

Are Stock Markets Driven More by Sentiments than Efficiency?

Mirosława Lasek¹ and Jan Lasek²

¹Professor, Department of Information Systems and Economic Analysis, Faculty of Economic Sciences, University of Warsaw, ul. Długa 44/50, 00-241 Warsaw, Poland.

²PhD Student, Interdisciplinary PhD Studies, Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, 01-248 Warsaw, Poland. E-mail: janek.lasek@gmail.com (corresponding author).

Production Management

Received August 12, 2014; received revision December 8, 2014; accepted January 8, 2015

Available online April 28, 2015

Abstract: In this paper, we investigate whether the sentiment contained in text of news items influences a company's stock price. With the means of sentiment analysis, a quantitative approach for measuring sentiment in text, we derive scores for positive and negative sentiment contained in news. Next, we relate these quantities to stock market prices. By regression analysis we show that the sentiment contained in text of news items affects a company's stock prices. This reveals that human sentiments influence capital markets and shows the potential of sentiment analysis as a quantitative approach to measure opinion with possible applications in e.g. public opinion tracking or brand and product management.

Keywords: Behavioral finance, capital market, news items, sentiment, sentiment analysis.

1. Introduction

Under the assumption that individuals act rationally and strive for utility maximization, many economic phenomena have plausible and reasonable explanations. However, the theory of behavioral economics prompts us to go far beyond these strict, and often incorrect, assumptions (Ariely, 2008; Kahneman, 2011). It provides insights into many phenomena that are not offered by the classical approach. The stock market in particular is a place, where psychology has proven to be an important factor of the valuation process for financial products.

In this paper, we investigate a specific form of impact of the Internet news on the capital market. By means of sentiment analysis, we perform an automated large-scale examination of text of news items concerning several financial entities. The method of sentiment analysis constitutes a convenient tool of evaluation of sentiment contained in particular piece of text. The goal of the article is to verify whether sentiments exhibited towards a company influence its daily returns (Lasek, 2013).

The main idea behind our research is the fact that the stock valuation is not a result of strict mathematical calculations, but rather it is based on mass psychology, at least to a certain extent. If a sufficient number of individuals share similar sentiment, common opinion, prejudice or expectation, we may observe that it affects an investigated phenomenon. We choose news items appearing on the Internet as a medium that directly influences the sentiment of a number of individual

investors. We carry out the evaluation and quantification of general sentiment in news items towards several chosen companies. Next, we relate these quantities to stock price changes. We believe that considerable media optimism or pessimism corresponds to analogous investor sentiment, which results in an upward or a downward pressure on stock prices.

The complexity of the capital market gives rise to many interesting research questions and a variety of theories for the explanation of its functioning. Among the laureates of the 2013 Nobel Prize in Economic Sciences we can find Eugene Fama and Robert Shiller for their influential research conducted in the field of finance (with Lars Hansen being the third laureate). Surprisingly, their views appear to be contradictory. Fama (1970) is known as the inventor of the Efficient Market Hypothesis. He claims that any potentially useful information is immediately included in asset prices and therefore stock prices are entirely unpredictable – the market is said to be efficient. In contrast, Shiller (2005) perceives the capital market as a place where psychology and “human factor” play a major role and therefore prices are irrationally settled. The research done so far in the field of finance delivers evidence on how complex the capital markets are. It reveals as well that forces acting on the stock markets are difficult to understand and that different theories may not be consistent with one another, even contradictory. There is a need of further investigation of dynamically changing and fascinating area of financial markets and much is to be discovered.

The article is structured as follows. In section two we discuss related work to our study. In the third section, we describe the method of sentiment analysis that serves to measure sentiment contained in a piece of text. In section four, we present the results of experiment with the use of these measurements and stock market data. Finally, in the last section, we summarize the main findings and conclusions of our work.

2. Related Work

There are two main fundamental approaches that are relevant for our work. The first one relates to the field of finance, theories of its functioning and the role of psychology on capital markets. The second originates from information technology and is concerned with automated text analysis in order to evaluate the sentiment contained in a given corpus of text documents (e.g. news items). First, we briefly discuss the economic environment of our study – the stock market.

The history of stock markets began in Amsterdam in early seventeenth century (den Heijer, 2002). Since then stock markets around the world developed greatly. Thanks to stock exchange, enterprises can acquire funding and their profits can be distributed among a group of investors. Hence stock market is a source of financial capital for e.g. an enterprise's production goals. On the other hand, a firm's performance on product markets, its managerial decisions and launch of new products affect its stock price (Chaney 1992, Mitchel and Stafford 2000, Peress 2010). We shall also note that the stock market is an inherently risky economic environment with stock prices varying continuously see, e.g. Guo and Whitelaw (2006) and references therein. Below we discuss two seemingly contradictory views on the stock market. The first one considers psychology as a major factor in financial asset valuation. According to the second one, the stock market is said to be efficient. We consider these views complementary rather than contradictory. Behavioral aspects of asset pricing were raised by many authors so far. Notably Keynes (1936), in his most prominent work, considers the capital market in terms of a beauty contest. In the contest, the goal is to indicate the prettiest face out of hundreds of photographs. The winner of the contest is the one whose choice is the closest to the preferences of the other competitors. Similarly, the goal of capital market investments is not to buy stocks that one considers profitable in the future but stocks regarded as such by other investors. Keynes also argues that "the market will be subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for reasonable calculation". Conventional valuation of assets is a result of mass psychology mechanisms rather than well-founded mathematical calculation. Kahneman and Tversky (1974) in their influential article describe how different types of heuristics can influence and bias people's decisions. In particular, if a description of a company is favorable or mediocre, one would rather predict high or low profits regardless the fact that such information may not be useful for forecasting whatsoever. Over the years many abnormalities in the functioning of capital markets were documented see, e.g. French (1980), Shiller (1981, 2005), Keim (1983), Reinganum (1983), De Bondt and Thaler (1985), Benzartzi and Thaler (1995), Johnson and Thaler (1999), Thaler (1999), Odean and Barber (1999), Schwert (2003), Zielonka (2004), Edmans et al. (2007) or

Kaplanski and Levy (2010). However, the evidence for their presence in real world market data is often ambiguous. What is more, any potential abnormality creates a possibility of arbitrage in the market and it should cancel out with time (Schwert, 2003).

