0.0122**	0.0069***	0.0002***
IBM.N	0.0358	0.8804	0.0358	0.8804	-0.3474	-0.4498	0.5659
p-value	0.3002	0***	0.3002	0***	0***	0***	0.1017
INTC.OQ	-0.0095	0.9737	-0.0095	0.9737	-0.1461	-0.2176	-0.146
p-value	0.8468	0***	0.8468	0***	0.145	0.0507*	0.767
IP.N	-0.0455	1.2945	-0.0455	1.2945	0.4107	0.4712	-0.9584
p-value	0.5245	0***	0.5245	0***	0.0048***	0.0036***	0.1797
JNJ.N	-0.0212	0.6452	-0.0212	0.6452	-0.3483	-0.2843	0.1249
p-value	0.4592	0***	0.4592	0***	0***	0***	0.6628
JPM.N	-0.0677	1.2259	-0.0677	1.2259	-0.2727	1.4613	0.2655
p-value	0.1609	0***	0.1609	0***	0.0056***	0***	0.582
KO.N	0.0152	0.6915	0.0152	0.6915	-0.4238	-0.327	0.1885
p-value	0.6353	0***	0.6353	0***	0***	0***	0.5565
MCD.N	0.0612	0.5887	0.0612	0.5887	-0.1886	-0.2796	-0.5405
p-value	0.0669*	0***	0.0669*	0***	0.0055***	0.0002***	0.1053
MMM.N	-0.0414	0.987	-0.0414	0.987	-0.0148	-0.1901	0.4027
p-value	0.2726	0***	0.2726	0***	0.8467	0.0259**	0.2861
MO.N	0.0616	0.5402	0.0616	0.5402	-0.2428	0.0177	-0.0894
p-value	0.0736*	0***	0.0736*	0***	0.0006***	0.8194	0.7948
MRK.N	-0.022	0.8244	-0.022	0.8244	-0.4675	0.0159	-0.0392
p-value	0.6006	0***	0.6006	0***	0***	0.8667	0.9255
MSFT.OQ	-0.0741	0.9367	-0.0741	0.9367	-0.2508	-0.3959	-0.05
p-value	0.0634*	0***	0.0634*	0***	0.002***	0***	0.9002
PG.N	-0.0035	0.5472	-0.0035	0.5472	-0.3768	-0.0968	-0.2765
p-value	0.9043	0***	0.9043	0***	0***	0.1436	0.3457
T.N	0.0009	0.6375	0.0009	0.6375	-0.3368	0.0135	0.625
p-value	0.9782	0***	0.9782	0***	0***	0.8573	0.0602*
UTX.N	-0.0301	1.0987	-0.0301	1.0987	-0.2441	-0.3058	0.5525
p-value	0.3726	0***	0.3726	0***	0.0004***	0.0001***	0.1015
WMT.N	0.0021	0.526	0.0021	0.526	-0.3069	-0.1767	0.0277
p-value	0.9525	0***	0.9525	0***	0***	0.0253**	0.937
XOM.N	0.0128	0.9842	0.0128	0.9842	-0.4595	-0.0064	0.3579
p-value	0.6876	0***	0.6876	0***	0***	0.9283	0.2597

P-values <= * show statistical significance at 90% confidence level, ** = 95%, *** = 1%. HAC standard errors.

4.1.2. Quantile Regression Results

Quantile Regression (QR) is a useful tool for studying the effects of the independent factors on asset's returns across various quantiles of the return distribution. We use QR to investigate the effect of daily DJIA sentiment scores on the asset returns across lower and higher quantiles. We quantify the effect of DJIA market sentiment scores during periods of relatively low and high stock returns using QR which permits the exploration of relationships in the tails of the distribution. We will first discuss the results obtained from the complete dataset, followed by the results from the two sub periods as discussed previously in the results from the OLS analysis.

The QR analysis evaluates five different quantile levels (5%, 25%, 50%, 75%, 95%) which range from low to high stock returns. Table-6 gives the results for Model-1 obtained for the complete dataset (2007-2012). The results in table-6 show that the $\beta_{i,s}$ is significant for all the assets in the lower 5% quantile as compared to the higher 95% quantile of the return distribution. These results are in accordance with the results from OLS where assets showed more significant $\beta_{i,s}$ during times of financial distress than in the comparatively normal times. These results suggest that the daily market news sentiment scores affect stock returns more at a time of losses than during a time of gains. Table-7 gives the QR results for the augmented Fama-French model, these results also prove that when combined with other three factors affecting the asset returns $\beta_{i,s}$ is more significant for the lower quantile than the higher quantile. In fact in the complete data sample all but one of the traded stocks have significant $\beta_{i,s}$ for the lower 5% quantile level. XOM.N is the only anomaly in this case which has significant $\beta_{i,s}$ for the higher 95% quantile. These results also show that $\beta_{i,s}$ or the sentiment effect is positive (directly related) to the stock returns for the lower quantile and negative (inversely related) for the higher quantile. This is an intuitive result as negative news sentiment in the stock market should result in declines in stock prices.

We further analyse the sentiment effect in a moving window of two years. Table-8 gives the results from Model-1 for the year 2007-2008, which is considered as a turbulent time in

this study. The results from Model-2 for this time period are reported in table-8. The results from table-8 when compared to the OLS results from Model-1 for the same period suggest a different factor effect across quantiles. OLS results for Model-1 (table-4) shows that all the stocks have significant sentiment effects which is not true across the quantiles as reported in table-8. This effect also varies across the quantiles. The sentiment effect ($\beta_{i,s}$) is significant across the stocks for lower quantile (5% and 25%) but its not significant for all the stocks in the higher 95% quantile. This result again indicates more influence of sentiment effects on low returns than on higher returns. This can be accounted for by more of investor decisions being based on public news around a time when the stock prices are declining than when stock prices are on the rise. The results in table-9 also support this as in the Model-2 there are more significant $\beta_{i,s}$ in the lower quantile than in the higher.

Figure-8 shows the plots of $\beta_{i,s}$ obtained from Model-2 as reported in table-9, these plots clearly show the trend of positive $\beta_{i,s}$ in the lower quantile to negative $\beta_{i,s}$ in the higher quantile.

