

Research Article

Special Issue on Econometrics and Business Analytics

Zibin Huang and Rustam Ibragimov*

Equity returns and sentiment

<https://doi.org/10.1515/demo-2022-0109>

received September 11, 2021; accepted March 29, 2022

Abstract: This paper analyzes approximately 100 Gigabytes of raw text data from Twitter with keywords “AAPL,” “S&P 500,” “FTSE100” and “NASDAQ” to explore the relationship between sentiment and the returns and prices on the Apple stock and the S&P 500, FTSE 100 and NASDAQ indices. The findings point to significant relationship and dependence between sentiment measures and the S&P 500 and FTSE 100 indices’ returns and prices. The econometric analysis of dependence between the aforementioned variables in the paper is presented in some detail for illustration of the methodology employed.

Keywords: sentiment, asset prices, asset returns, dependence, Granger causality, predictive regressions, autoregressive distributed lag models, GARCH models, volatility

MSC 2020: 62P20, 91B84

1 Introduction

Quantifying subjective opinion and using it as a predictor of stock market returns and prices have become an important topic of research and empirical analysis in academia and the industry. According to the efficient-market hypothesis (EMH) developed by Fama (see Fama et al. [11], Fama [12], and the references therein), information and thus the news are driving sources of equity prices as an “asset price reflects all available information.”

This paper focuses on the analysis of the question whether public sentiment provides valuable information that affects stock prices and on quantifying the significance of the effects of public opinion and sentiment for equity prices. The motivation for asking these questions is inspired by the research and the developments in behavioral economics and behavioral finance. In particular, research in these fields suggests that asset prices could be affected by human psychological or behavioral factors (Bollen et al. [4]). For instance, according to Gruhl et al. [17], Liu et al. [24], and Mishne and de Rijke [27], book, movie, and other products’ sales can be predicted by sentiment in social media such as blogs, Twitter posts, and so on. Hence, it is reasonable to assume that public feeling could affect the returns and prices of financial assets and indices. This is further consistent with the research in psychology, which indicates that emotion plays a significant role in human decision-making (Damasio [9]). Studies in behavioral finance and related fields also indicate that emotion and sentiment have a meaningful contribution to financial decisions and investors’ performance [25,28]. Moreover, research indicates that downward pressure on market prices is related to high media pessimism indicated by publications in the *Wall Street Journal* [33].

To study the effect of public mood on the stock price, we need to find a reliable, representative, and accessible proxy that can be used to construct time series on sentiment measures. Large-scale surveys for

* **Corresponding author: Rustam Ibragimov**, Imperial College Business School, London, United Kingdom; Centre for Econometrics and Business Analytics (CEBA) at St. Petersburg State University, St. Petersburg, Russia, e-mail: irustam@imperial.ac.uk

Zibin Huang: Imperial College Business School, London, United Kingdom, e-mail: aaahzb123@me.com

obtaining public mood are impractical, not only they are a waste of resources, but also they have great difficulty in producing a time-sensitive data. On the other hand, Twitter, a popular social media website that was launched in 2006, has millions of posts per day. The average number of monthly active Twitter users is over 330 million. Users of Twitter come from a variety of backgrounds, including CEOs, analysts, as well as the users' major component, the general public. Therefore, it is reasonable to choose the sentiment of Twitter posts as a proxy for the public mood [4]. Many papers in the literature have focused on the analysis of the relation between Twitter sentiment and stock market returns (see, among others, Behrendt and Schmidt [3], Corea [8], Groß-Klußmann et al. [16], Mittal and Goel [26], Ranco et al. [30], Washha et al. [34]).

In the first stage, this paper investigates the relationship between prices and public mood for three major indexes, S&P 500, NASDAQ, FTSE 100, and one corporation stock, Apple Inc, using a large database of collected posts on Twitter with the indices and the stock' tickers as keywords. In particular, the database includes approximately 8,029,963 posts for each keyword (3,000 randomly picked Tweets in each day) from January 1, 2008 to April 1, 2016.¹ We use Granger Causality tests to investigate whether a change in public sentiment can cause a change in stock prices.

In the second stage, we explore models that may be able to further explain the relationship between Twitter sentiment and the returns and prices on the S&P 500 and FTSE 100 indices. We employ and estimate Nonlinear Generalized AutoRegressive Conditional Heteroskedasticity (NGARCH) time series to explain and quantify the relationship and to model the effects of Twitter sentiment on volatility clustering in financial markets.

At the third stage, we focus on the analysis of the effects of Twitter sentiment on market volatility using the fitted GARCH models and Granger causality tests.

2 Data acquisition

This section discusses where and how the stock price and Twitter data are acquired and describes the methods for quantifying the sentiment in the text from Twitter posts and generating the data on sentiment measures that can be used in the analysis, tests, and model development.

The data on stock prices for the assets considered are downloaded from Yahoo Finance.

For the sentiment data, an R program has been developed by using Twitter Application Interface (API) to download an assigned and fixed number of Twitter posts (or tweets) with a particular keyword for each day throughout the time interval in consideration. In particular, we used the CRAN version (stable) *twitteR* library [14], which developed based on standard Twitter API (Twitter application access API at <http://dev.twitter.com/>). We used the "searchTwitter" function in this package to obtain data for a given date, keywords, maximum return tweets (the maximum number of returning tweets are limited by the API capability at the time), and other search strings. This program analyzed a total 100 GB of data from Twitter. Following Tetlock [33], to quantify sentiment, this paper uses various types of sentiment measures as suggested in the General Inquirer Categories in Harvard Psychological Dictionary [13].

The general analysis procedure employed in the paper is depicted in the flow graph (Figure 1).

2.1 Twitter mining algorithm

This section describes the algorithm that was used to convert Twitter text data into time series of different sentiment measures.

¹ The period considered thus does not include the beginning of the ongoing COVID-19 pandemic, and the analysis of structural breaks due to this and other shocks is beyond this paper and is left for future research.

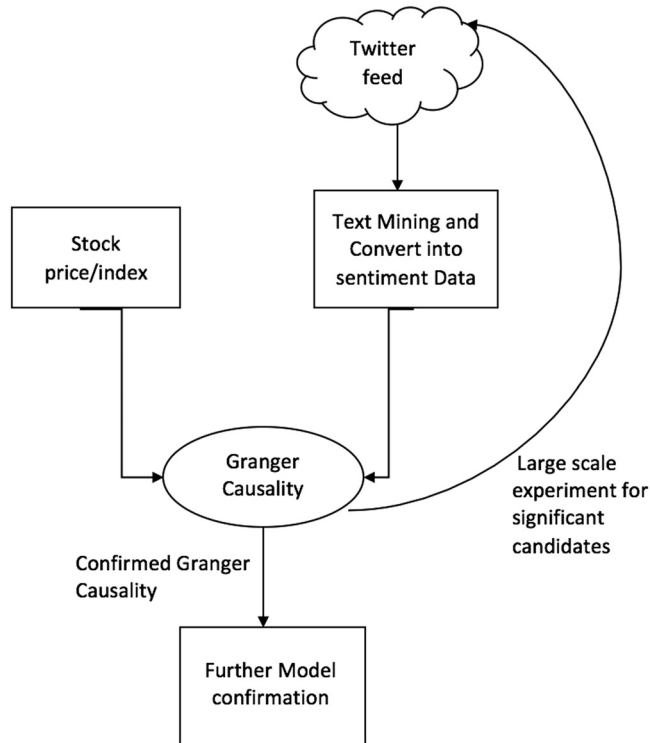


Figure 1: The analysis procedure.

