The Long-run Impact of Sentiment on Stock Returns*

Matthias W. Uhl†

May 2011

Abstract

We examine the explanatory and predictive power of fundamental macroeconomic and behavioral factors with regards to stock returns of the Dow Jones Industrials Index. With a novel sentiment dataset from over 3.6 million Reuters news articles, we find significant correlations between Reuters sentiment and stock returns. We show with vector autoregression and error correction models that Reuters sentiment can explain and predict changes in stock returns better than macroeconomic factors. Considering positive and negative sections of Reuters sentiment, we find that negative sentiment performs much better in simple trading strategies to predict stock returns than positive sentiment.

Keywords: Reuters sentiment, stock returns, out-of-sample forecasts

JEL classifications: G11, G14, G17

*Thanks go to Thomson Reuters for the sentiment dataset. I am grateful to Didier Sornette and Jan-Egbert Sturm for helpful comments and suggestions. Thanks also go to participants of the internal research seminar at MAN Investments for their feedback. I thank Edward Fridael for his IT support.

†KOF Swiss Economic Institute, ETH Zurich, 8092 Zurich, Switzerland, Tel. +41-44-632-2553, Fax +41-44-632-1218, E-mail: uhl@kof.ethz.ch

The Efficient Market Hypothesis (EMH), first introduced by Fama (1970), has been questioned widely on the grounds of psychological phenomena occurring in financial markets. Financial economists and psychologists alike have devoted time to research that relates sentiment among investors to financial market returns.

In this light, we want to introduce another way of explaining and predicting stock returns, undermining the EMH. In this paper, we test whether Reuters sentiment is able to explain changes in stock prices. With *Reuters sentiment,* we mean a (positive or negative) feeling, opinion, or emotion evoked among a reader while reading a certain Reuters news article. Tetlock's (2007) study and findings serve as motivation, as we identify the need to not only consider the predictive power of negative sentiment on stock returns, but also of positive sentiment as well as combined (positive and negative) sentiment. The dataset used in this study is novel and unique. Using sentiment in Reuters news and a macroeconomic indicator, we build Vector Error Correction Models (VECM) and simple trading strategies based on out-of-sample forecasts to test the predictive accuracy of the models. We find that negative sentiment predicts stock returns better than positive and combined sentiment, confirming Tetlock's (2007) findings that negative sentiment best predicts stock returns.

Section I gives an overview of the existing literature and lays out the motivation. Section II describes the dataset, while section III discusses the econometric modeling approach and the empirical results of the specified models. Section IV lay out simple trading strategies based on out-of-sample forecasts. Section V concludes.

I. Related Literature

Since the late 1980s, when the first studies emerged that postulated irrationality in financial markets, the domain of behavioral finance has introduced ways to explain that irrationality. Kahneman and Tversky (1981) find that subjects overreact to new information in making probabilistic judgments. Based on the same grounds, Shiller (1981) notes that financial markets display excess volatility and overreaction to new information. Summers (1986) then posed the question whether the stock market rationally reflects fundamental values and came to the conclusion that most tests of market efficiency have had little power to solidify the EMH, suggesting that excess volatility and negative autocorrelation can produce a deviation of the price in a rational fundamental market. Further, he elaborates, certain types of inefficiency in market valuations are not likely to be detected using standard methods. Thus, one should not conclude erroneously that market prices represent rational assessments of fundamental valuations based on the grounds that many studies have found that the EMH cannot be rejected. One of the first studies that attempted to link other exogenous variables to financial market returns was undertaken by De Bondt and Thaler (1985). They show that, based on research in experimental psychology, overreaction occurs mainly when unexpected and dramatic news events happen. A few years later, Cutler *et al* (1989) identified a link between news coverage and stock prices. Since then, studies have evolved that look at the potential influence that the media has on investor behavior.