Following the procedure in the OLS analysis, we also consider the results from the two sub-year period of 2010-2011. Table-10 gives the results for Model-1, where only the aggregate daily sentiment is an independent factor. The results again show that the $\beta_{i,s}$ is more significant in the lower tail and is not evident to the same degree in the higher tail. Considering this is a relatively normal period as compared to the year 2007-2008 and the previous OLS analysis this $\beta_{i,s}$ should get assimilated into the three Fama-French factors when evaluated in Model-2. Table-11 gives the results from Model-2 which shows that during less turbulent times the $\beta_{i,s}$ is not as significant as the other three factors.

Table 6: QR Results for Model-1 (2007-2012)

Assets	Model-1									
	5%		25%		50%		75%		95%	
	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$
.DJI	-1.8587	7.3299	-0.4955	3.619	0.0907	1.6264	0.6203	0.8009	2.0331	-2.5649
p-value	0	0	0	0	0.0001	0	0	0.007	0	0.0355
AA.N	-4.6837	17.9856	-1.4433	6.7786	0.0562	3.6219	1.5309	1.5416	4.8262	-1.9378
p-value	0	0	0	0	0.396	0	0	0.0734	0	0.4412
AXP.N	-3.7167	14.5267	-1.0954	6.8527	0.0613	2.8552	1.2753	0.6999	4.3685	-8.1966
p-value	0	0	0	0	0.2715	0	0	0.2828	0	0.0001
BA.N	-3.0221	9.4332	-0.9669	4.2223	0.0755	2.135	1.0696	0.7511	3.1373	-0.7739
p-value	0	0	0	0	0.1228	0	0	0.2135	0	0.3812
C.N	-5.2115	21.2655	-1.6277	9.2302	-0.1313	4.8361	1.5757	0.9988	5.7856	-12.5561
p-value	0	0	0	0	0.0379	0	0	0.3152	0	0
CAT.N	-3.3302	11.5016	-1.0043	5.1529	0.1407	2.9357	1.3167	2.2641	3.6138	-1.1898
p-value	0	0	0	0	0.0144	0	0	0.0002	0	0.4594
DD.N	-2.8608	11.1181	-0.9038	3.8709	0.1212	2.1193	1.1685	1.6196	3.0851	-4.5465
p-value	0	0	0	0	0.0175	0	0	0.0015	0	0.0009
DIS.N	-2.624	9.7005	-0.8504	3.5844	0.0727	2.3604	1.0021	1.3177	3.164	-4.3153
p-value	0	0	0	0	0.0817	0	0	0.012	0	0.0003
EK.N	-6.9554	10.2484	-2.0661	7.2127	-0.1466	2.7864	1.9167	0.56	7.4551	2.9946
p-value	0	0.0041	0	0	0.1106	0.0005	0	0.6481	0	0.5076
GE.N	-3.3167	13.6408	-0.8533	4.8502	0.0356	2.6272	1.0429	1.0029	3.21	-1.5362
p-value	0	0	0	0	0.4139	0	0	0.0744	0	0.2844
GM.N	-5.87	24.8914	-1.6676	8.6729	-0.0765	4.5749	1.581	0.5984	5.9529	-9.1736
p-value	0	0	0	0	0.3502	0	0	0.5199	0	0.0101
HD.N	-2.7522	8.1851	-0.8594	3.6735	0.0258	1.6813	1.0022	-0.4635	3.05	-4.4683
p-value	0	0	0	0	0.563	0	0	0.3868	0	0.0023
HON.N	-2.7638	9.6906	-0.7927	4.2735	0.0896	2.4153	1.0851	1.8875	3.1046	-3.1217
p-value	0	0	0	0	0.0816	0	0	0.0014	0	0.0297
HPQ.N	-2.9777	8.5965	-0.9339	4.611	0.0635	2.7088	1.0497	1.6102	2.8902	-3.819
p-value	0	0	0	0	0.1733	0	0	0.0077	0	0.0047
IBM.N	-2.0647	7.414	-0.6063	2.8847	0.0751	1.6872	0.8728	0.8616	2.3391	-2.7426
p-value	0	0	0	0	0.0375	0	0	0.0135	0	0.0122
INTC.OQ	-2.765	10.0109	-1.001	4.2887	0.11	2.2871	1.1127	0.774	3.0934	-4.4655
p-value	0	0	0	0	0.0195	0	0	0.1011	0	0.0002
IP.N	-3.9853	14.6914	-1.2636	5.6196	0.172	3.0051	1.483	1.2689	4.2139	-7.5317
p-value	0	0	0	0	0.0068	0	0	0.0543	0	0.0001
JNJ.N	-1.4795	4.6283	-0.4499	1.6627	0.0441	0.9829	0.5339	0.5566	1.5415	-1.1321
p-value	0	0	0	0	0.0395	0	0	0.042	0	0.1243
JPM.N	-3.8425	14.1696	-1.2175	7.2562	-0.0199	3.462	1.274	-0.1886	4.5398	-10.7536
p-value	0	0	0	0	0.72	0	0	0.754	0	0
KO.N	-1.6753	5.1193	-0.5275	2.556	0.0895	1.401	0.6558	0.173	1.904	-2.8162
p-value	0	0	0	0	0.001	0	0	0.5585	0	0.0012
MCD.N	-1.8214	5.1552	-0.5943	2.4112	0.1279	1.3688	0.7228	0.1974	2.0394	-1.61
p-value	0	0	0	0	0	0	0	0.5439	0	0.0172
MMM.N	-2.2748	7.9225	-0.6214	3.5014	0.0644	1.7169	0.8152	0.7536	2.3964	-2.6272
p-value	0	0	0	0	0.0649	0	0	0.1021	0	0.0188
MO.N	-1.7696	5.5259	-0.5053	2.2333	0.1263	1.0803	0.7512	0.3136	1.698	-2.971
p-value	0	0	0	0	0	0	0	0.2412	0	0
MRK.N	-2.3429	8.4846	-0.7888	2.4581	0.0307	1.3782	0.9103	0.5461	2.5637	-1.9318
p-value	0	0	0	0	0.4698	0.0002	0	0.2293	0	0.1386
MSFT.OQ	-2.6231	8.0995	-0.8182	3.5671	0.0397	2.2538	0.9251	0.65	2.7741	-3.9111
p-value	0	0	0	0	0.2969	0	0	0.1364	0	0
PG.N	-1.756	5.1165	-0.4524	1.7787	0.037	0.76	0.5395	-0.1491	1.8193	-2.073
p-value	0	0	0	0	0.1131	0.0003	0	0.6063	0	0.0319
T.N	-2.132	6.8669	-0.6644	2.9445	0.084	1.956	0.8051	1.1887	2.3964	-0.882
p-value	0	0	0	0	0.016	0	0	0.0012	0	0.4012
UTX.N	-2.291	8.1559	-0.7369	3.7168	0.0475	2.2421	0.929	0.5643	2.5922	-1.6433
p-value	0	0	0	0	0.2137	0	0	0.2199	0	0.1884
WMT.N	-1.8307	5.166	-0.5411	2.1503	0.0759	1.0793	0.6927	0.6463	1.9541	-1.2689
p-value	0	0	0	0	0.0205	0.0002	0	0.0179	0	0.2254
XOM.N	-2.4489	8.4748	-0.7108	2.9578	0.0995	1.7367	0.8874	1.1595	2.4499	-1.7852
p-value	0	0	0	0	0.0094	0	0	0.0104	0	0.0325

