Twitter Sentiment Analysis Applied to Finance: A Case Study in the Retail Industry

4 authors:

Thársis T. P. Souza
University College London
18 PUBLICATIONS   **123** CITATIONS

SEE PROFILE

Olga Kolchyna
University College London
4 PUBLICATIONS   **102** CITATIONS

SEE PROFILE

Philip Treleaven
University College London
103 PUBLICATIONS   **2,410** CITATIONS

SEE PROFILE

Tomaso Aste
University College London
218 PUBLICATIONS   **5,799** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project order and disorder in cellular patterns View project

Project How Did EU Immigrants Affect Income and Wealth Inequality in the UK? View project

Twitter Sentiment Analysis Applied to Finance: A Case Study in the Retail Industry

Thársis T. P. Souza[1], Olga Kolchyna[1], Philip C. Treleaven[12] and Tomaso Aste[12]

[1] Department of Computer Science, UCL, Gower Street, London, UK,
[2] Systemic Risk Centre, London School of Economics and Political Sciences, London, UK

Abstract. This paper presents a financial analysis over Twitter sentiment analytics extracted from listed retail brands. We investigate whether there is statistically-significant information between the Twitter sentiment and volume, and stock returns and volatility. Traditional newswires are also considered as a proxy for the market sentiment for comparative purpose. The results suggest that social media is indeed a valuable source in the analysis of the financial dynamics in the retail sector even when compared to mainstream news such as the Wall Street Journal and Dow Jones Newswires.

Keywords: retail, financial markets, data science, computational social science, social media, news analytics

1 Introduction

Twitter data has become an increasingly important source to describe financial dynamics. It provides a fine-grained real-time information channel that includes not only major news stories but also minor events that, if properly modelled, can provide ex-ante information about the market even before the main newswires. Recent developments have reflected this prominent role of social media in the financial markets. One major example is the U.S. Securities and Exchange Commission report allowing companies to use Twitter to announce key information in compliance with Regulation Fair Disclosure [10]. Twitter has also shown that can cause fast and drastic impact. In 2013, with the so-called *Hash Crash*, a hacked Twitter account of the American news agency Associated Press falsely disclosed a message about an attack on the White House causing a drop in the Dow Jones Industrial Average of 145 points in minutes [12].

Recent research also supports the hypothesis that Twitter data has statistical significant information related to financial indicators. As one of the first investigations analysing Twitter in the context of financial markets, [1] analyse the text content of daily Twitter feeds to identify two types of mood: (i) polarity (positive vs. negative) and (ii) emotions (calm, alert, sure, vital, kind, and

happy). They were able to increase the accuracy in the prediction of the DJIA index. More recently, [14] combined Information Theory with sentiment analysis to demonstrate that Twitter sentiment can contain statistically-significant ex-ante information on future prices of the S&P500 index and also identified a subset of securities in which hourly changes in social media sentiment do provide lead-time information. As a contribution to the to the field of event study research, [11] offer a methodology to analyse market reactions to combinations of different types of news events using Twitter to identify which news are more important from the investor perspective. In a similar way, [8] combine sentiment analytics with the identification of Twitter peaks in an event study approach to imply directions of market evolution. Further, exploring the social network structure from Twitter users, [13] provide empirical evidence of a financial community in Twitter in which users' interests align with the financial market.

In this work, we investigate the interplay between the Twitter sentiment extracted from listed retail brands, and financial stock returns and volatility. We verify whether there is statistically-significant information in this relationship and also compare it to a corresponding analysis using sentiment from traditional newswires. We consider volatility and log-returns as financial endogenous variables and we take Twitter sentiment and volume as exogenous explanatory variables in the financial dynamics of the selected stocks. We also consider traditional newswires as datasource for comparative purpose. Therefore, the main objectives are: (i) verify whether there is statistically-significant information between the Twitter sentiment, and stock returns and volatility and (ii) compare this interplay while using mainstream news as a proxy for the market sentiment. The main contribution of this work is an empirical evidence that supports the use of Twitter as a significant datasource in the context of financial markets in the retail industry even when compared to traditional newswires.

2 Dataset

Our analysis is conducted on a subset of five listed retail brands with stocks traded in the US equity market, which we monitor during the period from November 01, 2013 to September 30, 2014. The name of the investigated stocks with respective Reuters Instrument Codes (RIC) follow: ABERCROMBIE & FITCH CO. (ANF.N), NIKE INC. (NKE.N), HOME DEPOT INC. (HD.N), MATTEL INC. (MAT.N) and GAMESTOP CORP. (GME.N).

Given the companies selected, we consider three streams of time series data: (i) the market data, which is given at the daily stock price; (ii) news metadata supplied by [9], which consists in 10,949 news stories from Dow Jones Newswires, the Wall Street Journal and Barron's, and (iii) the social media data analytics provided by [5], which is based on 42,803,225 Twitter messages.

The news analytics data is provided in a metadata format where each news receives scores quantifying characteristics such as relevance, novelty and sentiment according to a related individual stock. Table 1 shows a sample of the news sentiment analytics data provided. The relevance score of the news ranges

between 0 and 100 and indicates how strongly related the company is to the underlying news story, with higher values indicating greater relevance. Usually, a relevance value of at least 90 indicates that the entity is referenced in the main title or headline of the news item, while lower values indicate references further down the story body. Here we filter the news stories with a relevance of 100. This increases the likelihood of the story considered being related to the underlying equity. Besides the novelty, we also consider the Event Sentiment Score (ESS). This measure indicates short-term positive or negative financial or economic impact of the news in the underlying company; higher values indicate more positive impact.