The growing evidence in the finance literature about news affecting investors and thus stock returns is key motivator for this study. DeLong *et al* (1990) are among the first to find that investors are subject to news. In their model, two sets of traders in the financial markets exist: professional arbitrageurs and unsophisticated traders, i.e. noise traders. The prevailing risk in the market, they find, is created by the unpredictability of the noise traders. Professional arbitrageurs respond to the behavior of noise traders rather than acting on fundamentals. In doing so, professional arbitrageurs consider pseudo signals such as volume and price patterns, but also news. With the growing importance of the media in financial markets globally, we can assume that the news effect is becoming more important. Assuming that markets are not efficient, examining under- and overreaction in stock prices due to news releases becomes then even more apparent. Barberis *et al* (1998) show that news can cause both over- and underreaction to stock prices by formulating a parsimonious model of investor sentiment. They claim that news are incorporated only slowly into stock prices. Their findings make the case for a lower frequency, i.e. monthly, analysis that we conduct in this study.

Other studies have identified a variety of behavioral aspects of stock investors with regards to news. For example, Klibanoff *et al* (1998) show that country-specific news reported on the front page of the *New York Times* affect the pricing of closed-end country funds. Huberman and Regev (2001) find that an article in the *Financial Times* on a biochemical firm made prices of that company soar. Antweiler and Frank (2004) consider the influence of Internet stock message boards. They find that stock messages predict market volatility. The above mentioned studies make the case for examining the impact of news closer, as news appear to have an effect on investors, which should be reflected in stock returns movements. We want to dig deeper and consider how news are written and portrayed.

In a journalistic study, Maier (2005) notes that 61% errors in local news and feature stories in the US, while subjective errors are considered most severe. Maier's results suggest that how a story is conveyed is at least as important as getting the facts straight. The results of these studies strongly speak for examining news reports for sentiment, and using the sentiment values to explain changes in stock prices. In their extensive study on the news media, Mullainathan and Shleifer (2005) identify that there are biases in economic and political news and that these are slanted towards the customers of the media outlet. Given these findings, it appears relevant that sentiment in news plays a crucial role in the decision process of investors who follow news.

Baker and Wurgler (2007) argue that the key nowadays for researchers is to find out how to measure investor sentiment and quantify its effects. Owing to the quest for more accuracy in explaining financial market returns from a behavioral point of view, studies have been aiming towards the quantification of sentiment recently. Thus, we introduce and test a new dataset that measures sentiment quantitatively in a systematic way, while trying to avoid subjectivity bias. With the growing importance of the media in the past decades, the obvious publicly available information are news, as De Bondt and Thaler (1985) as well as Cutler *et al* (1989) noted as early as a few decades ago. Based on these initial findings, we focus on news relevant to investors, such as Reuters news reports.

More recently, some researchers have looked at the quantification of sentiment in media reports. Tetlock (2007) is one of the first to quantitatively measure the interactions between the media and the stock market using daily content from a *Wall Street Journal* column. High media pessimism, he finds, predicts falling stock market prices followed by a reversion to fundamentals. Unusually high or low pessimism predicts high trading volume as well. In a follow-up to Tetlock's (2007) study, Tetlock *et al* (2008) use a simple quantitative measure of language to predict individual firms' accounting earnings and stock returns. Linguistic media content, they conclude, captures aspects of firms' fundamentals that are otherwise hard to quantify, which are quickly incorporated into stock prices. Fang and Peres (2009) investigate the cross-sectional relation between media coverage and expected stock returns. They find that stocks with no media coverage earn higher returns than stocks with high media coverage even after controlling for well-known risk factors. Their results are more pronounced among small stocks and stocks with high individual ownership, low analyst following as well as high idiosyncratic volatility. Given their findings, this suggests that the breadth of information dissemination affects stock returns. On a similar note, Livnat and Petrovits (2009) examine whether stock price reactions to earnings surprises and accruals vary systematically with the level of of investor sentiment. By formulating a monthly trading strategy, they find evidence that holding extreme good news firms following pessimistic sentiment periods earns significantly higher abnormal returns than holding extreme good news firms following optimistic sentiment periods. These results indicate that investor sentiment influences the source of excess returns from earnings-based trading strategies.

As Baker and Wurgler (2007) point out, it is no longer questionable whether sentiment affects investors and thus stock returns, but rather how to measure sentiment. Many studies have emerged in the past years attempting to tackle the issue of defining sentiment that influences stock markets and, more importantly, measuring it.[1] This study introduces a novel dataset and approach to measure sentiment in Reuters news. Therefore, we follow Tetlock's (2007) methodological approach of measuring sentiment in the media quantitatively. Tetlock uses the General Inquirer (GI), a quantitative content analysis program.[2] As explained in the appendix in Tetlock (2007), the GI has one major shortcoming: it is only able to distinguish between positive and negative words, or sentiment categories, but not between context. As opposed to Tetlock's (2007) dataset, the sentiment classifier used in this study is able to account for both individual words and context in the sentiment analysis through cutting-edge technology developed by Thomson Reuters.