P-values ≤ 0.1 show statistical significance at 90% confidence level or higher.

Table 8: QR Results for Model-1 (2007-2008)

Assets	Model-1									
	5%		25%		50%		75%		95%	
	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$
.DJI	-2.0986	10.0252	-0.609	5.5233	0.0355	2.9793	0.5885	1.4464	2.2562	-4.295
p-value	0	0	0	0	0.5221	0	0	0.0037	0	0.0017
AA.N	-5.0684	22.7581	-1.5524	11.6654	0.0298	7.102	1.4721	2.3235	4.9489	-6.4942
p-value	0	0	0	0	0.7845	0	0	0.1097	0	0.1993
AXP.N	-4.1957	16.2112	-1.3917	10.9367	-0.0859	6.0385	1.2952	1.0015	4.8885	-3.0654
p-value	0	0	0	0	0.4287	0	0	0.425	0	0.4571
BA.N	-3.0159	12.0822	-1.0311	5.868	-0.0068	2.9124	0.9991	0.1842	2.5565	-3.1541
p-value	0	0	0	0	0.9339	0	0	0.7623	0	0.1642
C.N	-5.5974	23.7692	-1.7171	11.0504	-0.3635	5.7336	1.1488	-1.5937	6.1331	-16.6706
p-value	0	0	0	0	0.0012	0	0	0.1082	0	0.0017
CAT.N	-3.0367	14.4477	-0.8973	7.694	0.1145	4.6862	1.144	1.9309	3.2055	-2.5741
p-value	0	0	0	0	0.1539	0	0	0.0162	0	0.2886
DD.N	-3.0887	13.8584	-0.8925	5.6405	0.0221	3.2564	0.9445	1.7988	3.1049	-7.0015
p-value	0	0	0	0	0.7902	0	0	0.0364	0	0.0201
DIS.N	-3.0879	11.5299	-0.951	5.2648	0.0436	2.9644	0.9491	1.4088	3.102	-6.3919
p-value	0	0	0	0	0.5712	0	0	0.0388	0	0.0054
EK.N	-4.028	15.7363	-1.3826	7.5854	-0.1561	5.211	1.1288	3.9506	4.1149	-0.2793
p-value	0	0	0	0	0.1214	0	0	0.0007	0	0.9532
GE.N	-3.301	15.705	-0.8661	6.0243	-0.0925	3.7754	0.815	0.9774	3.4193	-7.8217
p-value	0	0	0	0	0.1834	0	0	0.2588	0	0.0123
GM.N	-6.3169	28.2475	-2.2202	13.8359	-0.2188	8.6893	1.8749	2.3839	6.618	-8.0319
p-value	0	0	0	0	0.1416	0	0	0.0824	0	0.2231
HD.N	-3.3201	9.6384	-1.1369	6.1826	-0.2075	3.3815	1.2132	-1.5707	4.2049	-2.6216
p-value	0	0	0	0	0.021	0	0	0.2011	0	0.3329
HON.N	-2.8462	10.4708	-0.945	7.0082	0.0677	4.3099	1.0365	2.1957	2.8997	-5.004
p-value	0	0	0	0	0.4296	0	0	0.0048	0	0.1208
HPQ.N	-2.7607	9.2848	-0.931	6.0047	0.0419	2.8181	0.9851	0.757	3.023	-4.3161
p-value	0	0.0002	0	0	0.6084	0	0	0.4421	0	0.0836
IBM.N	-2.614	9.6866	-0.8584	4.7681	0.1655	3.5558	0.9883	0.6994	2.698	-3.6291
p-value	0	0	0	0	0.06	0	0	0.0725	0	0.1275
INTC.OQ	-3.2438	12.1089	-1.101	7.3408	0.1058	4.4863	1.1602	2.2607	3.7577	-7.6639
p-value	0	0	0	0	0.2973	0	0	0.0181	0	0.0138
IP.N	-3.7047	15.7728	-1.1469	7.1924	0.1218	4.9662	1.1643	1.457	3.4182	-7.4013
p-value	0	0	0	0	0.1633	0	0	0.0813	0	0.006
JNJ.N	-1.7668	6.7447	-0.5207	2.6053	0.0198	1.4631	0.5148	0.4821	1.7498	-3.9716
p-value	0	0	0	0	0.623	0	0	0.2599	0	0.006
JPM.N	-4.0884	15.7239	-1.3125	9.476	-0.1643	4.8028	1.2157	-0.9286	5.635	-18.8152
p-value	0	0	0	0	0.1058	0	0	0.5049	0	0.0005
KO.N	-1.8402	7.1761	-0.6217	3.9485	0.0367	2.5789	0.6757	0.8036	2.0893	-4.3801
p-value	0	0	0	0	0.4949	0	0	0.0967	0	0.0066
MCD.N	-2.2026	6.6707	-0.7661	4.1729	0.163	2.2648	0.9971	0.1589	2.3737	-2.3415
p-value	0	0	0	0	0.0324	0.0004	0	0.7693	0	0.2154
MMM.N	-2.2507	8.7321	-0.7106	5.1293	-0.0014	2.2452	0.6367	0.8864	2.6451	-2.6945
p-value	0	0	0	0	0.9796	0	0	0.1438	0	0.1615
MO.N	-2.3504	9.4397	-0.678	4.5042	0.0567	2.1808	0.7807	0.3763	2.083	-3.9542
p-value	0	0	0	0	0.3472	0	0	0.3879	0	0.0626
MRK.N	-2.7903	11.2622	-0.9827	4.6024	-0.1103	2.2403	1.0317	-0.423	3.1334	-1.9766
p-value	0	0	0	0	0.1783	0.001	0	0.6695	0	0.4303
MSFT.OQ	-2.9861	11.703	-0.976	5.7444	-0.036	3.6646	1.0625	2.3402	3.2304	-4.7537
p-value	0	0	0	0	0.6345	0	0	0.025	0	0.0337
PG.N	-1.8192	6.6554	-0.5137	2.5051	-0.0214	1.0825	0.5509	-0.6796	2.0602	-3.2834
p-value	0	0.0026	0	0	0.645	0.0007	0	0.0911	0	0.0958
T.N	-2.9714	8.8643	-1.0078	4.6346	0.0805	3.1442	1.0726	3.0089	3.3087	-2.0144
p-value	0	0.0002	0	0	0.3305	0	0	0.0017	0	0.236
UTX.N	-2.8071	12.5305	-0.878	5.5151	0.0042	3.6242	0.985	0.647	2.7885	-3.6985
p-value	0	0	0	0	0.9572	0	0	0.3152	0	0.2007
WMT.N	-2.0769	6.404	-0.7121	3.3569	0.0863	2.6276	0.8138	1.2483	2.9295	-0.8655
p-value	0	0	0	0	0.2249	0	0	0.0475	0	0.699
XOM.N	-3.1367	13.1961	-0.8793	4.792	0.1632	2.6374	1.0514	2.1621	3.0518	-4.5539
p-value	0	0	0	0	0.068	0.0004	0	0.0064	0	0.0721