Another way to look at the stock market is the study of its efficiency which we shall discuss here. Fama (1970), one of the most influential researchers in the area of stock market functioning, is known as the inventor of the Efficient Market Hypothesis. According to this hypothesis, all information available to investors is immediately included in stock prices. Hence there is no opportunity to make profit at the stock market as any piece of information is already accounted for in an asset price. There are close affinities between this theory and the Random Walk Hypothesis (Godfrey et al. 1964; Malkiel, 2007). According to it, stock prices follow a random walk, i.e. a process with independent and identically distributed increments. Stock price changes are entirely random, hence unpredictable. Both these prominent hypotheses for financial markets imply that it is impossible to predict stock prices movements with any kind of mathematical modelling, regardless of its complexity. There is evidence supporting each of the hypotheses, however, there is no consensus whether indeed the markets are efficient and the prices are unpredictable see, e.g. Borges (2010), Lo and MacKinlay (2011), Sewell (2011, 2012) and the preceding discussion on the role of psychology at the stock market. Several quantitative studies towards measuring general sentiment and its relation to the stock market prices have been conducted. To evaluate sentiment, the method of sentiment analysis is employed. This method measures opinion in a different manner than via surveys, interviews or certain market indicators, which are commonly used to measure investors' sentiment (Brown and Cliff, 2004; Baker and Wurgler, 2007). The method indirectly evaluates the sentiment contained in documents (e.g. news items) that presumably influence investors' opinions. An example of application of this method is the study conducted by Tetlock (2007). The author derives a so-called pessimism factor from a popular Wall Street Journal column. He shows that media pessimism is connected with both stock returns and trading volume. The author identifies their interrelation. Firstly, the proposed measure of media pessimism is a good predictor of the Dow Jones Industrial Average (DJIA). Secondly, lower stock returns in previous days contribute to deeper media pessimism. In a related research, Bollen et al. (2011) examine daily feeds from Twitter, a microblog portal, and investigate their predictive capabilities on the DJIA. They show that the data describing mood of society, measured by a sentiment analysis tools, can be useful in prediction of the direction of changes in DJIA. In the same paper they observe significant mood fluctuations around major social events, such as presidential elections or Thanksgiving. This reveals the potential of this method's application. Holum et al. (2013) investigate text mining methods for news articles of several companies listed in S&P500 index. The authors also draw attention to the fact that current results on the applicability of text mining methods are lacking consensus. Yet another example relevant to our study is so-called R-word index, an informal indicator of recession proposed by The Economist (<http://www.economist.com/node/21529079>, retrieved on December 7, 2014). By counting the number of appearances of the word 'recession' in major

newspapers, the recession index is derived, which appears to reflect relatively well fluctuations in the observed business cycle. Clearly, with the deepening crisis, more sources mention the word 'recession'. On the other hand, an interesting question that arises is whether an increase in the number of news items related to the recession really contributes to its very expansion. We will not pursue the discussion here.

The results of the research that has already been conducted in the field of sentiment analysis and the stock market prediction, as well as the theoretical foundations of the functioning of capital markets based on behavioral economics, are in line with the goal of our paper. There is also a great incentive to explore this area of research in the analysis of financial markets and their efficiency. Not to mention, many financiers and financial institutions use this approach as a tool for investing on the market. It should be obvious that algorithmic trading is being used extensively nowadays, which often causes market to behave in an uncontrolled way. For example, on April 23, 2013 Associated Press had its Twitter account hacked and a fake message on bomb assassination in The White House was posted. It caused a massive decline in the market as indicated by Dow Jones and S&P 500 indexes (<http://www.bloomberg.com/news/2013-04-23/dow-jones-drops-recovers-after-false-report-on-ap-twitter-page.html>, retrieved on December 7, 2014). The prices returned to its previous levels after the message was recognized as a fake one shortly after the dramatic drop.

In the next section we will introduce the underlying method of our research – the computational treatment of the sentiment contained in text, i.e. sentiment analysis.

3. Measuring Sentiment of Investors – The Method of Sentiment Analysis

Sentiment analysis can be regarded as a subfield of text mining in information technology, which emerged as an individual research area around the year 2000 (Pang and Lee, 2008; Liu, 2012). It has also been used in stock market prediction, as discussed above. This approach towards stock market can be regarded as relatively new as compared to the two methods of forecasting stock prices with the use of technical analysis and fundamental analysis. Predictive modelling based on sentiment analysis may be considered as an alternative approach. The main idea is to process text (in particular, news items), measure its sentiment and relate it to stock market prices.

People's opinions are an integral part of our perception of the world and the way we see ourselves. Usually, when we do not have sufficient information about a certain thing, we want to find out what other people do think about it. Along with the great expansion of the Internet, social media sites and blogs, people's views and opinions on a variety of matters became more accessible. The objective is to explore these opinions by means of computer algorithms.

