Contents lists available at ScienceDirect

Decision Support Systems

Do President Trump's tweets affect financial markets?

Peder Gjerstad [a], Peter Filip Meyn [a], Peter Molnár [a,b,c,d,*,1], Thomas Dowling Næss [a]

[a] *Norwegian University of Science and Technology, Trondheim, Norway*
[b] *University of Stavanger, Stavanger, Norway*
[c] *Prague University of Economics and Business, Prague, Czech Republic*
[d] *Nicolaus Copernicus University in Torun, Torun, Poland*

ARTICLE INFO

ABSTRACT

Frequent tweets of the former president of the United States, Donald Trump, provide a unique opportunity to study how financial markets respond to his statements. To do this, we utilize a precise timestamp of each tweet together with high-frequency financial data. We start by analyzing the impact of tweets in general, irrespective of their content. We find that tweets by President Trump are followed by increased uncertainty, increased trading and a decline in the US stock market. We utilize two methods in order to study whether the market reaction depends on the content of the tweets. First, classification of Trump's tweets depending on whether they contain a specific word reveals that market response is particularly negative for tweets containing the words "products" and "tariff". Second, we use Latent Dirichlet Allocation to affiliate tweets with distinct topics. We find that while most topics do not impact financial markets, the US stock market responds to tweets related to the topic of a "trade war" by price decline, increased trading volume and increased uncertainty. The "trade war" tweets affect other financial markets too, as the Chinese stock market responds to these tweets negatively, while the price of gold responds positively. We illustrate the practical importance of our approach by an automated trading system, which achieves positive abnormal returns.

1. Introduction

This paper analyzes the impact of President Donald Trump's tweets on broad financial markets. Since he was elected in November 2016, he used the social media platform Twitter in a way that is unparalleled by any previous high ranking official. During the period from May 2009 to October 2019, the number of tweets (including retweets) published by Mr. Trump amounts to 40,376, with more than 10,000 of them from his period as president of the United States. The tweet frequency has increased steadily throughout his presidency.

Social media generally, and Twitter in particular, provides real-time news and other information directly to the audience. Twitter data has been used in a wide range of applications, including the detection and location estimation of disasters [29] and the prediction of business trends [33] and sports matches outcomes [36,37]. Alongside an increasing usage of Twitter, a growing number of firms create corporate Twitter accounts to disseminate information to followers, stakeholders and investors [16,26]. This increases the relevance of Twitter as a source of valuable information to the financial markets. However, some

companies might be manipulating Twitter sentiment about their products [27]. Previous studies have found that Twitter sentiment can predict stock prices [2,8,34], and also cryptocurrency prices [25].

Frequent tweets by President Trump, together with the importance of his position, caused his tweets to be a subject of several studies. Mr. Trump has used Twitter to criticize, attack and praise various companies. Ge et al. [14], Born et al. [10], Juma'h and Alnsour [22], Ichev et al. [21], Ajjoub et al. [1] and Brans and Scholtens [11] investigate firm-specific tweets from Mr. Trump and mainly find a significant impact on stock returns of these companies. Researchers have studied the impact of Trump's tweets also on the overall US stock market [12,13,35], Chinese stock market [18] and European financial markets [24]. Bianchi et al. [6] show that FED-related tweets by President Trump have an impact on expectations about monetary policy.

All the above-mentioned papers except Bianchi et al. [6] are based on daily prices. It is very problematic to interpret correlation between daily price changes and Trump's tweets during that day in a causal way due to omitted variable problem. Consider the following example: New US unemployment rate is published, and it is much lower than expected.

Stock market goes up, and Donald Trump posts a tweet that he is a great president and because of him, unemployment went down. Therefore, it is completely possible to observe a strong association between Trump's tweets and financial markets in daily data, yet this association is caused by both financial markets and Donald Trump responding to same news. Thus, causality cannot be established.[2]

To overcome this issue, we utilize high-frequency financial data together with exact time stamps of Trump's tweets. This setup allows us to directly observe what happens to financial markets right after President Trump published his tweet. These results directly answer whether Trump's tweets influence financial markets, because it is unlikely that Donald Trump is responding to news faster than high-frequency traders. There exists a very good paper based on high-frequency data [6], but the focus is limited to tweets related to monetary policy and their impacts on the expectations about monetary policy.

Our next contribution is that while the above-mentioned papers study some subsets of Mr. Trump's tweets related to a particular subject, we use all the tweets of President Trump. Moreover, since these tweets cover a wide range of areas, we associate tweets with distinct topics in a fully automated way utilizing Latent Dirichlet Allocation (LDA). This allows us to obtain detailed insights whether the impact of President Trump's tweets on financial markets differs across these topics.

LDA, first presented in a machine learning context by Blei et al. [7], has become a popular method to extract topics from a set of documents. Under the correct assumptions, LDA extraction should capture natural topic structures in text documents that match human interpretation [17]. LDA has been used in various applications. For example, to find determinants of crowdfunding success [40] or to detect automobile insurance fraud [39]. The LDA approach is usually applied to longer text documents such as articles, while tweets, on the other hand, are limited to 280 characters. However, existing literature has proven that it is possible to successfully use the LDA on Twitter data [15,20,23,28,30,32], motivating our effort to apply the LDA analysis to President Trump's tweets.

We find that the topic of "trade war" has the strongest impact on the financial markets. These tweets were followed by price decline, increased trading volume and increased uncertainty in the US stock market. "Trade war" tweets have a negative impact on the Chinese stock market too but a positive impact on the gold price. We illustrate the practical importance of our approach by a simple, fully automated trading system. This trading system achieves both higher returns and lower volatility compared to a buy-and-hold strategy.

The rest of this paper is organized as follows. Section 2 describes data sources used in the paper and definitions of variables. Section 3 presents the approach for extracting topics from tweets. Section 4 describes the methodology, as well as the results and discussions. Section 5 concludes.

2. Data

We use "Trump Twitter Archive"[3] as the source of Trump's tweets. This dataset has been previously utilized e.g. by Rashid and de Leon [35] who create trading strategies based on Trump's twitter posts. The data contains a timestamp, tweet text content, source (kind of device used) and specifies whether or not the tweet is a retweet. As of September 22, 2019, this set consisted of 40,842 tweets. Since we are only interested in tweets from after President Trump got elected, we remove all tweets published prior to this date, which reduces the sample size to 10,911 tweets. We consider retweets to be less significant in being market-moving than tweets originating from President Trump himself (or one

Table 1
Overview of assets. Time period spans November 8, 2016 and September 15, 2019.

Asset	Type	Currency	Frequency	No. of obs.
S&P500	Futures	USD	1 min	993,843
WTI Crude Oil	Futures	USD	1 min	994,853
Hang Seng	Futures	HKD	1 min	575,010
Gold	Futures	USD	1 min	1,011,258
VIX	Futures	USD	1 min	607,483

of his advisors), hence we exclude retweets from the sample, resulting in the final sample containing 8,686 tweets.

We want to investigate whether any relationship exists between President Trump's tweets and movements of the financial markets, represented by the S&P 500 index, the VIX index, the oil price, the Hang Seng index and the gold price. S&P 500 index represents the performance of 500 large-cap companies listed on stock exchanges in the United States. The VIX index represents the market's 30-day expected volatility of S&P 500 and is calculated using the implied volatility from CBOE-traded SPX options. The index is widely used as a measure of investors' perception of risk. West Texas Intermediate (WTI) is a light and sweet grade of crude oil and is widely used as a benchmark in oil pricing. The Hang Seng index measures the stock performance of the 50 largest listed companies in Hong Kong. In general, this index is perceived as a good measure of the economy of Hong Kong, which is closely linked with the Chinese economy due to its status as a special administrative region of China. Due to the fact that Gold is highly liquid and has a low to negative correlation with other major financial asset classes, is a high-value metal often used as a hedge or as a safe haven in financial matters [4,5].

We utilize high-frequency (minute-by-minute) futures data obtained from Backtest Market.[4] An overview of the data sets is provided in Table 1. The data sets consist of the date, time, open, high, low, close and volume during all trading hours of the respective financial asset. The data set spans the period between April 1, 2007, and September 18, 2019. We are, however, only concerned with market data after November 8, 2016, which is the date of the 2016 United States presidential election. All data sets are adjusted for daylight saving time and converted into UTC. The futures contracts are rolled when the next one has become more traded than the expiring one, which generally happens a few days before expiration.

The $Return_{t, t+d}$ is calculated from an asset price $Price_t$ at the tweet time t, time interval d, as:

$$Return_{t,t+d} = \frac{Price_{t+d}}{Price_t} - 1 \qquad (1)$$

The trading volume varies greatly throughout the day with large spikes around the opening and closing trading auctions. Moreover, President Trump's tweets are not uniformly distributed throughout the day; they generally spike around lunchtime and late evening. We, therefore, introduce the abnormal trading volume (denoted simply as *Volume*), where we adjust the logarithm of the trading volume (*RawVolume*) each minute by subtracting the average logarithmic trading volume on that minute from the whole sample. In other words, when we calculate the abnormal trading volume for 10:13 on a particular day, we subtract the average logarithmic trading volume at 10:13 for all the days in the sample period. Formally:

$$Volume_{t+d} = logRawVolume_{t+d} - \frac{1}{N} \sum_{\hat{t} \in T:minute(t+d)=minute(\hat{t})} logRawVolume_{\hat{t}} \quad (2)$$

