U.PORTO

FEP FACULDADE DE ECONOMIA
UNIVERSIDADE DO PORTO

THE MICROBLOGGING EFFECT ON STOCK MARKET: FORECASTING TRADING VOLUME AND SHORT-TERM ABNORMAL RETURNS USING SENTIMENT ANALYSIS

Pedro Tiago Pombinho Alvarilhão

Dissertation
Master in Management

Supervised by

Professor Miguel Augusto Gomes Sousa, PhD

Professor João Manuel Portela da Gama, PhD

2018/2019

Biographic Note

Pedro Tiago Pombinho Alvarilhão, born on 3rd May 1996, in Porto, has a Bachelor in Economics at School of Management and Economics of the University of Porto (FEP) completed in 2017. He decided to start a Master in management at FEP, within the QTEM Master Network Program, to complement his academic education.

Pedro's academic path is characterized by his international exposure. During both programmes, he has been in Tampere, Nice and Oslo which he considers to be some of the richest experiences he has ever had.

In June 2018, Pedro was awarded with Rotary Club Porto-Foz / Vida económica Award attributed to the best student of the 1st year of the master in management at School of Management and Economics of the University of Porto.

Regarding his professional experience, Pedro was a Summer Intern at McKinsey during the 2018 summer in Lisbon where he was integrated on a digital transformation project. In addition, he did a curricular internship at CaetanoBus as Market Analyst in the beginning of 2019 being involved in zero-emission transport projects.

Acknowledgements

The present study is dedicated to those who crossed their path with mine until this moment and, in some type of form, taught and guided me to the present stage of my life.

An eternal "Thank you" for those that stay with me along this journey.

Abstract

This study's objective is to test the hypothesis that information regarding a company spreads faster than it is incorporated by the market on its fundamental stock price. The aim is to measure the effect of a "spike" of mentions on Twitter regarding a certain company in its stock trade volume and price. Contrary to other studies that use a daily or higher periodicity of sentiment analysis and forecast, this study looks to very short-term periods, as a "spike" is measured in 10-minute periods. This research uses Twitter posts containing the stocks tag of eight NASDAQ100 companies during a period of 79 days to measure those "spike" periods. The findings confirm a positive effect of microblogging posting on the traded volume, both on the period in which it occurs and on the following period as mentioned in previous studies. Additionally, the findings show that "spikes" with a positive content tend to have a positive and immediate effect (abnormal returns) on prices. However, the results suggest that an eventual overreaction or underreaction to new information by the market is corrected on the following ten minutes.

Keywords: forecasting techniques, stock market, abnormal stock returns, microblogging, sentiment analysis;

JEL-Codes: G17, G14;

Resumo

O objetivo deste estudo é testar a hipótese que a informação acerca uma dada empresa é disseminada mais rápido do que é incorporada pelos mercados financeiros no valor da empresa. O objetivo é quantificar o efeito de um "pico" de menções no twitter acerca de uma empresa no seu volume de troca e preço nos mercados de ações. Ao contrário dos outros estudos que utilizam periodicidades diárias ou superiores ao utilizarem análise de sentimentos, este estudo pretende estudar os períodos muito curtos dado que os picos são identificados em períodos de 10 minutos. Este estudo utiliza posts do Twitter the contêm o hashtag do código da ação de 8 empresas que pertencem ao NASDAQ100 durante o período de 79 dias para identificar os picos. Os resultados confirmam um efeito positivo do volume de *posts* online no volume de trocas, tanto no período em que o pico ocorre bem como seguinte. Além disso, os resultados demonstram que os picos classificados com conteúdo positivo possuem um efeito imediato na evolução do preço desse dado stock quando comparado ao preço do índice do mercado (existe um retorno anormal positivo). No entanto, os resultados sugerem que existe uma reacção exagerada pelos mercados face a picos negativos e uma reacção contida face a picos positivos nos primeiros instantes que é corrigida logo no período seguinte.

Palavras-chave: técnicas de previsão, mercado de ações, retornos anormais de ações, *microblogging*, analise de sentimentos;
Códigos JEL: G17, G14;

Index

List of Tables

List of Figures

1. Introduction

This study intends to test the hypothesis that information regarding companies disseminates faster than financial markets incorporate it on the companies' stocks fundamental value. The present research analyses microblogging data collected from Twitter to identify spikes, which are characterized by a 10-minute period where the number of tweets about a company hits the 99th percentile of mentionings. Then the effects of the spikes identified are analysed on stocks' trading volume and price. In addition, this study uses sentiment analysis on the content of the tweets analysed to test if it is feasible to predict the direction of the price change. If an effect is found either on stocks' volume or on price after the spike (positive and/or negative based on the sentiment analysis) occurs and if there is a way to quantify it, traders/investors with resources to identify the spike phenomenon in time can obtain abnormal stock returns in short periods of time.

The objective of this research is to develop further the current state of scientific knowledge regarding monetization of social media information in the very short-term (Tafti, Zotti, & Jank, 2016) since there is no prior research on this topic that uses sentiment analysis on periods of 10 minutes of microblogging data to forecast abnormal returns. To test the hypothesis that information is spread faster than markets incorporate it, public and free dataset containing all the tweets mentioning the stock's tag of the NASDAQ100 during a period of 79 days from 2016 March 28th to 2016 June 15th and the intraday stock prices of those companies in period of 10 minutes were used. The data processing and handling was done in Python due to the fact that its resources and libraries makes it possible to perform the data manipulation necessary and the sentiment analysis. With the date, hour and content, the 99th percentile periods where the company name was heavily mentioned ("spikes") were identified and the sentiment analysis of each tweet was computed.

This study aims to answer the following research questions:

1) Is it possible to forecast trading volume based on microblogging activity? Is it possible to quantify it?
2) Is it possible to predict how the stock's price behaves based on microblogging activity and sentiment analysis? Is it possible to quantify this effect?

The findings regarding trading volume showed a positive relationship between an abnormal microblogging posting volume and traded volume in the moment in which it occurs and in the following 10-minute period. Additionally, findings suggest a positive effect of “spikes” with positive sentiment on abnormal returns and proved that negative and positive “spikes” have different effects on abnormal returns.

Besides this section, this study is structured as follows: in Section 2, a literature review of the topic is made. Section 3 we will go over the methodological aspects, while Section 4 contains the descriptive analysis of the samples. Section 5 presents the results of the analysis, its discution and the critical interpretation of the results and a paralelism with the actual literature. Finally, Section 6 concludes and presents the main limitations of this study.