There are several terms used to name the field of automated opinion extraction from the text and its evaluation. The ones that are most commonly used are sentiment analysis or opinion mining, but terms such as subjectivity analysis, review mining or appraisal extraction and several others were also cited. Throughout the paper we will use the first term mentioned above.

Before we define the objective of sentiment analysis, we should formally define opinion (Liu, 2012).

Definition (Opinion):

An opinion is a quadruple (e, s, h, t) , where the subsequent elements denote:

e – opinion target, i.e. an entity about which an opinion is expressed,

s – sentiment about the target, which can be positive, negative, neutral or combination of these,

h – opinion holder,

t – time when the opinion was stated.

In general, every component in the definition above constitutes an important part. Undoubtedly, opinion target e is essential. Sentiment can be depicted as a category, i.e. a negative, neutral or positive opinion. It can also be presented on an ordinal scale as a rating, or a two-dimensional vector with positive and negative opinions recorded, or be of a more advanced form. The practice of quantifying sentiment on a two-dimensional continuous scale, positive and negative, is applied in the sentiment analyzer considered in this work. Opinion holder h is also essential, but depending on a domain it may be of greater importance. For example, when an opinion is expressed by a politician or an institution, it may be more significant depending on the reputation of the speaker. Time stamp t allows us to track how the opinions change in time. In our case, the opinion holder is assigned to the website which is the source of a particular news item. Time stamp is set to the moment in which the news item is published or downloaded, in case when publication date is not available.

The definition above is relatively concise, and it should suffice for our purposes. In such a form, retrieved opinion from unstructured text can be stored in a structured way in a relational database. However, the definition can be further refined. It expresses a general opinion about an entity e . It is easy to see that a particular opinion can be more detailed. For example, if we have an opinion about a camera, it can concern its different aspects: its size, the quality of pictures it takes, the properties of battery, etc.

Given the characterization of an opinion, we can formally define the goal of sentiment analysis (Liu, 2012):

The objective of sentiment analysis: Given an opinion document d , retrieve all the opinion quadruples (e, s, h, t) .

The opinion mining procedure is, in general, a difficult task. Every feature of the opinion from the definition needs to be retrieved, but a mistake in one of the parts disqualifies the correctness of the whole classification. Moreover, making a machine understand the sentiment value of a sentence is hard. Such abstract concepts as humor, irony or sarcasm are difficult to evaluate as even humans interpret them differently. For a more detailed treatment of the field of sentiment analysis we refer to Pang and Lee (2008) and Liu (2012).

Sentiment analysis may constitute a useful tool for many branches of science, as well as for business. It has a few interesting commercial applications. It may well be used in brand, product or project management as a

decision support system. It can be used in a review classification and summarization (Turney, 2002; Pang et al., 2002; Hu and Liu, 2004; Brody and Elhadad, 2010). Usually, before making a purchase, customers want to find out the opinion of other people who have already bought the product. Performing sentiment analysis of reviews is a fast way to summarize people's views on a particular product. Moreover, it can be used to monitor a brand's reputation on the Internet and identify strong points and weaknesses of goods. This can be done fast in an automatic way, which is infeasible for a human, or it would require much work. In turn, it can be used as a decision aid system in product and production management which signals the sentiment toward a particular good. Such information can be used to enhance the current product and engineer it in a way that meets customers' needs. Sentiment analysis may be also used to analyze and address critical points in social networks e.g. at enterprises via analysis of e-mails or periodic reports. This may be a quick way to identify strengths and weaknesses of projects and allow managers to address certain issues in advance (Bramsen et al., 2011; Prieto, 2013). This can be considered as an alternative approach to other ways of managing projects or constructing performance indicators via surveys or interviews see, e.g. Salloum (2013), Bouras and Bendak (2013). Another area of sentiment analysis application is in politics, where public opinion and the way media express opinions plays a major role (O'Connor et al., 2010). Tracking public sentiment towards acts and regulations is of major importance for politicians in the decision making process. Also, based on sentiment analysis of microblog services results of elections can be predicted (Tumasjan et al., 2010; Wang et al., 2012). Sentiment analysis may also be applied as a sub-component technology in recommendation systems and more accurate advertisement placements. An advertisement would be displayed (hidden) when positive (negative) feedback were received in the vicinity of a possible placement of the ad. Yet another area of sentiment analysis application is citation analysis, where it could be used to detect whether an author is supporting or dismissing of another author's work that he or she cites (Small, 2011).

We shall briefly describe the sentiment analyzer which was used to evaluate news items that contained sentiment in our study (Mulder et al., 2004). It is an unsupervised lexicon-based method toward sentiment evaluation. The system aims to exploit cognitive bias in a person's decision making process. More precisely, it aims at the affect and availability heuristics as a source of cognitive bias. The affect heuristic concerns the fact that a personal view might be biased due to subjectivity expressed in financial news. The availability heuristic states that people's judgments are biased by the ease to recall particular examples. With the Internet appearing as a widely used and influential medium nowadays, the affect and availability heuristics could bias people's view on the stock market prices and as a consequence, have impact on their investment decisions. In turn, these biased decisions influence prices of particular stocks.