Since the asset prices are available for each whole minute, the tweet

World Leader Meetings

begin peopl
govern enemi
north_korea
special
meet
head place
summit

TV Shows

last_night
foxandfriend
show
man amp
peopl great
thank rate
wonder

Corruption

right iran
presid said
go
say would
fire power
peopl

Trade War

dollar
trade billion
united_st countri
china
pay tariff
amp
product

Bragging

number
countri best
record economi
time
ever great
histori world

Making Deals

year great mexico
get
much would make
new
deal
done

North Korea/States

un chief game
play
gener relationship
kim_jong brave
state
one

US Military

border
love great
amp
america
strong militari
support full make

Achievements

hard
great
big work
congratul
win amp job
governor
much

Compassion

serv
famili
wonder peopl prayer
united_st men
women
amp
thought

Border Security

must
democrat
get republican
secur
want
need border
vote amp

Political Discussions

presid thing
made good
meet talk
well
mani
time happen

Undefined

join
today http
thank
great american
nation honor
co
whitehous

Border Wall/Board

plan state
built
secretari
california
wall build
pleas new
announc

Announcements

decis see
john
enjoy foxnew
interview
supreme_court crowd
justic tonight

Fake News

report news
peopl media cnn
amp
fake_new poll bad
stori

Russian Collusion

amp campaign
democrat
trump collus
total russia
witch_hunt
russian investig

Employment

year
job amp
busi american
back million
countri come
mani

Illegal Immigrants

fbi
countri
peopl
illeg must
mani law
stop
amp southern_bord

Presidents

obamacar elect sinc
trump
administr year
month
two
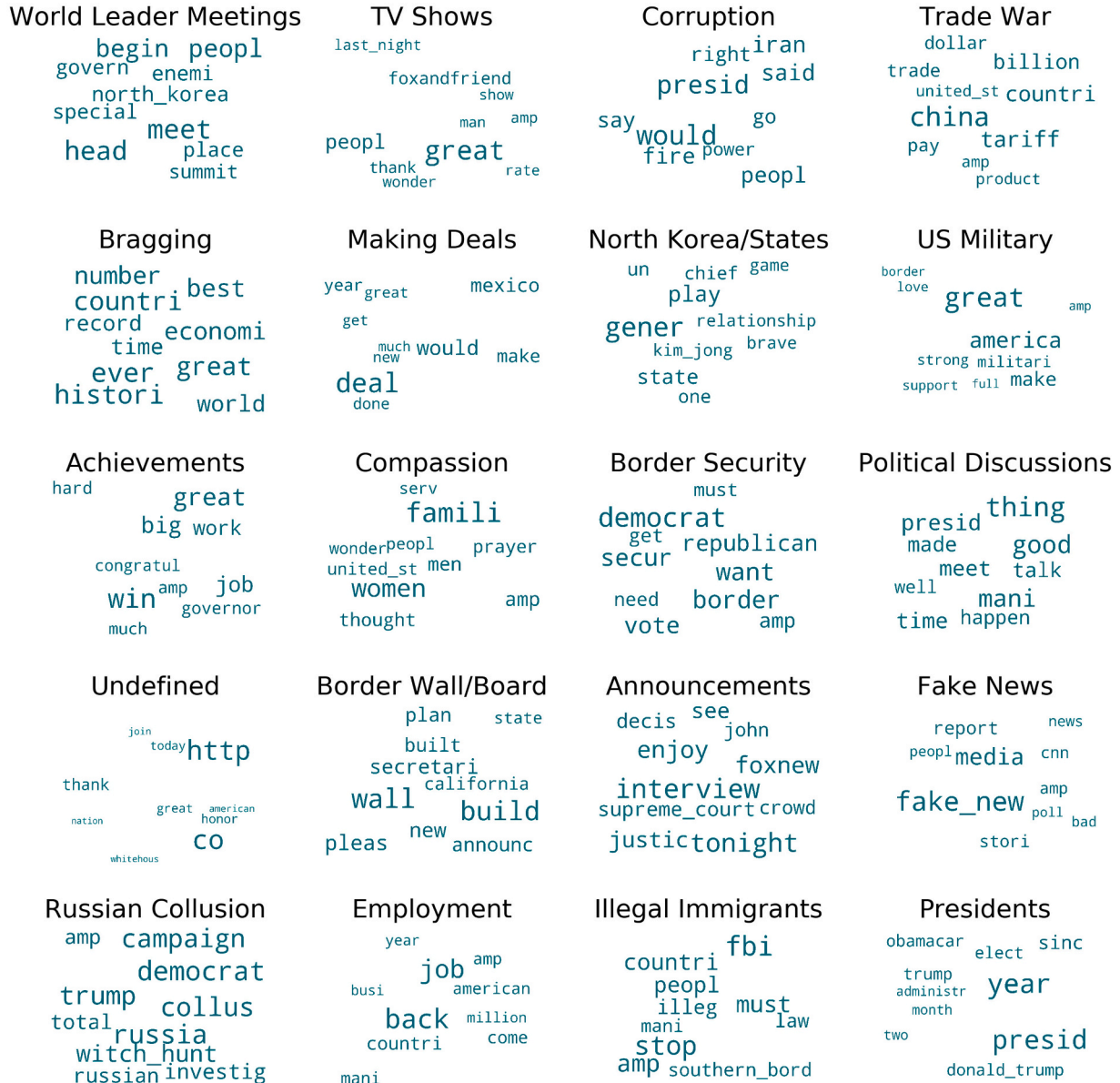presid
donald_trump

Fig. 1. Word cloud for each topic.

Table 2
Regression results for the impact of Trump's tweeting in general on the S&P 500 return.

	Dependent variable: Return$_{t, t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0001***	0.0009***	0.0019***	0.0031***	0.0042***	0.0054***	0.0096***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0000	−0.0005	−0.0016*	−0.0022**	−0.0010	−0.0010	−0.0015
	(0.941)	(0.415)	(0.052)	(0.038)	(0.467)	(0.542)	(0.564)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 3
Regression results for the impact of Trump's tweeting in general on the S&P 500 abnormal trading volume.

	Dependent variable: Volume$_{t, t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0076***	0.0057	0.0047	0.0053	0.0063	0.0022	0.0032
	(0.000)	(0.127)	(0.378)	(0.413)	(0.435)	(0.804)	(0.793)
Tweet	0.0612***	0.0426***	0.0424***	0.0559***	0.0644***	0.0621***	0.0575***
	(0.000)	(0.007)	(0.008)	(0.001)	(0.000)	(0.000)	(0.003)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

timestamp is rounded down to the nearest minute to ensure that the value at time t is completely unaffected by that tweet. For the return values, the "open" price is used. If the price at time t or at time $t + d$ is unavailable (if one of the two timestamps are outside trading hours), that tweet is ignored.

3. Topic extraction with Latent Dirichlet Allocation

3.1. Tweet preprocessing

Textual data available in natural language form is not suitable for use with data mining techniques. Natural language contains words that are inflected, i.e. words expressed in distinct grammatical terms. In order to produce representative insight from the text data, we need to use techniques and principles from text mining and natural language processing (NLP). Following Tripathi et al. [38], we use tokenization, stop words, filtering and stemming. Additionally, bigrams were created to capture two-word sequences. The following steps describe how the tweets were preprocessed:

1. **Tokenization:** Each tweet is divided into its fundamental structure (words).
2. **Make bigrams:** Bigram is an arrangement of two elements (words) that repeatedly occur together in a text document. "Fake news", for example, makes the bigram "fake_news".
3. **Stop words filtering:** Filtering out "stop words", i.e., words that are too common in the English language to contribute to separating the tweets into different topics. The Python package `nltk.corpus` includes a list of English "stop words", which we use for this purpose.
4. **Stemming:** Stemming is a text normalization technique used to reduce the inflection in words and makes it possible to map a collection of words to the identical root. "Hopefully" and "Hopeful" are, for instance, reduced to the root "Hope" to group them as the same entity. There are three common stemming methods in use today: Porter, Snowball, and Lancaster. We use Porter, as this is the most common method, and also the least "aggressive" of the three. For stemming, we use the Python package `nltk.stem.porter`.

3.2. The LDA method on tweets

Latent Dirichlet Allocation (LDA) is used to identify topics automatically and to infer the topic distribution of each tweet. The methodology and the terminology follows Blei et al. [7]. For the implementation, we use the python package `gensim.models.ldamodel`.

The LDA model does not inform us what the number of topics of a corpus should be. A necessary input of building our LDA model is, therefore, the number of topics, K. By utilizing "topic coherence", a measure of how well separated the topics are, our LDA model was run for different values of $K \in 060$ and calculate the score for each iteration. As we want the number of topics to be few enough to have distinct semantic interpretations, we also take this into consideration. For our data set, the highest coherence scores lie between values of 0.41 and 0.44, depending on K. $K = 20$ was chosen with a coherence score of 0.42 for our analysis.

The output of the LDA model is, as mentioned, a distribution of topics with size K, each containing a set of characteristic words. These topics arise from statistical properties only and do not necessarily have anything semantically in common. Human interpretation is therefore needed to generate topic names.

3.3. Tweet topics

The resulting topics are summarized in Fig. 1 as twenty word clouds containing the ten most important words for each topic. The relative size of each word represents the relative importance of that word in the given topic. We would like to emphasize that these twenty topics are generated in a fully automated way, while topic title are created by us after we inspected the word clouds as well as tweets that received high probability of belonging to these topics.

4. Results

We analyze short-term impact of Mr. Trump's tweets on financial markets. Our primary interest is the S&P 500 index because it measures the market movements in Trump's home country.

Since we work with high-frequency data, the return-series are very noisy and contain extreme outliers in both tails of the distribution. We therefore choose to be conservative and estimate robust linear models via iteratively reweighted least squares and use HuberT [19] as the robust criterion for downweighting outliers on the return- and VIX series. For the abnormal (log) volume-series, however, we use standard OLS, but with robust (heteroscedasticity-consistent (HC)) standard

Table 4
Regression results for the impact of Trump's tweeting in general on the VIX futures return.

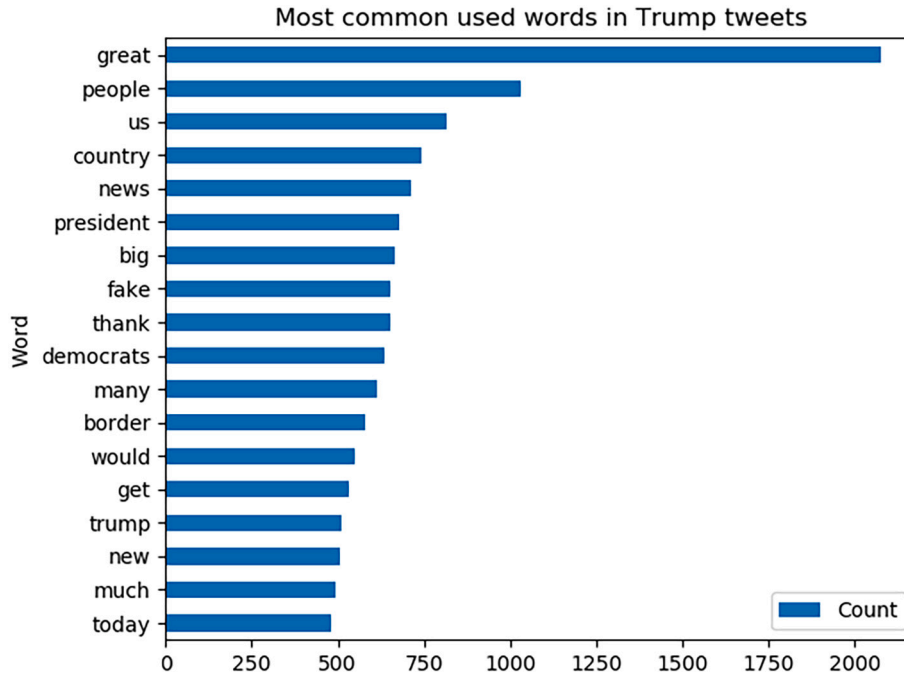|	Dependent variable: VIXreturn$_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	−0.0002	−0.0068***	−0.0212***	−0.0302***	−0.0460***	−0.0466***	−0.0890***
	(0.472)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0002	0.0103	0.0250**	0.0374***	0.0318*	0.0293	0.0293
	(0.945)	(0.188)	(0.020)	(0.004)	(0.055)	(0.129)	(0.340)
Observations	510,460	51,477	26,938	19,663	13,219	11,377	7171

Note: *p<0.1; **p<0.05; ***p<0.01.

Fig. 2. Most common words (excluding stop words) used by Mr. Trump in tweets.

errors of MacKinnon and White [31]. The robust linear model is implemented through the Python package `statsmodels.api.RLM`, and specifying `M = statsmodels.api.robust.norms.HuberT()`, while the OLS is implemented through the Python package `statsmodels.api.OLS` and specifying `cov_type = "HC1_se"`.

However, acknowledging that OLS is the most common approach, results for OLS on S&P and VIX return were also calculated. All conclusions remain the same for both approaches, yet as expected, the robust linear models report notably smaller, but more significant coefficients. For the sake of brevity, these results are not reported in this paper but are available upon request.

First we analyze what happens to financial markets after President Trump's tweets, irrespective of the content of the tweet. Next we analyze the impact of tweets containing certain words. Last, we analyze the impact of tweets belonging to topics determined by the LDA approach.

4.1. The general effect of Trump tweeting

The first method put into use to capture any effect from Mr. Trump's tweets is a linear regression model. The set of control samples is obtained by extracting all non-overlapping *d*-minute returns during Trump's presidency. To avoid control sample returns overlapping with

tweet sample returns, all control sample data points with timestamps were removed $t \in [\hat{t}-d, \hat{t}+d]$, where $\hat{t}$ denotes timestamp of the tweet. The regression is specified as follows:

$$Y_{t,t+d} = const + \beta Tweet_t + \epsilon_t \tag{3}$$

The $Y_{t,\ t+d}$ represents various measures of financial markets, e.g. *Return$_{t,\ t+d}$* or *Volume$_{t+d}$* of the US stock market. To estimate volatility, *Return$_{t,\ t+d}$* on the VIX, *VIXreturn$_{t,\ t+d}$* are used. The dummy variable *Tweet$_t$* takes the value of 1 for all periods *t* where a tweet has been posted, and 0 otherwise. The coefficient *const* can be interpreted as the averaged d-minute return in periods where Mr. Trump has not tweeted. The average return after Trump has tweeted is consequently *const* + β. The analysis is performed for a selection of values for *d*.

Table 2 shows that the S&P 500, on average, has a positive return during Trump's presidency. The dummy variable *Tweet* has a negative coefficient indicating that the S&P 500, in general, is negatively affected by Trump's tweets. In such a pooled regression, however, the effect is not very significant. The effect also seems to vanish within an hour.

In Table 3, the dummy variable *Tweet* has a significant, positive coefficient, indicating that the volume of S&P 500 increases in general after Trump has tweeted. Table 4 shows that VIX Futures, on average,

Table 5

Regression results for the impact of Trump's tweets containing the word "Great" on the S&P 500 return.

	Dependent variable: $Return_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0001***	0.0009***	0.0019***	0.0031***	0.0042***	0.0054***	0.0096***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0000	−0.0005	−0.0019**	−0.0024**	−0.0018	−0.0011	−0.0016
	(0.983)	(0.445)	(0.046)	(0.045)	(0.245)	(0.543)	(0.578)
"*Great*"	−0.0000	0.0001	0.0011	0.0008	0.0032	0.0005	0.0003
	(0.932)	(0.915)	(0.536)	(0.722)	(0.258)	(0.886)	(0.943)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 6

Regression results for the impact of Trump's tweets containing the word "Great" on the S&P 500 abnormal trading volume.