2. Literature review

In this section the main concepts of this study and their definitions according to the literature are presented. Afterwards, similar studies will be discussed and the relationship between them and this study is established. In the end, a critical analysis of the literature reviewed is presented.

2.1. Previous Studies

Studies on the predictive value of information gathered from social media are relatively new naturally due to the early age of digital platforms. The information used in this studies is considered microblogging[1] data and this form of blogging distinguished itself by the brevity of its content associated with platforms developed for it such as Twitter and Tumblr (Kte'pi, 2018). Its general text space is limited to 140 words and often has the characteristic of sparse features for its short text which leads to many errors and uncertainties in its processing (Li & Shen, 2017). In short, microblogging is a new form of communication where users share information and their interests in short posts (Jansen, Zhang, Sobel, & Chowdury, 2009), with their own expressions, features and language when compared to traditional text.

The first studies that looked to intraday market reactions used sentiment analysis[2] and news as sentiment's input (Groß-Klußmann & Hautsch, 2011; Schumaker, Zhang, Huang, & Chen, 2012). They have concluded that the intraday news' flow had an effect on volatility, volume and value traded (Groß-Klußmann & Hautsch, 2011) and that subjective news and articles with negative sentiment were easier to predict the stock price's reaction (Schumaker et al., 2012). The sentiment analysis' methodology consisted in sentiment analysis product (Groß-Klußmann & Hautsch, 2011) and generic lexicon dictionary (Schumaker et al., 2012).

More recently, several studies used microblogs as sentiment's data source, but its analysis' periodicity was daily. The first studies did not found evidence of effects from sentiment analysis on the returns but concluded that posting volume and sentiment were an important predictor of trading-related activities, such as two-day pump and two-day dump manipulation, and could improve daily trading volume forecasts (Oliveira, Cortez, & Areal,

[1] Several definitions of microblogging can be found in Appendix 1
[2] Several definitions of sentiment analysis can be found in Appendix 2

2013; Sabherwal, Sarkar, & Zhang, 2011; Sprenger, Tumasjan, Sandner, & Welpe, 2014). Other studies were able to conclude that tweets' sentiment can be associated with daily abnormal returns[3], market indexes and beat random investment strategies (Nasseri, Tucker, & de Cesare, 2015; Sprenger et al., 2014), which suggest that Twitter and StockTwits' data have valuable information for financial markets. The majority of the studies used supervised machine learning to perform their sentiment analysis (Nasseri et al., 2015; Oliveira et al., 2013; Sabherwal et al., 2011; Sprenger et al., 2014) but microblog financial lexicon was also used (Oliveira, Cortez, & Areal, 2016);(Oliveira, Cortez, & Areal, 2017).

With regard to intraday analysis using microblogs as a sentiment's source, Tafti et al. (2016) found a positive relationship between Twitter activity and traded volume in spike's following 40 minutes. However, they were not able to use this information to obtain monetary gains. The exploitation of tweets' sentiment data with the objective of monetization in the short period is suggested on their study as an opportunity for future research. The research gap identified is the one being developed further within this dissertation.

2.2. Critical analysis of the literature reviewed

Stock market returns can forecasted using sentiment analysis to some extent, however, the periodicity of studies using microblogging data is mostly daily or bigger due to the struggle to apply sentiment analysis to intraday content on certain specific topics like company's performance or stock. In this study the objective is not only to test previous findings related with the microblogging effect on traded volume but also to study abnormal returns behavior based on the sentiment analysis of the content disseminated globally within a period of less than 30 minutes. This study intends to go beyond the current state of knowledge and forecast and test the possibility of obtain monetary gains in very short-term periods.

[3] Several definitions of Abnormal Returns can be found in Appendix 3

3. Methodology

In this section the different phases of the study, the data collection process and treatment are detailed.

3.1. Data collection, Databases and Samples

For the sentiment analysis a public database available by www.followthehastag.com was used. This dataset is available freely on their website and contains all tweets mentioning any NASDAQ 100 Twitter Symbol, company by company in individual datasets. The information of the datasets is from 2016 March 28th to 2016 June 15th, which represents a total of 79 days. It contains about a million tweets that are related with the 100 companies during that period and has the day, time and content of each individual tweet. After a cut off due to the lack of availability of intraday stock prices data, the final sample consists of eight companies and a total of 283,036 tweets.

For the stock market information of the same period the website Finam.ru was used. This Russian website allows the extraction of intraday price information from U.S. Stock Market (BATS) and contrary to the majority of free service providers of intraday stock prices which provide information only from past week at maximum, Finam allows the extraction of this information from several months ago which enables the matching with dataset used to do the sentiment analysis from 2016. Due to the limited resources of data over the microblogging sample period, the final sample only includes eight companies: Adobe Systems Inc., Alphabet (Google), Apple Inc., Applied Materials Inc., CA Technologies Inc., Cisco Systems Inc., Microsoft Corp., and Yahoo Inc. For six of those companies there is a completed dataset (2,223 data entries) and for the remaining two despite not being complete, they still have a significant amount of entries that enables the hypothesis to be tested(2,222 for CA Technologies and 2,203 for Alphabet).

Table 1 presents the total number of tweets for each company during the period studied.

Table 1 – Datasets' characteristics regarding the companies featured in the study

Company	Number of Tweets
Adobe Systems Inc.	2,937
Alphabet (Google)	37,642
Apple Inc.	166,631
Applied Materials Inc.	3,490
CA Technologies Inc.	1,316
Cisco Systems Inc	13,283
Microsoft Corp.	43,060
Yahoo Inc.	17,614
Total:	283,036

Source: author own elaboration

3.2. Data Treatment

3.2.1 Microblogging Data

Spike Identification

Firstly, the microblogging data was filtered for weekdays from 10:00 AM to 4:00 PM in spite of the stock market's operating hours being from 9:30 AM to 4:00 PM. The purpose of this selection is to avoid the abnormal trade volume from the first thirty minutes of activity from stock markets

Afterwards, the database was divided in periods of 10 minutes and by using a count function it enabled to have knowlegde of how many tweets existed for each company in every 10-minute period. Then a "spike period" was identified as a 10-minute period in which the number of tweets was higher than the 99th percentile of each company.