In his recent study, Tetlock (2011) tests whether investors distinguish between old and new information about firms, or, what he calls the "staleness of news." A firm's return on the day of stale news negatively predicts its return in the following week, which speaks for the fact that individual investors overreact to stale information, leading to temporary movements in firms' stock prices. In our dataset, we are able to account for the issue of stale news, as every news item is coded accordingly in order to avoid this pitfall.

[1] See, for example, Cao and Wei (2005), Edmans et al (2007), Hirshleifer (2001), Hirshleifer and Shumway (2003), Kamstra et al (2003), and Yuan et al (2006), among others.

[2] See *The General Inquirer Home Page*, available at http://www.wjh.harvard.edu/~inquirer/, last accessed 23 November 2010.

II. Dataset

As opposed to Tetlock's (2007) dataset, we want to analyze both positive and negative sentiment in relation to stock returns. The sentiment scores are not only obtained through simply coding positive and negative words according to a database. Owing to new technological advance in text mining, Thomson Reuters is able to undertake a sentiment analysis that takes the context into account. For example, the sentiment algorithm is able to distinguish between negative words and negations of positive words. "Good" would be categorized as positive in the sentiment analysis, but "not good" would be classified as negative. This has not been possible so far in textual mining programs that are based on a pre-defined databases of positive and negative words only. Thus, we want to contribute to the literature with a more precise methodological approach as opposed to earlier studies.

Based on this dataset, we introduce the concept of measuring sentiment in Reuters news articles quantitatively in order to explain stock returns. Every Reuters news article is coded with positive $\{1\}$, neutral $\{0\}$, or negative $\{-1\}$ sentiment. In the past, most solutions have come from the text mining industry that caters to the financial markets industry, in which news texts can be scanned in great quantities and a short amount of time for sentiment with specific sentiment algorithms. Thomson Reuters is one of the few providers of sentiment classified news.[3] The dataset at hand consists of high-frequency (tick data) sentiment rated Thomson Reuters news pieces, classified from a wide list of topics for the US market.[4] For this study, we filter all Reuters news items for sentiment from the Equities topic codes section.[5] In order to test and validate Tetlock's (2007) findings that negative words predict falling stock returns, we extract both positive and negative sentiment values in order to form two independent time-series. We then aggregate the tick sentiment scores to monthly values. Also, the dataset can account for the issue of staleness as described in

[3] See Thomson Reuters News Analytics, http://thomsonreuters.com/products_services/financial/financial_products/quantitative_research_trading/news_analytics, last accessed 7 September 2010.

[4] The topics range from financial market to economic and political news, categorized into topic codes. See *Reuters Codes - A quick guide*, available at https://customers.reuters.com/training/trainingCRMdata/promo_content/ReutersCodes.pdf, last accessed 9 December 2010.

[5] We filter for "U" in the product code section, and for "DIV, MRG, RES, RESF, RCH, STX" in the topic code section. These codes mean that we filter for news related to dividends, ownership changes, broker research, corporate results, results forecasts and stock markets for North American companies.

Tetlock (2011) because the sentiment algorithm is able to tag each news item with a unique time stamp and topic identifier, so that repeatedly reported news items are not considered again in the analysis.

Table I shows the number of news pieces that were tagged; in total, over 3.6 million Reuters news items were coded for sentiment from January 2003 to December 2010.

[insert table I about here]

Monthly price return data for the Dow Jones Industrials stock index were obtained from Thomson Reuters Datastream. The corresponding monthly volume data for the Dow Jones stock index are from MasterData.[6] To capture the real macroeconomic development, we use a time series of the Conference Board Leading Economic Indicators Index. This index consists of a combination of leading indices, such as production, employment, monetary, and consumer data for the US.[7] The advantage over using many different indicators is that one variable is easier to handle in our subsequent model than multiple variables. Given that we attempt to explain stock returns with non-conventional measures - inconsistent with the EMH - such as sentiment, we need to include fundamental facts that are consistent with the EMH to capture all possible channels of influence on the stock index, and to compare the fundamental to the behavioral. The Conference Board Leading Economic Indicators Index appears the most suited for "summarizing" macroeconomic factors in one variable. Monthly data for this indicator were obtained from Thomson Reuters Datastream.