We proceed to description of the system. For a chosen monitored entity (i.e. a company) and a particular piece of news, the sentiment analyzer operates by searching for relevant sentences in the text, i.e. sentences which mention the entity. Next, it measures sentiment in the sentence toward that entity. The way it determines sentiment

is based on a dictionary with affective words. The words are assigned weights ranging from 1 (maximal positive sentiment) to -1 (maximal negative sentiment). The words in the dictionary are assigned weights in accordance with the system inventors' subjective evaluation of sentiment contained in them. In total, the dictionary has roughly 1400 entries. The system accounts for negations, which change the sign of a particular weight value. Further, the distance (measured by the number of words) between the name of an entity and an affective word (a word appearing in the dictionary with a non-zero weight) is taken into account – the greater the distance, the less impact the word has on the emotional tone of a sentence. The total sentiment value of a given sentence is the sum of sentiment values of the words appearing in it. Affective values aggregated from a piece of news yield two scores, which indicate its positive and negative sentiment. At the final stage these scores are normalized so that they are contained in the unit interval [0, 1]. It should be clear that scores derived by the system include substantial noise. The data we are working with is not perfect. Nevertheless, we shall treat it as the first approximation of sentiment toward the companies under consideration.

Having introduced the key method of our analysis, we present experimental results in the next section.

4. Quantitative Study of Sentiment and its Relation to Stock Returns

In this section we present the data, setup of our experiment and results. We start with describing data. Next, we present the results of regression analysis of stock returns against sentiment data.

4.1. Data

In our analysis we focus on several companies whose shares are traded on Euronext stock exchange. Financial data are obtained from the Yahoo! Finance website (<http://finance.yahoo.com/>, retrieved on December 7, 2014.). The period of the analysis covers four years and a quarter, from January 1, 2009 to March 31, 2013. The data, derived by means of sentiment analysis discussed in the previous chapter, we refer to as sentiment data, i.e. sentiment scores derived from news items for particular monitored entities (data were provided by Sentient Information Systems, <http://www.sentient.nl>). A pair of a positive and negative score we refer to simply as scores. The starting and closing date of the analysis are dictated by the availability of sentiment data.

The size of a company, its popularity and the impact on the economy constitute factors that influence the media interest. Therefore, some companies receive more attention than others. In order to obtain meaningful results, we decided to choose the companies with the highest number of relevant news items (articles), i.e. the ones containing a company name at least once. The number of relevant news items per each company during the period of investigation is summarized in Table 1.

Table 1 reveals that there is significant variation in the number of relevant news items per company.

Table 1. The amount of relevant news items per company

No.	Company	Number of relevant news items
1	AEGON	22210
2	Ahold	56504
3	Akzo Nobel	4006
4	ASML	9749
5	Corporate Express	531
6	DSM	9754
7	Hagemeyer	166
8	Heineken	21870
9	ING	69784
10	KPN	38234
11	Philips	31333
12	Randstad	22561
13	Royal Dutch Shell	2663
14	SBM Offshore	6599
15	TNT	19918
16	TomTom	19069
17	Unibail Rodamco	770
18	Unilever	37691
19	Vedior	388
20	Wolters Kluwer	3587

The largest number of articles regards ING – 69784 – and the smallest number Hagemeyer – only 166. In the database, there are given scores for 1541 out of 1551 days during the investigated period between January 1, 2009 and March 31, 2013. For a few days scores are not present (from June 27, 2009 till June 30, 2009, November 30, 2009 and from February 24, 2010 till February 28, 2010). This gives an average number of 13.55 relevant articles per company per day.

We decided to focus on the following list of 10 companies: AEGON, Ahold, Heineken, ING, KPN, Philips, Randstad, TNT, TomTom and Unilever. Our choice is motivated by the availability of sentiment scores for these entities. On average, for this set of companies, we have the number of 22 scored articles per monitored entity per day.

Along with the number of articles per day we were also interested in free float of a company's shares. It is defined as the fraction of all shares available to the public (i.e. individual investors who own less than 5% of the total number of shares), excluding company insiders and institutional investors. It is an important factor taken into consideration in our analysis. The idea behind our research problem is that media sentiment influences investors' sentiment. Our goal is to choose companies whose shares are distributed among many financiers who presumably are following the news and consider it as an element in the decision making process of portfolio management. Therefore, we choose the companies with high values

of free float. Low value of free float indicates that the shares are owned by relatively few investors. Their decisions influence prices to a greater extent, which is an undesired effect; volatility of a stock can be attributed to decisions of few major investors. We concluded that on average free float remains constant or changes insignificantly over time. But for Heineken and TomTom with free float around 0.45, all the companies under our consideration have free float over 0.90 in the span of our analysis. Having in mind the possible influence of low values of free float on the volatility of a stock, we will come back to this issue in the subsequent section.

Having retrieved and preliminarily presented the scores, we proceed to create attributes that aggregate the information they provide. The features extracted in this section will be used in model building in later part. Before we proceed to describe derived features, we define what does a single day for data aggregation in our analysis means. Euronext stock market, where the shares of the companies under our consideration are traded, operates on daily basis between 9:00 and 17:30 CET. We decide to establish a 24-hour cycle, starting at 17:00, as a single day for aggregating data in our analysis. For explanatory analysis, we focus on the estimation of daily returns at day using the data from last several days (up to day t) and for predictive modelling we perform prediction of the next day's stock price changes (at day $t+1$) using the same data.

We engineered the following basic features based on sentiment scores:

1. Sum of positive/negative scores – summation over all positive (negative) the scores on a given day.
2. Difference between positive and negative scores – given a score (p, n), where p stands for a positive score and n for a negative one, we derive the difference $p-n$. It can provide an additional insight into the difference between the positive and negative sentiment.
3. z -score for positive/negative/difference scores – this attribute is intended to measure the recent changes of a given score. For a chosen sentiment variable (as defined in point 2. above) on a given day d it is computed according to Eq. (1):

$$z_d = \frac{s_d - \mu_m}{\sigma_m} \quad (1)$$

where s_d denotes the value of the sentiment variable, μ_m and σ_m are the sample mean and standard deviation of the sentiment variable computed with the use of m recent days. This attribute is simply standardized sentiment score.