First, for each day in the period from January 1, 2008 to January 4, 2016, a random sample of 3,000 Tweets was extracted. All Tweets on the same day were collected to form a large text file that was used as a proxy for public comments on Twitter. For each of the downloaded daily text file, all the punctuation and other symbols (e.g. “https://”) were removed to form a crude corpus. In the crude corpus, we applied a further filtration for removing any meaningless (for the purpose of sentiment quantification) words, such as “is,” “this,” etc., to form the final daily corpus.

Second, such daily generated corpus was checked using the Harvard Psychological Dictionary’s General Inquirer Categories [13] with four broad classes Positive, Negative, Active, and Passive, and also their subclasses Affiliation, Hostile, Strong, and Weak, by counting the frequency of words in the corpus that fall into a particular category. Hence, for each day, eight values for the word frequency in each of the group were obtained from the collected Tweets. The process has been repeated for every day in time interval dealt with. Following the procedure, time series of the raw sentiment data from Twitter were generated.

The example in Appendix A1 demonstrates how the algorithm works for randomly acquired five Tweets on a particular date. The Tweets considered in Appendix A1 are not sentimentally neutral and contain polarized oriented words such as “drop,” “Active,” “unable,” etc. that indicates their sentiment orientation to some extend. This observation also provides a logical justification for using the categorization method in the paper.

Third, to get a time series for testing the Granger causality between the Twitter sentiment and the stock price, several different sentiment measures are used in the paper. The sentiment measures considered are inspired by the analysis by Zhang and Skiena [35] and are summarized in the following formulas, with #Positive and #Negative, etc. standing for the number of words in the positive, negative, and other corresponding sentiment categories. The polarity sentiment measure is defined as follows:

$$\text{Polarity} = \frac{\# \text{Positive} - \# \text{Negative}}{\# \text{Positive} + \# \text{Negative}}. \quad (2.1)$$

Obviously, the Polarity measure is not guaranteed to always be positive similar to asset prices as the number of positive words in the Twitter posts considered is not necessarily larger than the word count of

negative words. To ensure positivity of the sentiment measures considered, without loss of generality, Zhang and Skiena' Polarity measures are modified as the "Relative positive" measures, which are defined as follows:

$$\text{Relative_positive} = \frac{\# \text{Positive}}{\# \text{Positive} + \# \text{Negative}}. \quad (2.2)$$

In a similar way, we define Relative Affiliation, Relative Active, and Relative Strong measures as follows.

$$\text{Relative_Active} = \frac{\# \text{Active}}{\# \text{Active} + \# \text{Passive}}. \quad (2.3)$$

The categories Affiliation and Strong are subclasses of the Positive and Active categories. Their relative sentiment measures are defined as follows:

$$\text{Relative_Affiliation} = \frac{\# \text{Affiliation}}{\# \text{Positive} + \# \text{Negative}}, \quad (2.4)$$

$$\text{Relative_Strong} = \frac{\# \text{Strong}}{\# \text{Active} + \# \text{Passive}}. \quad (2.5)$$

3 Granger causality analysis

As the sentiment data have been acquired, Granger causality tests were performed to investigate whether there is a causality between the Twitter sentiment and the stock price, which also gives a partial answer to the question on the relationship between these variables. The Ganger causality tests are applied to time series confirmed to be stationary $I(0)$ processes.

3.1 (Non-)stationarity analysis

First, we conduct augmented Dickey-Fuller (ADF) tests to investigate the presence of unit root in the processes considered. Testing for a unit root in a time series X_t is based on the ordinary least squares (OLS) regression:

$$X_t = \beta_0 + \alpha t + \delta X_{t-1} + \gamma_1 \Delta X_{t-1} \dots + \gamma_p \Delta X_{t-p} + \varepsilon_t,$$

where β_0 is a constant, α is a time trend coefficient, and ε_t is the innovation process with zero mean. Under $\mathbf{H}_0 : \delta = 0$, the process is nonstationary, and $\mathbf{H}_a : \delta < 0$ corresponds to stationarity in X_t (Dickey and Fuller [10], Ch. 17 in Hamilton [19] and Section 15.7 in Stock and Watson [31]).

The results of the ADF tests indicate that all the processes considered, including all the time series of sentiment measures dealt with and the logarithms of stock prices, are unit root $I(1)$ processes. The analysis and tests of Granger causality in the paper are therefore based on stationary $I(0)$ first differences of the aforementioned processes. In other words, in the following section, we focus on testing of Granger causality between the changes of the log price and the changes in the Twitter sentiment measures considered.

3.2 Autoregressive distributed lag models in Granger causality tests

A process X_t is said to Granger cause a process Y_t if the lags of X_t have useful predictive content for forecasting Y_t , above and beyond the other regressors, e.g., the lags of Y_t itself, in the model (see, among

others, Section 15.4 in [31]). The Granger causality test is usually carried out using autoregressive distributed lag (ADL) models

$$Y_t = \beta_0 + \sum_{j=1}^J \beta_j Y_{t-j} + \sum_{k=1}^K \gamma_k X_{t-k} + u_t. \quad (3.1)$$

The Granger causality test is carried out using an F -test on all the coefficients on the lags of X_t . The null hypothesis in this test is $\mathbf{H}_0 : \forall k \in \{1, \dots, K\}, \gamma_k = 0$, which equivalently means that X_t is not a useful predictor of Y_t , given the lags of the latter process. The alternative hypothesis in the test is $\mathbf{H}_a : \exists k \in \{1, \dots, K\}, \gamma_k \neq 0$, which corresponds to the property that the lags of X_t do have some useful predictive content for forecasting Y_t , beyond the lags of Y_t itself.