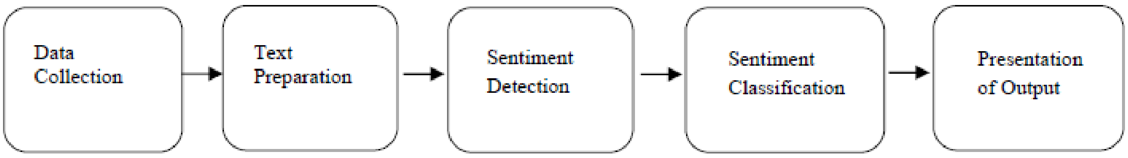
Sentiment Analysis

The database content is in its raw form, meaning that it contains the actual messages of the tweets on it. Therefore, a sentiment analysis was conducted directly on the data gathered.

Sentiment analysis is a data mining technique which extracts content of interest from textual data using different features (Rambocas & Gama, 2013). This analysis enables the extraction of feelings, attitudes and emotions towards a subject (Liu, 2010). It is an automated analysis using as resource machine language and computer programing (Akshi Kumar & Sebastian, 2012).

The sentiment analysis in this research was done using Python. The simple interface and Textblob package enabled to perform this analysis following the five phases suggested by Rambocas and Gama (2013) of the sentiment analysis process which is presented in Figure 1.

Figure 1 – Sentiment Analysis process



Source: (Rambocas & Gama, 2013)

The first phase, Data Collection, consists in the collection of the tweets that are going to be analysed. The next phase, Text Preparation, involves the treatment and cleaning of the dataset: The non-textual content (for example, the hashtags, points, and commas) is eliminated as is also the tweets whose content is not relevant for the course of the study (for example, spamming messages). In Sentiment Detection the dataset continues to be filtered based on the tweet's content. The objective is to help in the analysis stage and increase the precision of the sentiment classification by removing the tweets transmitting facts and objective communication. On the fourth phase, Sentiment Classification, each tweet is classified as a negative, neutral or positive message towards the company in specific. To compute the classification a general lexicon dictionary was chosed to be used because lexicons establish the connection between language and knowledge transmitted by that language (Rambocas & Gama, 2013). This classification output will ultimately be used as input for the segmentation of spike's effects on companies' stock.

3.2.2 Financial Data

Volume normalization

To allow the volume's comparison between the eight different companies the volume traded by period was transformed into the normalized volume. This was obtained by using the usual normalization function detailed below.

$$VNorm_{(t;i)} = (Volume_{(t;i)} - Volume_{(mean;i)}) \,/\, Volume_{(StdDev;i)} \qquad (1)$$

Where:

$VNorm_{(t;i)}$= Normalized volume from firm i on the period t

$Volume_{(t;i)}$ = Volume transacted by firm i on the period t

$Volume_{(mean;i)}$ = Average volume transacted by firm i

$Volume_{(StdDev;i)}$= Standard deviation of the volume transacted by firm i

Abnormal Returns

Research on abnormal returns (AR) has been done previously, but was mostly focussed on the long-run perspective (Barber & Lyon, 1997; Lyon, Barber, & Tsai, 1999). Despite the difference between this study – short-term analysis - and previous studies – long-run analysis -, the previous AR definition remains up to date.

The AR should be treated as the difference between a firm buy-and-hold return and a reference portfolio or a control company buy-and-hold return (Barber & Lyon, 1997). This second acts as the expected return for the firm buy-and-hold return (Lyon et al., 1999). The AR calculation can be written as the following expression:

$$AR_{(i,t)} = R_{(i,t)} - E(R_{(i,t)}) \qquad (2)$$

Where:

$AR_{(i,t)}$ – Is the abnormal return from firm i on period t;

$R_{(i,t)}$ – is the buy-and-hold return from firm i on period t;

$E(R_{(i,t)})$ – is the expected buy-and-hold return from firm i on period t.

The abnormal return should oscillate around zero since it is the difference between the effective return of firm i and its own expected return. If that happens not to be the case, it means that the event influences stock price (MacKinlay, 1997), either on a positive or on a negative way.

According to the market model, the expected return of one stock can be written as following:

$$E\left(R_{(i,t)}\right) = \alpha_{(i)} + \beta_{(i)} R_{(m,t)} \qquad (3)$$

Where:

$E\big(R_{(i,t)}\big)$ - expected return of the share of acquiring firm i on period t;

$\alpha_{(i)}$- intercept, measure of the average return of firm i stock that is not explained by the market return;

$\beta_{(i)}$ - coefficient or slope, measure of the sensibility of the volatility of firm i stock towards the market volatility;

$R_{(m,t)}$- return of the market index on period t.

Replacing the previous function on the expression number 2 the abnormal returns calculation can be written as:

$$AR_{(i,t)} = R_{(i,t)} - (\alpha_{(i)} + \beta_i R_{(m,t)}) \tag{4}$$

Where:

$AR_{(i,t)}$- abnormal return of the share of acquiring firm i for period t;

$R_{(i,t)}$ - observed or actual return of the share of acquiring firm i for period t;

$\alpha_{(i)}$- intercept, measure of the average return of firm i stock that is not explained by the market return;

β_i- coefficient or slope, measure of the sensibility of the volatility of firm i stock towards the market volatility;

$R_{(m,t)}$- return of the market index on period t.

The abnormal returns were calculated using the Market-adjusted model (MAM). The MAM is a model with restrictions such as asset returns being independent and identically distributed throught time (MacKinlay, 1997) but is simpler because it does not adjust the market's stock price to the firm i. As a result, the expected return for the company i is the market return ($\alpha = 0$ and $\beta = 1$).

In conclusion, the expression for the Abnormal Returns ends up looking like this:

$$AR_{(i,t)} = R_{(i,t)} - R_{(m,t)} \tag{5}$$

Where:

$AR_{(i,t)}$ - abnormal return of the share of acquiring firm i for period t;

$R_{(i,t)}$ - observed or actual return of the share of acquiring firm i for period t;

$R_{(m,t)}$ - return of the market index on period t.

For the market index return, the S&P500 index was choosen since it is usually used by investors as benchmark. Additionally, this index enables gathering of a complete intraday sample for this study analysis' period.

