Research Report

Analyzing the Relationship between Differentiated Online Sentiment and Company-Specific Stock Prices

PRACTITIONERS AND RESEARCHERS ALIKE INCREASINGLY USE SOCIAL MEDIA MESSAGES AS AN ADDITIONAL SOURCE OF INFORMATION WHEN DEALING WITH STOCKS. BASED ON EMOTION THEORY AND AN ESTABLISHED SENTIMENT LEXICON, WE DEVELOP AND APPLY AN OPEN SOURCE DICTIONARY FOR THE ANALYSIS OF SEVEN DIFFERENT EMOTIONS IN 5.5 MILLION TWITTER MESSAGES ON 33 S&P 100 COMPANIES. WE FIND VARYING EXPLANATORY POWER OF DIFFERENT EMOTIONS (ESP. HAPPINESS, AND DEPRESSION) FOR COMPANY-SPECIFIC STOCK PRICE MOVEMENTS OVER A PERIOD OF THREE MONTHS.

Marten Risius

Fabian Akolk

Roman Beck

Introduction

Comprehensive and immediate information plays a crucial role in stock price analysis. In this regard, researchers and practitioners alike increasingly consider online user-generated content as an additional source of information for investment decision-making. Events like the loss of approximately USD 136 billion in equity market value within three minutes due to a fake tweet from the hacked Associated Press Twitter account demonstrate the interwovenness between Social Media and stock markets. Social Media also affects the stock market on a more regular basis, seeing that the New York Stock Exchange (NYSE) introduced an automated sentiment analysis of Social Media platforms to provide investors with real-time information on indices, industry sectors, and specific companies.

Company-specific tweet sentiments have been found to intervene with investor decision-making, affect market prices through information leakage, and have explanatory power for daily stock price changes (e.g., Bollen et al., 2011). So far, the majority of research has ignored the more complex, multi-dimensional structure of human emotions and only considered aggregated sentiment measures. The few studies that considered differentiated emotions found differential effects for positive and negative messages on stock prices (e.g., Sprenger et al., 2014). However, this research is generally limited to general market indices, included very few emotion words, has neglected the predictive value of specific emotions, and withheld specifics on the operationalization of the emotions.

In this study, we address this research gap by developing an open source emotion-specific dictionary which is derived from the established SentiStrength word list. It enables us to assess seven different emotions whose operationalization is based on the model of the hierarchical structure of the affective domain developed by Ekkekakis (2013). Subsequently, this sentiment analysis is applied to 5.5 million Twitter messages on 33 S&P 100 companies which we collected over a three-month period. Ultimately, we conduct a lagged panel regression of the differentiated emotion strength on the company-specific NYSE stock price movements obtained from Yahoo!Finance. Overall, in this study, we investigate how differential emotions correspond to company-specific stock price movements.

Developing an Emotion-Specific Sentiment Analysis

When applying automated sentiment analysis, Information Systems researchers have predominantly focused on measuring the average emotionality (positive vs. negative). However, the undifferentiated dimensional approach implies a lower degree of specificity which can only be overcome through the assessment of distinct emotional states. In a critical reconsideration of emotion theory research, Ekkekakis (2013) integrated different emotion concepts into one model of the hierarchical structure of the affective domain. In this study, we rely on the model's refined emotional states in addition to the basic valence (positive vs. negative) dimension.

To develop the differentiated sentiment analysis tool, we draw on the established "SentiStrength 2" dictionary which provided the emotion words with over 2,300 sentiment words (Thelwall et al., 2012). Three independent coders classified these words into the seven different emotions (affection, happiness, satisfaction, fear, anger, depression, and contempt) described by Ekkekakis (2013). The underlying description of the classification can be found in Table 1. We provide open access to the differentiated sentiment analysis with the emotion specific coding here: http://bit.ly/1BpocLl.