Table 1: News Sentiment Analytics. Each line represents a news story related to a company. The metadata considered consists of the relevance and sentiment scores and a timestamp.

Story	Company	Date	Hour	Relevance	Event Sentiment Score
1	NIKE INC.	20140104	210130	33	64
2	MATTEL INC.	20140105	41357	100	50
3	NIKE INC.	20140105	145917	93	88
4	NIKE INC.	20140105	150523	100	61
5	GAMESTOP CORP.	20140105	193507	44	50
6	GAMESTOP CORP.	20140106	170040	99	44
7	MATTEL INC.	20140106	222532	100	61
8	GAMESTOP CORP.	20140107	32601	100	50
9	MATTEL INC.	20140107	172628	55	40
10	NIKE INC.	20140110	204027	100	67

For the Twitter data analytics, the dataset from [5] provides the number of positive, negative and neutral English messages related to the underlying company on a daily basis. Also, the daily total number of messages related to a company is also provided, regardless of the language. [5] employed a new approach of sentiment classification based on the combination of existing common used techniques (lexicon-based and machine learning based) which outperformed standard benchmarks.

Table 2 shows an example of the Twitter sentiment analytics for the company MATTEL INC. Also, Table 3 shows a summary description of the selected companies with the number of stories considered.

Table 2: Twitter Sentiment Analytics. Sample of metadata information for the company MATTEL INC. In a daily basis we consider the number of positive, negative and neutral English Twitter messages related to the company and also the total number of messages regardless of the language.

Date	CompanyID	Volume	#Positive	#Negative	#Neutral
01/11/2013	MATTEL INC.	1,980	8	4	485
02/11/2013	MATTEL INC.	1,750	12	2	339
03/11/2013	MATTEL INC.	1,700	8	1	518
04/11/2013	MATTEL INC.	2,720	19	2	429
05/11/2013	MATTEL INC.	1,980	11	8	793
06/11/2013	MATTEL INC.	1,580	11	4	470
07/11/2013	MATTEL INC.	1,770	7	1	498
08/11/2013	MATTEL INC.	1,900	5	4	288
09/11/2013	MATTEL INC.	1,260	16	2	236
10/11/2013	MATTEL INC.	1,700	7	8	313

Table 3: Summary table of selected companies. The five retails brands selected for the analysis along with their market capitalization. Also, we present the total number of news and tweets in the selected period. The relevant news represent the news filtered with the highest relevance score (100).

Company	RIC	Market Cap.* ($Billions)	Total No. of News	Relevant News	No. of Tweets
ABERCROMBIE & FITCH CO.	ANF.N	2.86	1,608	174	1,352,643
NIKE INC.	NKE.N	67.39	2,881	178	38,033,900
HOME DEPOT INC.	HD.N	111.57	3,835	241	1,593,204
MATTEL INC.	MAT.N	15.02	1,508	125	613,798
GAMESTOP CORP.	GME.N	6.41	1,117	167	1,209,680

(*) Market Capitalization as in October 31, 2013. Source: Thomson Reuters Ikon.

3   Financial and Sentiment Variables

In this section, we define both the financial variables derived from the market data and the sentiment variables from news and Twitter. The former is comprised by log-returns and volatility while the latter is described not only by sentiment scores but also by the daily volume of stories.

3.1 Definition of the Variables

Let $P(t)$ be the closing price of an asset at day t and $R(t) = \log P(t) - \log P(t-1)$ its daily log-return. We consider the Excess of Log-return[3] of the asset over the return of the market index $\widehat{R}$ as:

$$ER(t) = R(t) - \widehat{R}(t). \tag{1}$$

As a proxy for the daily volatility of a stock, we define:

$$VOL(t) = 2\frac{P_{high}(t) - P_{low}(t)}{P_{high}(t) + P_{low}(t)}, \tag{2}$$

where $P_{high}(t)$ and $P_{low}(t)$ are the highest and the lowest price of the stock at day t, respectively.

In terms of Twitter analytics, we define the number of positive $(G(t))$ and negative $(B(t))$ English messages at day t mentioning a giving company, and we compute the variables [6]:

$$S_A(t) = G(t) - B(t), \quad S_R(t) = \frac{G(t) - B(t)}{G(t) + B(t)} \in [-1, 1] \tag{3}$$

as the absolute and relative, respectively, sentiment of that company in a given day. Notice that $S_R(t_0) = +1$, represents a day t_0 with the highest positive sentiment for the company considered; conversely $S_R(t_0) = -1$ indicates the highest negative sentiment, whereas we consider neutrality when $S_R(t_0) = 0$. We also define $V(t)$ as the total number of stories observed at day t, regardless of their language.