The returns were calculated as:

$$R_{(i,t)} = \ln(CP_{(i,t)}) - \ln(OP_{(i,t)}) \tag{6}$$

$$R_{(m,t)} = \ln(ICP_{(t)}) - \ln(IOP_{(t)}) \tag{7}$$

Where:

CP – market closing price of the share on period t of firm i;

OP – market opening price of the share on period t of firm i;

ICP – market closing price of the market index I on period t;

IOP – market opening price of the market index I on period t.

The abnormal returns' effects were analysed across three different periods: (T-1, T), (T) and (T, T+1) where T is the period where the spike materializes. For instance, (T-1, T) means the return since the beginning of the 10-minute period before the "spike" (T-1) and the end of the 10-minute period when the spike occurred (T).

3.3. Data analyses

3.3.1 Differences-in-Differences

To avoid autocorrelation problems due to the data gathered being in panel, it was opted to study the "spike" effect on normalized volume using a differences-in-differences (DiD) methodology. As explained before, it is not only studied the effect on the period where the "spike" occurred but also the following ten-minute period.

To execute this methodology, a comparable control group was create to compare it with the "spike" group (treatment group). The treatment-pair match was made using the normalized volume from the spikes' previous periods of ten minutes. All observations selected for the control group were from the same company as the treatment and they did not had a "spike" on the previous or following nine periods to avoid contamination from the control group.

From the treatment group the spikes identified at 10:00 AM and at 3:50 PM were removed because either the 'previous period' or the "following period" used to select the pair would have occurred during the first 30 minute of each day at the stock market (an abnormal trade period, as explained before).

The Differences-in-Differences model used in the present research is represented below:

$$VolumeNorm_{i,t} = \beta_0 + \beta_1 * Treat + \beta_2 * Period + \beta_3 * Period * Treat + \mu_i + \varepsilon_{i,t} \tag{8}$$

Where:

$VolumeNorm_{i,t}$ – Volume Normalized from company i on period t;

$Treat$ – Dummy variable: 0 – control group, 1 – treatment group;

Period – Dummy variable: 0 – period 1, 1 – period 2

μ_i – fixed effect from company i;

ε_t – noise on period t;

In this model the fixed effect from each company was clustered in order to make the analysis sounder. This studies' variable of interest is β_3 which is the estimator of spike's effect on volume traded.

3.3.2 Abnormal returns significance tests

To test the significance of the abnormal returns both a parametric test and a non-parametric test were used. Besides complementing each other, the use of the two different tests allow a better and well-rounded interpretation of the results.

Additionally, the sentiment analysis made possible the spikes' segmentation by their sentiment (negative, neutral or positive). Afterwards, these segmentations were used to compare both the mean ($H_{0:}$: $\mu_i = \mu_j$) using the T-test and the median ($H_{0:}$: $med_i = med_j$) using the Wilcoxon test between samples.

Parametric test: T-test

Assuming the abnormal returns have a normal distribution and are independently and identically distributed over time the test statistics for the null hypothesis, this is, the sample abnormal returns' average equals to zero, has a t-Student distribution (Brown & Warner, 1985).

The rejection of the null hypothesis (H0: Abnormal Returns' Average [ARA] = 0) verifies an impact of the spike abnormal returns for investors and shareholders.

$$ARA \sim N(0, \sigma) \tag{9}$$

$$T_{stat} = \frac{ARA}{\hat{S}(AR)} \tag{10}$$

$$\hat{S}(AR) = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (AR_i - ARA)^2} \tag{11}$$

Where:

T_{stat} - t-student test statistic with $n-2$ degrees of freedom for the market model;

ARA - abnormal returns' average;

AR_i - abnormal return of spike i, i = 1, ..., N;

N - total number of spikes with available abnormal returns;

$\hat{S}(AR)$ - standard deviation of AR, an unbiased estimator of standard deviation of population (O).

Non-parametric test: Wilcoxon signed ranked-test

The non-parametric test chosen is the Wilcoxon signed-rank test (Wilcoxon, 1945) and is used to test the significance of the AR – the variable of analysis. This test considers both the sign and the magnitude of abnormal returns as fundamental (Serra, 2004) and therefore considers as the null hypothesis median equals to zero ($H_{0:}$ median = 0). The Wilcoxon signed rank is not affected by outliers and assumes the test statistics is like a normal distribution when the number of observations is large.

$$z \sim N(0, \sigma) \tag{12}$$

$$E(z) = \frac{N(N+1)}{4} \tag{13}$$

$$\sigma^2(z) = \frac{N(N+1)(2N+1)}{24} \qquad (14)$$

$$z = \sum_{i=1}^{N} r_i \qquad (15)$$

Where:

z - statistics for the Mann-Whitney/Wilcoxon rank test;

r_i - abnormal return for the observation i;

N- number of spikes in the sample.

4. Descriptive Analysis

In this chapter the descriptive statistics will be presented. The number of spike periods and its distribution by company, day of the week and by hour, the comparison between the statistics of the treatment periods and the control periods determined for the normalized volume analysis and finally the abnormal returns' descriptive statistics of the spikes on the durations previously determined, i.e., T-1 to T, T to T and T to T+1 when a spike happens on period T are described below.

4.1. Spike identification

A total of 141 spikes was identified on the data collected. Only one out of the eight companies had less than 15 spikes (CA Technology Inc. with 11 spikes) due to the low sample of tweets gathered which caused the company's 99^{th} percentile identified to be equal to several period's number of tweets and therefore not classified as "spike". Their exact distribution is detailed in the Table 2. Cisco Systems Inc was the company with the biggest number of spikes identified – 21.

Table 2 – Spikes identified by company

Company	Number of Spikes
Adobe Systems Inc.	16
Alphabet (Google)	16
Apple Inc.	19
Applied Materials Inc.	19
CA Technology Inc.	11
Cisco Systems Inc	21
Microsoft Corp.	20
Yahoo Inc.	19
Total:	141

Source: author own elaboration

According to the sample, spikes tend to be more frequent at the end of the week. Thursdays and Fridays were the days with the most spikes identified with 24,8% (35) and 28,4% (40) of the spikes respectively. Table 3 shows the distribution of the spikes identified in the dataset over the days of the week.

Table 3 – Spikes identified by day of the week

Day of the week	Number of Spikes Identified
Monday	30
Tuesday	15
Wednesday	21
Thursday	35
Friday	40
Total:	141

Source: author own elaboration

The spike's distribution during the day is incremental until 2:00 PM where it reaches its peak. Afterwards the number of spikes decreases nearly to half of its max. The first two hours of the day studied have the lowest number of spikes identified. The exact distribution is described in Table 4.