Let P_t denote the logarithm of the price of a stock/index considered and let Sem denote a sentiment measure. The test whether the Twitter sentiment Granger causes the stock price is based on the following model:

$$\Delta P_t = \beta_0 + \sum_{j=1}^J \beta_j \Delta P_{t-j} + \sum_{k=1}^K \gamma_k \Delta \text{Sem}_{t-k} + u_t. \quad (3.2)$$

Similarly, the test that the stock price Granger causes the Twitter sentiment is based on the model

$$\Delta \text{Sem}_t = \beta_0 + \sum_{m=1}^M \beta_m \Delta \text{Sem}_{t-m} + \sum_{n=1}^N \gamma_n \Delta P_{t-n} + \varepsilon_t. \quad (3.3)$$

3.3 Determination of the number of lags using the BIC criterion

As shown in Eqs. (3.1)–(3.3), implementation of Granger causality tests requires determination of the number of lags J, K, M, N for both of the processes considered. We use the Bayesian information criterion (BIC) for determination of the number of lags of processes X_t and Y_t (ΔP_t and ΔSem) in the ADL models dealt with. More precisely, as usual, first, the number J of lags in autoregressive (AR) models for the process Y_t

$$Y_t = \beta_0 + \sum_{j=1}^J \beta_j Y_{t-j} + u_t$$

is determined based on the BIC, and then the criterion is applied to determine the number of lags K of the potential predictor X_t in model (Eq. (3.1)) with the estimated number J .

The results of the lag length selection on the basis of the BIC for Granger causality tests are provided in Appendix A2.

3.4 The results of Granger causality tests

The results of tests of Granger causality between the considered asset returns and the corresponding Twitter sentiment are provided in Tables 1 and 2.

Similar to Section 3.2, the null hypotheses in the tables are that the changes in sentiment do not Granger cause the change in the log prices, that is, the returns, and *vice versa*.

The results in Table 2 indicate that, somewhat surprisingly, that the changes in (log) prices apparently do not Granger cause the Twitter sentiment. These conclusions are in contrast to the conventional belief that the changes in asset prices affect public sentiment.

Further, the results in Tables 1 and 2 point to the conclusion that the change of Twitter sentiment related to S&P 500 and FTSE 100 indices Granger causes their price changes but not *vice versa*. In particular,

Table 1: *F*-statistics for the hypothesis H_0 : the changes in sentiment do not Granger cause the asset returns

	Analyzed tweets	Relative positive	Relative active	Relative affiliation	Relative strong
S&P500	9,116,537	0.0024253***	0.010023**	0.01043**	0.0106**
FTSE100	9,099,000	0.012089**	0.022854**	0.017246**	0.011853**
NASDAQ	8,826,775	0.85884	0.44155	0.83846	0.64976
AAPL	7,608,000	0.40384	0.86842	0.26875	0.13011

Notes: ***indicates the 1% significance and **indicates the 5% significance.

Table 2: *F*-statistics for the hypothesis H_0 : the asset returns do not Granger cause the changes in sentiment

	Analyzed tweets	Relative positive	Relative active	Relative affiliation	Relative strong
S&P500	9,116,537	0.88878	0.47615	0.23448	0.94569
FTSE100	9,099,000	0.92718	0.33062	0.57001	0.80284
NASDAQ	8,826,775	0.75628	0.66484	0.54052	0.91209
AAPL	7,608,000	0.86636	0.62385	0.67329	0.69855

according to Table 1, among all the sentiment measures and the assets considered, the effect of the Relative Positive measure on S&P 500 returns appears to be the most significant, with significance of the test statistics at the 1% level. In contrast, the returns on the NASDAQ index and the Apple stock appear not to be Granger caused by the respective Twitter sentiment of price.

3.5 S&P 500: Granger causality tests using the big data

To further evaluate and confirm the results on Granger causality in the previous section, we conduct the tests of Granger causality between the S&P 500 returns and the respective sentiment measures using a very large-scale database.

Different from the first stage analysis (3,000 randomly selected with target keywords, by limiting the maximum return tweet in twitterR: searchTwitter function), in this section, we do not give a limit to the maximum return tweets in twitterR: searchTwitter function, just go for the max number that Twitter application access API can provide in one request. Tweet with the indices and the stock tickers as keywords are applied in the search. Also, The twitterR Library are based on Twitter Application access API. The Twitter Application access API not only limit the maximum number of tweets in each request but also limit a certain amount requests in a time period with a given IP. To acquire large data for analysis, We have registered multiple accounts (Hence multiple tokens) and switch IP each time when a request were sent. The analysis is based on all the acquirable data from the Twitter API for the index in the time period considered. The analysis is not conducted for the FTSE 100 index as there are much fewer posts related to it as compared to the S&P 500. As indicated earlier, for the S&P 500, as many data as possible were extracted and analyzed for each day in the time period dealt with, with approximately 14,005 Tweets per day (this number is a mathematical average estimation based on all obtained data with the target keywords) and 42,478,072 Tweets in total.

The general observation was that the number of obtained tweets with corresponding keyword in each day is increasing with time. This is consisted with the fact that tweet is getting more and more well known and more people are posting their thoughts on tweet over time. For example, some keys have only 5,000 tweets per day in 2008, but obtained tweets number per day gets more in the follow year. Eventually the obtained daily number of tweets by a specific keyword is limited by the Twitter API we used.

The results of Granger causality tests for the S&P 500 returns and the sentiment measures using the large-scale database are provided in the second rows of Tables 3 and 4. For comparison, we also provide, in the first rows of the tables, the Granger causality test statistics for the S&P 500 from the previous section.

Table 3: F -statistics for the hypothesis H_0 : the changes in sentiment do not Granger cause the S&P 500 returns

	Analyzed tweets	Relative positive	Relative active	Relative affiliation	Relative strong
1st attempt	9,116,537	0.0024253***	0.010023**	0.01043**	0.0106**
2nd attempt	42,478,072	0.0054962***	0.0042141***	0.0036191***	0.0086103***

Notes: ***indicates the 1% significance and **indicates the 5% significance.

Table 4: F -statistics for the hypothesis H_0 : the S&P 500 returns do not Granger cause the change in sentiment

	Analyzed tweets	Relative positive	Relative active	Relative affiliation	Relative strong
1st attempt	9,116,537	0.88878	0.47615	0.23448	0.94569
2nd attempt	42,478,072	0.78777	0.10034	0.67303	0.72055

The results in Tables 3 and 4 using the large-scale data confirm the results in the previous section that the Twitter sentiment related to the S&P 500 index appears to Granger cause its returns and (log) price changes but not *vice versa*.

In conclusion, the returns and the prices of the S&P 500 and FTSE 100 indices appear to be Granger caused by the public sentiment. On the other hand, according to the results in this and the previous section, the changes in prices of the assets considered appear not to Granger cause the respective sentiment.

4 Causality modeling: ADL and GARCH models

As discussed in the previous section, the Twitter sentiment appears to Granger cause the returns on the S&P 500 and FTSE 100 indices. In this section, we focus on the analysis of models for the relationship between the returns on the indices and the respective sentiment. In particular, we evaluate the ADL models for the relationship and further fit GARCH time series to model the effects of sentiment volatility on market volatility.

4.1 Volatility clustering

We first focus on the estimation of ADL models for the relationship between the returns on the S&P 500 and FTSE 100 indices and the lags of the sentiment measures. Similar to the analysis in the previous section, following the results in Section 3.1, the models are estimated for the stationary changes in the log prices – the returns – and the stationary changes in the measures of sentiment dealt with.