For the news analytics, the metadata information from [9] provides the Event Sentiment Score for each news' story. This score ranges between 0 and 100 where higher values indicate more positive sentiment while lower values below 50 show negative sentiment. We then normalize this score such as it ranges between -1 and 1 and consider its daily mean as the relative sentiment for news $S_R(t) \in [-1, 1]$. We label a news' story as positive if $S_R(t) > 0$ and as negative if $S_R(t) < 0$. Thus, we count the number of positive $G(t)$ and negative $B(t)$ stories to get the relative sentiment for News as $S_A(t) = G(t) - B(t)$. The volume of news $V(t)$ is also defined as the total number of news stories observed at day t. It is important to mention that the news dataset here considered contains only English stories whereas the Twitter data has no such limitation.

Fig. 1 shows a sample of the calculated variables from Twitter for the company Home-Depot Inc. It is interesting to notice a spike in volume and decrease in sentiment at the end of the period which follows a corresponding drop in excess of log-return. Further, Fig. 2 depicts the distribution of values of the relative sentiment obtained from Twitter and News. We observe that both present

---

[3] An alternative approach is to examine the alpha generation as the excess return of the underlying stock relative to its benchmark adjusted for a given level of risk as in the market model described in [3].

a skewed distribution while news has a more neutral-centred distribution compared to Twitter. It is important to notice that the sentiment provided by the Twitter analytics presents a distinct proxy for sentiment compared to News as each company analysed depicts different positive/negative sentiment tones, e.g., the NIKE's Twitter sentiment is highly positive while the news' sentiment has a mean around a neutral point.
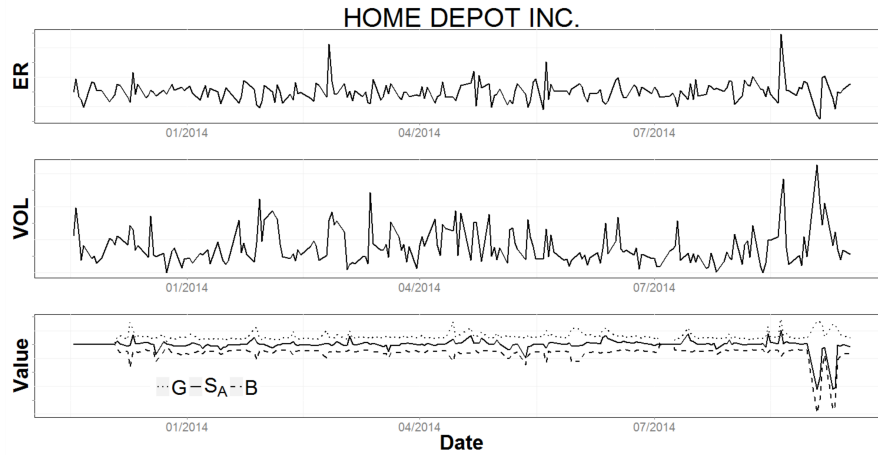

Fig. 1: Sample Descriptive Analysis for Home-Depot Inc. Variables: Excess of log-return, ER; volatility, VOL; absolute sentiment, S_A; number of positive messages, G and number of negative messages, B. There is a spike in volume and decrease in sentiment at the end of the period which follows a corresponding drop in excess of log-return.
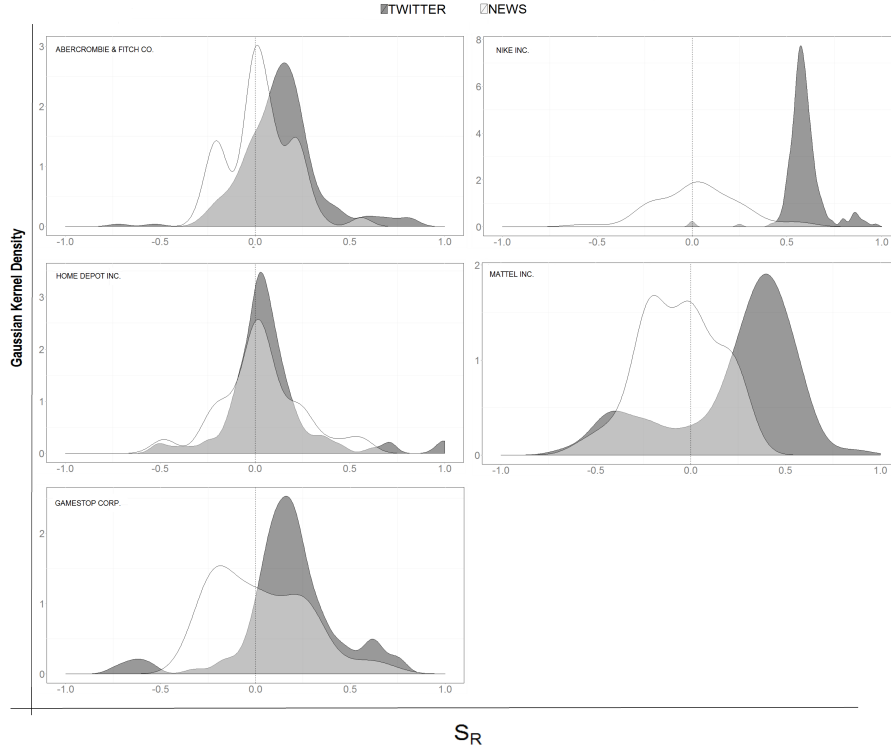
Fig. 2: Distribution of relative sentiment $S_R(t)$ from Twitter and News for the companies: ABERCROMBIE & FITCH CO., GAMESTOP CORP., HOME DEPOT INC., MATTEL INC. and NIKE INC. It is clear that the sentiment provided by Twitter is a distinct proxy for market sentiment compared to News as each company analysed depicts different distributions of sentiment.