Table 4 – Spikes identified by hour of the day

Hour	Number of Spikes Identified
10 AM – 11 AM	6
11 AM – 12 AM	9
12 AM – 1 PM	22
1 PM – 2 PM	50
2 PM – 3 PM	26
3 PM – 4 PM	28
Total:	141

Source: author own elaboration

4.2. Treatment and Control group

Table 5 exhibits the differences between the treatment and the control group. Each group is composed by 134 observations (141 minus the two spikes at 10:00 AM and five spikes at 3:50 PM).

The pairing has been done based on the normalized volume of the period previous to the spike. Despite of the difference of 0.002 between the means, this difference was not significative (*p-value* = 0.519)and therefore the trade volume of the spike's previous period is similar on both treatment and control groups.

Table 5 – Descriptive statistics from DiD sample

Variable: Normalized Volume	**Treatment Group (n = 134)**			**Control Group (n = 134)**		
Period	Previous	"Spike"	Following	Previous	"Spike"	Following
Mean	0.376	0.407	0.537	0.374	0.132	0.232
Std. Deviation	1.175	1.168	1.627	1.180	0.890	1.217
Median	0.043	0.092	0.117	0.047	-0.092	-0.102

Source: author own elaboration

By analysing the period where the spike occurred, interesting differences between both groups can be observed. While the mean and median on the treatment group increase from previous period to the spike period, the control group has a different behavior. The behaviour from the spike period to the following period is similar for both groups' mean - an increase around 0.1 on traded volume.

Comparing the traded volume between both groups in the spike's following period some discrepancy can be observed which suggests that the spike had a positive effect on traded volume both on the period it occurs and in the following period, as expected by taking into consideration the literature review(Tafti et al., 2016).

4.3. Abnormal Returns

Table 6 summarises the statistics regarding the abnormal returns for the periods of analysis. The period (T-1, T) has 139 observations due to the two spikes that occurred at 10:00 AM and the period (T, T+1) has 136 observations due to the five spikes at 3:50 PM. As expected, the period (T) is composed by the whole sample of spikes identified.

Table 6 – Abnormal Returns descriptive statistics from the different periods of analysis

Period	(T-1, T)	(T)	(T, T+1)
Number of observations	139	141	136
Mean	0.000232	0.000064	0.000195
Median	0.000378	0.000077	0.000226

$*p < .1$. $**p < .05$. $***p < .01$.

Source: author own elaboration

T-tests were runned for each sample ($H_0: Mean = 0$) but significant differences were not found. The Wilcoxon signed-rank test ($H_0: Median = 0$) was performed and the same results were obtained. However, the period (T-1, T) was very close to show statistical significance with a *p-value* of 0.1043 in the Wilcoxon signed-rank test. In addition, it is interesting that both the abnormal returns' mean and median from the period (T-1, T) are higher than the ones from (T, T+1).

The results obtained in this descriptive analysis might suggest that the market is reacting to the spike even before it happens on Twitter and therefore signal some level of inside information from the big players on the spike's prior period of trade. Despite this assumption, the findings can also indicate that reacting to the spike when it happens might not be as lucrative as anticipating the spike. If anticipation to the spike was possible it could represent the existence of privileged inside information regarding a company or event in the market.

Although the values from period (T) appear rather small when compared with the other two periods, this period represents half of other periods' length in time which might be taken into account. Therefore, a more accurate measure of comparison would be using the expression: $1 + Ret(T))^2 - 1$ because period (T) has half the duration from the other two periods . However, this expression did not truly represented the return obtained in this period and hence would envy or distort this analysis.

5. Results

In this section, the results achieved by the application of the different methodologies explained previously are presented and discussed. The objective is to answer the main questions of this dissertations:

1) Is it possible to forecast trading volume based on microblogging activity? Is it possible to quantify it?
2) Is it possible to predict how the stock's price behaves based on microblogging activity and sentiment analysis? Is it possible to quantify this effect?

5.1. Normalized Volume

Table 7 represents the results from the DiD model between spike's previous period and spike's period on normalized volume.

Table 7 – Group A's DiD model results

Variable	Coefficient Estimator	Standard Error
Constant	-0.0385	0.033
Treatment	0.0013	0.002
Period	-0.2419**	0.106
Period*Treatment	**0.2732*** **	**0.104**
Alphabet	0.4237***	0.000
Apple	-0.0893***	0.000
Applied Materials	1.7337***	0.000
CA Inc.	0.1130***	0.000
Cisco	0.2837***	0.000
Microsoft	0.5244***	0.000
Yahoo	0.1306***	0.000

*p < .1. **p < .05. ***p < .01. n = 536

Source: author own elaboration

The results show no significant differences between the treatment and the control group as mentioned before (treatment's *p-value* = 0.519). The variable of interest, β_3, has a coefficient estimator of 0.2732 (Period*Treatment *p-value* = 0.009) and therefore one can mention that the spike has a positive effect on the trading volume in the period it materializes with a 99 percentage of confidence level. This means that when the number of comments about a company's stock during a 10-minute period is very high (a spike occurs), the trade volume of the stock increases by 27% of its standard deviation.

On the second regression, similar results were found. The spike's effect on trade volume in the following 10-minute period of its occurrence can be observed in Table 8. In the spike's following period, it is expected the normalized volume to be higher than 30% of its standard deviation when compared to its expected volume traded when it does not occur.

Table 8 – Group B's DiD model results

Variable	Coefficient Estimator	Standard Error
Constant	-0.0249	0.033
Treatment	0.0013	0.002
Period	-0.1428	0.098
Period*Treatment	**0.3042****	**0.132**
Alphabet	0.5054***	0.000
Apple	-0.1130***	0.000
Applied Materials	1.8252***	0.000
CA Inc.	0.0150***	0.000
Cisco	0.1737***	0.000
Microsoft	0.5161***	0.000
Yahoo	-0.0685***	0.000

*$p < .1$. **$p < .05$. ***$p < .01$. n = 536

Source: author own elaboration

The coefficient Period*Treatment (0.3042) of the following 10-minutes is higher than the coefficient Period*Treatment of the previous 10-minutes when compared to the spike's period (0.2732).This might be explained by the expected reaction on the volume traded both when the spike occurs and in the following periods as the information is disseminated to additional market players. The following period also presents a lower level of statistical confidence (Period*Treatment *p-value* = 0.021 with the 95% confidence interval of [0.046; 0.562]) when compared with the previous period (Period*Treatment *p-value* = 0.009 with the 95% confidence interval of [0.069; 0.478]).