Differential Sentiment Analysis and Stock Price Movements

Online user-generated content is increasingly relevant as a source of information for investment decision-making. Social Media content in general and company-specific tweet sentiments in particular have been found to inter-

Emotion: Valence	Emotion: State	Description	Emotion Word
Positive	Affection	Genuine fondness and liking that is attributed to a particular person or object.	Love, Adoration
	Happiness	Amplified enthusiasm and excitement about attaining something desired or desirable.	Joy, Terrific
	Satisfaction	Proud acknowledgement of and contentment with reaching a predetermined goal.	Pride, Contentment
Negative	Fear	Anticipatory horror or anxiety in unpredictable or potentially harmful situations.	Horror, Anxiety
	Anger	Animated animosity towards malice that can motivate rectification.	Hate, Outrage
	Depression	Impeding sadness evoked by an aversive event that may hinder activity.	Sadness, Hopeless
	Contempt	Revulsion to something considered socially offensive or unpleasant.	Guilt, Disgust

Table 1: Overview of the Seven Different Emotions and Their Operationalization

vene with investor decision-making, affect market prices through information leakage, and have predictive power for daily stock price changes (e.g., Li et al., 2014). However, recently researchers discovered differential effects for positive and negative messages on stock prices and an improvement of the predictive validity on global market indices by considering specific emotions while the undifferentiated sentiment only shows a poor correlation with company stocks (e.g., Sprenger et al., 2014). Thus, we hypothesize:

Hypothesis 1: The average message sentiment about a company is unrelated to the company's stock prices.

Comparisons of the differential effects of positive and negative sentiment valence on Social Media platforms show that negative messages spread more easily than positive news and receive more attention. Consequently, company-specific negative user-generated content is more thoroughly processed and has more explanatory power of abnormal returns than positive information. Accordingly, we conclude:

Hypothesis 2: The stronger the negative valence of the message sentiment about a company, the lower its stock prices, whereas the positive sentiment is unrelated to company stock prices.

Only a small share of research has investigated the specific effects of different emotions on stock prices. These findings, however, can be generally summarized in the sense that different emotions (mostly the negative ones) have differential effects on stock prices. Especially fear, depression, and anger have been linked with a downward pressure on stock prices and trading volume. We therefore expect:

Hypothesis 3: The strength of the differentiated message sentiment about a company is related to company stock price variations.

Empirical Investigation

In order to empirically analyze the relationship between different emotions and company-specific stock prices, we collected 5.5 million tweets and 61 NYSE daily closing values on a random sample of 33 S&P 100 companies over a three-month period. We computed the average sentiment (Model 1), the average strength of positive and negative emotions (Model 2), and the average strength of the differentiated emotions (Model 3) as well as the daily closing value difference per company for each day. Afterwards, we conducted fixed effects panel regressions with robust standard errors to test our hypotheses (Table 2).

The average sentiment (Hypothesis 1) of company-related tweets showed no explanatory power for the stock price movements.

The emotional valence specific analysis (Hypothesis 2) shows that the stronger the negative sentiment towards a company, the lower its stock price, while the strength of the positive sentiment was unrelated to stock price movements. This finding – as opposed to the case of an aggregated sentiment – shows the necessity for a more differentiated sentiment analysis especially when analyzing company-specific effects.

The subsequent simultaneous consideration of the distinct emotions (Hypothesis 3) shows more precisely that depression and happiness are significantly associated with stock price movements. Considering that happiness has a significant effect while the average positivity is not associated with stock prices shows again the importance of investigating single emotions separately.

Generally, our results support our underlying assumption that the differentiated sentiment offers additional explanatory power for the company-specific stock price developments.

Discussion of the Results

The goal of this study was to analyze the explanatory power of differentiated emotions expressed in tweets for company-specific stock prices. Specifically, we focused separately on emotions with positive (affection, happiness, satisfaction) and negative valence (fear, anger, depression, contempt). Based on established emotion research (Ekkekakis, 2013) and sentiment analysis (Thelwall et al., 2012), we developed and applied an open source emotion-specific dictionary that also considers the underlying valence and activity dimensions. By analyzing daily closing values of 33 S&P 100 companies over the period of three months, this study provides three key findings: (1) the differentiated emotions are more strongly associated with company-specific stock price changes than the undifferentiated average sentiment, (2) negative emotions generally have a higher explanatory power, and (3) especially the strength of emotions referring to specific events