4 Method

4.1 Granger Causality

We are interested in investigating the statistical causality between sentiment and the financial variables. In this sense, [4] introduced a concept of cause-effect dependence where the cause not only should occur before the effect but also should contain unique information about the effect. Therefore, we say that X Granger-cause Y if the prediction of Y can be improved using both information from X and Y as compared to only utilizing Y.

In a Vector Auto-Regressive (VAR) framework, we can assess the Granger causality performing a F-test to verify the null hypothesis that Y is not Granger-caused by X and measure its probability of rejection within a confidence level.

Hence, assuming the VAR models:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_k Y_{t-k} + \beta_1 X_{t-1} + \ldots + \beta_k X_{t-k} + \epsilon_t, \qquad (4)$$

$$X_t = \gamma_0 + \gamma_1 X_{t-1} + \ldots + \gamma_k X_{t-k} + \theta_1 Y_{t-1} + \ldots + \theta_k Y_{t-k} + \widehat{\epsilon}_t, \qquad (5)$$

we take the null hypothesis in equation (6) and test it against its alternative one in equation (7). Thus, a rejection of the null hypothesis implies that Y Granger-cause X.

$$\mathcal{H}_0 : \beta_1 = \beta_2 = \ldots = \beta_k = 0 \qquad (6)$$

$$\mathcal{H}_1 : \exists \beta_\tau, \ 0 \leq \tau \leq k : \quad \beta_\tau \neq 0 \qquad (7)$$

In the same way, the test for X Granger-cause Y can be done considering the equation (5) and taking the hypotheses from equations (6) and (7) in an analogous way.

Here we will test the Grange-causality between the Excess of Log-return ER and the number of positive stories G, the number of negative stories B, the relative sentiment S_R and the absolute sentiment S_A. For the volatility VOL, we will consider the total volume of stories V in addition to previously mentioned variables. Furthermore, we will perform the Granger-causality test over the normally standardized versions of the time series analysed such as they have zero mean and standard deviation 1. To perform the F-statistics of the Granger-causality tests we use the function anova.lm from the package stats of the R Project for Statistical Computing [7].

For the only purpose of visualization of the Granger-causality results we create a Granger-causality graph $G = [V, E]$, where V is a node set, and E is an edge set. A node $u \in V$ represents a variable in the causality test and an edge $e = (u, v)$ indicates that u Granger-causes v within a pre-defined significance level. Further, we define $p(e)$ as an attribute of the edge and if C is the set of companies in which we have a causality between u and v, then we set $p(e) = C$. Fig. 3 shows an example of a Granger-causality graph that indicates that u Granger-causes v for the set of companies C.
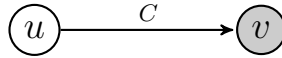

Fig. 3: Granger-causality graph. The variable u Granger-causes the variable v for the set of companies C.

4.2 Predictive Analysis

To evaluate the predictive power of sentiment we consider two auto-regressive models, with and without sentiment, and conduct a 1-step ahead prediction

analysis:

$$\mathcal{M}_0 : X(t) = \alpha + \sum_{\tau=1}^{k} \beta_\tau X(t-\tau) + \epsilon_t, \qquad (8)$$

$$\mathcal{M}_1 : X(t) = \alpha + \sum_{\tau=1}^{k} \beta_\tau X(t-\tau) + \sum_{i=1}^{k} \gamma_i Y(t-\tau) + \widehat{\epsilon}_t \qquad (9)$$

where,

$$X(t) \in \{ER(t), VOL(t)\}, \qquad (10)$$
$$Y(t) \in \{G(t), B(t), S_R(t), V(t)\}. \qquad (11)$$

As $S_A(t)$ is a linear combination between $G(t)$ and $B(t)$, we will not consider it in the linear regression for any dataset. Moreover, we consider only 1 day lag for the sentiment variables and a lag of 2 days for the financial variables[4]. Again, we will consider the normally standardized versions of the time series analysed.

Hence, we will consider the following regression models for the excess of log-return prediction:

$$\mathcal{M}_0 : ER(t) = \alpha + \beta_1 ER(t-1) + \beta_2 ER(t-2) + \epsilon_t, \qquad (12)$$
$$\mathcal{M}_1 : ER(t) = \alpha + \beta_1 ER(t-1) + \beta_2 ER(t-2) \qquad (13)$$
$$+ \gamma_1 G(t-1) + \gamma_2 B(t-1) + \gamma_3 S_R(t-1) + \widehat{\epsilon}_t$$