Fig. 1 illustrates the relations that we look for. For Unilever, daily return and 5-days z -scores for differences in sum of positive and negative sentiment scores are presented on the same plot (for a sample of thirty-seven trading days between May 15, 2012 and July 4, 2012). Returns are expressed in percentage values.

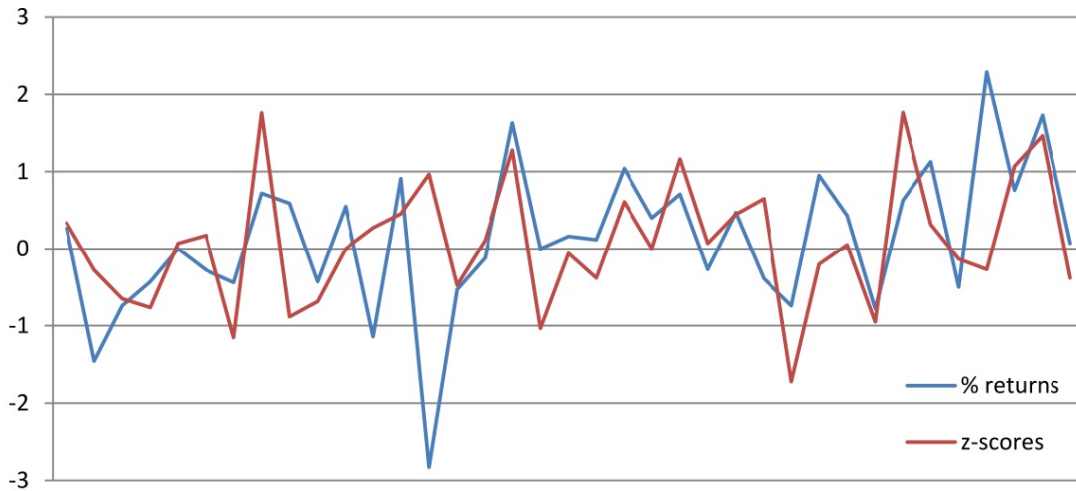


Fig. 1. Daily returns and 5-day z-scores for Unilever

We see that both time series exhibit similar behavior to a certain degree. On several occasions the two series share the pick values at the same moment. It is hard to conclude that one of them determines the other though. There is a moderate positive correlation between the two series equal to 0.35 (at the edge of significance level of 0.05). However, if we correlate one-day lagged returns with z-scores and one-day lagged z-scores with returns, we arrive at correlations equal to 0.08 and -0.07 respectively, which are not significantly different from 0.

4.2. Regression analysis of stock returns against sentiment

To perform regression, we select a set of variables that are suggested by the theory and that are relevant to our study. We use several variables along with their lagged values. To verify whether daily returns can be explained to a certain extent by sentiment in news appearing on the Internet we suggest performing the following regression model of stock returns:

$$r_t = \beta_0 + \beta_1 L(r_t) + \beta_2 p_t + \beta_3 n_t + \beta_4 L(p_t) + \beta_5 L(n_t) + \beta_6 m + \beta_7 f + \beta_8 z_{t,k} + \epsilon_t, \tag{2}$$

where:

r_t – denotes return on day,

p_t, s_t – stand for the sum of positive and the sum of negative scores on a given day t respectively,

m, f - are equal to 1 if day t is Monday or Friday respectively, and 0 otherwise,

$z_{t,k}$ – denotes calculated z-score with the use of last k days for computing the mean and standard deviation,

ϵ_t – is independently and identically distributed Gaussian error term.

Moreover, $L(*)$ is the lag operator, $L(x_t) = x_{t-1}$ and β_i are the coefficients to estimate, $i = 0, 1, \dots, 8$.

The set of chosen variables is motivated by the theory and the goal of our paper. We include previous day’s return to investigate whether the returns are dependent (to remove any possible autocorrelation). The two indicator variables for Monday and Friday were introduced to account for the day of the weekend-effect

(French, 1980). The most interesting variables, from the point of view of our research, are the sums of positive and negative scores. We introduced to the model their values lagged by one day. We believe that older values of the sentiment features are irrelevant (this was also supported by insignificant coefficient estimates for models with additional lags). Finally, we introduced z-score to capture the deviation of day t sentiment with respect to the recent days. This variable is correlated with both positive and negative sentiment scores. This means that it carries similar information. We inspected several values of k , the number of days to compute mean and standard deviation in Eq. (1) as described in previous subsection. However, different choices yielded analogous quantitative and qualitative conclusions with a tendency for diminishing significance of coefficient estimates for larger values of k . Therefore we decided to set $k = 5$.

We set the significance level to 0.05 for inference about significance of estimates. The results of estimation procedure are presented in Table 2.

Table 2. Result of least squares estimation of the linear model given by Eq. (2)

Variable	Coefficient estimate (β)	t-statistic	p-value
intercept	-0.0002707	-0.600	0.548
L(rt)	0.0066005	0.615	0.539
m	0.0002636	0.341	0.733
f	-0.0002145	-0.285	0.775
pt	0.0015399	14.713	<2e-16
nt	-0.0019647	-13.167	<2e-16
L(pt)	-0.0005150	-4.807	1.56e-06
L(nt)	0.0006941	4.593	4.43e-06
zt,5	0.0004621	1.202	0.229

The variables in the model are jointly significant, as indicated by F -test statistic. However, only four variables

are significant individually (in bold in Table 2). Furthermore, we observe no significant day of the week effect. Moreover, the intercept and lagged value of return are insignificant. This is in line with the Efficient Market Hypothesis in the weak form. It states that historical values of stock prices cannot be used for prediction of future values. The significant variables are the sum of positive and negative sentiment scores: p_t , n_t and their values lagged by one day.