	Dependent variable: $Volume_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0076***	0.0057	0.0047	0.0053	0.0063	0.0022	0.0032
	(0.000)	(0.127)	(0.378)	(0.413)	(0.435)	(0.804)	(0.793)
Tweet	0.0471***	0.0264	0.0273	0.0506***	0.0377*	0.0548***	0.0512**
	(0.006)	(0.141)	(0.130)	(0.006)	(0.052)	(0.004)	(0.015)
"*Great*"	0.0570	0.0659*	0.0609*	0.0211	0.1078***	0.0293	0.0256
	(0.106)	(0.065)	(0.076)	(0.542)	(0.002)	(0.412)	(0.465)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 7

Regression results for the impact of Trump's tweets containing the word "Great" on the VIX futures.

	Dependent variable: $VIXreturn_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	−0.0002	−0.0068***	−0.0212***	−0.0302***	−0.0460***	−0.0466***	−0.0890***
	(0.472)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0026	0.0030	0.0192	0.0281*	0.0168	0.0118	−0.0134
	(0.499)	(0.740)	(0.116)	(0.056)	(0.363)	(0.584)	(0.690)
"*Great*"	0.0093	0.0297*	0.0237	0.0381	0.0615*	0.0724*	0.1827***
	(0.226)	(0.091)	(0.313)	(0.171)	(0.067)	(0.058)	(0.001)
Observations	510,460	51,477	26,938	19,663	13,219	11,377	7171

Note: *p<0.1; **p<0.05; ***p<0.01.

have a negative return. This is in accordance with existing literature [3,9]. The dummy variable Tweet, however, has a small positive coefficient indicating that investor fear is higher after tweets by President Trump are published. The tendency is most significant for 30 min.

There seems to be a time interval where the reaction of the market to Mr. Trump's tweets is largest, and which is different for *Return*, *Volume* and the volatility (*VIX* futures). Tables 2–4 indicate that the effect on S&P 500 return lasts at most 45 min, while the effect on volume lasts at least an hour and possibly even longer. The duration of the impact on the VIX futures is similar to the duration of the effect on the return, with the effect becoming insignificant within the first hour. It is important to emphasize that predictability of the S&P 500 index or VIX futures translates directly into possible trading profit, whereas no such opportunity exists for trading volume. The results, therefore, mean that even though the impact of Mr. Trump's tweets persists for at least two hours in terms of trading volume, profitable trading opportunities seem to exist only approximately within 30 min after his tweets. On the other hand, even though trading volume is increased right after the tweet, the S&P 500 index and VIX futures do not move significantly within first 10 min.

4.2. A Naïve approach based on single words

Next we investigate whether a tweet including a specific word differs significantly from other tweets with regard to its impact on financial markets. We therefore add a binary variable $Target_t$ to the regression 3, which takes the value 1 if President Trump posted a tweet containing the target word during that minute, and 0 otherwise. To control for the effect of a general tweet identified in the previous subsection, we continue including the dummy variable $Tweet_t$, which has the same interpretation as in Regression 3. Regressing the *d*-minute effect after the President publishes tweets containing the chosen words against other *d*-minute intervals where he does not tweet, we obtain the average effect on the financial markets of those tweets. We perform Regression 4 for selected values of $d \in [1,120]$.

$$Y_{t,t+d} = const + \beta_1 Tweet_t + \beta_2 Target_t + \epsilon_t \qquad (4)$$

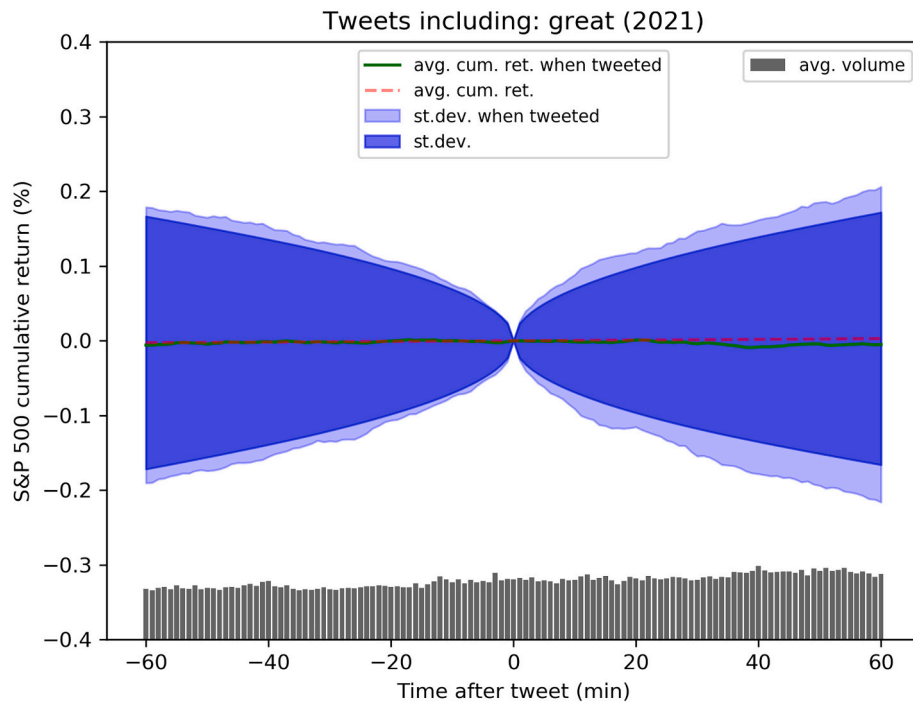In addition, to better illustrate how the same selection of tweets

Fig. 3. S&P 500 Cumulative Return and Volume for tweets including "Great".

Table 8
Regression results for the impact of Trump's tweets containing the word "Products" on the S&P 500 return.

	Dependent variable: $Return_{t, t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0001***	0.0009***	0.0019***	0.0031***	0.0042***	0.0054***	0.0096***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	0.0000	−0.0003	−0.0014*	−0.0019*	−0.0007	−0.0008	−0.0016
	(0.986)	(0.564)	(0.092)	(0.066)	(0.615)	(0.654)	(0.560)
"*Products*"	−0.0080***	−0.0461***	−0.0592***	−0.0827***	−0.0989***	−0.0961***	0.0045
	(0.008)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.887)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 9
Regression results for the impact of Trump's tweets containing the word "Products" on the S&P 500 abnormal trading volume.

	Dependent variable: $Volume_{t, t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0076***	0.0057	0.0047	0.0053	0.0063	0.0022	0.0032
	(0.000)	(0.127)	(0.378)	(0.413)	(0.435)	(0.804)	(0.793)
Tweet	0.0570***	0.0396**	0.0410***	0.0537***	0.0624***	0.0606***	0.0578***
	(0.000)	(0.012)	(0.010)	(0.001)	(0.000)	(0.001)	(0.003)
"*Products*"	0.9698***	0.6947***	0.3110	0.4801**	0.4630***	0.3185**	−0.0643
	(0.000)	(0.000)	(0.192)	(0.018)	(0.005)	(0.041)	(0.691)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 10

Regression results for the impact of Trump's tweets containing the word "Products" on the VIX futures.

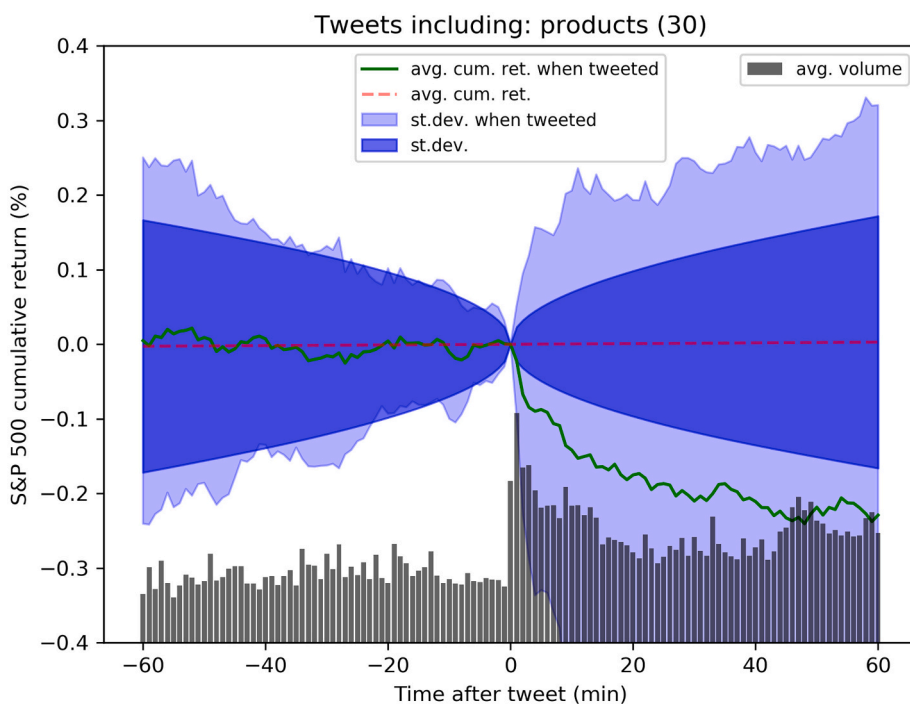|	Dependent variable: $VIXreturn_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	−0.0002	−0.0068***	−0.0212***	−0.0302***	−0.0460***	−0.0466***	−0.0890***
	(0.472)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0005	0.0085	0.0228**	0.0349***	0.0295*	0.0275	0.0293
	(0.893)	(0.278)	(0.034)	(0.008)	(0.075)	(0.155)	(0.342)
"Products"	0.0553	0.3843***	0.5965***	0.4553***	0.4367**	0.4708**	0.0077
	(0.191)	(0.000)	(0.000)	(0.002)	(0.013)	(0.025)	(0.979)
Observations	510,460	51,477	26,938	19,663	13,219	11,377	7171

Note: *p<0.1; **p<0.05; ***p<0.01.

Fig. 4. S&P 500 Cumulative Return and Volume around tweets including "products".

Table 11

Regression results for the impact of Trump's tweets containing the word "Tariff" on the S&P 500 return.

	Dependent variable: $Return_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0001***	0.0009***	0.0019***	0.0031***	0.0042***	0.0054***	0.0096***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	0.0001	−0.0003	−0.0013	−0.0015	−0.0004	−0.0005	−0.0015
	(0.710)	(0.627)	(0.119)	(0.157)	(0.793)	(0.788)	(0.580)
"Tariff"	−0.0064***	−0.0134***	−0.0198***	−0.0461***	−0.0445***	−0.0371***	−0.0035
	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.820)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 12

Regression results for the impact of Trump's tweets containing the word "Tariff" on the S&P 500 abnormal trading volume.

	Dependent variable: $Volume_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0076***	0.0057	0.0047	0.0053	0.0063	0.0022	0.0032
	(0.000)	(0.127)	(0.378)	(0.413)	(0.435)	(0.804)	(0.793)
Tweet	0.0473***	0.0336**	0.0324**	0.0474***	0.0575***	0.0569***	0.0547***
	(0.002)	(0.034)	(0.042)	(0.004)	(0.001)	(0.001)	(0.005)
"*Tariff*"	0.7588***	0.4954***	0.5334***	0.4567***	0.3708***	0.2706***	0.1461
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.134)
Observations	992,560	102,694	53,371	37,998	26,182	21,300	13,342

Note: *p<0.1; **p<0.05; ***p<0.01.

Table 13

Regression results for the impact of Trump's tweets containing the word "Tariff" on the VIX futures.