The estimated ADL models include all the sentiment measures that were shown in the previous section to Granger cause the returns on the indices. The estimated models thus have the following form:

$$\Delta P_t = \beta_0 + \sum_{k=1}^K \beta_k \Delta P_{t-k} + \sum_{i=1}^I \zeta_i \Delta \text{Pos}_{t-i} + \sum_{j=1}^J \gamma_j \Delta \text{Act}_{t-j} + \sum_{l=1}^L \eta_l \Delta \text{Aff}_{t-l} + \sum_{m=1}^M \alpha_m \Delta \text{Str}_{t-m} + \varepsilon_t, \quad (4.1)$$

where $\{\Delta P_t\}$, $\{\Delta \text{Pos}_t\}$, $\{\Delta \text{Act}_t\}$, $\{\Delta \text{Aff}_t\}$, and $\{\Delta \text{Str}_t\}$ are the time series of the change in the logarithm of the prices of the indices, and the time series of the changes in the Relative Positive, Relative Active, Relative Affiliation and Relative Strong sentiment measures. The form of the ADL models is motivated by accounting for different classes and subclasses of sentiment in Harvard Psychological Dictionary's General Inquirer Categories [13] and related measures in Section 2.1 used in the sentiment analysis in the paper. The numbers I , J , L , and M of lags included in the models are those determined by the BIC criterion as discussed in Section 3.3 and Appendix A2.

ADL models (4.1) estimated by the OLS result in a poor fit for the time series of the returns on both the S&P 500 and FTSE 100 indices. Further, the plots of the residuals from the ADL regressions point to pronounced volatility clustering in the errors in the estimated linear models.

The results on the poor fit of linear models for the returns and the presence of volatility clustering in the ADL regression errors and the returns are in accordance with the well known stylized facts of the absence of linear dependence and the presence of nonlinear dependence in financial returns (see Cont [7]).

Following the results, in the next section, we thus focus on the models capturing the volatility clustering in the ADL model errors and the returns on the indices considered.

4.2 Modeling Granger causality and volatility clustering: ADL models with NGARCH errors

To model Granger causality between the sentiment and the returns on the S&P 500 and FTSE 100 indices accounting for volatility clustering in the returns, as usual, we employ GARCH-type time series. As is well known, GARCH-type processes can be used to capture and model the most of the stylized facts of financial returns, including the absence of linear autocorrelations, the presence of volatility clustering and autocorrelations in squared returns, heavy-tailedness and conditional heavy-tailedness, and the leverage effect (see, among others, Alberg et al. [2], Christoffersen [6], and Cont [7]). $\mathcal{F}_t$ denotes the filtration that contains all the information up to time t , and $N(0, 1)$ and $t(\nu)$ denote the standard normal and (heavy-tailed) Student- t distribution with ν degrees of freedom, respectively.

The following ADL models with NGARCH errors exhibiting the aforementioned stylized facts are estimated using the maximum likelihood (ML):

$$\Delta P_t = \omega_0 + \sum_{k=1}^K \omega_k \Delta P_{t-k} + \sum_{i=1}^I \zeta_i \Delta \text{Pos}_{t-i} + \sum_{j=1}^J \gamma_j \Delta \text{Act}_{t-j} + \sum_{l=1}^L \eta_l \Delta \text{Aff}_{t-l} + \sum_{m=1}^M \alpha_m \Delta \text{Str}_{t-m} + u_t, \quad (4.2)$$

where

$$u_t = \varepsilon_t \sigma_t, \quad \varepsilon_t | \mathcal{F}_{t-1} \sim N(0, 1) \text{ or } t(\nu), \quad (4.3)$$

with i.i.d. innovations ε_t that have a standard normal or Student- t distributions, and the volatility dynamics is given by the NGARCH model in the following form:

$$\sigma_t^2 = \alpha_0 + \alpha_1 (u_{t-1} - \kappa_1)^2 + \beta_1 \sigma_{t-1}^2. \quad (4.4)$$

The NGARCH specification for the errors in ADL models for index returns considered accounts for the properties of absence of linear autocorrelations, volatility clustering, heavy-tailedness, and leverage effect in returns time series.

The models impose stationary, that is, the condition

$$\alpha_1 + \beta_1 < 1,$$

on the GARCH parameters.

The results of the ML estimation of the aforementioned models are provided in the following sections.

4.2.1 S&P 500

As shown in the results in Table 5, unlike linear ADL models estimated by the OLS, the ADL models with NGARCH errors described in the previous section provide an exceptional fit for the S&P 500 returns.

The results in Table 5 further confirm that the changes in sentiment measures are useful predictors of the changes of the index prices and returns. In particular, in the case of the ADL models with NGARCH

Table 5: AGARCH model for S&P500

Error distribution	Normal $N(0,1)$		Student- t	
	Coefficient	t -Prob	Coefficient	t -Prob
$\Delta Price$	-0.0512350	0.014*	-0.0228679	0.082*
$\Delta Rel_Positive$	685.612	0.088*	539.484	<0.001***
ΔRel_Active	-328.132	0.594	-108.797	<0.001***
$\Delta Rel_Affiliation$	73.6310	0.854	192.964	<0.001***
ΔRel_Strong	-234.034	0.361	-188.109	<0.001***
Constant	0.721184	<0.001***	0.567255	<0.001***
α_0	2.84304	0.002***	7.78590	0.006***
α_1	0.0674871	<0.001***	0.0789915	<0.001***
β_1	0.917239	<0.001**	0.921008	<0.001***
Asymmetry (κ_1)				
Degree of freedom (Student- t)	N/A	N/A	2.37587	<<0.001***

Notes: ***indicates the 1% significance, **indicates the 5% significance and *indicates the 10% significance.

errors and heavy-tailed Student- t innovations, the lags of the changes in all the sentiment measures appear to be highly significant, with the corresponding p -values less than 0.001. Further, even in the case of ADL-NGARCH model errors with standard normal innovations, one of the sentiment measures, Relative Positive, exhibits statistical significance in predictive models for the S&P 500 returns.

4.2.2 FTSE 100

Similar to the S&P 500 case, the estimation results in Table 6 for predictive ADL regression models for FTSE 100 returns with NGARCH errors demonstrate statistical significance of the changes in the sentiment measures.

The results in Table 6 indicate statistical significance of the regressors, including all the sentiment measures considered, in the predictive ADL models for FTSE 100 returns both in the case of normal and nonnormal heavy-tailed Student- t innovations in NGARCH models for the regression errors. Similar to the previous section, the results confirm predictive power of the sentiment measures for prediction of the returns and further confirm the presence of volatility clustering and other stylized facts in the ADL regression errors and the returns dealt with, the properties not captured by ADL models estimated by the OLS.