P-values ≤ 0.1 show statistical significance at 90% confidence level or higher.

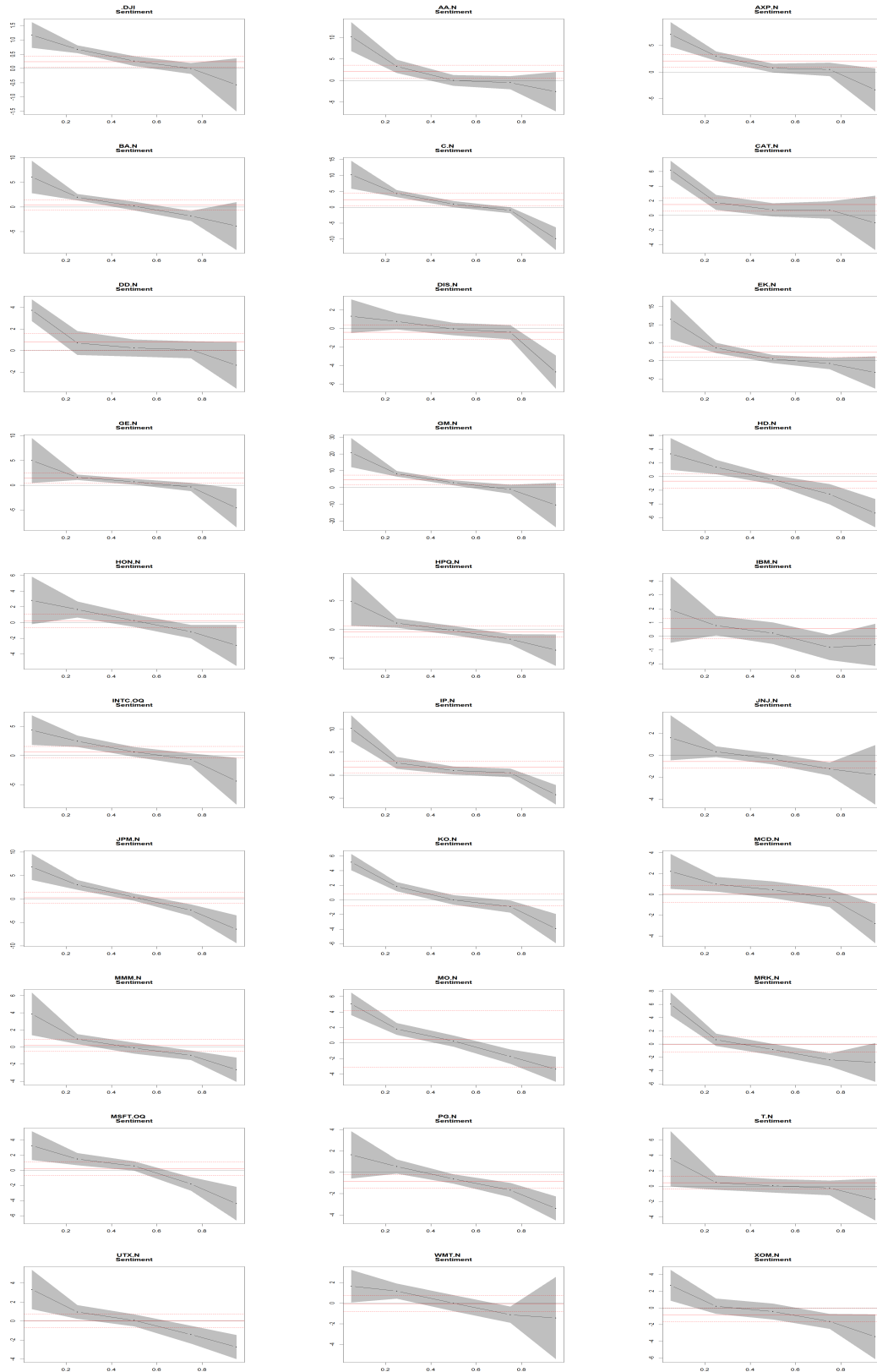


Figure 8: Daily sentiment effect as obtained from augmented Fama French model (2007-2008)