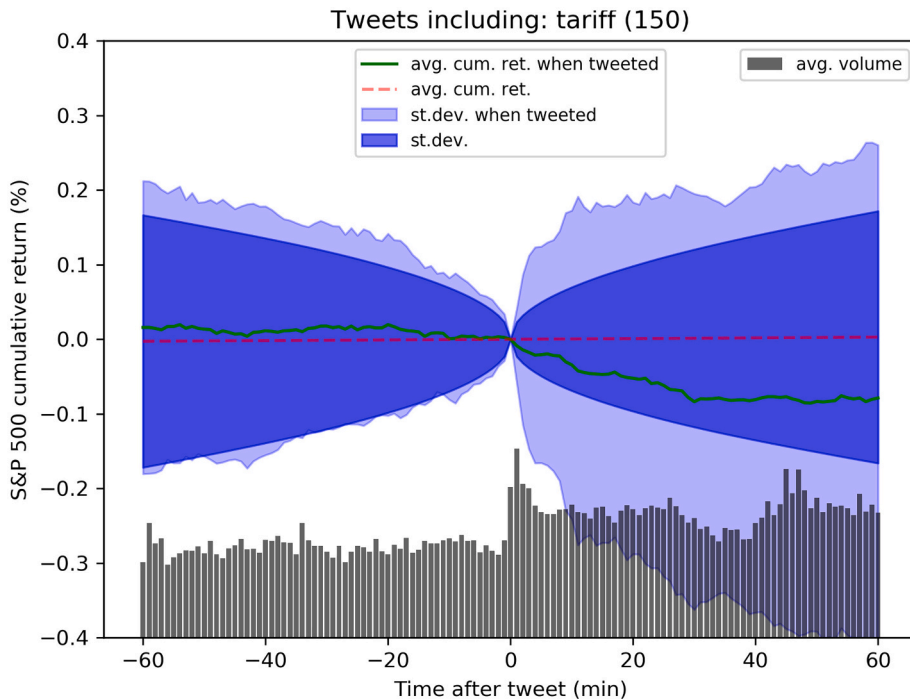|	Dependent variable: $VIXreturn_{t,\ t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	−0.0002	−0.0068***	−0.0212***	−0.0302***	−0.0460***	−0.0466***	−0.0890***
	(0.472)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Tweet	−0.0008	0.0077	0.0212*	0.0319**	0.0264	0.0254	0.0265
	(0.810)	(0.331)	(0.051)	(0.015)	(0.113)	(0.192)	(0.391)
"*Tariff*"	0.0329	0.1421***	0.2316***	0.3519***	0.3066***	0.2331**	0.1758
	(0.157)	(0.008)	(0.001)	(0.000)	(0.002)	(0.043)	(0.274)
Observations	510,460	51,477	26,938	19,663	13,219	11,377	7171

Note: p<0.1; **p<0.05; ***p<0.01.

Fig. 5. S&P 500 Cumulative Return and Volume for tweets including "tariff".

Table 14
Regression results for the impact of tweets with various topics on the S&P 500 return.

	Dependent variable: $Return_{t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0001***	0.0010***	0.0020***	0.0032***	0.0044***	0.0063***	0.0120***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
WorldLeaderMeetings	0.0021	0.0013	−0.0142	−0.0035	−0.0109	−0.0254	−0.0222
	(0.476)	(0.868)	(0.224)	(0.807)	(0.540)	(0.233)	(0.487)
TVShows	−0.0001	−0.0101**	−0.0065	−0.0045	−0.0041	−0.0038	−0.0116
	(0.958)	(0.032)	(0.339)	(0.588)	(0.687)	(0.757)	(0.529)
Corruption	−0.0005	−0.0008	−0.0027	0.0113	0.0114	−0.0005	−0.0409*
	(0.801)	(0.884)	(0.724)	(0.233)	(0.334)	(0.971)	(0.051)
TradeWar	−0.0054**	−0.0153***	−0.0263***	−0.0422***	−0.0488***	−0.0533***	−0.0277
	(0.011)	(0.008)	(0.001)	(0.000)	(0.000)	(0.000)	(0.216)
Bragging	−0.0008	0.0060	−0.0071	0.0006	0.0052	0.0097	−0.0151
	(0.691)	(0.284)	(0.376)	(0.950)	(0.670)	(0.502)	(0.483)
MakingDeals	0.0011	0.0082	0.0119	0.0009	0.0374***	0.0155	0.0117
	(0.583)	(0.132)	(0.123)	(0.921)	(0.001)	(0.271)	(0.581)
NorthKorea/States	0.0030	−0.0015	0.0004	−0.0096	−0.0160	0.0181	0.0599
	(0.450)	(0.890)	(0.981)	(0.617)	(0.497)	(0.521)	(0.158)
USMilitary	−0.0014	−0.0017	0.0008	−0.0024	−0.0037	0.0001	0.0111
	(0.335)	(0.674)	(0.886)	(0.737)	(0.678)	(0.995)	(0.494)
Achievements	−0.0016	−0.0002	0.0080	0.0086	0.0034	−0.0101	−0.0105
	(0.333)	(0.969)	(0.212)	(0.270)	(0.724)	(0.385)	(0.549)
Compassion	0.0022	0.0110	0.0099	0.0197	0.0216	0.0107	0.0055
	(0.425)	(0.151)	(0.369)	(0.146)	(0.199)	(0.593)	(0.856)
BorderSecurity	0.0007	−0.0021	−0.0158***	−0.0200***	−0.0137*	−0.0056	0.0317**
	(0.591)	(0.538)	(0.001)	(0.001)	(0.073)	(0.538)	(0.020)
PoliticalDiscussions	0.0028	0.0052	0.0048	0.0020	−0.0014	−0.0009	−0.0252
	(0.131)	(0.298)	(0.502)	(0.816)	(0.894)	(0.944)	(0.198)
Undefined	−0.0002	0.0001	0.0026	0.0062	0.0060	0.0087	0.0009
	(0.787)	(0.970)	(0.437)	(0.139)	(0.249)	(0.167)	(0.924)
BorderWall/Board	0.0039	−0.0009	−0.0041	−0.0002	−0.0171	−0.0107	0.0191
	(0.209)	(0.920)	(0.748)	(0.989)	(0.369)	(0.632)	(0.574)
Announcements	−0.0016	−0.0169**	−0.0079	−0.0198	−0.0292*	−0.0417**	−0.0317
	(0.530)	(0.017)	(0.446)	(0.112)	(0.059)	(0.024)	(0.249)
FakeNews	0.0016	−0.0023	0.0002	−0.0012	0.0007	0.0045	0.0156
	(0.274)	(0.561)	(0.978)	(0.863)	(0.937)	(0.664)	(0.317)
RussianCollusion	−0.0025*	−0.0005	−0.0040	−0.0076	−0.0131	−0.0092	−0.0029
	(0.100)	(0.910)	(0.494)	(0.290)	(0.143)	(0.385)	(0.856)
Employment	0.0008	0.0053	0.0057	0.0179*	0.0180	0.0295*	−0.0041
	(0.723)	(0.377)	(0.503)	(0.084)	(0.164)	(0.057)	(0.861)
IllegalImmigrants	0.0025	0.0056	−0.0043	−0.0197*	0.0098	0.0083	−0.0293
	(0.261)	(0.355)	(0.621)	(0.061)	(0.453)	(0.594)	(0.207)
Presidents	−0.0011	−0.0037	0.0061	0.0012	−0.0067	−0.0123	−0.0352
	(0.682)	(0.600)	(0.547)	(0.924)	(0.666)	(0.505)	(0.201)
Observations	974,450	93,473	45,827	31,057	20,367	15,847	9020

Note: *p<0.1; **p<0.05; ***p<0.01.

impact S&P 500, we have plotted the average cumulative return and the average volume for every t minute, $t \in -6060$. We let $t = 0$ represent the time when President Trump publishes tweets that contain the word being investigated. We also include one standard error band for both the control- and examined samples.

To select the words to analyze, the most used words from the tweet data set were inspected, see Fig. 2. The word "Great", which appears in more than 2000 tweets from the sample, was chosen for analysis. Since Trump has been engaged in trade war(s), particularly with China, we also investigate words that are related to this subject: "products" and "tariff". These words are often included in tweets in which he comments on the trade war and writes about imposing *tariffs* on particular *products*. Since increased tariffs and intensification of the trade war have a negative impact on the economy, we expect these words to have a negative impact on the broad financial market. Next we present the results for each of the three selected words.

4.2.1. The word "great"

Table 5 shows that the impact of tweets which include the word *great*

on the S&P 500 is insignificant for all time intervals d. The impact on the abnormal volume (Table 6) is small and positive but only significant for some values d. Table 7 shows that the VIX return coefficient is small and positive, but as was the case for volume, only significant for some values d. Fig. 3 confirms that tweets including the word *great* on average have negligible impacts on the index return. Average cumulative return drops, on average, approximately one basis point during the 60 min following a tweet publication. This is probably due to the general effect of Trump tweeting, as 2,021 tweets include the word *great*. We also observe that neither average volume nor the standard deviation seem to change notably, despite some significant coefficients in Tables 6 and 7.

4.2.2. The word "products"

Table 8 shows that the impact of tweets including the word *products* on the S&P 500 is significantly negative. The coefficient size increases throughout the first hour, but most of the effect appears to vanish after two hours. Table 9 shows that the impact on abnormal trading volume is large and positive for the first ten minutes. The effect remains positive during the first hour, but it diminishes with time. Table 10 shows that

Table 15

Regression results for the impact of tweets with various topics on the S&P 500 abnormal trading volume.

	Dependent variable: Volume$_{t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0060***	−0.0013	−0.0079	−0.0114	−0.0144	−0.0229**	−0.0335*
	(0.000)	(0.751)	(0.174)	(0.121)	(0.141)	(0.042)	(0.069)
WorldLeaderMeetings	−0.3066	−0.2724	0.2600	0.0899	−0.0777	0.1601	−0.1361
	(0.185)	(0.246)	(0.203)	(0.684)	(0.709)	(0.458)	(0.555)
TVShows	0.1976	0.1706	0.0719	0.2269*	0.2046	0.1768	0.2298*
	(0.132)	(0.236)	(0.601)	(0.100)	(0.129)	(0.190)	(0.078)
Corruption	0.1789	0.2912**	0.3338**	0.0978	0.1884	0.3197**	0.0809
	(0.219)	(0.041)	(0.014)	(0.515)	(0.164)	(0.020)	(0.594)
TradeWar	0.7746***	0.4097**	0.0625	0.2315	0.2126	0.0937	0.2485
	(0.000)	(0.011)	(0.709)	(0.139)	(0.175)	(0.570)	(0.106)
Bragging	−0.2791*	−0.4251***	−0.4469***	−0.4758***	−0.4531***	−0.3362**	−0.0542
	(0.063)	(0.004)	(0.005)	(0.001)	(0.005)	(0.030)	(0.722)
MakingDeals	0.1979	0.3088**	0.0710	0.1188	0.1340	0.3593***	0.1948
	(0.178)	(0.025)	(0.588)	(0.398)	(0.320)	(0.004)	(0.156)
NorthKorea/States	0.2757	−0.4696	0.1309	0.1174	−0.1752	0.4553	−0.0017
	(0.341)	(0.146)	(0.690)	(0.682)	(0.584)	(0.129)	(0.995)
USMilitary	−0.0523	−0.0448	−0.0890	0.0277	−0.0450	−0.0276	−0.0019
	(0.641)	(0.716)	(0.446)	(0.817)	(0.716)	(0.830)	(0.988)
Achievements	−0.0725	−0.1412	0.0295	−0.1416	0.1031	0.1400	−0.0346
	(0.555)	(0.280)	(0.811)	(0.282)	(0.438)	(0.284)	(0.793)
Compassion	−0.1801	−0.2281	−0.0162	0.1789	0.1954	−0.3087	−0.0486
	(0.404)	(0.295)	(0.943)	(0.418)	(0.369)	(0.178)	(0.840)
BorderSecurity	0.2852***	0.2596***	0.4464***	0.3578***	0.1969**	0.0676	0.2008**
	(0.002)	(0.005)	(0.000)	(0.000)	(0.037)	(0.473)	(0.035)
PoliticalDiscussions	0.1076	0.0731	−0.1002	0.0765	0.2171	0.0188	0.1587
	(0.459)	(0.599)	(0.480)	(0.571)	(0.126)	(0.887)	(0.243)
Undefined	0.0440	0.1443**	0.0626	0.1210*	0.1681**	0.1811***	0.0861
	(0.508)	(0.029)	(0.366)	(0.080)	(0.018)	(0.009)	(0.276)
BorderWall/Board	−0.0205	−0.1071	−0.0926	−0.2684	0.0779	−0.0851	0.0128
	(0.933)	(0.667)	(0.717)	(0.331)	(0.743)	(0.732)	(0.956)
Announcements	0.0347	0.0966	−0.0203	0.2959	0.0953	0.0799	0.0694
	(0.867)	(0.669)	(0.926)	(0.221)	(0.679)	(0.729)	(0.735)
FakeNews	−0.0050	−0.1356	−0.1154	0.0826	0.0766	0.1546	0.3233***
	(0.964)	(0.220)	(0.321)	(0.436)	(0.520)	(0.148)	(0.002)
RussianCollusion	−0.1293	0.0003	0.0734	−0.0573	−0.1447	−0.0953	0.0942
	(0.264)	(0.997)	(0.500)	(0.592)	(0.217)	(0.371)	(0.355)
Employment	0.4746***	0.2685	0.3035*	0.1051	0.2658	0.2160	0.1805
	(0.005)	(0.106)	(0.068)	(0.481)	(0.136)	(0.213)	(0.291)
IllegalImmigrants	−0.0911	0.1756	0.0934	0.1509	0.4448***	0.3306**	0.0123
	(0.597)	(0.268)	(0.541)	(0.349)	(0.002)	(0.035)	(0.935)
Presidents	−0.1917	−0.0399	−0.0533	0.0189	−0.1780	−0.1257	0.0497
	(0.356)	(0.849)	(0.768)	(0.922)	(0.378)	(0.520)	(0.788)
Observations	974,450	93,473	45,827	31,057	20,367	15,847	9020

Note: *p<0.1; **p<0.05; ***p<0.01.

the impact on the VIX futures is large and positive during the first hour but is insignificant in the first minute after a tweet as well as after 120 min. This indicates that there is some initial lag before investors react, and that the effect vanishes within two hours.