Traded volume findings of of trade volume increase on the spike's period and on spike's following period are in line with the ones from Groß-Klußmann and Hautsch (2011);

Oliveira et al. (2013, 2017); Sabherwal et al. (2011); Sprenger et al. (2014) and Tafti et al. (2016), more specifically with findings of Groß-Klußmann and Hautsch (2011) and Tafti et al. (2016) where the authors found a positive effect from spikes on the traded volume in the spike's following 40-minute period.

5.2. Abnormal Returns by market sentiment

5.2.1 Period (T)

Table 9 contains the mean and median abnormal returns for the period (T) according to the market sentiment analysis.

Table 9 – Abnormal Returns descriptive statistics during period (T)

Type of sentiment	Negative	Neutral	Positive	Total
Number of observations	19	17	105	141
Mean	-0.000468	-0.000058	0.000180	0.000064
Median	-0.000607	-0.000111	0.000176	0.000077

**p* < .1. ***p* < .05. ****p* < .01.

Source: author own elaboration

When performing the Wilcoxon signed rank test ($H_0: Median = 0$), despite the fact of not being statistic significant both the negative and the positive sample presented a *p-value* around 0.15.

It is important to remember that the values above represent the returns in a 10-minute period and therefore, the mean abnormal return of 0.000180% of positive spikes represents a daily return of 2.625%.

The first descriptive analysis on abnormal returns did not bring significant insights or contributions for this study. Nonetheless, when the analysis with the segmentation by spike's sentiment was done, the interpretations on all periods of analysis became richer.

Although no statistical significance was found on the period (T), the mean and the median values can be interpreted. As expected, the abnormal return is negative when the sentiment of the spike is identified as negative and becomes positive when the spike's sentiment is positive as well. Neutral spike's abnormal returns are between the values from the other two groups which is in line with the theory. This raises the possibility that positive and negative spikes have a different effect on their periods' expected abnormal return.

5.2.2 Period (T-1, T)

By analysing the abnormal returns since the beginning of the 10-minute period before the spike until the end of the 10-minute period when the spike occurs (T-1, T), the results became more interesting. Significant results on the positive spike's abnormal returns were found when performing both the t-test and the Wilcoxon signed-rank test as detailed in Table 10.

Table 10 – Abnormal Returns descriptive statistics during period (T-1, T)

Type of sentiment	Negative	Neutral	Positive	Total
Number of observations	19	16	104	139
Mean	-0,000297	-0,000237	0,000400*	0.000232
Median	-0,000133	-0,000287	0.000503**	0.000378*

*p < .1. **p < .05. ***p < .01.

Source: author own elaboration

The results suggest that a company return will outperform the market return from the spike's previous period to the period where the spike occurs when its content is classified as positive (t-test *p-value* = 0.0955 and Wilcoxon *p-value* = 0.0369).

In this period, the increase on the mean and median abnormal return both on positive and negative segment needs to be noticed. This might indicate that previously to the spike, the market is outperformed by the specific stock. Regarding the positive spikes, it is understandable to consider the existence of privileged information to explain the data obtained. However, in terms of the negative spikes, that hypothesis does not hold because the stock outperforms the market in period T-1.This fact might also be a coincidence from the dataset.

5.2.3 Period (T, T+1)

On the period (T, T+1) the abnormal returns' resultsare described on Table 11 but interesting insights as the ones previously discovered have not been found.

Despite no statistically significant results have been found, the direction of the abnormal returns categorized by sentiment is according to the expected from theory.

Table 11 – Abnormal Returns descriptive statistics during period (T, T+1)

Type of sentiment	Negative	Neutral	Positive	Total
Number of observations	18	14	104	136
Mean	-0.000386	0.000379	0.000270	0.000195
Median	-0.000259	0.000569	0.000342	0.000226

**p* < .1. ***p* < .05. ****p* < .01

Source: author own elaboration

Comparing the values from period (T, T+1) with the abnormal returns from period (T) both the mean and median from negative and positive spikes increase. This might indicate that the reaction to the companies' stocks might suffer an overreaction when the spikes' content is negative and a conservative reaction by investors when the spikes' content is positive. The correction of this phenomenon appears to happen right on the following period of trade since the levels of significance found on period (T)'s positive spikes dissapear on period (T, T+1)'s positive spikes.

5.2.4 Negative Spikes vs Positive Spikes

Based on the previous results, the analysis and tests were extended on both the negative and positive sample of the three periods of analysis to evaluate if there are significative differences between them. The Table 12 shows the results of this analysis.

Table 12 – Abnormal Returns comparison between positive and negative spikes on the different periods of analysis

Period	(T-1, T)	(T)	(T, T+1)
Mean Difference	0.000697	0.000648	0.000656
Median Difference	0.000636*	0.000783*	0.000601

**p* < .1. ***p* < .05. ****p* < .01.

Source: author own elaboration

Although the t-test did not show any significant differences, the Wilcoxon ranked-test provided curious results.

The negative and positive abnormal returns samples' median from periods (T-1, T) (*p-value* = 0.0951) and (T) (*p-value* = 0.0538) is different with a confidence level of 90%. This result indicates that the spike's sentiment causes different effects on abnormal returns in a 10-minute period.

The interpretation of negative and positive samples' analysis indicate that investors might be able to take some advantage if they are able to anticipate the spike manifestation. Nonetheless, the interesting fact is that the significant effect disappears from period (T) and period (T, T+1). This suggests that the effect from the spike's sentiment segmentation on abnormal returns disappears right on the period after it happened. Therefore, it means the opportunity window to exploit this effect is only of one period of ten minutes – the one where the spike occurs. Afterwards the market corrects an eventual overreaction or underreaction as previously stated which is in line with the conclusion from previous chapter.

6. Conclusions

This study's objective was to determine the microblogging impact on volume traded and abnormal returns in the short-term on financial markets.

This dissertation's results support the findings about microblogging effect on traded volume (Groß-Klußmann & Hautsch, 2011; Tafti et al., 2016) on the short-term. Additionally, it provides a step forward for the literature regarding the use of sentiment analysis on abnormal returns' studies, especially in the short-term.