Table 6: AGARCH Model for FTSE 100

Error distribution	Normal $N(0,1)$		Student- t	
	Coefficient	t -Prob	Coefficient	t -Prob
$\Delta Price$	-0.0447849	0.020**	-0.0204596	0.150
$\Delta Rel_Positive$	-4356.07	<0.001***	360.002	<0.001***
ΔRel_Active	9606.08	<0.001***	1836.06	<0.001***
$\Delta Rel_Affiliation$	4863.51	<0.001***	149.458	<0.001***
ΔRel_Strong	-1255.61	<0.001***	-1380.03	<0.001***
Constant	-0.343865	0.678	0.824210	0.081*
α_0	4.83313×10^{-115}	<<0.001***	1.90373×10^{-14}	<<0.001***
α_1	0.0512398	<0.001***	0.111713	<<0.001
β_1	0.918548	<0.001***	0.888287	<0.001***
Asymmetry (κ_1)	40.3643	<0.001***	52.3062	<0.001***
Degree of freedom (Student- t)	N/A	N/A	2.33052	<0.001***

Notes: ***indicates the 1% significance and **indicates the 5% significance.

5 Causality between asset price volatility and sentiment volatility

The results in Sections 3 and 4 indicate the presence of volatility clustering in the errors from the predictive ADL models, with the dynamics that can be modeled using NGARCH time series. In this section, we focus on the analysis of causality between the volatilities of the returns and sentiment processes. Similar to Patton [29], the analysis is based on Granger causality tests using the residuals from fitting GARCH-type models to both of the processes considered.

5.1 Models for causality between volatilities

Consider two time series $\{X_t\}$ and $\{Y_t\}$ and, as mentioned earlier, denoted by $\mathcal{F}_t$ the filtration containing the information up to time t . The analysis of causality between the volatilities of the processes is based on Granger causality tests for innovations-residuals z_t and z'_t from the GARCH processes fitted to $\{X_t\}$ and $\{Y_t\}$:

$$X_t = \sigma_t z_t, \quad z_t | \mathcal{F}_t \sim N(0, 1), \quad (5.1)$$

$$Y_t = \delta_t z'_t, \quad z'_t | \mathcal{F}_t \sim N(0, 1) \quad (5.2)$$

and

$$\sigma_t^2 = w + \alpha \cdot X_{t-1}^2 + \beta \cdot \sigma_{t-1}^2 + \varepsilon_t, \quad (5.3)$$

$$\delta_t^2 = w' + \alpha' \cdot Y_{t-1}^2 + \beta' \cdot \delta_{t-1}^2 + \eta_t. \quad (5.4)$$

More precisely, the estimates of the GARCH parameters are obtained using maximum likelihood estimation (MLE), and then tests of Granger causality are conducted for the GARCH model residuals/standardized processes $\hat{\varepsilon}_t = X_t / \hat{\sigma}_t$ and $\hat{\eta}_t = Y_t / \hat{\delta}_t$. Granger causality testing described in Section 3 is used to investigate whether there is a causality between $\{\varepsilon_t\}$ and $\{\eta_t\}$. In particular, if the tests indicate that $\{\varepsilon_t\}$ Granger causes $\{\eta_t\}$, then this implies that the information contains in the past volatility of $\{X_t\}$ is useful for forecasting the volatility of $\{Y_t\}$. In the following analysis, the approach is applied to the time series $\{X_t\}$ and $\{Y_t\}$ being the processes of asset returns and the measures of Twitter sentiment considered.

5.2 Data preparation

The analysis of Granger causality between volatilities is based on standardized returns and sentiment measures.

More precisely, given the observations on a time series (e.g., that of the returns or the sentiment measures) $\{X_t\}$, we consider its standardized version:

$$\text{STD}(\{X_t\}) := \left\{ \frac{X_t - \bar{X}}{s_X}, \quad \forall t \right\}, \quad (5.5)$$

where, as usual, $\bar{X}$ and s_X denote the sample mean and standard deviation of the time series observations.

The analysis is based on the standardized time series $\text{STD}(\{r_t\})$ and $\text{STD}(\{\text{Sem}_t\})$ for the returns and sentiment processes $\{r_t\}$ and $\{\text{Sem}_t\}$, respectively.

A few (approximately 5 out of 3,500) large outliers are observed in the standardized sentiment measure time series $\text{STD}(\{\text{Sem}_t\})$. The presence of the outliers may be due to a relatively large number of reposts of polarized sentiment-oriented Tweets. This is similar to observations and the discussion in Tetlock [32] on some news having textual similarity with others. In the case of the Twitter, outliers caused by reposts may not be representing the real sentiment, as people can actually write their own comments along with their reposts. Some of the people's comments along with their reposts could have a sentiment that is totally

Table 7: F -statistics for the hypothesis H_0 : the sentiment volatility does not Granger cause the asset return volatility

Ganger causality: Volatility of stock index caused by sentiment volatility?				
	Adj_Positive	Adj_Active	Adj_Strong	Adj_Affiliation
S&P500	0.20012	0.49954	0.5639	0.64188
FTSE100	$1.503 \times 10^{-07***}$	$6.994 \times 10^{-08***}$	$8.891 \times 10^{-08***}$	$1.263 \times 10^{-07***}$

Notes: ***indicates the 1% significance.

opposite to the reposted Tweets. Further, the presence of such large outliers may severely affect the fitting of GARCH models, in part due to the analyzed sentiment measures being squared in the analysis.

To deal with the outliers, we make the assumption that the maximum change in the standardized sentiment is the same as the maximum change in the standardized return and estimate the following model with GARCH errors (see also Carnero et al. [5]):

$$\text{Sem}_t = k \cdot \text{Sem}_t \cdot 1_{\{|\text{Sem}_t| > \max(\{|\eta_t|\})\}} + u_t, \quad (5.6)$$

$$u_t = \sigma_t z_t, \quad z_t | \mathcal{F}_{t-1} \sim N(0, 1)$$

and

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \varepsilon_t. \quad (5.7)$$

The estimation results in Appendices 4 and 5 indicates the coefficient k being statistically significant and equal to 1 for all the sentiment measures considered. Hence, we further estimate the GARCH model for the following process:

$$\text{Sem}_t \cdot (1 - 1_{\{|\text{Sem}_t| > \max(\{|\eta_t|\})\}}) = \sigma_t z_t, \\ z_t | \mathcal{F}_t \sim N(0, 1),$$

and the dynamics of the volatility σ_t^2 follows (Eq. (5.7)).

5.3 Granger causality tests for sentiment and return volatilities

The analysis and tests for Granger causality is similar to the discussion in Section 3. The Granger causality tests are provided for the volatility of the standardized returns and the adjusted sentiment measures for different categories as discussed in Section 5.2. The results of the tests are as follows.

The results in Table 7 indicate that the volatility of the FTSE 100 index return appears to be Granger caused by the volatility of all the sentiment measures related to the index, while this is not the case for the S&P 500 return volatility. In addition, according to the results in Table 8, the volatility of the Affiliation sentiment measure related to the FTSE 100 index appears to be Granger caused by the index return volatility. The results in the tables further point to the absence of Granger causality between the return and sentiment volatility for the S&P 500.