Fig. 4 confirms that tweets including the word *products* on average yield negative returns. Average cumulative return drops, on average, approximately 30 basis points during the 60 min following a tweet. There is also a significant spike in average volume immediately after the tweets, indicating that investors might disagree about how the information contained in the tweet should be interpreted. Additionally, there is a considerable increase in volatility.

4.2.3. The word "tariff"

In Table 11, we see tweets including the word *tariff* have a significantly negative impact on the S&P 500. As is the case for tweets including *products*, the coefficient size for *Tariff* increases throughout the first hour, but practically all of the effect seems to vanish after the first

120 min. Table 12 shows that the abnormal volume coefficient *Products* is large and positive for all time intervals, but similarly to the return, the effect seems to diminish within 120 min. Table 13 contains the coefficients for VIX returns, and shows very similar patterns for the tweets that include the word *Products*.

Fig. 5 shows that the average cumulative return becomes negative at $t = 0$ and drops approximately ten basis points on average during the first 60 min after a tweet containing "tariff" is published. For the case of average volume, it increases considerably during the first hour after the tweets. Another interesting observation is the notable increase in the standard deviation for returns after the tweets.

4.3. A topic based approach

We continue our analysis by employing the LDA approach in which we first identify a set of distinct topics, and then run the regression on the degree of affiliation each tweet has with every topic. In this model

Table 16

Regression results for the impact of tweets with various topics on the VIX futures.

	Dependent variable: VIX_{t+d}						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	−0.0001	−0.0071***	−0.0229***	−0.0335***	−0.0501***	−0.0556***	−0.1050***
	(0.638)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
WorldLeaderMeetings	−0.1084**	−0.1023	0.2007	−0.0136	0.1952	−0.1029	0.2300
	(0.034)	(0.371)	(0.194)	(0.941)	(0.377)	(0.683)	(0.537)
TVShows	0.0165	−0.0552	−0.0010	0.0153	0.0737	−0.0424	−0.0680
	(0.587)	(0.424)	(0.991)	(0.888)	(0.568)	(0.770)	(0.750)
Corruption	−0.0416	0.0120	−0.0049	−0.0962	−0.0805	−0.0291	0.2625
	(0.209)	(0.875)	(0.961)	(0.420)	(0.582)	(0.855)	(0.255)
TradeWar	0.1050***	0.2105***	0.2678**	0.3106**	0.2826*	0.3700**	0.2735
	(0.003)	(0.008)	(0.010)	(0.013)	(0.059)	(0.032)	(0.276)
Bragging	0.0442	0.0503	0.0671	0.1548	0.1319	0.0080	0.1875
	(0.204)	(0.525)	(0.527)	(0.220)	(0.383)	(0.962)	(0.456)
MakingDeals	−0.0711**	−0.0334	−0.0093	0.0611	−0.1616	−0.1781	−0.0719
	(0.032)	(0.662)	(0.926)	(0.598)	(0.256)	(0.274)	(0.761)
NorthKorea/States	0.0231	0.0915	0.1209	0.2699	0.2134	−0.2803	−0.6485
	(0.739)	(0.551)	(0.569)	(0.282)	(0.474)	(0.411)	(0.199)
USMilitary	0.0169	0.0528	0.0416	0.0985	0.1763	0.2245*	0.3382*
	(0.514)	(0.375)	(0.599)	(0.288)	(0.127)	(0.081)	(0.084)
Achievements	−0.0120	−0.0138	−0.0217	−0.0546	0.1231	0.4016***	0.4458**
	(0.670)	(0.832)	(0.801)	(0.592)	(0.322)	(0.004)	(0.030)
Compassion	−0.0042	−0.0733	−0.1083	−0.1818	−0.0202	0.2287	0.2862
	(0.929)	(0.484)	(0.438)	(0.270)	(0.920)	(0.317)	(0.421)
BorderSecurity	0.0109	0.0132	0.1579**	0.1374*	0.0944	0.0583	−0.1616
	(0.619)	(0.785)	(0.013)	(0.072)	(0.311)	(0.582)	(0.292)
PoliticalDiscussions	−0.0589*	−0.0046	−0.0616	0.0272	−0.0940	0.0218	0.0251
	(0.060)	(0.948)	(0.506)	(0.805)	(0.477)	(0.884)	(0.908)
Undefined	0.0186	−0.0335	−0.0593	−0.0104	−0.0246	−0.1048	−0.0488
	(0.183)	(0.292)	(0.166)	(0.840)	(0.697)	(0.152)	(0.672)
BorderWall/Board	0.0697	0.2047*	0.2325	0.2123	0.3309	0.3492	−0.1566
	(0.186)	(0.098)	(0.156)	(0.268)	(0.184)	(0.197)	(0.699)
Announcements	−0.0075	0.0275	0.0389	0.0266	0.2279	0.0634	−0.3152
	(0.880)	(0.804)	(0.794)	(0.882)	(0.286)	(0.801)	(0.382)
FakeNews	0.0176	0.0035	−0.0355	0.0229	−0.1255	−0.0239	0.0497
	(0.511)	(0.952)	(0.647)	(0.801)	(0.259)	(0.846)	(0.780)
RussianCollusion	0.0120	0.0125	−0.0785	0.0204	−0.0308	−0.0916	−0.3177*
	(0.651)	(0.834)	(0.319)	(0.824)	(0.781)	(0.468)	(0.084)
Employment	0.0219	0.1013	0.1451	−0.0157	−0.0026	0.0098	0.1808
	(0.553)	(0.231)	(0.189)	(0.903)	(0.987)	(0.956)	(0.498)
IllegalImmigrants	−0.0704*	−0.0551	0.0614	0.1308	−0.1492	−0.0198	0.1834
	(0.072)	(0.529)	(0.588)	(0.330)	(0.378)	(0.915)	(0.495)
Presidents	−0.0552	0.0451	0.0199	−0.0212	0.1760	0.1812	0.1377
	(0.205)	(0.640)	(0.878)	(0.891)	(0.339)	(0.401)	(0.656)
Observations	476,707	43,920	21,481	14,790	9425	7737	4517

Note: *p<0.1; **p<0.05; ***p<0.01.

we modify regression 4 by replacing the variable $Target_t$ with a set of variables $Topic_{i,\,t}$ which take the value of the weight in $Topic_i$ for a tweet posted during minute t. As $\sum_i Topic_i = 1$, the variable $Tweet_t$ is removed to avoid issues concerning multicollinearity.

$$Y_{t,t+d} = \alpha + \{\beta_i\}_{i=1}^{K}\{Topic_{i,t}\}_{i=1}^{K} + \epsilon_t, \quad K \in [1,2,\dots,20] \qquad (5)$$

We continue the practice that $Y_{t,\,t+d}$ can be either $Return_{t,\,t+d}$, $Volume_{t+d}$, or $VIXreturn_{t,\,t+d}$ or, and we perform regression (5) for a selection of values of $d \in [1;120]$.

Table 14 shows the coefficients and the corresponding *p*-values for the different topics at a selection of time intervals within $d \in [1;120]$ on the S&P 500 index. Tables 15 and 16 show the regression results for abnormal (log) volume and volatility (as measured by VIX return), respectively. Additionally, we illustrate a selection of the most interesting findings in Figs. 6, 7 and 8. In these figures, we plot the coefficients on the topic "Trade war" for various time intervals (*d*) together with one standard error band.

The impact of tweets with words "Trade war" on the S&P 500 index is, as we expected, negative and large, indicating that investors consider tweets from this topic to be bad news for the stock market. Interestingly, the topic "Making deals" has positive coefficients but is only significant at the $d = 45$ interval. This may indicate that investors interpret these tweets to be good news for the financial markets acting in harmony with the negative and significant coefficients on the "Trade war" topic. Additionally, the "Border security" topic is also negative and significant. We think a likely explanation is that these tweets also tend to include statements related to cross-border trade and that some of the effects from this topic are shared with the "Trade war" topic. We include two examples of this phenomenon in Table 17.

The coefficients on abnormal (log) volume are significant for several topics. As one might expect, the topics "Trade war" and "Border security" have large, positive coefficients across most time intervals. Interestingly, we also observe negative and significant coefficients on the topic "Bragging", but we are unable to find a reasonable explanation for this phenomenon.
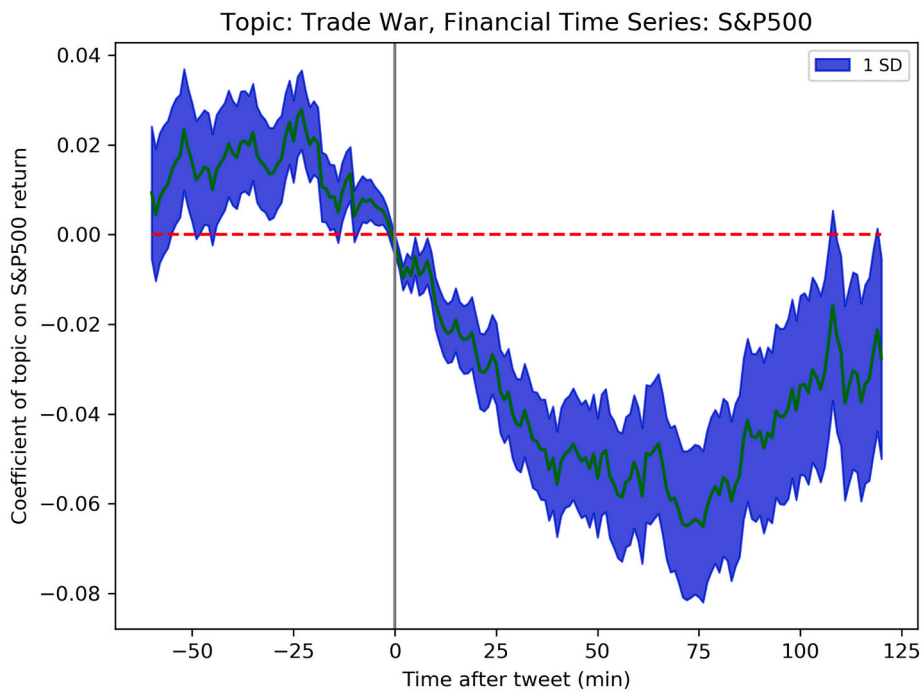
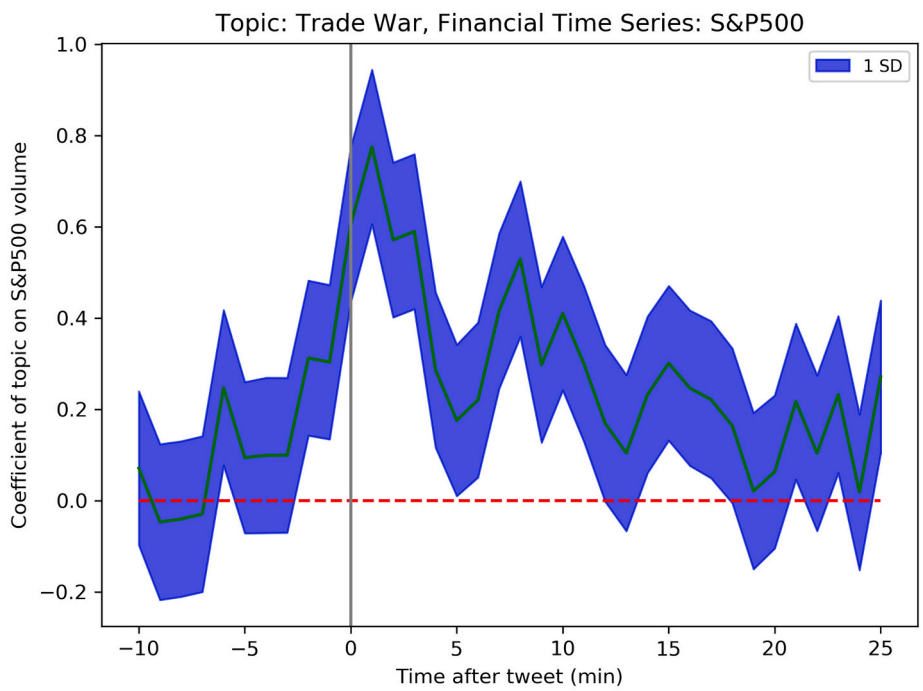**Fig. 6.** Impact of the topic "Trade War" in Trump's tweets on the S&P 500 return.