For the volatility prediction using news as datasource we will not include the volume time series $V(t)$ as an explanatory variable in the regression because of its high correlation with the number of positive and negative news already taken into account in the model. Notice that, for the Twitter case, the volume time series consider also non-English messages which are not taken into account by the time series in $G(t)$ and $B(t)$. Therefore, we keep $V(t)$ as an explanatory variable in the Twitter model. As a result, we have for news:

$$\mathcal{M}_0 : VOL(t) = \alpha + \beta_1 VOL(t-1) + \beta_2 VOL(t-2) + \epsilon_t, \qquad (14)$$
$$\mathcal{M}_1 : VOL(t) = \alpha + \beta_1 VOL(t-1) + \beta_2 VOL(t-2) \qquad (15)$$
$$+ \gamma_1 G(t-1) + \gamma_2 B(t-1) + \widehat{\epsilon}_t$$

and the corresponding model for Twitter:

$$\mathcal{M}_0 : VOL(t) = \alpha + \beta_1 VOL(t-1) + \beta_2 VOL(t-2) + \epsilon_t, \qquad (16)$$
$$\mathcal{M}_1 : VOL(t) = \alpha + \beta_1 VOL(t-1) + \beta_2 VOL(t-2) \qquad (17)$$
$$+ \gamma_1 G(t-1) + \gamma_2 B(t-1) + \gamma_3 V(t-1) + \widehat{\epsilon}_t.$$

---

[4] A model selection approach can be also used in order to find an optimal lag for the explanatory variables, examples of selection's criteria are: the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and Mallow's Cp. See [2].

Forecasting accuracy is measured by comparing the two residuals ϵ_t and $\widehat{\epsilon}_t$ in terms of Residual Standard Error:

$$\hat{\sigma} = \sqrt{\frac{\sum\limits_{i=1}^{T} \left(y_i - \widehat{y}_i\right)^2}{n}} = \sqrt{\frac{\sum\limits_{i=1}^{T} \hat{\epsilon}_i^2}{n}} \qquad (18)$$

where, T is the total number of points, n is the number of degrees of freedom of the model, $\widehat{y}_i$ are the predicted values and y_i are the observed ones.

5   Results and Discussion

Here we present the results from the Granger-causality tests and the predictive analysis between the financial variables and sentiment data from Twitter and news. The sentiment predictive power and its Granger-causality tests are fulfilled in 1-step ahead fashion. We investigate the statistical significance of the sentiment variables in regards to movements in returns and volatility and we compare the Twitter results with news. We provide empirical evidence that Twitter is moving the market in respect to the excess of log-returns for a subset of stocks. Also, Twitter presents a stronger relationship with stock returns compared to news, in the selected retail companies. On the other hand, the Twitter sentiment analytics showed a weaker relationship with volatility compared to news.

5.1   Excess of Log-Returns

We analyse the dynamics of excess of log-returns of the stocks considered in relation to absolute and relative sentiments and also with the number of positive and negative stories. We consider the S&P 500 as the market index in the computation of excess of log-returns in the equation (1).

Fig. 4 shows the Granger-causality graph that summarises the significant Granger-causalities (p-value < 0.05) between the excess of log-return and the sentiment variables for both news and Twitter. See Table 5 for the detailed results. We observe that the Twitter's sentiment analytics present more significant points compared to news. The Twitter's relative sentiment and its number of positive messages Granger-cause log-returns, respectively, for the companies GAMESTOP CORP. and MATTEL INC. The Twitter's absolute sentiment also Granger-causes returns for MATTEL INC. while having a two-way significant (p-value < 0.01) Granger-causality for the company HOME DEPOT INC. The number of negative stories alone has no significant relationship with returns but combined with the number of positive stories in the form of relative and absolute sentiment it shows to be an important measure. The News' analytics have only one significant relationship that it is observed in the number of positive news Granger-causing the excess of log-returns for the company GAMESTOP CORP.

The solution of the multiple regression analysis in Table 6 agrees with the Granger-causality tests as it shows Twitter with a higher number of significant
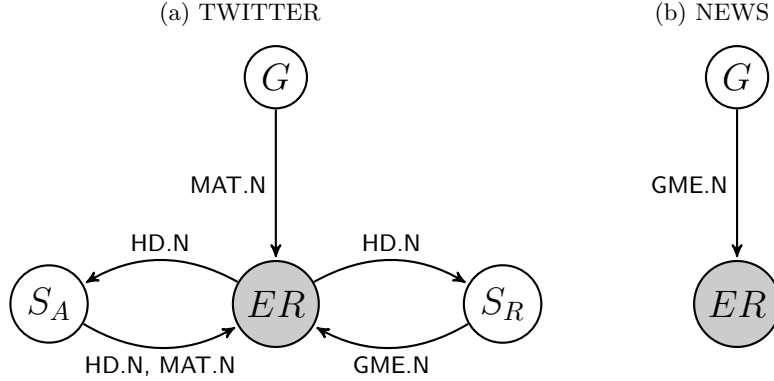
Fig. 4: Granger-causality graph for (a) Twitter and (b) news. It shows the significant points in the the Granger-causality test between excess of log-returns (ER) and the sentiment analytics: number of positive stories (G), number of negative stories (B), absolute sentiment (S_A) and relative sentiment S_R. Sentiment variables that presented no significant causality are not shown in the graph.

sentiment coefficients compared to news. The company MATTEL INC. particularly presents all sentiment coefficients with high significance (p-value < 0.01) suggesting that the Twitter sentiment analytics is indeed relevant in the prediction of the next day excess of log-return. The companies HOME DEPOT INC. and GAMESTOP CORP. also presented significant sentiment coefficients. For the News analytics, the sentiment was significant only for the company GAMESTOP CORP. Further, analysis of the Residual Standard Error of the models with and without sentiment variables in Table 4 shows that the use of the Twitter sentiment variables reduced the error of the model with only market data for the companies MATTEL INC., HOME DEPOT INC. and GAMESTOP CORP. while the news' sentiment improved the prediction only for the company GAMESTOP CORP.