Table 8: F -statistics for the hypothesis H_0 : the asset return volatility does not Granger cause the sentiment volatility

Ganger Causality: Volatility of sentiment caused by stock index volatility?				
	Adj_Positive	Adj_Active	Adj_Strong	Adj_Affiliation
S&P500	0.43677	0.7713	0.73771	0.87481
FTSE100	0.65901	0.28048	0.12179	0.014483***

Notes: ***indicates the 1% significance.

6 Conclusion

This paper has focused on the problems of quantifying subjective sentiment and the analysis of its use as a predictor for asset returns and prices. The study is based on a vast first amount of data (approximately 100 GB) acquired from Twitter using text mining and quantification of the sentiment according to the General Inquirer Categories of the Harvard psychological dictionary. The relation between the sentiment and the returns on the stock indices is analyzed using Granger causality tests and GARCH-type modeling.

The results of the analysis indicate that the Twitter sentiment apparently has predictive power for the returns on the S&P 500 and FTSE 100 financial indices. The results of the study further indicate that the volatility of the sentiment measures related to the FTSE 100 index appears to Granger cause the index return volatility, while this is not the case for the S&P 500 index.

An important problem that is left for further research is that of structural breaks in models for the relation and dependence between asset prices and returns and the sentiment, including the structural breaks due to the beginning of the on-going COVID-19 pandemic.

The paper uses Harvard Psychological Dictionary for sentiment analysis; further research may focus on applications of more recent sentiment analysis methods using artificial neural networks and other machine learning, such as Bidirectional Encoder Representations from Transformers (BERT) technique for natural language processing.

Due to the fact that the sentiment appears to Granger cause the returns on the financial indices considered, further analysis may also focus on predictive models incorporating the sentiment and other predictors, including the factors used in predictive regressions for financial returns and also on using the sentiment as a signal for trading. The analysis may be based on the widely used econometric methods as well as machine learning approaches.

As is common in the literature dealing with the analysis of dependence between financial and economic time series exhibiting volatility clustering such as financial returns and foreign exchange rates (see, among others, Patton [29]), the analysis of Granger causality in the paper is based on estimated volatilities. Further research may focus on the development and the use of methods that account for the uncertainty introduced in the first stage of the analysis by volatility estimation. In particular, the use of robust t -statistic approaches to inference under heterogeneity, dependence and heavy-tailedness developed by Ibragimov and Müller [21,22] (see also Section 3.3 in the study by Ibragimov et al. [23]), and their extensions appears to be perspective in the context of econometric inference using general two-stage procedures as the approaches do not require consistent estimation of limiting variances of estimators dealt with/their standard errors (see, in particular, Ibragimov and Müller [21] for the discussion of applicability and the properties of t -statistic approaches in inference in two-stage instrumental variable regressions and general GMM models and Abduraimova [1], for applications of the approaches in IV regressions for the analysis of effects of contagion on the tail risk in complex financial networks).

Acknowledgments: We thank two anonymous referees and the participants at the seminar series at the Centre for Econometrics and Business Analytics (CEBA) at St. Petersburg State University for helpful comments and suggestions.

Conflict of interest: The authors state no conflict of interest.

Appendix

A1 Text mining algorithm: An illustration

This appendix provides an example of five random selected Tweets from all the Twitter posts on August 2, 2016, with the keyword “S&P 500” and the results of quantification, using the text mining algorithm, of the

frequency (word count) of words that belong to a particular class in the Harvard Psychological Dictionary (Table A1).

Searching the keyword: “S&P 500 since: 2016-08-02 until: 2016-08-02”

Five randomly obtained tweets:

- [1] “AdCapPeru: The S&P 500 posted its biggest drop since June, as Oil was unable to hold gains, closing below \$40.00, ahead of weekly #AdCapPeru #oil”
- [2] “BuyETFs: Wonder how stocks do in election years? Average S&P 500 performance in election years since 1928 is +11.25%. #Elections2016 #stocks”
- [3] “dass5981: Cisco (NASDAQ:CSCO) drops 0.36% on Tuesday:Among top most actives on S&P 500 [\\$CSCO](https://t.co/vXhfhxcr) #Nasdaq <https://t.co/p484eBH0ux>”
- [4] “dass5981: General Electric drops 0.32% on Tuesday : Among top 10 most actives on S&P 500 [\\$GE](https://t.co/Qv9bbNVg5b) #NYSE <https://t.co/tUP5aemp2c>”
- [5] “dass5981: Apple (NASDAQ:AAPL) drops 1.48% on Tuesday:Among top most actives on S&P 500 [\\$AAPL](https://t.co/gyGA7E1O3F) #Nasdaq <https://t.co/4nezU7PCAC>”

Text mining result:

Table A1: Frequency (word count) of words in different sentiment classes

Sentiment category	Word counts
Positive	2
Negative	1
Affiliation (subclass of positive)	0
Hostile (subclass of negative)	1
Active	5
Passive	3
Strong (subclass of active)	2
Weak (subclass of passive)	3

A2 Lag length selection for the Granger causality tests

This appendix provides the results of the lag length selection for Granger causality tests using the BIC (Table A2).

A3 Volatility clustering in the residuals of the ADL linear models estimated by the OLS

Figures A1 and A2 provide the plots of the time series of the residuals from the ADL linear predictive models for the FTSE 100 and S&P 500 returns estimated by the OLS. According to the plots, pronounced volatility clustering is present in the time series of regression errors in ADL models discussed in Section 4.1.

A4 GARCH model for FTSE 100 returns

The estimates of the parameters of the GARCH model fitted to the FTSE 100 returns are provided in the following table.

Table A2: BIC lag length selection

	Sentimental cause price		Price cause sentimental	
	BIC_Stock	BIC_Sentiment	BIC_Stock	BIC_Sentiment
BIC lag selection of S&P 500				
Relative Positive	1	1	1	3
Relative Active	1	1	1	2
Relative Strong	1	1	1	6
Relative Affiliation	1	1	1	2
BIC lag selection of FTSE100				
Relative Positive	1	1	1	11
Relative Active	1	1	1	12
Relative Strong	1	1	1	11
Relative Affiliation	1	1	1	11
BIC lag selection of NASDAQ				
Relative Positive	2	1	1	1
Relative Active	2	1	1	1
Relative Strong	2	1	1	1
Relative Affiliation	2	1	1	1
BIC lag selection of AAPL				
Relative Positive	1	1	1	4
Relative Active	1	1	1	8
Relative Strong	1	1	1	4
Relative Affiliation	1	1	1	3