Fig. 7. Impact of the topic "Trade War" in Trump's tweets on the S&P 500 abnormal volume.
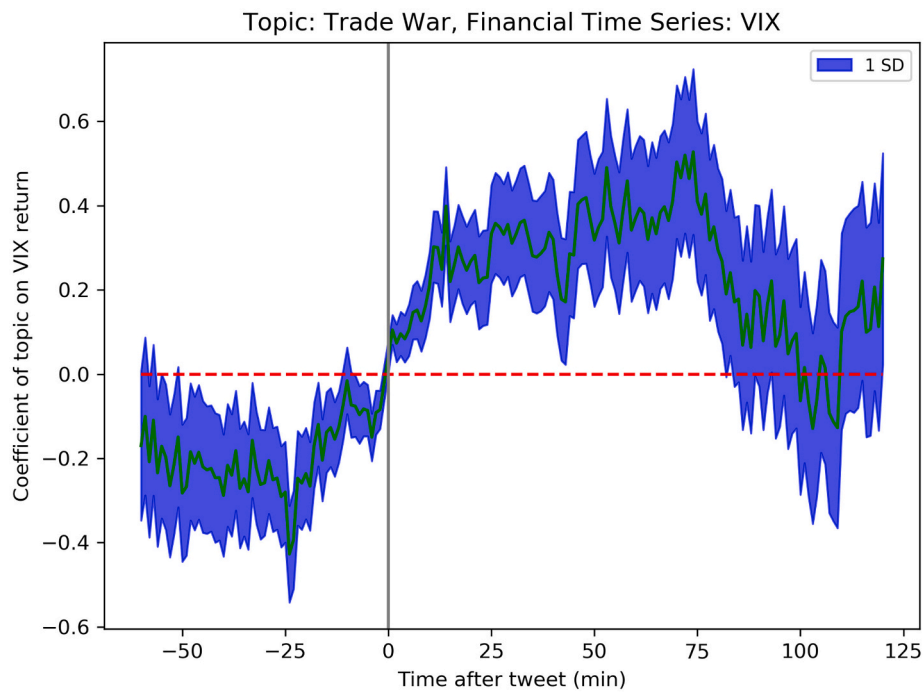
Fig. 8. Impact of the topic "Trade War" in Trump's tweets on the VIX futures.

Table 17
Example of tweets containing both the topics "Border security" and "Trade war".

Date	Tweet	Main topic
Apr 7, 2019	"....Mexico must apprehend all illegals and not let them make the long march up to the United States, or we will have no other choice than to Close the Border and/or institute Tariffs. Our Country is FULL!"	Border security
Jul 11, 2019	Mexico is doing great at the Border, but China is letting us down in that they have not been buying the agricultural products from our great Farmers that they said they would. Hopefully they will start soon!	-26emBorder security

Table 18
Example of Trump twittering about same topic consecutively on Sep 4, 2019.

Time	Tweet	Main topic
13:44:02	....We can't have a system where we run our entire economy for the benefit of other countries, which have long charged us big tariffs. Don't keep ducking the reality. The U.S. has been subject to Tariff Terrorism for 50 years. (...)	Trade war
13:52:58	"U.S. Winning Trade War With China In Dollars." CNBC	Trade war

When it comes to uncertainty, "Trade war" once again stands out as the topic with the clearest results. We also note that the topic "Making deals" has a significant negative short-term effect, along with "Political discussions" and "World leader meetings". The fact that the topic named "Trade war" is distinctly more significant than the others might not come as a surprise, as we observe that both of the words "tariff" and "products" from the naïve approach also appear in this topic (see Fig. 1). The same results are also recognized in most of the other financial data

sets, and as a result, we choose to focus our discussion on this topic.

As Figs. 6–8 show, we generally observe that a (positive or negative) effect seems to start slightly before Trump has tweeted, as if there has been some kind of information leakage. This could be an explanation, but we find it unlikely. A more likely explanation is related to the fact that Trump often tweets about the same topic consecutively. An example is illustrated in Table 18. We only merge tweets that take place within the same minute. However, multiple tweets occurring within a matter of minutes will have a similar influence, as long as the tweet's topics have enough in common.

Fig. 6 illustrates the S&P 500 return coefficient is negative and highly significant for the first two hours following the tweets. We also observe that there seems to be a "sweet spot", around 75 min after a tweet is published where the results are the most significant. In Fig. 8, the VIX return coefficient follows a similar pattern but with the opposite sign. The hypothesis that "Trade War" tweets lead to more insecurity in financial markets is therefore supported by the results from the analysis of the VIX index. Fig. 7 shows that the volume increases significantly in the first couple of minutes directly following a "Trade War"-tweet publication. The increase in trading volume indicates that investors tend to disagree about what the tweets imply about the correct pricing of the assets. While some investors may put considerable emphasis on official statements from the President (net sellers), others may regard the tweets as noise. These investors may consider the link between the President's attitudes towards a trade war and real-life implementation to be overestimated by financial markets (net buyers).

Given that the "Trade war" topic mostly refers to the (ongoing) disagreements between the US and China, we anticipate that the Hang Seng index, which is closely related to the Chinese economy, is also negatively affected by tweets regarding the this topic. We have, therefore, conducted a similar analysis for the Hang Seng index, see Table 19 and Fig. 9. Although the effect is less significant than it is for the S&P 500, we still see the coefficient behave similarly, suggesting that investors also consider the tweets from Mr. Trump about the trade war to

Table 19

Regression results for the impact of tweets with various topics on the Hang Seng return.

	Dependent variable: Return$_{t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0000	0.0011***	0.0021***	0.0027**	0.0047***	0.0097***	0.0163***
	(0.230)	(0.001)	(0.002)	(0.013)	(0.007)	(0.000)	(0.001)
WorldLeaderMeetings	0.0053	0.0070	−0.0011	0.0172	−0.0125	−0.0220	0.0193
	(0.375)	(0.707)	(0.965)	(0.595)	(0.750)	(0.619)	(0.766)
TVShows	0.0013	0.0028	−0.0097	−0.0200	−0.0161	−0.0337	−0.0396
	(0.713)	(0.791)	(0.510)	(0.272)	(0.471)	(0.172)	(0.242)
Corruption	−0.0043	−0.0152	−0.0126	−0.0086	−0.0028	−0.0060	−0.0610
	(0.261)	(0.203)	(0.451)	(0.676)	(0.911)	(0.831)	(0.127)
TradeWar	−0.0064	−0.0316**	−0.0329*	−0.0375*	−0.0454*	−0.0587**	−0.0253
	(0.113)	(0.012)	(0.062)	(0.081)	(0.087)	(0.046)	(0.533)
Bragging	0.0059	0.0200*	0.0224	0.0310	0.0411	0.0333	0.0090
	(0.129)	(0.098)	(0.187)	(0.133)	(0.108)	(0.244)	(0.820)
MakingDeals	−0.0003	0.0008	0.0132	−0.0036	0.0197	−0.0006	−0.0132
	(0.926)	(0.942)	(0.418)	(0.855)	(0.416)	(0.983)	(0.732)
NorthKorea/States	0.0068	0.0125	0.0235	0.0127	0.0083	0.0106	−0.0217
	(0.381)	(0.609)	(0.488)	(0.762)	(0.874)	(0.854)	(0.782)
USMilitary	−0.0004	−0.0003	−0.0269**	−0.0039	0.0009	−0.0001	0.0391
	(0.886)	(0.978)	(0.042)	(0.812)	(0.966)	(0.997)	(0.227)
Achievements	−0.0028	−0.0013	−0.0001	0.0300*	0.0095	−0.0318	−0.0008
	(0.374)	(0.899)	(0.995)	(0.077)	(0.654)	(0.174)	(0.981)
Compassion	−0.0020	0.0192	0.0098	0.0020	−0.0207	−0.0247	−0.0410
	(0.756)	(0.338)	(0.730)	(0.955)	(0.635)	(0.609)	(0.578)
BorderSecurity	−0.0006	−0.0064	−0.0197*	−0.0186	−0.0263	−0.0271	0.0152
	(0.801)	(0.396)	(0.063)	(0.151)	(0.101)	(0.132)	(0.549)
PoliticalDiscussions	0.0006	0.0091	0.0039	0.0003	−0.0086	0.0068	−0.0138
	(0.864)	(0.401)	(0.799)	(0.987)	(0.706)	(0.788)	(0.700)
Undefined	−0.0008	−0.0039	0.0007	0.0142	0.0326**	0.0380**	−0.0017
	(0.674)	(0.541)	(0.936)	(0.204)	(0.023)	(0.022)	(0.948)
BorderWall/Board	0.0067	0.0126	0.0064	0.0302	0.0466	0.0749	0.0644
	(0.350)	(0.566)	(0.838)	(0.430)	(0.324)	(0.166)	(0.390)
Announcements	0.0064	−0.0153	0.0052	−0.0036	−0.0401	−0.0550	−0.0385
	(0.292)	(0.417)	(0.841)	(0.910)	(0.307)	(0.210)	(0.534)
FakeNews	0.0017	0.0022	0.0010	0.0018	0.0095	−0.0005	−0.0166
	(0.525)	(0.796)	(0.935)	(0.900)	(0.590)	(0.980)	(0.549)
RussianCollusion	−0.0003	0.0016	−0.0020	−0.0129	−0.0285	−0.0256	0.0147
	(0.927)	(0.853)	(0.870)	(0.377)	(0.113)	(0.202)	(0.602)
Employment	0.0016	0.0002	0.0171	0.0143	0.0315	0.0215	0.0114
	(0.721)	(0.988)	(0.374)	(0.544)	(0.282)	(0.510)	(0.801)
IllegalImmigrants	0.0001	−0.0056	−0.0307*	−0.0330	−0.0237	−0.0364	−0.0848**
	(0.987)	(0.664)	(0.086)	(0.136)	(0.387)	(0.237)	(0.043)
Presidents	−0.0015	0.0087	0.0006	0.0044	−0.0057	−0.0266	−0.0490
	(0.776)	(0.580)	(0.978)	(0.871)	(0.864)	(0.472)	(0.353)
Observations	559,787	53,147	25,315	15,914	10,752	7449	4622

Note: *p<0.1; **p<0.05; ***p<0.01.

negatively affect the Chinese stock market.

We also examine how Trump tweeting about a trade war affects the traditional safe-haven, gold [5]. The results are reported in Table 20, and a more detailed view about the impact of the "Trade war" topic in Trump's tweets is provided in Fig. 10. This, interestingly, shows that gold prices do, in general, increase after Trump tweets about the topic "Trade war", albeit with moderate significance. This change indicates that gold becomes a relatively more attractive investment when Trump's "Trade war" tweets impose insecurity on the financial markets.

4.3.1. Trading strategy

To test some of our findings, we evaluate a basic trading strategy on the S&P 500 index. We consider a large, institutional investor capable of continuously monitoring Trump's Twitter feed, and is additionally able to adjust his or her position within a negligible time delay after a tweet triggers a certain criterion.

We choose a trading strategy that involves being long in the S&P 500 index, but immediately entering an offsetting short position on the S&P 500 index the minute after a trigger tweet is published. In reality, this position can be taken almost instantaneously, but we choose to be somewhat conservative, as we only have observations for every minute in the data. The trigger criterion we use is that the tweet has a likelihood of more than 10% of belonging to the "Trade war" topic. We keep this position for 75 min, which is precisely the sweet spot we have identified in our analysis, and that can be observed in Fig. 6. Transaction costs are disregarded, as we assume that the large investor is already trading with high frequency and is able to implement the strategy by only making adjustments to his or her regular series of trades. We compare this portfolio to a basic buy-and-hold strategy on the S&P 500, as well as a pure shorting strategy that only enters short positions for a period of 75 min after each trigger tweet. To illustrate the strategy, Fig. 11 shows how the strategy would play out on a real example from August 1, 2019.