Table 4: Residual standard error improvement in prediction of excess of log-return $ER(t)$ using $S_R(t), G(t)$ and $B(t)$ compared to the model with only market data.

Company	Error Reduction (%)	
	NEWS	TWITTER
NIKE INC.	-2.41	-0.58
ABERCROMBIE & FITCH CO.	-1.26	-0.60
HOME DEPOT INC.	-0.99	1.23
MATTEL INC.	-0.48	2.82
GAMESTOP CORP.	8.34	1.10

Table 5: Statistical significance (p-values) of Granger-causality analysis between excess log-return and $S_A(t), S_R(t), G(t)$ and $B(t)$ for the companies: ABERCROMBIE & FITCH CO. (ANF.N), NIKE INC. (NKE.N), HOME DEPOT INC. (HD.N), MATTEL INC. (MAT.N) and GAMESTOP CORP. (GME.N).

	TWITTER ANALYTICS					NEWS ANALYTICS				
	HD.N	MAT.N	GME.N	NKE.N	ANF.N	HD.N	MAT.N	GME.N	NKE.N	ANF.N
$S_A \to ER$	0.003***	0.046**	0.140	0.888	0.477	0.303	0.411	0.140	0.621	0.707
$ER \to S_A$	0.006***	0.404	0.231	0.354	0.937	0.423	0.451	0.230	0.546	0.281
$S_R \to ER$	0.449	0.497	0.032**	0.680	0.591	0.747	0.977	0.696	0.816	0.814
$ER \to S_R$	0.024**	0.855	0.196	0.995	0.875	0.942	0.314	0.162	0.564	0.213
$G \to ER$	0.182	0.016**	0.885	0.400	0.685	0.203	0.228	0.014**	0.304	0.231
$ER \to G$	0.050*	0.305	0.957	0.380	0.197	0.388	0.382	0.171	0.199	0.518
$B \to ER$	0.327	0.559	0.267	0.344	0.763	0.681	0.976	0.920	0.796	0.398
$ER \to B$	0.219	0.792	0.538	0.166	0.480	0.855	0.646	0.894	0.769	0.863

Significance codes: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.1: *

Table 6: Summary statistics of multiple regression. Prediction of excess of log-return (ER) using sentiment (S_R, G, B) for the companies: ABERCROMBIE & FITCH CO. (ANF.N), NIKE INC. (NKE.N), HOME DEPOT INC. (HD.N), MATTEL INC. (MAT.N) and GAMESTOP CORP. (GME.N).

| | TWITTER ANALYTICS | | | | | NEWS ANALYTICS | | | | |
	HD.N	MAT.N	GME.N	NKE.N	ANF.N	HD.N	MAT.N	GME.N	NKE.N	ANF.N
$ER_{(t-1)}$	4.414018e-04	1.192513e-01*	-1.704137e-01**	-3.001408e-02	1.203011e-02	2.592302e-01	6.129545e-01***	-9.060641e-02	4.018227e-02	-1.457410e-01
$ER_{(t-2)}$	-9.806531e-02	9.192829e-03	1.535305e-02	9.752118e-02	-4.741730e-02	-5.695452e-02	-2.674834e-01*	4.458446e-01***	4.363861e-02	-6.310047e-02
$SR_{(t-1)}$	-2.258381e-02	3.718939e-01***	1.821209e-01**	1.817089e-02	2.558886e-02	-1.092573e-01	-8.211806e-02	-2.852138e-01	4.487921e-02	9.364799e-02
$G_{(t-1)}$	2.536661e-01***	-4.286924e-01***	-1.229142e-02	-3.956882e-02	4.445273e-02	2.600786e-01	2.683852e-01	4.962514e-01***	-2.144345e-01***	-2.279505e-01
$B_{(t-1)}$	-2.612111e-01**	4.801024e-01***	-1.159965e-01	-1.459096e-02	-3.347246e-02	-9.858689e-02	3.625886e-03	-5.315150e-02	-6.179474e-02	3.457372e-02

Significance codes: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.1: *

5.2 Volatility

Here we analyse the interplay between message's volume and sentiment with stock's volatility. As volume measures we consider: the number of positive and negative English stories and also the total volume of stories, regardless of the language. As sentiment analytics we consider: absolute sentiment and relative sentiment.