Our next step is to simplify the model given by Eq. (2). We test the hypothesis on the joint insignificance of the intercept and other individually insignificant variables in the model. To this end, we estimated the restricted version of the model.

$$r_t = \beta_2 p_t + \beta_3 n_t + \beta_4 L(p_t) + \beta_5 L(n_t) + \varepsilon_t. \quad (3)$$

The results of the estimation are given in Table 3.

Table 3. Result of least squares estimation of the restricted model

Variable	Coefficient estimate (β)	t-statistic	p-value
pt	0.001570	16.857	<2e-16
nt	-0.002057	-15.435	<2e-16
L(pt)	-0.000541	-5.274	1.37e-07
L(nt)	0.000697	4.764	1.93e-06

Based on the estimated models, we compute F statistic and test whether the restriction can be imposed. The p -value of the statistic is equal 0.76, which means that the hypothesis on joint insignificance of reminder variables cannot be rejected. The R^2 statistic for the model given by Eq. (3) is equal to 0.0403, which means that the model explains roughly 4% of the total variance of daily returns. This is a relatively small fraction. However, similar number was reported in an analogous study by Tetlock (2007). Also, comparing both the Akaike and Bayesian information criteria, the simplified model given by Eq. (3) should be preferred to the model given by Eq. (2). We also tried to simplify the model further, in order to aid interpretation. For example, we examine a hypothesis whether we can put $\beta_2 = -\beta_3$ and $\beta_4 = -\beta_5$ in the model given by Eq. (3). However, it was rejected as being too restrictive.

We turn to the interpretation of the results. First of all, the qualitative interpretation of the coefficients associated with variables p_t and n_t is in line with our intuition. The total sum of positive scores contributes to the increase of daily return. Just the opposite applies to the total sum of negative scores. However, the coefficients for the lagged values of these variables, $L(p_t)$ and $L(n_t)$, do not have such an intuitive interpretation. They are also much smaller in magnitude in comparison to the estimates at day t , which suggest that their influence on returns is smaller. A quantitative interpretation of daily returns yields that an increase in the total sum of positive scores on the current day (p_t) by 1 yields 15.6 basis points increase in daily returns.

We shall emphasize an important point about the lexical sentiment analyzer discussed in previous parts. In the dictionary of affective words, there are entries such

as ‘gain’ or ‘profit’ and ‘loss’ or ‘losing’ with non-zero affective values. Although words associated with gaining or losing are relatively scarce in the list of roughly 1400 affective words, this point may partially explain our findings. On the other hand, news regarding these events may well intensify them on a self-feedback loop basis (Shiller, 2005).

The next step is to examine the robustness of our methodology. First of all, the regression of stock returns was performed jointly for ten companies. Because different stocks can exhibit dramatically different behavior, we apply the regression model given by Eq. (3) for each of the companies separately. Table 4 below summarizes the results. Single and double stars denote the estimates significant at the level of 0.05 and 0.01 respectively. In each case, excluding Ahold, the four variables were jointly significant at any standard level of significance (p-value of F less than 0.001). In case of Ahold, p-value of statistic was equal to 0.0107.

When we carry out multiple regressions, we observe consistent qualitative findings for the variables p_t and n_t . The total sum of positive scores is associated with an increased return and the total sum of negative scores with lower returns. For the lagged values of these variables the result are not consistent anymore (in terms of sign of coefficient estimates) and in most cases insignificant. We conclude that the sentiment aggregates at day t (variables p_t and n_t) are correlated with returns. The direction of the relation is in line with intuition. However, their lagged values do not share this property. It supports the claim that any potentially useful information is immediately included in stock prices. Finally, in the previous section we discussed the issue of a low free float. Heineken and TomTom are the companies with the lowest free float of 0.45. Nevertheless, low value of this statistic does not appear to perturb the results and our conclusions.

Regression analysis in this section provided us with a preliminary insight into sentiment data and its connection with daily returns. We observed its correlation with the total sentiment (i.e. the sum of the scores) and daily return. In this way we verified that the sentiment toward a company influences its returns. Though, the relation does not seem to be long-lasting. When we tried to predict next day’s change in returns (increase or decrease) with current day’s financial and sentiment data, the accuracy of predictions were no better than that of a coin flip.

Table 4. Regression model of stock returns against sentiment variables per company

Company	Coefficient estimate for the variable			
	p_t	n_t	$L(p_t)$	$L(n_t)$
AEGON	0.002213**	-0.003522**	-0.000677	0.001814**
Ahold	0.000278*	-0.000604**	-0.000221	0.000257
Heineken	0.001051**	-0.001175**	-0.000295	0.000267
ING	0.002106**	-0.002366**	-0.000723*	0.000845*
KPN	0.000717**	-0.001281**	0.000016	0.000010
Philips	0.001268**	-0.001852**	-0.000294	0.000867*
Randstand	0.001716**	-0.001946**	-0.000011	-0.000557
TomTom	0.003345**	-0.003266**	-0.001656**	-0.000554
TNT	0.008096**	-0.007361**	-0.003482*	0.002720
Unilever	0.001073**	-0.001727**	-0.000446**	0.000959**

5. Conclusion

In this paper we investigated the relation between the news items appearing in the Internet and stock market. With the use of sentiment analysis, positive and negative sentiments in news items were measured. This allowed us to arrive at so-called sentiment scores. The goal of our research was to verify whether sentiment (measured by the scores) towards a chosen company influences its daily stock returns. To verify this claim, we estimated a regression model of stock returns against sentiment scores. We found that there is a relation between the sentiment included in news items and the same day company's returns. More precisely, the sum of positive scores is associated with an increased return, and a reverse relation applies to the sum of negative scores. Given these results we conclude that the sentiment does influence stock returns. The more positive the sentiment, the greater the stock returns. It seems that the sentiment is one of the factors that are taken into account in financial products valuation. The psychology does seem to affect stock prices and to play role on the capital market.