Table 10: QR Results for Model-1 (2010-2011)

Assets	Model-1									
	5%		25%		50%		75%		95%	
	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$	α_i	$\beta_{i,s}$
.DJI	-1.764	5.7	-0.5165	3.1196	0.0803	0.7523	0.5721	0.5056	1.7788	-0.8263
p-value	0	0.0001	0	0	0.0589	0.0899	0	0.4147	0	0.6862
AA.N	-4.1137	13.3513	-1.4604	4.9523	-0.0233	3.505	1.3618	1.836	3.8744	-0.2751
p-value	0	0.0001	0	0.0003	0.8373	0.0024	0	0.2475	0	0.9295
AXP.N	-3.0468	9.0467	-0.9103	2.725	0.0814	1.9479	1.1304	0.811	3.2335	-3.8985
p-value	0	0	0	0.054	0.3149	0.0215	0	0.4465	0	0.1027
BA.N	-2.9299	8.0889	-1.0203	3.3617	0.12	2.5353	1.1359	1.5435	3.2575	-0.5778
p-value	0	0.0005	0	0.0011	0.1911	0.0073	0	0.164	0	0.7547
C.N	-4.7248	12.4372	-1.418	7.504	-0.0207	3.8864	1.468	2.1417	4.5701	-5.8678
p-value	0	0.0001	0	0	0.8444	0.0004	0	0.275	0	0.0544
CAT.N	-3.5075	8.561	-1.0085	3.9549	0.2077	1.6427	1.3619	1.0911	3.7178	-0.6667
p-value	0	0.0005	0	0.0022	0.0484	0.1234	0	0.3346	0	0.8441
DD.N	-2.6285	7.4459	-0.9252	3.232	0.1215	1.5292	1.1932	1.5931	2.7753	-0.0234
p-value	0	0.0002	0	0.0046	0.1628	0.0928	0	0.0721	0	0.9915
DIS.N	-2.5873	9.6162	-0.8748	3.2077	0.0051	0.9776	0.9887	1.5052	2.7746	0.5729
p-value	0	0	0	0.0005	0.944	0.1942	0	0.1352	0	0.855
EK.N	-7.8329	8.8233	-2.4918	5.2683	-0.3015	1.8767	2.0227	2.5349	7.358	-12.4108
p-value	0	0.5114	0	0.0254	0.1244	0.3368	0	0.3103	0	0.2266
GE.N	-2.9946	6.8633	-0.8776	4.2808	0.0675	3.1244	1.0797	1.5225	2.8311	-0.4203
p-value	0	0.0018	0	0	0.3885	0.0001	0	0.1151	0	0.8346
GM.N	-4.0887	9.0228	-1.3988	4.2992	-0.0878	2.9068	1.2867	-1.5588	3.3673	-4.9019
p-value	0	0.1025	0	0.0143	0.5494	0.0402	0	0.3046	0	0.3186
HD.N	-2.3775	4.3608	-0.7228	2.7605	0.0501	1.3009	0.9105	-0.5997	2.6093	-1.539
p-value	0	0.0917	0	0.0025	0.4364	0.0513	0	0.5148	0	0.487
HON.N	-2.7995	8.1111	-0.8034	3.5626	0.0663	2.457	1.0682	1.5798	3.0702	-1.3461
p-value	0	0.0027	0	0.0001	0.4485	0.0066	0	0.1055	0	0.5482
HPQ.N	-3.2198	9.7403	-1.0225	4.1095	0.0324	1.7229	1.0072	2.1437	2.8592	-2.5156
p-value	0	0	0	0.0007	0.693	0.0427	0	0.0274	0	0.2539
IBM.N	-1.9041	6.2195	-0.6105	2.3463	0.1137	0.7219	0.7511	1.0513	2.1477	-1.8759
p-value	0	0	0	0.0011	0.0514	0.2282	0	0.1336	0	0.2138
INTC.OQ	-2.6388	5.2472	-1.0068	2.1322	0.0188	0.8358	1.0954	-0.2615	2.7486	-1.0555
p-value	0	0.0049	0	0.0168	0.8141	0.3179	0	0.7882	0	0.5751
IP.N	-3.9409	7.7698	-1.4591	4.0261	0.1214	1.5821	1.5055	0.4534	4.0703	-5.9951
p-value	0	0.0744	0	0.0141	0.3452	0.2256	0	0.7343	0	0.0455
JNJ.N	-1.4875	4.1566	-0.4962	1.4532	-0.0025	0.5978	0.5163	0.1145	1.5775	0.4551
p-value	0	0	0	0.0391	0.9448	0.0983	0	0.8319	0	0.7635
JPM.N	-3.4753	8.3704	-1.2745	5.8126	-0.0483	3.21	1.2065	0.8883	3.4002	-4.6761
p-value	0	0.0043	0	0	0.6089	0.0012	0	0.5532	0	0.1213
KO.N	-1.683	5.4805	-0.4899	1.2841	0.1038	0.8861	0.6214	-0.5751	1.7167	0.4578
p-value	0	0.0014	0	0.0423	0.0109	0.035	0	0.3853	0	0.6874
MCD.N	-1.6025	1.6404	-0.384	0.7898	0.105	0.261	0.5816	-0.1224	1.8192	0.0382
p-value	0	0.0572	0	0.1516	0.012	0.5439	0	0.8203	0	0.972
MMM.N	-2.3964	7.3262	-0.7626	3.2203	0.0485	1.3402	0.8931	1.6415	2.3364	-4.4255
p-value	0	0.0049	0	0.0001	0.4324	0.0375	0	0.0509	0	0.0281
MO.N	-1.6064	3.8565	-0.4722	1.571	0.1593	0.7806	0.6937	0.2664	1.5705	-1.5376
p-value	0	0.0207	0	0.0188	0.0002	0.0782	0	0.6535	0	0.0041
MRK.N	-1.9556	7.155	-0.7549	1.566	0.0275	0.6954	0.8171	0.6051	2.1141	-2.2254
p-value	0	0.0006	0	0.0229	0.7043	0.3523	0	0.3062	0	0.1179
MSFT.OQ	-2.317	4.5977	-0.8792	1.3971	0.0221	1.3507	0.8169	1.0119	2.2071	-3.0959
p-value	0	0.0306	0	0.0602	0.7297	0.0395	0	0.2164	0	0.1654
PG.N	-1.4934	1.8703	-0.4066	1.1107	0.0571	0.4569	0.5334	-0.3924	1.4113	0.2215
p-value	0	0.1507	0	0.0365	0.1721	0.2923	0	0.4333	0	0.8461
T.N	-1.7716	3.0521	-0.6103	2.2873	0.081	1.3597	0.7125	1.0517	1.6945	1.2703
p-value	0	0.048	0	0.0002	0.0889	0.0061	0	0.1196	0	0.3154
UTX.N	-2.4218	5.3155	-0.701	3.0257	0.0357	1.0249	0.7772	0.2652	2.5642	-0.2818
p-value	0	0.046	0	0.0001	0.504	0.0661	0	0.7425	0	0.8987
WMT.N	-1.5412	3.1365	-0.4532	1.1923	0.0748	0.4104	0.5532	0.1705	1.4643	-2.0327
p-value	0	0.0004	0	0.0613	0.1149	0.3662	0	0.7368	0	0.0406
XOM.N	-2.1617	6.4947	-0.6457	2.8112	0.0683	1.506	0.8245	0.3809	2.2555	0.0263
p-value	0	0	0	0	0.2887	0.0234	0	0.6282	0	0.9896