A positive effect on the volume traded caused by the spike on the period it happens and in the following period was found. The dimension of this effect on the spike's period is an increase on the volume traded equal to 27% of its standard deviation. On the following period, the effect is higher and it is expected the volume traded to increase by 3⁄10 of its standard deviation.

Furthermore, the results on the abnormal returns brought interesting insights in spite of the spike sample being composed mostly by positive spikes. The company stock outperforms the market index when positive spikes occurs between the spike's previous period and its own period. This might suggest the existence of privilege information and that the key players of the market probably are reacting to the new information before it is made public. Looking at the situation from another perspective, the questions regarding the spike being the trigger for abnormal returns or the abnormal returns triggering the microblogging posting volume can be raised. In this perspective, the spike can be interpreted as a lagged event from the market and therefore information diffusion is slower than it is incorporated by the markets.

In addition, the expected abnormal return differs between positive and negative spikes. This can be concluded only from spike's preceding period and spike's period and within the period where the spike takes place. This possibly suggest that an eventual overreaction or underreaction to new information by the market is corrected in the following ten minutes.

Overall, the dissertation's findings support previous research regarding the microblogging effect on trading activity and brought intriguing insights about the use of sentiment analysis on the study of short-term abnormal return. This study went one step forward on the literature gap identified and the work done in this dissertation is relevant and might be the catalyst for new and additional research about the use of microblogging data in financial market for monetization in the short-term in the future.

6.1. Limitations and future research

The interpretation of this study's results should have into account its limitations.

Firstly, this study used free of charge resources which implied the use of data that can be considered outdated since it is from 3 years ago. This might be relevant in the current fast and volatile world, especially in terms of social networks and information's diffusion, which might implicate the results found in the study. A replication of this study with more actual data might be pertinent.

Secondly, The collection of tweets was restristed to the ones mentioning the stock cash-tag and not all tweets related with a specific company. The difficulty of this task due to the twitter post's unstructured content is understood but might be a point to further development.

In addition, the data gathered seems insufficient for this type of study, both regarding the number of companies (eight presented in this study) and the data's temporal length (79 days). A longer period of analysis would be beneficial, by allowing the identification of additional spikes and a more accurate identification of the percentile 99th. Additional companies in the sample would also be beneficial and could lead to studies about industries' sensibility to spikes or even specific companies.

Lastly, a general lexicon dictionary was used when performing the sentiment analysis. The use of a microblogging financial lexicon would be more pertinent because the same word can have different meanings according to the topic it is inserted, as example, "unexpected" has a positive connotation in cinema language and negative connotation in financial language while the word "blasting" has a positive connotation in financial language and a negative connotation in the general day-to-day language. The use of microblogging financial lexicon would make the sentiment analysis' results more robust because it is in line with the type of content extracted and used as the primary resource. The work done by Oliveira et al. (2016) has created a finantial lexicon dictionary that can be exploited as a resource in the future.

6.2. Theoretical contributions

This dissertation brings essentially three theoretical contributions.

Firstly, the results strengthen recent research done on the impact of microblogging data on the forecast of traded volume in the short-term period.

Secondly, new knowledge and information for future research has been developed about the microblogging data's forecasting capacity in the short-term with sentiment analysis' support. The spike's effect on abnormal return is different between negative and positive spikes on period (T-1, T) and period (T). Simultaneously, the dilution of this effect on period (T, T+1) can be observed.

Lastly, suggestions for further and deeper research on microblogging data predictive power of abnormal return were made and the study also raised new hypothesis and questions for further research.

7. References

Akshi Kumar, & Sebastian, T. M. (2012). Sentiment Analysis: A Perspective on its Past, Present and Future. *IJISA*, 4, 1-14.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, 43(3), 341-372.

Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics*, 14(1), 3-31.

Grace, J. H., Zhao, D., & boyd, d. (2010). Microblogging: what and how can we learn from it? Paper presented at the *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, Atlanta, Georgia, USA.

Groß-Klußmann, A., & Hautsch, N. (2011). When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance*, 18(2), 321-340.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *J. Am. Soc. Inf. Sci. Technol.*, 60(11), 2169-2188.

Kte'pi, B. M. A. (2018). Microblogging. In.

Li, Y., & Shen, B. (2017). Research on sentiment analysis of microblogging based on LSA and TF-IDF.

Liu, B. (2010). Sentiment Analysis and Subjectivity. Handbook of Natural Language Processing. 2nd ed Chapman and Hall: Florida.

Lyon, J. D., Barber, B. M., & Tsai, C. L. (1999). Improved methods for tests of long-run abnormal stock returns. *Journal of Finance*, 54(1), 165-201.

MacKinlay, A. C. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature*, 35(1), 13-39.

Nasseri, A. A., Tucker, A., & de Cesare, S. (2015). Quantifying StockTwits semantic terms' trading behavior in financial markets: An effective application of decision tree algorithms. *Expert Systems with Applications*, 42(23), 9192-9210.

Oliveira, N., Cortez, P., & Areal, N. (2013). On the Predictability of Stock Market Behavior Using StockTwits Sentiment and Posting Volume, Berlin, Heidelberg.

Oliveira, N., Cortez, P., & Areal, N. (2016). Stock market sentiment lexicon acquisition using microblogging data and statistical measures. *Decision Support Systems*, 85, 62-73.

Oliveira, N., Cortez, P., & Areal, N. (2017). The impact of microblogging data for stock market prediction: Using Twitter to predict returns, volatility, trading volume and survey sentiment indices. *Expert Systems with Applications*, 73, 125-144.

Rambocas, M., & Gama, J. (2013). Marketing Research: The Role of Sentiment Analysis. *Working Papers (FEP)* -- Universidade do Porto(489), 1-21.

Sabherwal, S., Sarkar, S. K., & Zhang, Y. (2011). Do Internet Stock Message Boards Influence Trading? Evidence from Heavily Discussed Stocks with No Fundamental News. *Journal of Business Finance & Accounting*, 38(9-10), 1209-1237.

Schumaker, R. P., Zhang, Y., Huang, C.-N., & Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, 53(3), 458-464.

Serra, A. P. S. F. M. (2004). Event Study Tests: A Brief Survey. *Gestão.Org-Revista Electrónica de Gestão Organizacional*, Vol.2, No.3, pp.248-255.

Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and Trades: the Information Content of Stock Microblogs. *European Financial Management*, 20(5), 926-957.