To get a first understanding of the data, we look at the variables graphically in fig. 1. The Dow Jones stock index shows a pattern, in which we can make out the bull market from 2003 to 2008 and the subsequent crash when the financial crisis hit global capital markets in 2008. As of March 2009, prices have recovered until the end of the period examined. The volume chart shows more or less an inverse pattern to stock prices. This suggests a negative correlation between stock prices and volume. Tetlock (2007) finds that a high level of pessimism in the media predicts falling market prices. The Reuters sentiment graph shows that the stock indices follow Reuters sentiment with a certain lag. Most prominently, the trough in Reuters sentiment occurred around December

[6] See www.masterdatacsv.com, last accessed 15 October 2010.

[7] See *Global Business Cycles Indicators* for more detailed information at http://www.conference-board.org/economics/bci, last accessed 7 December 2010.

2008, whereas the stock market bottomed in March 2009. The Conference Board Index shows a similar movement as the Dow Jones Industrials index. We thus undertake further empirical tests to find out whether a combination of fundamental data, i.e. the Conference Board Index, and behavioral data, i.e. Reuters sentiment, can explain changes in stock prices.

Fig. 2 shows cross-correlations of the Dow Jones stock index returns and volume, the Conference Board Index and Reuters sentiment. As graphically anticipated, stock index volume has a negative correlation with the Dow Jones Industrials stock index at most lags. The Conference Board Index has a strong correlation with Dow Jones stock returns, greatest at lag zero. This observation makes sense when considering the common belief that stock markets price in immediately any real macroeconomic development; especially for monthly data, the effect should be already priced in. The Reuters sentiment variable is positively correlated with stock prices, with the highest correlation at lag 1. This means that Reuters sentiment moves one month "ahead" of stock markets. In fig. 3, we consider the cross-correlations between stock returns and positive and negative sentiment scores. Positive and negative sentiment both show the highest correlation at lag one, whereas positive sentiment has a positive correlation and negative sentiment a negative correlation with stock returns, as one would expect.

In the next section, we proceed by constructing a model to test our initial observations.

III. Modeling

By constructing a Vector AutoRegression model (VAR), we tackle possible endogeneity issues. Since we have unit roots in most of the variables, we test for cointegration according to Johansen (1991) first. We find one cointegrating relation. Thus, we formulate a Vector Error Correction Model (VECM) according to the reduced rank (RR) estimation procedure as in Johansen (1995) to account for nonstationarity and cointegration in the data as follows:

$$\triangle y_t = \alpha \beta^{*'} \left[D_{t-1}^{co} \right] + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_p \Delta y_{t-p} + C D_t + u_t, \tag{1}$$

where y_t refers to the endogenous variables, which are the Dow Jones Industrials stock index, Reuters sentiment, Dow Jones stock index volume, and the Conference Board Index, D_t refers to the deterministic term (here: a constant

C), D^{co}_{t-1} is the cointegrating relation, u_t is the error term, and β^* is the cointegration matrix.

In total, we construct three VECMs: first, a model that includes all variables named above with the Reuters sentiment variable that includes all scores, namely positive, neutral and negative. Second, one model comprises only the negative sentiment scores plus the Conference Board Index and stock index volume and, third, one that incorporates positive sentiment, also with the Conference Board Index and stock index volume. To find an optimal lag structure of the models, we perform lag length selection tests according to the Akaike Info Criterion, as shown in table II. For two of the three models, we obtain an optimal number of lags of four, and for one model, which incorporates negative sentiment, an optimal lag length of two. Given our graphical interpretation as well as the insights from the cross-correlograms, which show that sentiment has leading characteristics over stock returns, it appears suited to use a lag structure in the models.