P-values ≤ 0.1 show statistical significance at 90% confidence level or higher.

4.2. Results from the Seven Day SMA of the Daily Sentiment Scores

Here we present the results obtained from using a seven day SMA of daily sentiment scores. We repeat the analysis with this SMA measure to account for the publicly released news on the weekends and non trading days and associated sentiment. The 7 day simple SMA takes into account these days which should get captured in the sentiment effect (hereby called $\beta_{i,s7}$) from the two models. The major point of investigation here is to see if accounting for the non-trading days by taking the SMA of the daily scores changes the results considerably. Table-12 shows the results as obtained from both the models using the SMA sentiment scores. The Model-1 results are again all significant but there is a drop in significant $\beta_{i,s7}$ numbers for Model-2 from 8 to 5 as compared to the same OLS results obtained from daily sentiment scores. This indicates that the effect of news sentiment on non-trading days is not as significant as the effect of daily news sentiment.

Considering figure-9 which shows the frequency of news reported in year 2008 divided by days, it is clear that the number of news reported on weekends is considerably less as compared to the number of news reported on the weekdays (usually the trading days). This can be accounted as the major reason why the SMA scores do not improve on the results from the daily sentiment scores. We conducted the analysis for all the periods using both OLS and QR and found similar inferences, hence we will not report the rest of the result for the sake of brevity⁵.

⁵The analysis is also repeated for various lags (1 to 5 days) and SMA (14,28,60,90 days) periods but all other specifications do not work as well as the one with daily sentiment scores.

Table 12: OLS Results for 2007 to 2012 for 7 Day SMA

Assets	Model-1		Model-2				
	α_i	$\beta_{i,s7}$	α_i	$\beta_{m,i}$	s_i	h_i	$\beta_{i,s7}$
DJI	-0.0108	0.9134	-0.0108	0.9134	-0.1938	-0.0495	0.114
p-value	0.1624	0**	0.1624	0***	0***	0***	0.0909*
AA.N	-0.0468	1.7124	-0.0468	1.7124	-0.2307	-0.1589	1.4576
p-value	0.4769	0***	0.4769	0***	0.0099***	0.0843*	0.0112**
AXP.N	-0.0263	1.3555	-0.0263	1.3555	-0.2543	1.0084	0.0118
p-value	0.628	0***	0.628	0***	0.0006***	0***	0.9801
BA.N	-0.0144	1.0272	-0.0144	1.0272	-0.0908	-0.2184	0.5029
p-value	0.7384	0***	0.7384	0***	0.1206	0.0003***	0.1807
C.N	0.0882	1.5459	0.0882	1.5459	-0.3187	2.6855	2.1529
p-value	0.6727	0***	0.6727	0***	0.2614	0***	0.2377
CAT.N	0.0342	1.2395	0.0342	1.2395	0.3106	-0.0905	0.6902
p-value	0.4449	0***	0.4449	0***	0***	0.1495	0.0779*
DD.N	-0.0202	1.1317	-0.0202	1.1317	0.1093	0.0278	0.1848
p-value	0.5645	0***	0.5645	0***	0.0219**	0.5708	0.5462
DIS.N	0.0008	1.0875	0.0008	1.0875	-0.179	-0.0409	0.0005
p-value	0.9809	0***	0.9809	0***	0.0002***	0.408	0.9986
EK.N	-0.2198	1.0644	-0.2198	1.0644	0.2145	0.9658	1.262
p-value	0.205	0***	0.205	0***	0.3627	0.0001***	0.4047
GE.N	-0.0167	1.0129	-0.0167	1.0129	-0.053	0.7179	0.7228
p-value	0.719	0***	0.719	0***	0.4007	0***	0.0749*
GM.N	-0.1639	1.3693	-0.1639	1.3693	0.0856	0.9737	2.5066
p-value	0.3408	0***	0.3408	0***	0.686	0***	0.0702*
HD.N	-0.0136	0.9158	-0.0136	0.9158	0.1555	0.1689	-0.3621
p-value	0.7416	0***	0.7416	0***	0.0055***	0.0035***	0.3143
HON.N	0.0216	1.0943	0.0216	1.0943	0.0759	-0.1382	0.4786
p-value	0.54	0***	0.54	0***	0.114	0.0053***	0.1209
HPQ.N	-0.0871	1.0166	-0.0871	1.0166	-0.0903	-0.3766	0.2479
p-value	0.082*	0***	0.082*	0***	0.1841	0***	0.5705
IBM.N	0.0411	0.8025	0.0411	0.8025	-0.0972	-0.2721	0.2367
p-value	0.1832	0***	0.1832	0***	0.0206**	0***	0.3803
INTC.OQ	0.0005	1.1234	0.0005	1.1234	-0.0714	-0.4341	0.3611
p-value	0.9904	0***	0.9904	0***	0.2005	0***	0.3139
IP.N	0.0267	1.2776	0.0267	1.2776	0.124	0.635	0.8088
p-value	0.6863	0***	0.6863	0***	0.1676	0***	0.1614
JNJ.N	-0.0182	0.5736	-0.0182	0.5736	-0.2868	-0.2472	-0.1397
p-value	0.4359	0***	0.4359	0***	0***	0***	0.4931
JPM.N	-0.0426	1.1789	-0.0426	1.1789	-0.059	2.266	-0.4733
p-value	0.4337	0***	0.4337	0***	0.4251	0***	0.3193
KO.N	-0.0316	0.6263	-0.0316	0.6263	-0.1764	-0.3286	0.0834
p-value	0.6198	0***	0.6198	0***	0.0417**	0.0002***	0.8808
MCD.N	0.0298	0.5967	0.0298	0.5967	-0.0749	-0.211	0.117
p-value	0.3432	0***	0.3432	0***	0.0792*	0***	0.6697
MMM.N	-0.0224	0.852	-0.0224	0.852	0.0202	-0.0746	-0.0896
p-value	0.4593	0***	0.4593	0***	0.6237	0.0789***	0.7348
MO.N	-0.0968	0.59	-0.0968	0.59	-0.3062	-0.3975	-0.1642
p-value	0.3544	0***	0.3544	0***	0.0312**	0.0067***	0.8572
MRK.N	-0.003	0.8515	-0.003	0.8515	-0.3868	-0.3936	0.1989
p-value	0.947	0***	0.947	0***	0***	0***	0.6128
MSFT.OQ	-0.0165	1.0508	-0.0165	1.0508	-0.2728	-0.5585	0.1966
p-value	0.6688	0***	0.6688	0***	0***	0***	0.5601
PG.N	-0.0172	0.5789	-0.0172	0.5789	-0.3079	-0.0989	-0.1271
p-value	0.5282	0***	0.5282	0***	0***	0.0097***	0.5937
T.N	-0.0113	0.8036	-0.0113	0.8036	-0.4718	-0.0489	0.0787
p-value	0.7339	0***	0.7339	0***	0***	0.2931	0.7861
UTX.N	0.0068	0.9893	0.0068	0.9893	-0.0391	-0.1626	0.2683
p-value	0.8191	0***	0.8191	0***	0.331	0.0001***	0.2996
WMT.N	0.0057	0.5664	0.0057	0.5664	-0.1195	-0.2447	-0.286
p-value	0.8632	0***	0.8632	0***	0.0074***	0***	0.3186
XOM.N	-0.0199	1.0742	-0.0199	1.0742	-0.6039	-0.4841	-0.1155
p-value	0.518	0***	0.518	0***	0***	0***	0.6669