Our paper also shows the potential of application of the sentiment analysis method in order to automatically process text for sentiment evaluation. The analysis revealed that the sentiment in news items appears to influence returns of companies. This could be used for short-term investing in the capital market. However, we note that frequent trading yields high transaction costs. Moreover, since we gathered current day's news items and returns, we would need to make a decision on investing based on the analysis of content of news gathered earlier, perhaps using less data than we did (all the news items appearing before 17:00 prior to closing of the market at 17:30).

There is a growing interest in application of advanced data analysis methods to financial markets. A number

of researchers, as well as individual investors and financial institutions, investigate the potential of trading based on sentiment analysis and predictive power of data derived in this manner. As it happened with previously discovered anomalies that cancelled out, we may observe similar phenomenon regarding such a form of investing in financial markets. If sentiment analysis based trading strategies yield profits, their increased usage should arbitrage off possible gains and increase market efficiency with time.

6. Acknowledgements

The authors would like to thank Sentient Information Systems (Singel 160, 1015 AH Amsterdam, The Netherlands, info@sentient.nl, http://www.sentient.nl/) for providing data for this research.

Jan Lasek would like to acknowledge the support by the European Union from resources of the European Social Fund, Project PO KL "Information technologies: Research and their interdisciplinary applications", agreement UDA-POKL.04.01.01-00-051/10-00 via the Interdisciplinary PhD Studies Program.

References

- Ariely, D. (2008). *Predictably Irrational. Hidden Forces That Shape Our Decisions*. New York, USA: HarperCollins Publishers.
- Baker, M. and Wurgler, J. (2007). Investor Sentiment in the Stock Market, *Journal of Economic Perspectives*, 21(2), 129-152.
- Benzartzi, S. and Thaler, R. (1995). Myopic Loss Aversion and the Equity Premium Puzzle. *Quarterly Journal of Economics*, 110(1), 73-92.
- Bollen, J., Mao, H., Zeng, X.-J. (2011). Twitter Mood Predicts the Stock Market, *Journal of Computational Science*, 2(1), 1-8.