Predictor Variables	Model Statistics		
	Coefficient	Standard error	t-Value
Model 1	$F_{10,32} = 63.81^{***}$, $R^2_{within} = 35.8\%$		
Average	.002	(.001)	1.55
Model 2	$F_{11,32} = 56.46^{***}$, $R^2_{within} = 36\%$		
Positive	.0001	(.002)	-0.12
Negative	-.003**	(.002)	2.11
Model 3	$F_{16,32} = 58.24^{***}$, $R^2_{within} = 36.26\%$		
Affection	.0003	(.003)	0.1
Happiness	-.007*	(.004)	-1.83
Satisfaction	.004	(.003)	1.29
Fear	.0006	(.002)	0.29
Anger	.001	(.002)	0.66
Depression	.009***	(.003)	2.89
Contempt	.003	(.004)	0.79

Notes: Model 1 = Average Sentiment; Model 2 = Average Strength of Positive and Negative Sentiment;
Model 3 = Average Strength of Differentiated Emotions
Each model controlled for weekdays, mean S&P 500 return, pre-holidays, and earnings releases

Table 2: Results of the Fixed Effects Panel Regressions to Test our Hypotheses

(depression and happiness) account for price movements.

Considering the theoretical foundation of these measures, it seems that – while general stock market indices are influenced by the anticipation of hypothetical aversive events (i.e., fear) – company stocks are only influenced by events that have actually occurred. This would also explain why happiness – which constitutes the conceptual opposite of depression – also affects stock prices significantly. The surprisingly negative effect of happiness could be explained by the distinction between immediate and expected emotions. The respective literature suggests that positive emotions might make investors more risk-avoidant by trying to avoid a disturbance of positive feelings. In a similar vein, researchers found evidence that the amount of expressed emotions and not the specific type (i.e., fear, worry, and hope) causes a market index decrease.

Our study offers substantial contributions to research and practitioners alike. The differentiated sentiment analysis developed in this work overcomes existent limitations of the few other differentiated sentiment analyses which have not considered the strength of an emotion, do not respect the exclusiveness of emotion states, or withhold detailed insights into the classification of emotions (e.g., Bollen et al., 2011). On the contrary, we provide access to the dictionary for practitioners to apply.

The evidence presented for the necessity of a more differentiated sentiment analysis is equally relevant for practice considering that NYSE does provide sentiment scores on stocks and industry branches which, however, are limited to the binary emotional valence.

The implications of this study must be considered in the light of their limitations that also provide a basis for future research. The results are limited in their generalizability to microblogging platforms and to the western culture since we only considered tweets in English and NYSE stock prices. Moreover, it could be assumed that the larger number of negative than positive words present on our dictionary might cause bias towards the bigger influence of negative emotionality (e.g., Nielsen, 2011). Future research will need to compare the impact of single actually identified words and the number of words within sets of message.

Furthermore, Sprenger et al. (2014) found time-related effects for positive and negative emotions. Future research should analyze potential intraday and day outlasting effects of differential emotions. Also, the interplay of different emotions needs to be considered as, for example, depression has been found to have a competing effect to anger on risk-taking. Lastly, we intend to analyze whether different emotions are more important in other environments such as customer care, where anger might be expressed more openly.

References

Bollen, J.; Mao, H.; Zeng, X.:
Twitter mood predicts the stock market.
In: Journal of Computational Science, 2 (2011) 1, pp. 1-8.

Ekkekakis, P.:
The measurement of affect, mood, and emotion: A guide for health-behavioral research. Cambridge University Press, Cambridge, 2013.

Li, Q.; Wang, T.; Li, P.; Liu, L.; Gong, Q.; Chen, Y.:
The effect of news and public mood on stock movements.
In: Information Sciences, 278 (2014) 25, pp. 826-840.

Nielsen, F. Å.:
A new anew: Evaluation of a word list for sentiment analysis in microblogs.
In: Proceedings of the ESWC2011, Workshop on Making Sense of Microposts: Big Things come in Small Packages, Crete, Greece, 2011.

Sprenger, T. O.; Sandner, P. G.; Tumasjan, A.; Welpe, I. M.:
News or noise? Using twitter to identify and understand company-specific news flow.
In: Journal of Business Finance & Accounting, 41 (2014) 7-8, pp. 791-830.

Thelwall, M.; Buckley, K.; Patoglou, G.:
Sentiment strength detection for the social web.
In: Journal of the American Society for Information Science and Technology, 63 (2012) 1, pp. 163-173.