[insert table II about here]

We empirically test the above models to obtain further clues whether Reuters sentiment as well as other variables can explain and/or predict stock returns. Table III shows the results of the VECM estimation with Reuters sentiment, allowing for up to four lags, as specified above. The estimated cointegration relation shows statistically significant values for volume and sentiment, both with correctly specified coefficient signs. Interestingly, the Conference Board Index coefficient is not statistically significant, although the coefficient sign is correct. In the cointegration relation, a negative coefficient sign means that there is a positive relationship with stock returns, and vice versa. For the lagged endogenous term results, the coefficients of sentiment are statistically significant at lags one and three, whereas the Conference Board Index is not statistically significant. These observations lead to assuming that Reuters sentiment has more statistical power to explain stock returns than the Conference Board Index in our model. Macroeconomic factors might thus not be as relevant as behavioral aspects for stock markets in the longer term.

[insert table III about here]

We consider these results in more detail by looking at positive and negative Reuters sentiment individually. Table IV shows the VECM estimation results with Reuters negative sentiment values, allowing up to two lags. The estimated cointegration relation results show highly statistically significant coefficients for volume and negative sentiment, whereas the Conference Board Index coefficient is not statistically significant. Furthermore, the coefficient sign for Reuters sentiment is correctly specified. The lagged endogenous term results show that the negative sentiment coefficient is statistically significant at lag two. The Conference Board Index coefficient is highly statistically significant at lag one, whereas volume is statistically significant at lags one and two. In this model, both Reuters sentiment and the Conference Board Index are statistically significant, so that we can assume that this model is good to explain changes in stock returns better.

[insert table IV about here]

Table V shows the VECM estimation results with Reuters positive sentiment. The coefficients of volume, positive sentiment and the Conference Board index of the estimated cointegration relation are all highly statistically significant. However, the coefficient sign of Reuters positive sentiment is not correctly specified. Furthermore, the coefficients in the lagged endogenous term estimation of Reuters positive sentiment are not statistically significant. The Conference Board index coefficients are statistically significant at lags one and two. These results suggest that positive sentiment is not as well suited as general and negative sentiment. This is in line with Tetlock's (2007) finding that negative sentiment predicts falling stock returns.

[insert table V about here]

To analyze the dynamic interactions between the endogenous variables of the VEC process, we draw on the impulse response analysis so that we can analyze the dynamic interactions between the endogenous variables of the VEC(p) process. A structural vector error correction (SVEC) analysis appears suited in this case.[8] The SVEC model is used to identify the shocks to be traced in

[8] See Appendix A.1 for a detailed discussion of Impulse Responses in VEC(p) processes,

an impulse response analysis by imposing restrictions on the matrix of long-run effects of shocks and the matrix B of contemporaneous effects of the shocks.[9]

Fig. 4 shows the results of the impulse response functions based on the SVEC model. We focus on the first row of the impulse response graphs because we want to identify possible impacts of sentiment, volume, and macroeconomic facts on stock returns. The graphs show an effect of the Conference Board Index as well as Reuters sentiment on stock returns, while stock index volume does not seem to have a significant impact on the Dow Jones Industrials stock index. Stock returns show the greatest response to Reuters sentiment after one month, and to the Conference Board index after two months. Fig. 5 shows the impulse responses based on the SVEC model with Reuters negative sentiment. As one would expect, the response of stock returns to Reuters negative sentiment is negative and greatest after one month. The response to the Conference Board Index is positive and also greatest after one month. In fig. 6, we get a similar pattern with Reuters positive sentiment. The response of stock returns to Reuters positive sentiment is positive and greatest after one month. The same applies for responses of stock returns to the Conference Board Index. Hong and Stein (1999) make similar findings. They show that prices underreact in the short run, suggesting that this should ultimately lead to overreaction in the long run. In this study, we consider the longer term with our monthly data analysis, in which we also find an overreaction to sentiment. In a recent study, Livnat and Petrovits (2009) account for a post-earnings announcement drift among investor sentiment. They find evidence that holding firms with extremely good news following pessimistic sentiment periods earns significantly higher abnormal returns than holding firms with extreme good news following optimistic sentiment periods. Similarly, they show that holding low accrual firms following pessimistic sentiment periods earns significantly higher abnormal returns than holding low accrual firms following optimistic sentiment periods. Chan (2003) also finds evidence of a post-news drift. These findings are in line with our results, as we experience a longer lasting response of stock returns to Reuters sentiment that remains for months, although the response is most pronounced after one month.

We further test how much impact each variable has on stock returns in relation to another. To do this, we draw on the forecast error variance decomposition

and the case for a structural vector error correction (SVEC) model.