We include evaluation conducted on both the original data set used elsewhere in this paper (described in Section 2) and on a new and disjoint out-of-sample data set from the period between September 19, 2019 and December 5, 2019 (see Table 21). The analysis of the out-of-sample data reveals whether or not results are likely to be valid also outside the analyzed data.
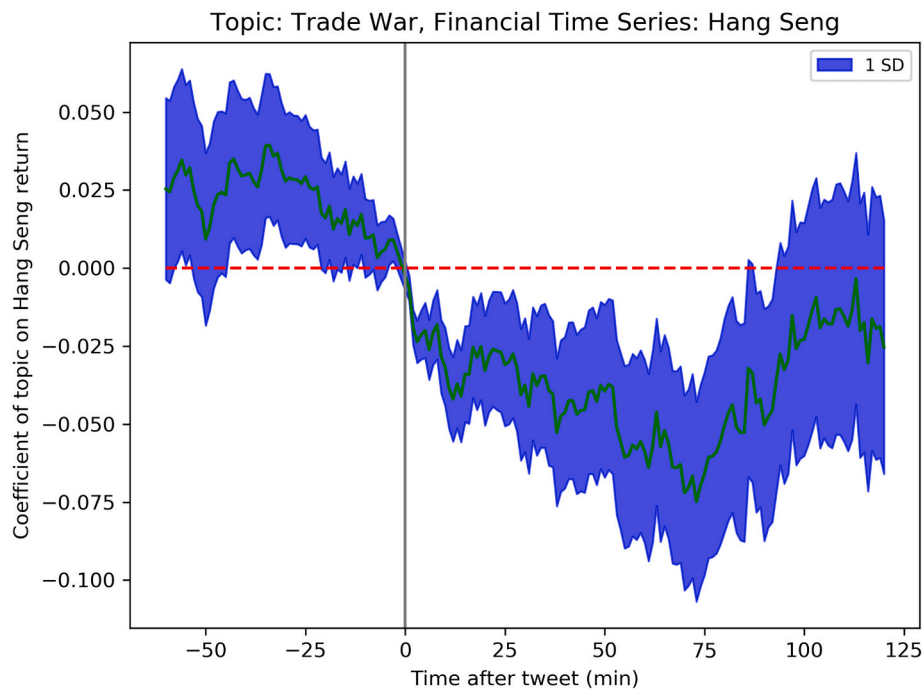
Fig. 9. Impact of the topic "Trade War" in Trump's tweets on the Hang Seng return.

The in-sample data contains 1,188 tweets that trigger the short criterion. Fig. 12 shows the cumulative return on the different trading strategies. Despite the fact that the S&P 500 has a 25% annualized return, we note that having short positions right after "Trade War" tweets seems to be a profitable strategy. The combined strategy of offsetting the long position after trigger tweets is even more promising. This strategy offers an annualized return of 34% compared to 25% for the buy-and-hold strategy, despite keeping the volatility slightly lower (Table 22). The Sharpe ratio[5] of our portfolio during the presidency period is 1.75, compared to 1.21 on the buy-and-hold strategy. Graphically, we observe in Fig. 12 that the shorting limits some of the losses in periods with large, negative returns, such as during the last quarter of 2018.

Fig. 13 shows the cumulative portfolio returns for the out-of-sample period. This period contains 180 tweets that trigger the short criterion. We note that the pattern is the same for the training sample. The hedged portfolio achieves an annualized return of 21%, compared to only 18% for the buy-and-hold strategy. Also, here, the hedging strategy has slightly lower volatility than a buy-and-hold strategy. The resulting Sharpe ratio is 1.71 for the hedging strategy, compared to 1.38 for the buy-and-hold strategy.

5. Conclusions

For the first time in history, the president of the United States, one of the most influential leaders in the world, frequently posts his opinions on Twitter about various topics, such as the ongoing negotiations regarding the trade war with China. We study how tweets from

President Trump affect financial markets, utilizing high-frequency (minute-by-minute) financial data, combined with a data set containing over 8000 of Mr. Trump's tweets.

We start our analysis by evaluation of the impact of Mr. Trump's tweets in general, irrespective of the actual content of the tweet. We find that the US stock market tends to decline the first 30 min following his publishing of a tweet and that trading volume and volatility are significantly higher after Trump has tweeted.

Next, we classify tweets depending on whether they contain a specific word. We find that tweets containing the words "products" and "tariff" have a negative impact on the stock market, i.e. these tweets are followed by decline in the market, increased trading volume and increased uncertainty (level of VIX futures).

Finally, we use Latent Dirichlet Allocation as a topic extraction technique, which assigns each tweet a degree of affiliation to a set of topics without any human input or interpretation. We analyze the short-term impact of each topic on financial markets. We find that tweets affiliated with the topics "Trade war" and "Border Security" are followed by negative returns on the S&P 500 index, increased volatility and increased trading volume. Furthermore, we find a significant decrease in the Hang Seng index, and a significant increase in the gold price after these tweets were published. This is consistent with previous results, as tweets which imply possible restrictions in international trade are often negative news for affected countries, whereas gold is safe haven and often appreciates in uncertain periods.

To illustrate the economic significance of the topic analysis, we implement a simple trading strategy of holding the S&P 500 index, except for 75-min periods following the President Trump's tweets related to the "Trade war" topic. In other words, the investor holds the S&P 500 index and then sells it whenever a tweet related to the "Trade

---

[5] We assume an annual risk-free rate of 3% when calculating the Sharpe ratio

Table 20

Regression results for the impact of tweets with various topics on the Gold return.

|	Dependent variable: $Return_{t+d}$						
	(1 min)	(10 min)	(20 min)	(30 min)	(45 min)	(60 min)	(120 min)
Constant	0.0000	0.0000	0.0000	0.0003	−0.0001	0.0006	−0.0002
	(0.431)	(0.804)	(0.876)	(0.550)	(0.892)	(0.591)	(0.938)
WorldLeaderMeetings	−0.0044	−0.0065	−0.0217*	−0.0142	−0.0272	−0.0447*	−0.0822**
	(0.125)	(0.454)	(0.079)	(0.356)	(0.158)	(0.052)	(0.017)
TVShows	0.0002	−0.0001	0.0054	0.0022	−0.0037	−0.0081	−0.0183
	(0.887)	(0.979)	(0.460)	(0.807)	(0.743)	(0.543)	(0.359)
Corruption	−0.0008	0.0010	−0.0054	−0.0197*	−0.0108	−0.0241	0.0183
	(0.681)	(0.866)	(0.517)	(0.054)	(0.398)	(0.112)	(0.419)
TradeWar	0.0048**	0.0039	0.0186**	0.0174	0.0256*	0.0401**	0.0381
	(0.020)	(0.523)	(0.034)	(0.110)	(0.060)	(0.013)	(0.115)
Bragging	−0.0017	−0.0085	−0.0132	−0.0166	−0.0120	−0.0184	0.0035
	(0.397)	(0.152)	(0.123)	(0.115)	(0.365)	(0.241)	(0.879)
MakingDeals	−0.0020	−0.0038	−0.0072	−0.0045	−0.0267**	−0.0393***	−0.0255
	(0.297)	(0.506)	(0.380)	(0.661)	(0.037)	(0.010)	(0.261)
NorthKorea/States	−0.0071*	0.0028	−0.0124	−0.0241	0.0050	−0.0135	−0.0532
	(0.069)	(0.809)	(0.458)	(0.245)	(0.847)	(0.658)	(0.247)
USMilitary	−0.0013	−0.0040	−0.0157**	−0.0224***	−0.0117	−0.0138	0.0020
	(0.364)	(0.351)	(0.012)	(0.004)	(0.231)	(0.233)	(0.908)
Achievements	0.0037**	0.0089*	−0.0002	−0.0009	−0.0063	−0.0035	−0.0067
	(0.019)	(0.061)	(0.972)	(0.911)	(0.547)	(0.777)	(0.723)
Compassion	0.0021	0.0064	0.0025	−0.0134	−0.0090	−0.0001	0.0372
	(0.441)	(0.431)	(0.833)	(0.357)	(0.622)	(0.997)	(0.250)
BorderSecurity	0.0004	−0.0009	0.0106**	0.0135**	0.0151*	0.0221**	0.0193
	(0.729)	(0.800)	(0.047)	(0.041)	(0.068)	(0.024)	(0.190)
PoliticalDiscussions	−0.0006	−0.0036	0.0221***	0.0318***	0.0322***	0.0414***	0.0211
	(0.749)	(0.500)	(0.004)	(0.001)	(0.007)	(0.003)	(0.318)
Undefined	−0.0000	−0.0026	−0.0036	−0.0063	−0.0112**	−0.0138**	−0.0290***
	(0.976)	(0.300)	(0.313)	(0.157)	(0.045)	(0.040)	(0.005)
BorderWall/Board	−0.0028	0.0091	0.0217*	0.0399**	0.0144	0.0029	0.0077
	(0.358)	(0.318)	(0.099)	(0.014)	(0.478)	(0.905)	(0.834)
Announcements	0.0028	−0.0002	0.0003	0.0125	0.0034	0.0166	−0.0058
	(0.260)	(0.981)	(0.980)	(0.348)	(0.841)	(0.409)	(0.842)
FakeNews	0.0003	−0.0003	0.0013	0.0054	0.0133	0.0211*	0.0085
	(0.823)	(0.952)	(0.834)	(0.475)	(0.159)	(0.059)	(0.610)
RussianCollusion	0.0004	−0.0001	0.0033	0.0038	0.0027	0.0133	0.0222
	(0.805)	(0.985)	(0.606)	(0.631)	(0.782)	(0.246)	(0.201)
Employment	0.0049**	0.0022	0.0040	−0.0012	0.0035	0.0071	0.0309
	(0.021)	(0.727)	(0.661)	(0.917)	(0.802)	(0.671)	(0.214)
IllegalImmigrants	−0.0038*	−0.0112*	−0.0138	−0.0192*	−0.0222	−0.0241	−0.0327
	(0.078)	(0.083)	(0.135)	(0.093)	(0.119)	(0.154)	(0.193)
Presidents	0.0001	0.0063	−0.0190*	−0.0139	−0.0117	−0.0174	0.0052
	(0.974)	(0.401)	(0.080)	(0.301)	(0.487)	(0.381)	(0.861)
Observations	991,103	95,928	47,313	31,376	20,840	16,006	9165

Note: *p<0.1; **p<0.05; ***p<0.01.
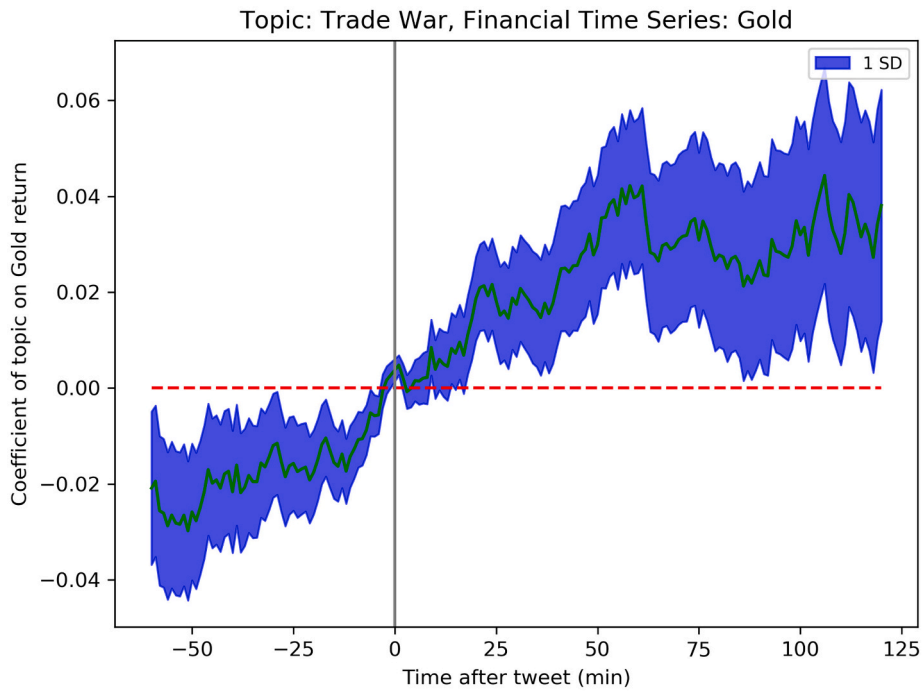
Fig. 10. Impact of the topic "Trade War" in Trump's tweets on the Gold return. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
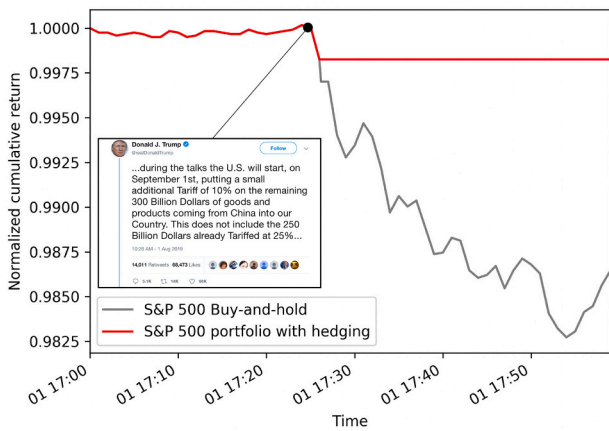