Tafti, A., Zotti, R., & Jank, W. (2016). Real-Time Diffusion of Information on Twitter and the Financial Markets. *Plos One*, 11(8), 16.

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometrics Bulletin,* 1(6), 80-83.

8. Appendixes

8.1. Appendix 1: Examples of Microblogging Definitions

Definitions	**Author (Year)**
Microblogging is a type of blogging distinguished by the brevity of its content, typically associated with blogging platforms specifically designed for such, the best known of which are Twitter and Tumblr, with Facebook's status update essentially serving the same purpose within a richer social media environment.	**(Kte'pi, 2018)**
Microblogging, the act of broadcasting short, real-time messages, is a relatively new communication practice allowing people to share information they are less likely to express using existing technologies (e.g. email, phone, IM or weblogs).	**Grace, Zhao, and boyd (2010)**
Microblogging is a broadcast medium that exists in the form of blogging. It is different from the traditional blog, for its general space is limited to 140 words. Compared with the traditional text, microblogging text often has the characteristic of sparse features for its short text content, which leads to many errors and uncertainties in the processing of microblogging text.	**Li and Shen (2017)**
Microblogging is a new form of communication in which users can describe things of interest and express attitudes that they are willing to share with others in short posts (i.e., microblogs).	**Jansen et al. (2009)**

Source: author own elaboration

8.2. Appendix 2: Examples of Sentiment Analysis definitions

Definitions	**Author (Year)**
Sentiment analysis is a data mining technique that uses natural language processing, computational linguistic and text analytics to identify and extract content of interest from a body of textual data.	**Rambocas and Gama (2013)**
Process that categorizes a body of textual information to determine feelings, attitude and emotions towards a particular issue or object.	**Liu (2010)**
Automated subjectivity analysis similar to opinion mining and appraisal extraction which focuses on extracting and classifying texts with machine language and computer programming	**Akshi Kumar and Sebastian (2012)**

Source: author own elaboration

8.3. Appendix 3: Examples of Abnormal Returns definitions

Definition	Formula	Author (Year)
The long-horizon buy-and-hold abnormal returns is the difference between the return for a security I in period T and the expected return for the same security I during period in question. Where ARi, is the T period buy-and-hold abnormal return for security i, Ri, is the r period buy-and-hold return, and E(Ri,) is the T period expected return for security i	$AR_{IT} = R_{IT} - E(R_{IT})$	**Lyon et al. (1999)**
Researchers should calculate abnormal returns as the simple buy-and-hold return on a sample firm less the simple buy-and-hold return on a reference portfolio or control firm. Where R_T is the month t simple return on a sample firm, $E(R_{IT})$ is the month t expected return for the sample firm and AR_{IT} is the abnormal return in month t.	$AR_{IT} = R_{IT} - E(R_{IT})$	**Barber and Lyon (1997)**

Source: author own elaboration

8.4. Appendix 4: Similar studies comparison: Aims and Findings

Author (Year)	Aim of the Study	Findings
Groß-Klußmann and Hautsch (2011)	Examine high-frequency market reactions to an intraday stock-specific news flow.	Found distinct responses in returns, volatility, trading volumes and bid-ask spreads due to news arrivals.
Schumaker et al. (2012)	Investigate if the choice of words and tone used by the authors of financial news articles correlate to measurable stock price movement and if the magnitude of price movement be predicted using these same variables	Found that subjective news articles were easier to predict in price direction obtaining 3.30% return on those articles. Found that articles with a negative sentiment were easiest to predict in price direction and obtain a 3.04% return on them.
Sabherwal et al. (2011)	Understand the effect of online postings on trading activities and reduce the error due to stocks with small message board followings.	Conclude that message board sentiment is an important predictor of trading-related activities.
Oliveira et al. (2013)	Explore data from StockTwits.	Found no evidence of return predictability using sentiment indicators, and of information content of posting volume for forecasting volatility. However, they found evidence that posting volume can improve the forecasts of trading volume.
Sprenger et al. (2014)	Explore whether and to what extent stock microblogs reflect and affect financial market developments.	The sentiment of tweets to be associated with abnormal stock returns and message volume to predict next-day trading volume.
Nasseri et al. (2015)	Predict an intelligent trading support mechanism to screen out the most significant and profitable trading terms or combination of terms from StockTwits data that may help investors to make correct and accurate (selling, buying or	Achieved a promising performance and the tweet term trading strategies dramatically outperform random investment strategies. Confirm that StockTwits postings contain valuable information and lead trading activities in capital markets.

	holding) decisions in capital markets.	
Oliveira et al. (2017)	Assess the value of microblogging data to the forecasting of stock market variables.	Found that Twitter sentiment and posting volume were relevant for the forecasting of returns of S&P 500 index, portfolios of lower market capitalization and some industries.

Source: author own elaboration

8.5. Appendix 5: Similar studies' methodology

Study (Date)	Sentiment's Data Source	Sentiment Analysis' Method	Financial Data Periodicity	Number of companies studied	Data length
Groß-Klußmann and Hautsch (2011)	News	Sentiment Analysis Product	Intraday	39	18 Months
Schumaker et al. (2012)	News	Generic Lexicon	Intraday	500 (S&P500)	23 Days
Sabherwal et al. (2011)	Message Boards	Supervised Machine Learning	Daily and Intraday	64	13 Months
Oliveira et al. (2013)	Microblogs	Supervised Machine Learning	Daily	5 companies and 1 index	28 Months
Sprenger et al. (2014)	Microblogs	Supervised Machine Learning	Daily	99	6 Months
Nasseri et al. (2015)	Microblogs	Supervised Machine Learning and General Lexicon	Daily	0 – used DJIA index (30 companies)	13 Months
Oliveira et al. (2017)	Microblogs and Surveys	Microblog Financial Lexicon	Daily	5 indexes, 13 portfolios (by market cap and by industry), SMB [4], HML[5] and MOM[6]	35 Months
This Study	Microblogs	General Lexicon	Intraday	8	79 days

Source: author own elaboration

[4] SMB - Small Minus Big: a Fama and French factor corresponding to the difference in returns between small and large firms.

[5] HML - High Minus Low: a Fama and French factor that is equal to the difference in returns between value (i.e., high book- to-market ratios) and growth (i.e., low book-to-market ratios) stocks.

[6] MOM - Momentum Factor: spread in returns between high prior return portfolios and low prior return portfolios.