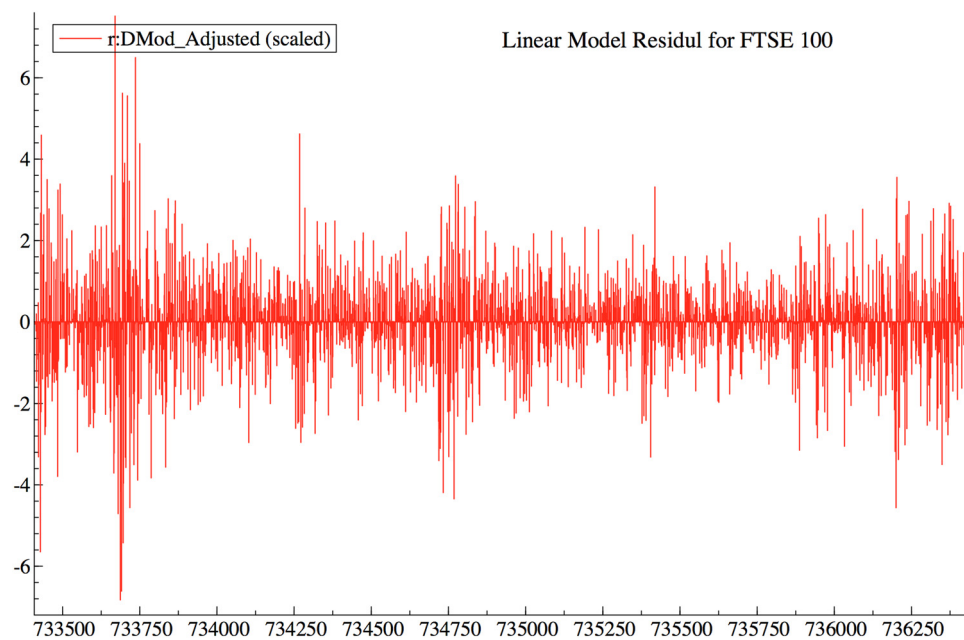


Figure A1: Residuals from the ADL linear predictive model for FTSE 100 returns.

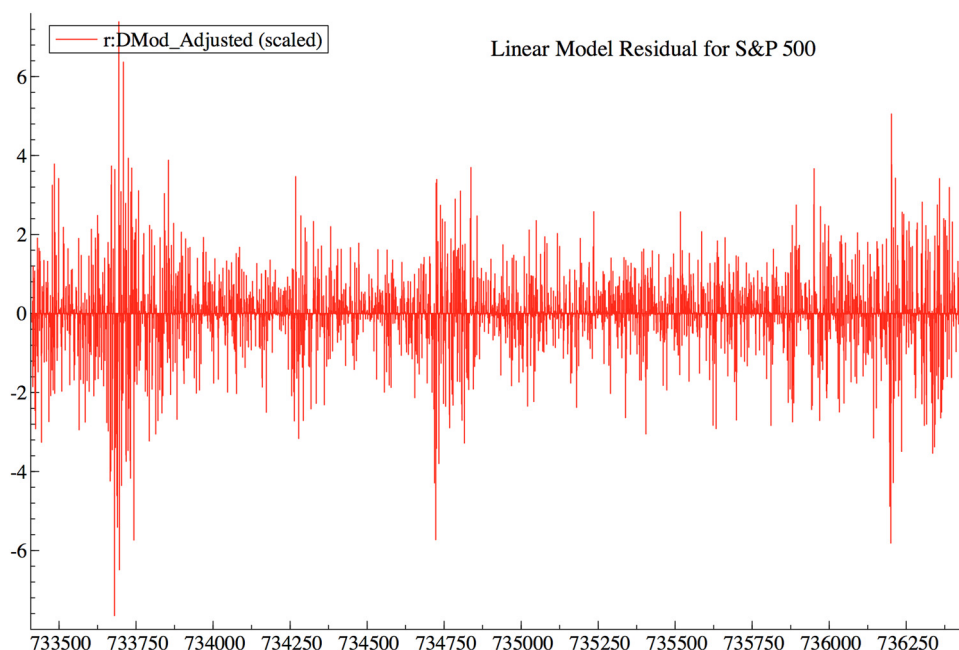


Figure A2: Residuals from the ADL linear predictive model for FTSE 100 returns.

Table A3: FTSE 100 returns: GARCH model estimates

Modelling	std_rel_pos	by	restricted	GARCH(1,0)		
				robust-SE	t-value	t-prob
adj_rel_pos	X	1	0.00551	0	-.inf	0.000
alpha_0	H	0.082809	0.002253	0.018657	4.44	0.000
alpha_1	H	0.349394	0.0478	0.1266	2.76	0.006
Modelling	std_rel_act	by	restricted	GARCH(1,1)		
				robust-SE	t-value	t-prob
adj_rel_act	X	1	0.00866	0	+.inf	0.000
alpha_0	H	0.002089	0.000255	0.001293	1.62	0.106
alpha_1	H	0.028631	0.0027	0.01239	2.31	0.021
beta_1	H	0.967491	0.00261	0.01076	89.9	0.000
Modelling	std_rel_str	by	restricted	GARCH(1,1)		
				robust-SE	t-value	t-prob
adj_rel_str	X	1	0.004448	0	+.inf	0.000
alpha_0	H	0.001237	0.000144	0.001091	1.13	0.257
alpha_1	H	0.033111	0.003252	0.01155	2.87	0.004
beta_1	H	0.961003	0.003203	0.02043	47	0.000
Modelling	std_rel_aff	by	restricted	GARCH(1,0)		
				robust-SE	t-value	t-prob
adj_rel_aff	X	1	0.007232	0	+.inf	0.000
alpha_0	H	0.133714	0.003615	0.0179	7.47	0.000
alpha_1	H	0.228286	0.03251	0.07408	3.08	0.002

Modelling	D_Price	by	restricted	GARCH(1,1)		
		Coefficient	Std.Error	robust-SE	t-value	t-prob
alpha_0	H	37.0421	8.444	12.57	2.95	0.003
alpha_1	H	0.050623	0.006162	0.01045	4.84	0.000
alpha_1	H	0.937318	0.007247	0.01154	81.2	0.000

Note: adj_^* denotes the series of outliers $\text{Sem}_t \cdot 1_{\{\text{Sem}_t > \max(\{|r_t|\})\}}$ in (5.6).

A5 GARCH model for S&P 500 returns

The following table provides the estimates of the parameters of the GARCH model for the S&P 500 returns.