Fig. 11. Illustration of how the hedging strategy plays out with a real example on August 1, 2019.
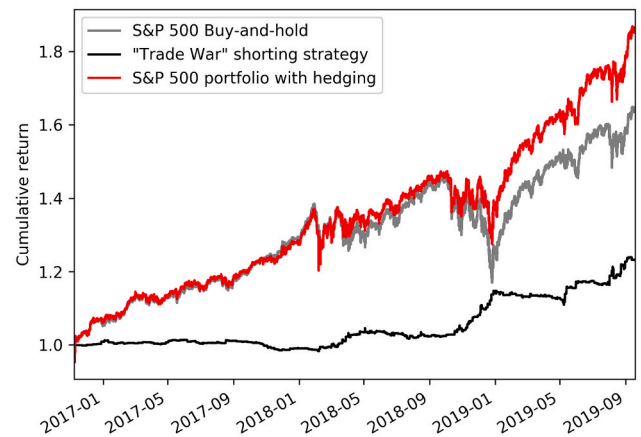

Fig. 12. Simulating trading on training sample data set.

Table 21
Sample data comparison.

Sample	No. of Tweets	No of Obs.
In-sample	8686	993,843
Out-of-Sample	1665	74,883

war" topic is published, but then the investor buys it back 75 min later. We evaluate this strategy and find that it successfully beats a buy-and-hold portfolio in terms of both higher return and lower volatility. Since this trading strategy is very simple, it is likely that more sophisticated forecasting techniques, such as artificial neural networks, would perform even better.

Table 22
Comparison of the trading strategies in the training and test sample. Values are annualized.

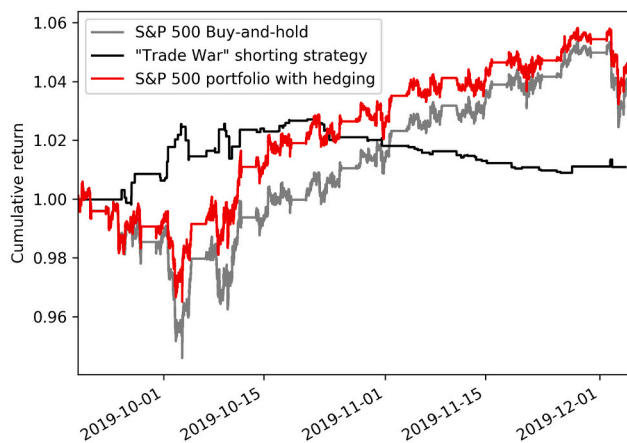|	Standard Dev	Avg return	Sharpe ratio
Training sample			
S&P 500 Buy-and-hold	0.1793	0.2466	1.2085
"Trade War" shorting strategy	0.0463	0.0830	1.1443
S&P 500 portfolio with hedging	0.1792	0.3428	1.7459
Test sample			
S&P 500 Buy-and-hold	0.1114	0.1832	1.3754
"Trade War" shorting strategy	0.0436	0.0520	0.5037
S&P 500 portfolio with hedging	0.1062	0.2112	1.7065

Fig. 13. Simulating trading on out-of-sample data set.

References

[1] C. Ajjoub, T. Walker, Y. Zhao, Social media posts and stock returns: The Trump factor, Int. J. Manag. Financ. 17 (2) (2020) 185–213.

[2] E. Bartov, L. Faurel, P.S. Mohanram, Can twitter help predict firm-level earnings and stock returns? Account. Rev. 93 (2017) 25–57.

[3] M. Bašta, P. Molnár, Long-term dynamics of the VIX index and its tradable counterpart VXX, J. Futur. Mark. 39 (2019) 322–341.

[4] D.G. Baur, B.M. Lucey, Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold, Financ. Rev. 45 (2010) 217–229.

[5] D.G. Baur, T.K. McDermott, Is gold a safe haven? International evidence, J. Bank. Financ. 34 (2010) 1886–1898.

[6] F. Bianchi, H. Kung, T. Kind, Threats to Central Bank Independence: High-Frequency Identification with Twitter. Technical Report, National Bureau of Economic Research, 2019.

[7] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent Dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.

[8] J. Bollen, H. Mao, X. Zeng, Twitter mood predicts the stock market, J. Comput. Sci. 2 (2011) 1–8.

[9] C. Bordonado, P. Molnár, S.R. Samdal, VIX exchange traded products: Price discovery, hedging, and trading strategy, J. Futur. Mark. 37 (2017) 164–183.

[10] J.A. Born, D.H. Myers, W.J. Clark, Trump tweets and the efficient market hypothesis, Algorithmic Fin. 6 (2017) 103–109.

[11] H. Brans, B. Scholtens, Under his thumb the effect of president Donald Trump's twitter messages on the us stock market, PLoS One 15 (2020), e0229931.

[12] T. Burggraf, R. Fendel, T.L.D. Huynh, Political news and stock prices: evidence from Trump's trade war, Appl. Econ. Lett. 27 (2020) 1485–1488.

[13] C. Colonescu, et al., The effects of Donald Trump's tweets on us financial and foreign exchange markets, Athens J. Bus. Econ. 4 (2018) 375–388.

[14] Q. Ge, A. Kurov, M.H. Wolfe, Do Investors care about presidential company-specific tweets? J. Financ. Res. 42 (2019) 213–242.

[15] M.S. Gerber, Predicting crime using twitter and kernel density estimation, Decis. Support. Syst. 61 (2014) 115–125.

[16] S. Gong, J. Zhang, P. Zhao, X. Jiang, Tweeting as a marketing tool: a field experiment in the TV industry, J. Mark. Res. 54 (2017) 833–850.

[17] T.L. Griffiths, M. Steyvers, Finding scientific topics, Proc. Natl. Acad. Sci. 101 (2004) 5228–5235.

[18] S. Guo, Y. Jiao, Z. Xu, Trump's effect on the chinese stock market, J. Asian Econ. 72 (2021) 101267.

[19] P.J. Huber, Robust regression: Asymptotics, conjectures and Monte Carlo, Ann. Stat. 1 (1973) 799–821.

[20] N.F. Ibrahim, X. Wang, A text analytics approach for online retailing service improvement: evidence from twitter, Decis. Support. Syst. 121 (2019) 37–50.

[21] R. Ichev, M. Marinc, N. Massoud, Social media and political power: Their impact on financial markets, in: SSRN Scholarly Paper ID 3334386, Social Science Research Network, Rochester, NY, 2019.

[22] A. Juma'h, Y. Alnsour, Using social media analytics: The effect of president Trump's tweets on companies' performance. Juma'h, Ahmad H., and Yazan Alnsour. Using Social Media Analytics: The Effect of President Trump's Tweets On Companies' Performance, J. Account. Manag. Informa. Syst. 17 (2018) 100–121.

[23] Y. Kim, K. Shim, TWILITE: a recommendation system for twitter using a probabilistic model based on latent Dirichlet allocation, Inf. Syst. 42 (2014) 59–77.

[24] J. Klaus, C. Koser, Measuring Trump: the volfefe index and its impact on european financial markets, Financ. Res. Lett. 101447 (2020).

[25] O. Kraaijeveld, J. De Smedt, et al., The predictive power of public twitter sentiment for forecasting cryptocurrency prices, J. Int. Financ. Mark. Inst. Money 65 (2020).

[26] D. Lee, K. Hosanagar, H.S. Nair, Advertising content and consumer engagement on social media: Evidence from Facebook, Manag. Sci. 64 (2018) 5105–5131.

[27] S.Y. Lee, L. Qiu, A. Whinston, Sentiment manipulation in online platforms: an analysis of movie tweets, Prod. Oper. Manag. 27 (2018) 393–416.

[28] S.L. Lo, R. Chiong, D. Cornforth, Ranking of high-value social audiences on twitter, Decis. Support. Syst. 85 (2016) 34–48.

[29] C. Loynes, J. Ouenniche, J. De Smedt, The detection and location estimation of disasters using twitter and the identification of non-governmental organisations using crowdsourcing, Ann. Oper. Res. (2020) 1–33.

[30] M.G. Lozano, J. Schreiber, J. Brynielsson, Tracking geographical locations using a geo-aware topic model for analyzing social media data, Decis. Support. Syst. 99 (2017) 18–29.

[31] J.G. MacKinnon, H. White, Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties, J. Econ. 29 (1985) 305–325.

[32] D.E. Pournarakis, D.N. Sotiropoulos, G.M. Giaglis, A computational model for mining consumer perceptions in social media, Decis. Support. Syst. 93 (2017) 98–110.

[33] L. Qiu, H. Rui, A. Whinston, Social network-embedded prediction markets: the effects of information acquisition and communication on predictions, Decis. Support. Syst. 55 (2013) 978–987.

[34] Rao, T., Srivastava, S., 2012. Analyzing stock market movements using twitter sentiment analysis, in: Proceedings of the 2012 International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2012), IEEE Computer Society, Washington, DC, USA. pp. 119–123.

[35] R. Rashid, A. de Leon, Making Trading Great Again : Trump-Based Stock Predictions Via Doc 2 Vec Embeddings, 2019.

[36] R.P. Schumaker, A.T. Jarmoszko, C.S. Labedz Jr., Predicting wins and spread in the premier league using a sentiment analysis of twitter, Decis. Support. Syst. 88 (2016) 76–84.

[37] R.P. Schumaker, C.S. Labedz Jr., A.T. Jarmoszko, L.L. Brown, Prediction from regional angst–a study of NFL sentiment in twitter using technical stock market charting, Decis. Support. Syst. 98 (2017) 80–88.

[38] P. Tripathi, S.K. Vishwakarma, A. Lala, Sentiment analysis of English tweets using rapid miner, in: 2015 International Conference on Computational Intelligence and Communication Networks (CICN), 2015, pp. 668–672.

[39] Y. Wang, W. Xu, Leveraging deep learning with lda-based text analytics to detect automobile insurance fraud, Decis. Support. Syst. 105 (2018) 87–95.

[40] H. Yuan, R.Y. Lau, W. Xu, The determinants of crowdfunding success: a semantic text analytics approach, Decis. Support. Syst. 91 (2016) 67–76.

Peder Gjerstad has an M.Sc. in Industrial Economics and Technology Management from the Norwegian University of Science and Technology and Harvard University, specializing in computer science, artificial intelligence, and finance. He currently works as a consultant at Boston Consulting Group, focusing primarily on quantitative analysis for the consumer industry and public sector. He has previously worked within private equity and venture capital, as well as with asset management strategies at the Norwegian Ministry of Finance.

Peter Filip Meyn has an M.Sc. in Industrial Economics and Technology Management from the Norwegian University of Science and Technology and University of New South Wales, specializing in electrical energy engineering and finance. He currently works as an investment analyst at Anthon B Nilsen, focusing primarily on valuations, strategies and analysis of potential investments. He has previously interned as an equity research analyst for Skandinaviska Enskilda Banken (SEB).

Peter Molnár works as Associate Professor at University of Stavanger Business School in Norway. He has a PhD in Finance from Norwegian School of Economics. Dr. Molnár is an accomplished researcher with over 40 publications. His research about financial markets is often focused on predictions based on sentiment analysis, textual analysis, machine learning, and other methods from data science.

Thomas Dowling Næss has an M.Sc. in Industrial Economics and Technology Management from the Norwegian University of Science and Technology and University of New South Wales, specializing in mechanical engineering and finance. He currently works as an analyst at Sparebank 1 Markets, focusing on equity research within the renewables sector. He has previously worked within consultancy field.