- Borges, M. R. (2010). Efficient Market Hypothesis in European Stock Markets. *The European Journal of Finance*, 16(7), 711-726.
- Bouras, A. and Bendak, S. (2014). Managing IT Projects in Public Companies: A Case Study. *Journal of Engineering, Project, and Production Management*, 4(2), 74-80.
- Bramsen, P., Escobar-Molano, M., Patel, A., Alonso, R. (2011). Extracting Social Power Relationships from Natural Language. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, Portland, Oregon, USA, 773-782.
- Brody, S. and Elhadad, N. (2010). An Unsupervised Aspect-Sentiment Model for Online Reviews. *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Los Angeles, California, USA, 804-812.
- Brown, G. W. and Cliff, M. T. (2004). Investor Sentiment and the Near-Term Stock Market. *Journal of Empirical Finance*, 11(1), 1-27.
- Chaney, P. K. Devinney, T. M. (1992). New Product Innovations and Stock Price Performance. *Journal of Business Finance and Accounting*, 19(5), 677-695.
- De Bondt, W. and Thaler, R. (1985). *Does the Stock Market Overreact?* *Journal of Finance*, 40(3), 793-805.
- den Heijer, H. (2002). The VOC and the Exchange. *How the VOC Laid the Foundation for the World's First Stock Exchange*. Amsterdam, The Netherlands: Aksant Academic Publishers.
- Edmans, A., Garcia, D., Norli Ø. (2007). Sports Sentiment and Stock Returns. *Journal of Finance*, 62(4), 1967-1998.
- Fama, E. (1970). Efficient Capital Markets: a Review of Theory and Empirical Work. *Journal of Finance*, 25(2), 383-417.
- French, K. R. (1980). Stock Returns and the Weekend Effect. *Journal of Financial Economics*, 8(1), 55-69.
- Godfrey, M. D., Granger, C. W. J., Morgenstern, O. (1964). The Random-Walk Hypothesis of Stock Market Behavior. *Kyklos*, 17(1), 1-30.
- Guo, H. and Whitelaw, R. (2006). Uncovering the Risk-Return Relation in the Stock Market. *Journal of Finance*, 61(3), 1433-1463.
- Hollum, A. T. G., Mosch, B. P., Szlavik, Z. (2013). Economic Sentiment: Text-Based Prediction of Stock Price Movements with Machine Learning and WordNet. *Recent Trends in Applied Artificial Intelligence*, Lecture Notes in Computer Science, 7906, 322-331.
- Hu, M. and Liu, B. (2004). Mining and Summarizing Customer Reviews. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Seattle, Washington, USA, 168-177.
- Johnson, E. and Thaler, R. (1999). Gambling with the House Money and Trying to Break Even. *Management Science*, 36(6), 643-660.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, USA: Macmillan.
- Kahneman, D. and Tversky, A. (1974). *Judgment under Uncertainty: Heuristics and Biases*, Science, 185(4157), 1124-1131.
- Kaplanski, G. and Levy, H. (2010). Exploitable Predictable Irrationality: The FIFA World Cup Effect on the U.S. Stock Market, *Journal of Financial and Quantitative Analysis*, 45(2), 535-553.
- Keim, D. B. (1983). Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence, *Journal of Financial Economics*, 12(1), 13-32.
- Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*, London, United Kingdom: Macmillan.
- Lasek, J. (2013). Is the Stock Market More Sentimental than Efficient? *Master Degree Thesis in Information Technology and Econometrics*, Faculty of Economic Sciences, University of Warsaw, Warsaw, Poland.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Retrieved from <http://www.cs.uic.edu/~liub/FBS/SentimentAnalysis-and-OpinionMining.pdf> on December 7, 2014.
- Lo, A. W. and MacKinlay, A. C. (2011). *A Non-Random Walk Down Wall Street*. Princeton, New Jersey, USA: Princeton University Press.
- Malkiel, B. (2007). *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing*. New York, USA: W. W. Norton & Company.
- Mitchell, M. and Stafford, E. (2000). Managerial Decisions and Long-Term Stock Price Performance. *Journal of Business*, 73(3), 287-329.
- Mulder, M., Nijholt, A., den Uyl, M., Terpstra, P. (2004). A Lexical Grammatical Implementation of Affect. Text, Speech and Dialogue, *Lecture Notes in Computer Science*, 3206, 171-177.
- Odean, T. and Barber, B. (1999). The Courage of Misguided Convictions, *Financial Analyst Journal*, 41-55.
- O'Connor, B., Balasubramanyan, R., Routledge, B., Smith, N. (2010). *From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series*, Proceedings of the International AAAI Conference on Weblogs and Social Media, Washington, USA, 122-129.
- Pang, B. and Lee, L. (2008). *Opinion Mining and Sentiment Analysis, Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Pang, B., Lee, L., Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification Using Machine Learning Techniques, *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 79-86.
- Peress, J. (2010). Product Market Competition, Insider Trading, and Stock Market Efficiency. *Journal of Finance*, 65(1), 1-43.
- Prieto, B. (2013). Utilization of Project Sentiment Analysis as a Project Performance Predictor. *PM World Journal*, 2(11), 1-9.
- Reinganum, M. R. (1983). The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects, *Journal of Financial Economics*, 12(1), 89-104.
- Salloum, M. (2013). Explaining the Evolution of Performance Measures - A Dual Case-Study Approach. *Journal of Engineering, Project, and Production Management*, 3(2), 99-106.
- Sewell, M. (2011). *History of the Efficient Market Hypothesis*. Research note retrieved from http://www.cs.ucl.ac.uk/fileadmin/UCL-CS/images/Research_Student_Information/RN_11_04.pdf on December 7, 2014.
- Sewell, M. (2012). The Efficient Market Hypothesis: Empirical Evidence. *International Journal of Statistics and Probability*, 1(2), 164-178.

- Schwert, G. W. (2003). Anomalies and Market Efficiency. *Handbook of the Economics of Finance*, 1(1), 939-974.
- Shiller, R. J. (1981). Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends? *American Economic Review*, 71(3), 421-436.
- Shiller, R. J. (2005). *Irrational Exuberance*. Princeton, New Jersey, USA: Princeton University Press.
- Small, H. (2011). Interpreting Maps of Science Using Citation Context Sentiments: A Preliminary Investigation. *Scientometrics*, 87(2), 373-388.
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance*, 62(3), 1139-1168.
- Thaler, R. H. (1999). The End of Behavioral Finance. *Financial Analysis Journal*, 55(6), 12-17.
- Tumasjan, A., Sprenger, T., Sandner, P., Welpe, I. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, Washington, USA, 178-185.
- Turney, P. D. (2002). Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, Philadelphia, Pennsylvania, USA, 417-424.
- Wang, H., Can, D., Kazemzadeh, A. Bar, F., Narayanan, S. (2012). A system for real-time Twitter sentiment analysis of 2012 U.S. Presidential Election Cycle. In *Proceedings of the ACL 2012 System Demonstrations*, Jeju Island, Korea Republic, 115-120.
- Zielonka, P. (2004). Technical Analysis as the Representation of Typical Cognitive Biases. *International Journal of Financial Analysis*, 13(2), 217-225.



Professor Mirosława Lasek passed away in 2014. She was Professor at University of Warsaw, Poland. For many years she was the head of Department of Information Systems and Economic Analysis at the Faculty of Economic Sciences. Professor Mirosława Lasek is an author of numerous publications in areas of business process analysis, design and development of enterprise information systems, application of statistical, econometric and data mining models in decision support and forecasting at enterprises and banks, to mention a few expertise areas. Being also a kind, generous and willing to help person Professor Lasek was appreciated by her colleagues and her students who eagerly attended lectures given by her. She also mentored and promoted many PhD and Master students at the Faculty of Economic Sciences.



Jan Lasek is a PhD student at Interdisciplinary PhD Studies at Institute of Computer Science, Polish Academy of Sciences, Warsaw, Poland. He holds Master Degree in Economics from University of Warsaw and Master Degree in Mathematics from VU University Amsterdam and University of Warsaw which he completed during Joint Master's Program between the two universities. His current research focuses on preference modelling, construction of evaluation metrics for information resources and ranking algorithms. He is also interested in applications of statistical, data mining and machine learning models in diverse areas as finance or sport analytics.