Table A4: S&P 500 returns: GARCH model estimates

Modelling	D_Price	by	restricted	GARCH(1,1)		
		Coefficient	Std.Error	robust-SE	t-value	t-prob
alpha_0	H	1.03×10^{-6}	2.20×10^{-7}	3.65×10^{-7}	2.81	0.005
alpha_1	H	0.072306	0.007532	0.0124	5.83	0
alpha_1	H	0.91821	0.007949	0.01291	71.1	0

Modelling	STD_REL_POS	by	restricted	GARCH(1,1)		
		Coefficient	Std.Error	robust-SE	t-value	t-prob
ADJ_REL_POS	X	1	0.01808	0	+.Inf	0
alpha_0	H	0.025963	0.003048	0.01459	1.78	0.075
alpha_1	H	0.089888	0.01005	0.02847	3.16	0.002
beta_1	H	0.878273	0.01113	0.03816	23	0

Modelling	STD_REL_ACT	by	restricted	GARCH(1,1)		
		Coefficient	Std.Error	robust-SE	t-value	t-prob
ADJ_REL_ACT	X	1	0.01598	0	+.Inf	0
alpha_0	H	0.027205	0.003211	0.008323	3.27	0.001
alpha_1	H	0.140739	0.01377	0.04142	3.4	0.001
beta_1	H	0.812659	0.01518	0.04097	19.8	0

Modelling	STD_REL_STR	by	restricted	GARCH(1,1)		
		Coefficient	Std.Error	robust-SE	t-value	t-prob
ADJ_REL_STR	X	1	0.05309	0	+.Inf	0
alpha_0	H	0.025642	0.002967	0.009229	2.78	0.005
alpha_1	H	0.106837	0.01016	0.02371	4.51	0
beta_1	H	0.868708	0.01018	0.02616	33.2	0

Modelling	STD_REL_AFF	by	restricted	GARCH(1,1)		
				robust-SE	t-value	t-prob
ADJ_REL_AFF	X	1	0.05168	0	+.Inf	0
alpha_0	H	0.010921	0.001602	0.003624	3.01	0.003
alpha_1	H	0.037582	0.00432	0.01187	3.17	0.002
beta_1	H	0.94712	0.005295	0.01253	75.6	0

Notes: adj_* denotes the series of outliers $\text{Sem}_t \cdot 1_{\{\text{Sem}_t > \max(\{r_t\})\}}$ in (5.6).

References

- [1] Abduraimova, K. (2019). Contagion and tail risk in complex financial networks. Ph.D. Thesis, Imperial College Business School.
- [2] Alberg, D., Shalita, H., & Yosef, R. (2008). Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics*, 18, 1201–1208.
- [3] Behrendt, S., & Schmidt, A. (2018). The Twitter myth revisited: Intraday investor sentiment, Twitter activity and individual-level stock return volatility. *Journal of Banking and Finance*, 96, 355–367.
- [4] Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2, 1–8.
- [5] Carnero, M., Pena, D., & Ruiz, E. (2012). Estimating GARCH volatility in the presence of outliers. *Economics Letters*, 114, 86–90.
- [6] Christoffersen, P. F. (2012). Elements of Financial Risk Management. 2nd edn., Waltham, MA: Academic Press.
- [7] Cont, R. (2001). Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, 1, 223–236.
- [8] Corea, F. (2016). Can twitter proxy the investors' sentiment? The case for the technology sector. *Big Data Research*, 4, 70–74.
- [9] Damasio, A. R. (1994). *Descartes' error: Emotion reason, and the human brain*. New York, NY: Putnam.
- [10] Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427–431.
- [11] Fama, E. F., Fischer, L., Jensen, M. C., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review*, 10, 1–21.
- [12] Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance*, 46, 1575–1617.
- [13] Harvard (2002). *General inquirer categories*. Cambridge, MA: Harvard University. <http://www.wjh.harvard.edu/inquirer>.
- [14] Gentry, J. (2015). *CRAN - Package twitterR*. <https://cran.r-project.org/web/packages/twitterR/index.html>.
- [15] Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with Applications*, 40, 6266–6282.
- [16] Groß-Klußmann, A., König, S., & Ebner, M. (2019). Buzzwords build momentum: Global financial Twitter sentiment and the aggregate stock market. *Expert Systems with Applications*, 136, 171–186.
- [17] Gruhl, D., Guha, R., Kumar, R., Novak, J., & Tomkins, A. (2005). The predictive power of online chatter. In: *KDD '05: Proceeding of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining* (pp. 78–87). New York, NY: ACM Press.
- [18] Gu, C., & Kurov, A. (2020). Informational role of social media: Evidence from Twitter sentiment. *Journal of Banking and Finance*, 121, 105969.
- [19] Hamilton, J. D. (1994). *Time series analysis*. Princeton, New Jersey: Princeton University Press.
- [20] Hanck, C., Arnold, M., Gerber, A., & Schmelzer, M. (2020). *Introduction to Econometrics in R*. <https://www.econometrics-with-r.org/index.html>
- [21] Ibragimov, R., & Müller, U. K. (2010). *t*-statistic based correlation and heterogeneity robust inference. *Journal of Business and Economic Statistics*, 28, 453–468.
- [22] Ibragimov, R., & Müller, U. K. (2016). Inference with few heterogeneous clusters. *Review of Economics and Statistics*, 98, 83–96.
- [23] Ibragimov, M., Ibragimov, R., & Walden, J. (2015). *Heavy-tailed distributions and robustness in economics and finance*. Vol. 214. Lecture notes in statistics. Heidelberg: Springer.
- [24] Liu, Y., Huang, X., An, A., & Yu, X. (2007). ARSA: A sentiment-aware model for predicting sales performance using blogs. In: *SIGIR '07: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 607–614). New York, NY: ACM.
- [25] Lo, A. W., Repin, D. V., & Steenbarger, B. N. (2005). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review*, 95, 352–359.

- [26] Mittal, A., & Goel, A. (2011). *Stock prediction using Twitter sentiment analysis. Working paper*. Stanford University. Available at <https://cs229.stanford.edu/proj2011/GoelMittal-StockMarketPredictionUsingTwitterSentimentAnalysis.pdf>.
- [27] Mishne, G., & de Rijke, M. (August 2006). Capturing global mood levels using blog posts. In: N. Nicolov, F. Salvetti, M. Liberman, & J. H. Martin (Eds.), *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs* (pp. 145–152). Menlo Park, CA: The AAAI Press/Stanford University.
- [28] Nofsinger, J. R. (2005). Social mood and financial economics. *Journal of Behaviour Finance*, 6, 144–160.
- [29] Patton A. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47, 527–556.
- [30] Ranco, G., Aleksovski, D., Caldarelli, G., Grcccar, M., & Mozeticcc, I. (2015). The effects of Twitter sentiment on stock price returns. *PLOS One*, 10(9), e0138441.
- [31] Stock, J. H., & Watson, M. W. (2019). *Introduction to econometrics, Global edition*. 4th ed., Harlow: Pearson.
- [32] Tetlock, P. (2010). All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies*, 24, 1481–1512.
- [33] Tetlock, P. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62, 1139–1168.
- [34] Washha, M., Qaroush, A., Mezghani, M., & Sedes, F. (2019). Unsupervised collective-based framework for dynamic retraining of supervised real-time spam tweets detection model. *Expert Systems with Applications*, 135, 129–152.
- [35] Zhang, W., & Skiena, S. (2010). Trading strategies to exploit blog and news sentiment. In: *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*. <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/viewFile/1529/1904>.