[9]See Appendix A.2 for the derivation of matrix B.

(FEVD).[10] The FEVD of Dow Jones Stock index returns is depicted in fig. 7. Interestingly, the impact of the economic factors, in the form of the Conference Board Index, makes up around 5% of the variance of the forecast error of stock returns. The largest share has Reuters sentiment, making up around 15-20% of the variance of the forecast error of stock returns. Volume only attributes to about 5% of the variation in stock returns. This is in line with our empirical results from the VECM and the impulse response functions, strongly speaking for Reuters sentiment as a relevant variable to explain stock returns.

Overall, we can say that both fundamental, i.e. the Conference Board Index, and behavioral, i.e. Reuters sentiment, factors can explain stock returns. Other factors that we have accounted for, such as stock index volume, do not explain stock returns too well, but Reuters negative sentiment appears to have more explanatory power to stock returns than positive sentiment. In the next section, we test how our models perform in a forecasting environment.

IV. Forecasting

Tetlock (2007) shows that one can use negative words in news articles to predict quarterly earnings. Negative words, he finds, consistently predict lower earnings, regardless of the measure and the newspaper. Based on a systematical analysis, a measure of media content specifically tied to either negative investor sentiment or risk aversion, he constructs a hypothetical zero-cost trading strategy using negative words to predict returns of the Dow Jones Industrials Stock Index that yields excess returns (7.3% p.a.). He notes, however, that since this strategy neither accounts for transaction costs nor for slippage and bid-ask spreads while trading daily, it is questionable whether this strategy would remain profitable in a real-world setting. Inspired by his findings, we formulate a simple trading strategy that only requires to trade once per month, given our low-frequency (monthly) data analysis, so that we do not have to account for transaction costs. We attempt to formulate a similar strategy by hypothesizing that Reuters sentiment, i.e. both positive and negative as well as individually, can predict stock returns.

To practically test the predictive power of our models, we construct forecasts. The forecasts are derived from the previously formulated VECMs in (1) based

[10] See Appendix A.3 for a more detailed explanation of the FEVD.

on conditional expectations assuming independent white noise u_t.[11] The vector y_t, incorporating the endogenous variables Dow Jones Industrials stock index returns and volume, the Conference Board Index as well as Reuters Sentiment, is altered for the forecasts to test which variables add forecasting power, and which ones do not. We estimate the out-of-sample forecasts with values from January 2003 to December 2009. Then, we perform step-by-step $t+1$ forecasts for each month of 2010, simulating a real-world trading environment. In total, we estimate seven different models according to results of the Johansen test and the Akaike Info Criterion test. Depending on the test results, we use VARs or VECMs and differing endogenous lag structures. Table VI shows the results.

[insert table VI about here]

The first row shows the absolute performance of the Dow Jones Industrials stock index in 2010: almost 8%. This is our benchmark to which we compare the performance of each trading strategy. Based on the predicted values of the model, we formulate a simple long-short strategy. If the forecast is above the month-end closing price of the stock index, the strategy goes long at the beginning of the forecast month. If the forecast is below the month-end closing price of the stock index, the strategy goes short. The position is closed at the end of each month at the closing price and adjusted in the direction if the forecast assumes a reversal. For simplicity reasons, the available equity is always invested in full at the beginning of each month.

The first model that we build our trading strategy on has the same variables and characteristica as the initial VECM in (1), from which the results are outlined in table III. The model contains stock returns and volume, the Conference Board Index and Reuters sentiment (all values), allowing up to four endogenous lags. The annual performance of the strategy is less than 4%, so that it underperforms the benchmark by over 4%. The success rate is above 50%, indicating that the trading direction whether the index went up or down was predicted correctly in over 6 months for the year.[12] With the next strategy, we want to test how well the model performs without Reuters sentiment, so that we estimate a VECM with stock returns and volume, and the Conference Board Index as endogenous variables. According to the Akaike Info Criterion test,

[11] See Appendix A.4 for a more detailed description of the forecasting model as in Lütkepohl (1991).