Fig. 5 shows the significant links (p-value < 0.05) of the Granger-causality test between the volatility and the sentiment variables. See Table 8 for the detailed results. Overall, there are more significant points of causality for the News sentiment analytics compared to Twitter. We observe that the number of positive stories and the total volume both Granger-cause volatility for Twitter as well as for news but more companies are affected by news. The absolute sentiment Granger-causes volatility only for news, observed in the company ABERCROMBIE & FITCH CO. (ANF.N). The relative sentiment and the number of negative stories are not causing volatility, on the other hand volatility is Granger-causing negative news for the company GAMESTOP CORP. (GME.N).
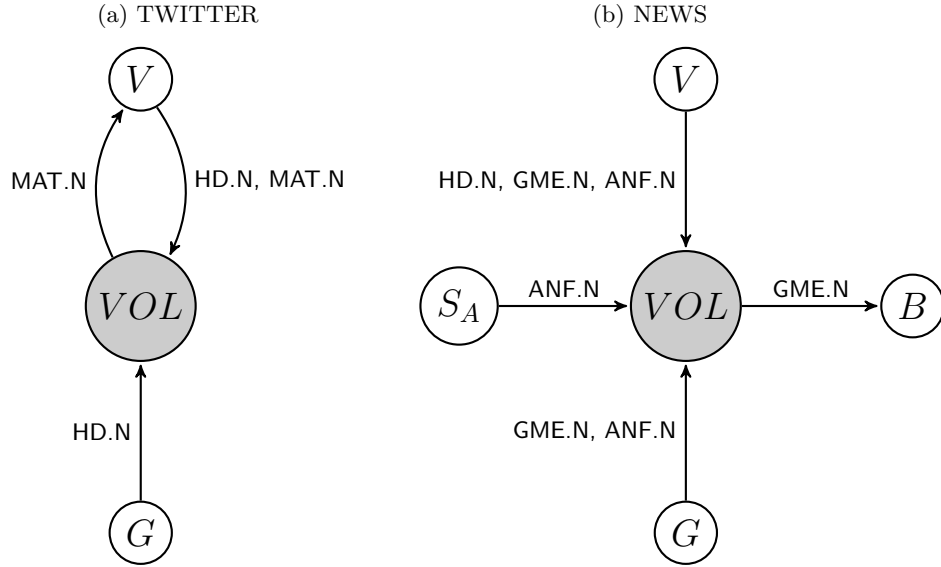

Fig. 5: Granger-causality graph for (a) Twitter and (b) news. It shows the significant points in the Granger-causality test between volatility (VOL) and the sentiment analytics: total number of stories (V), number of positive stories (G), number of positive stories (B), absolute sentiment (S_A) and relative sentiment (S_R). Sentiment variables with no significant causality are not shown in the graph.

The solution of the multiple regression analysis in Table 9 shows that the number of positive stories is a significant variable for both news and Twitter, being more significant for the former. The number of negative stories showed no relevance in both regressions. The total volume of Twitter was relevant only for the company NKE.N. Moreover, analysis of the Residual Standard Error of the models with and without sentiment variables in Table 7 shows that both Twitter and News were able to reduce the error in prediction for a subset of the companies. In the cases where the model was improved with sentiment, News provided a higher reduction of error compared to Twitter.

Table 7: Residual standard error improvement in the prediction of volatility $VOL(t)$ using $G(t), B(t)$ and $V(t)$ compared to the model with only market data.

Company	Error Reduction (%)	
	NEWS	TWITTER
NIKE INC.	1.36	1.08
ABERCROMBIE & FITCH CO.	4.03	-0.52
HOME DEPOT INC.	2.46	1.10
MATTEL INC.	-2.21	-0.36
GAMESTOP CORP.	14.99	0.20

Table 8: Statistical significance (p-values) of Granger-causality analysis between volatility ($VOL(t)$) and $G(t)$, $B(t)$ and $V(t)$ for the companies: ABERCROMBIE & FITCH CO. (ANF.N), NIKE INC. (NKE.N), HOME DEPOT INC. (HD.N), MATTEL INC. (MAT.N) and GAMESTOP CORP. (GME.N).

	TWITTER ANALYTICS					NEWS ANALYTICS				
	HD.N	MAT.N	GME.N	NKE.N	ANF.N	HD.N	MAT.N	GME.N	NKE.N	ANF.N
$V \rightarrow VOL$	0.025**	0.004***	0.751	0.976	0.729	0.029**	0.320	0.020**	0.846	0.001***
$VOL \rightarrow V$	0.228	0.016**	0.980	0.611	0.924	0.307	0.453	0.053*	0.462	0.560
$G \rightarrow VOL$	0.004***	0.072*	0.678	0.397	0.228	0.060*	0.341	0.031**	0.961	<0.001***
$VOL \rightarrow G$	0.560	0.273	0.668	0.330	0.690	0.146	0.390	0.211	0.859	0.615
$B \rightarrow VOL$	0.053*	0.118	0.720	0.636	0.884	0.522	0.301	0.204	0.526	0.540
$VOL \rightarrow B$	0.420	0.477	0.363	0.729	0.967	0.484	0.620	0.026**	0.391	0.801
$S_R \rightarrow VOL$	0.539	0.305	0.489	0.786	0.220	0.416	0.925	0.510	0.441	0.056*
$VOL \rightarrow S_R$	0.944	0.867	0.791	0.611	0.623	0.854	0.352	0.590	0.184	0.340
$S_A \rightarrow VOL$	0.274	0.736	0.437	0.142	0.216	0.435	0.194	0.497	0.652	0.010**
$VOL \rightarrow S_A$	0.458	0.900	0.747	0.120	0.950	0.173	0.352	0.187	0.707	0.996

Significance codes: p-value < 0.01: ***; p-value < 0.05: **; p-value < 0.1: *

Table 9: Summary statistics of multiple regression. Prediction of volatility (VOL) using sentiment (G, B) and volume (V) for the companies: ABERCROMBIE & FITCH CO. (ANF.N), NIKE INC. (NKE.N), HOME DEPOT INC. (HD.N), MATTEL INC. (MAT.N) and GAMESTOP CORP. (GME.N).