P-values <= * show statistical significance at 90% confidence level, ** =95%, *** =1%. HAC standard errors

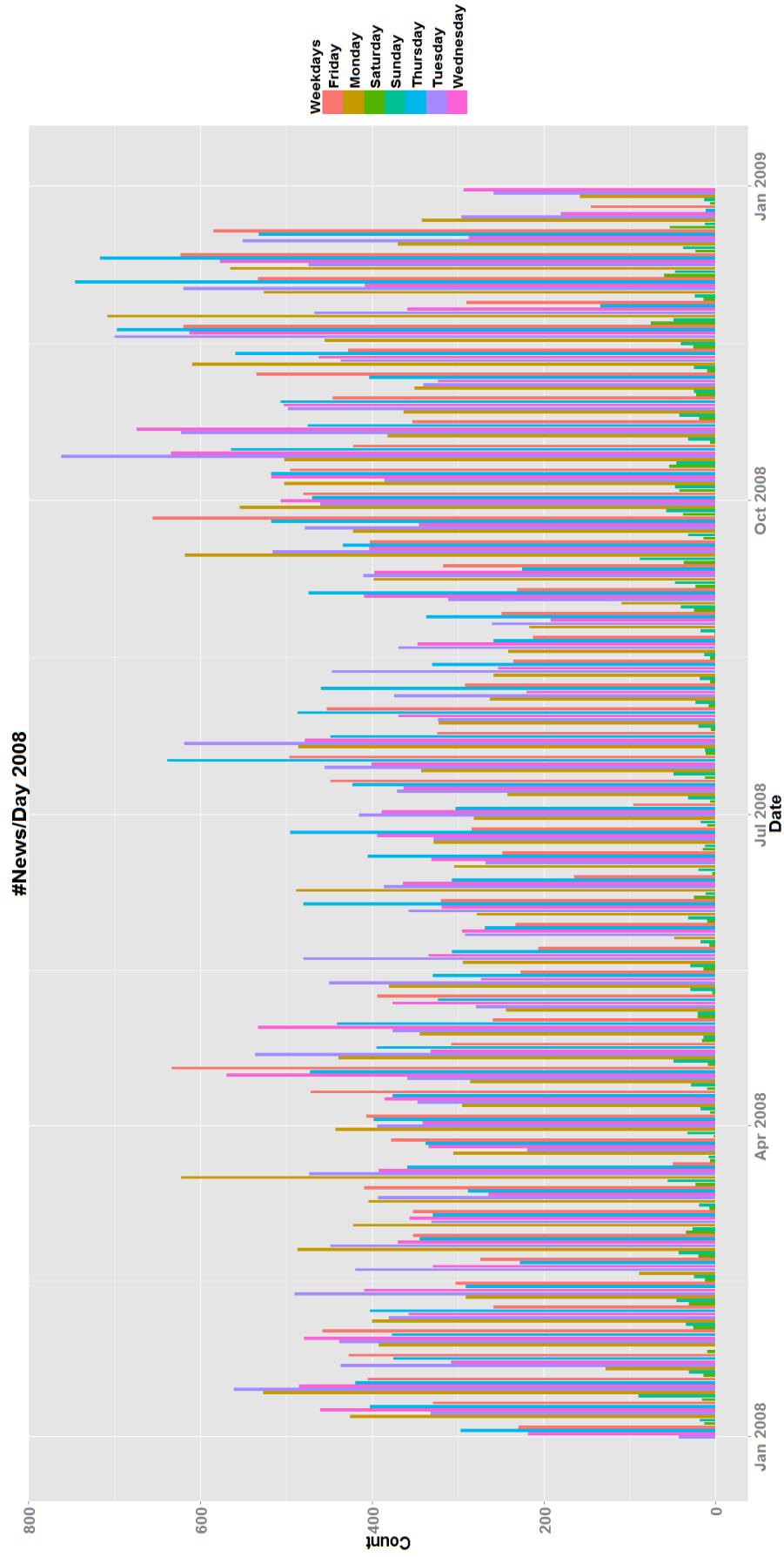


Figure 9: No of news sentiment reported per day for year 2008.