	TWITTER ANALYTICS					NEWS ANALYTICS				
	HD.N	MAT.N	GME.N	NKE.N	ANF.N	HD.N	MAT.N	GME.N	NKE.N	ANF.N
$VOL_{(t-1)}$	2.305851e-01***	2.211374e-01***	1.019004e-01	1.893666e-01***	1.308395e-01*	3.891e-01*	6.265e-02	9.587e-02	3.785e-01**	3.792e-01**
$VOL_{(t-2)}$	1.250733e-01*	1.238580e-01*	-6.173133e-02	4.307688e-02	8.735957e-02	-9.120e-02	-4.451e-03	5.092e-02	3.093e-02	1.054e-01
$G_{(t-1)}$	1.845026e-01*	1.073941e-01	-2.807417e-01*	-1.586209e-03	3.533443e-02	3.018e-01**	2.772e-03	6.046e-01***	1.274e-01	-3.329e-01*
$B_{(t-1)}$	1.420223e-03	-7.759929e-02	4.461084e-02	-2.181354e-03	-8.569212e-02	9.294e-02	-1.105e-01	1.184e-01	-1.400e-01	-1.300e-01
$V_{(t-1)}$	-1.468387e-03	-1.862391e-02	1.920558e-01	1.901394e-01**	5.742508e-02	N/A	N/A	N/A	N/A	N/A

Significance codes: p-value < 0.01: ***, p-value < 0.05: **, p-value < 0.1: *

6   Conclusion

We showed that measures of the Twitter sentiment extracted from listed retail brands have statistically-significant relationship with stock returns and volatility. While analysing the interplay between the excess of log-return and the Twitter sentiment variables we conclude that: (i) Twitter presented a stronger Granger-causality with stock's returns compared to news; (ii) positive tweets and Twitter's sentiment Granger-cause excess of log-returns for a subset of companies; (iii) Twitter's sentiment analytics is indeed relevant in the prediction of the next day excess of log-return even compared to traditional newswires. Moreover, in the volatility analysis we found that: (i) Twitter's analytics showed a weaker relationship with volatility compared to the one observed with returns; (ii) the number of positive tweets and the total volume both Granger-cause volatility for some companies but present reduced Granger-causality compared to news; (iii) the number of positive tweets is a significant variable for the 1-step ahead prediction of volatility while the number of negative messages showed no relevance. Overall, the Twitter sentiment analytics showed to be a distinct and complementary proxy of market's sentiment compared to news in the analysis of the financial dynamics of retail brands' stocks. Surprisingly, the Twitter's sentiment presented a relatively stronger relationship with the stock's returns compared to traditional newswires. The results suggest that social media analytics have a prominent role in the dynamics of the retails sector in the financial markets.

Acknowledgments

References

1. Bollen, J., Mao, H., Zeng, X.: Twitter mood predicts the stock market. Journal of Computational Science 2(1), 1 – 8 (2011)
2. Box, G., Jenkins, G.: Time series analysis: forecasting and control. Holden-Day series in time series analysis and digital processing, Holden-Day (1976)
3. Fama, E.F.: Efficient capital markets: A review of theory and empirical work. The Journal of Finance 25(2), 383–417 (1970), http://dx.doi.org/10.1111/j.1540-6261.1970.tb00518.x
4. Granger, C.W.: Testing for causality: a personal viewpoint. Journal of Economic Dynamics and control 2, 329–352 (1980)
5. Kolchyna, O., Souza, T.T.P., Treleaven, P., Aste, T.: Twitter sentiment analysis applied to retail brand sales (2015), http://arxiv.org/abs/1507.00955

6. Lillo, F., Miccichè, S., Tumminello, M., Piilo, J., Mantegna, R.N.: How news affect the trading behavior of different categories of investors in a financial market (Jul 2012), `http://arxiv.org/abs/1207.3300`
7. R Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria (2014), `http://www.R-project.org/`
8. Ranco, G., Aleksovski, D., Caldarelli, G., Mozetič, I.: Investigating the relations between twitter sentiment and stock prices (2015), `http://arxiv.org/abs/1506.02431`
9. Ravenpack: Ravenpack official website. http://www.ravenpack.com/ (2015), last accessed on Jan 29, 2015
10. SEC: U.S. Securities and Exchange Commission. SEC says social media ok for company announcements if investors are alerted. http://1.usa.gov/1zFxUPa (Apr 2013), last accessed on Jan 29, 2015
11. Sprenger, T.O., Sandner, P.G., Tumasjan, A., Welpe, I.M.: News or Noise? Using Twitter to Identify and Understand Company-specific News Flow. Journal of Business Finance & Accounting 41(7-8), 791–830 (09 2014)
12. WSJ: Wall Street Journal. False AP twitter message sparks stock-market selloff. http://on.wsj.com/12ms85v (Apr 2013), last accessed on Jan 29, 2015
13. Yang, S.Y., Mo, S.Y.K., Zhu, X.: An empirical study of the financial community network on twitter. 2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (2014)
14. Zheludev, I., Smith, R., Aste, T.: When Can Social Media Lead Financial Markets? Scientific Reports 4 (Feb 2014)