5. Conclusion

In this paper we have used the Financial News Sentiment scores reported by TRNA to study the effects of daily aggregated market news sentiment score series that we constructed on the stock returns traded in the DJIA. We proceeded to build a daily sentiment score for the DJIA by aggregating the sentiment scores for all the traded stocks in the DJIA using probability weights. This daily sentiment score and its SMA was then used to quantify the effect of market news sentiment on stock returns. The analysis evaluated two regression models with only sentiment scores as the independent variable and then adopted an augmented Fama French three factor model. We presented a comparative analysis of the two regression models using OLS and QR. The results from the empirical analysis clearly demonstrate that stock prices are significantly affected by the financial news sentiment generated during the trading day. The results obtained from QR strongly suggest that there is a more significant $\beta_{i,s}$ effect on the lower stock returns than on the higher stock returns which is one of the major inferences drawn from the analysis. This is also consistent with the fact that the mean sentiment score is negative. The results from this study show that the financial news sentiment factor adds significantly to the traditional asset pricing model and can be a relevant additional factor in asset pricing. Similar results were found in the recent studies by Cahan et al. (2009) and Hafez and Xie (2012b) using the Ravenpack sentiment dataset. There are no previous studies either using the Ravenpack or TRNA datasets which study the effect of daily market sentiment on stock price returns using OLS and QR. Although comprehensive the analysis could be further extended by using a different weighting scheme for generating the market sentiment series and also by expanding the frequency of the analysis from daily to a higher frequency 'within the day' analysis.

Acknowledgements

The authors gratefully acknowledge the support of the QUANTVALLEY/FdR: 'Quantitative Management Initiative'. We are also grateful to SIRCA for providing the TRNA

data sets. For financial support, the first author wishes to thank the Australian Research Council, and the second author wishes to acknowledge the Australian Research Council and the National Science Council, Taiwan.

References

- Allen, D., Powell, R., and Singh, A. (2011). Quantile regression as a tool for portfolio investment decisions during times of financial distress. *Annals of Financial Economics*, 6(01):1150003.
- Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4):1645–1680.
- Barber, B. M. and Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly journal of Economics*, pages 261–292.
- Barber, B. M. and Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies*, 21(2):785–818.
- Bhattacharya, U., Galpin, N., Ray, R., and Yu, X. (2009). The role of the media in the internet ipo bubble. *Journal of Financial and Quantitative Analysis*, 44(03):657–682.
- Borovkova, S. and Mahakena, D. (2015). News, volatility and jumps: the case of natural gas futures. *Quantitative Finance*, (ahead-of-print):1–26.
- Cahan, R., Jussa, J., and Luo, Y. (2009). Breaking news: How to use news sentiment to pick stocks. *Macquarie US Equity Research*.
- Da, Z., Engelberg, J., and Gao, P. (2011). In search of attention. *The Journal of Finance*, 66(5):1461–1499.

- diBartolomeo, D. and Warrick, S. (2005a). *Linear Factor models*, chapter Making covariance based portfolio risk models sensitive to the rate at which markets reflect new information, pages Chapter–12. Elsevier Finance.
- diBartolomeo, D. and Warrick, S. (2005b). Making covariance based portfolio risk models sensitive to the rate at which markets reflect new information" chapter 12 in linear factor models edited. knight, j. and satchell, s., elsevier finance.
- Dougal, C., Engelberg, J., Garcia, D., and Parsons, C. A. (2012). Journalists and the stock market. *Review of Financial Studies*, page hhr133.
- Dzielinski, M. (2012). Which news resolves asymmetric information? *Available at SSRN 2102512*.
- Dzielinski, M., Rieger, M. O., and Talpsepp, T. (2011). Volatility asymmetry, news, and private investors. *The Handbook of News Analytics in Finance*, pages 255–270.
- Engelberg, J. E. and Parsons, C. A. (2011). The causal impact of media in financial markets. *The Journal of Finance*, 66(1):67–97.
- Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns, journal of finance (june).
- Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56.
- Groß-Klufmann, A. and Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance*, 18(2):321–340.
- Hafez, P. and Xie, J. (2012a). Ravenpack sentiment and its relationship with macro-economic indicators. *RavenPack International SL*.

- Hafez, P. A. and Xie, J. (2012b). Factoring sentiment risk into quant models. *Available at SSRN 2071142*.
- Huynh, T. D. and Smith, D. R. (2013). News sentiment and momentum. *FIRN Research Paper*.
- Koenker, R. (2005). Quantile regression. econometric society monograph series 38.
- Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.
- Leinweber, D. and Sisk, J. (2011). Relating news analytics to stock returns the handbook of news analytics in finance (pp. 147-172).
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37.
- Mitra, L. and Mitra, G. (2011). Applications of news analytics in finance: A review. *The Handbook of News Analytics in Finance*, 596:1.
- Mitra, L. R., Mitra, G., and Di Bartolomeo, D. (2008). Equity portfolio risk (volatility) estimation using market information and sentiment. *Available at SSRN 1425624*.
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of economic theory*, 13(3):341–360.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The journal of finance*, 19(3):425–442.
- Sinha, N. R. (2010). Underreaction to news in the us stock market. *Available at SSRN 1572614*.
- Smales, L. A. (2014). News sentiment in the gold futures market. *Journal of Banking & Finance*, 49:275–286.

- Storckenmaier, A., Wagener, M., and Weinhardt, C. (2012). Public information in fragmented markets. *Financial Markets and Portfolio Management*, 26(2):179–215.
- Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.
- Tetlock, P. C. (2010). Does public financial news resolve asymmetric information? *Review of Financial Studies*, 23(9):3520–3557.
- Tetlock, P. C., SAAR-TSECHANSKY, M., and Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, 63(3):1437–1467.