[12] Success Rate = number of correctly forecast trading direction (i.e. up or down) months divided by number of total forecast months.

the optimal endogenous lag structure is one. This strategy obtains a negative performance in 2010 of almost -17%, a great underperformance to the index. This lets suggest that Reuters sentiment does add value in forecasting models of stock returns. Further, we want to test Reuters sentiment individually to predict stock returns. According to the Johansen test, we do not find a cointegrating relation, so that we apply a VAR model as opposed to a VECM. The performance of this strategy is quite high with a total outperformance over the index of 23%. Confirming our earlier assumption and in line with our findings from the FEVD, Reuters sentiment is a good variable to predict stock returns.

For the next strategies, we consider the VECM results from tables IV and V with Reuters negative and positive sentiment individually. The strategy with Reuters negative sentiment returns over 22% with a high success rate of 75%. The Sharpe Ratio, a measure that puts returns in relation to volatility, is quite high with a score of 1.62.[13] The strategy that includes stock index volume, the Conference Board index, and Reuters positive sentiment is not as successful as the previous one, as it returns 19% with a much lower success rate and lower Sharpe Ratio. Nevertheless, this strategy is more successful than the first strategy with all values of Reuters sentiment. We can thus infer that "directional" sentiment, i.e. positive or negative, has more power to predict stock returns than combined sentiment from Reuters news pieces. This might also hail from the fact that the combined sentiment also contains neutral sentiment, i.e. ambiguous and indiscernible statements without clear sentiment status, which might blurr the sentiment score, although more words and context have been coded. Therefore, it is a clear advantage to consider only the positive and negative shares of the coded sentiment.

The last two strategies that we consider are based on VAR models with solely negative and positive sentiment, respectively, so that we can test directly whether positive or negative sentiment is the better predictor for stock returns. The strategy with negative sentiment returns over 47% in 2010, whereas the strategy with positive sentiment returns 15%. The difference between the two strategies gets more imminent when looking at the success rates: 83% vs. 50%. This makes negative sentiment clearly the better indicator for stock returns than positive sentiment. This finding is in line with Tetlock's (2007) study, and it extends his findings with the result that although positive sentiment has some predictive power, negative sentiment in Reuters news is more suited to

[13] See Appendix A.5 for a detailed calculation of the Sharpe Ratio.

predicting stock returns. Also, the absolute annual performance of this strategy is higher than the one from Tetlock (2007).

According to the various tests and analyses that we have undertaken, we stress three major findings. First, we confirm the EMH by Fama (1970) to the extent that fundamental factors, accounted for by the Conference Board Index, can partly explain stock returns of the Dow Jones Industrials stock index. This finding is pronounced in both the impulse response functions and the variance decomposition analysis, in which the Conference Board Index makes up less of the variance of stock returns than sentiment. We also find that volume plays a minor role in the model. Second, we reject the EMH on the grounds that behavioral factors can explain a great share of stock returns, in particular to a greater extent than fundamental factors (i.e. the Conference Board Index) can. Reuters sentiment appears to capture investor sentiment quite well, entailing strong predictive power for stock returns. Third, even among sentiment there is a difference in the predictive power, as we discern between positive and negative sentiment. As in Tetlock (2007), we find that negative sentiment is a much better predictor for stock returns than positive sentiment. We thus reject the EMH on the same grounds as Tetlock (2007) by confirming and extending his results through an extension of the analysis to positive and negative sentiment, a more sophisticated approach as well as a more extensive dataset with over 3.6 million Reuters news articles.

V. Conclusion

Based on the EMH by Fama (1970), we examine whether fundamental and/or behavioral factors influence US stock returns. To account for fundamental factors, we use the Conference Board Index that comprises of a basket of various macroeconomic variables and indicators. We use stock index volume to control for possible market depth and liquidity constraints. To account for behavioral factors, we use a novel dataset with sentiment values that is obtained from over 3.6 million Reuters news articles. Tetlock's (2007) approach serves as inspiration for this study, as the use of his textual analysis tool, the General Inquirer (GI), seems limited, given that it is only able to distinguish between positive and negative words, but not between the context of the article.

We confirm Tetlock's (2007) findings that sentiment has an impact on stock returns, rejecting the EMH by Fama (1970) on the same grounds, given that we find positive correlations between negative media sentiment and declines in stock returns as well as between positive media sentiment and gains in stock returns. We show with impulse response functions and a forecast error variance decomposition (FEVD) analysis of a Vector Error Correction Model (VECM) that behavioral factors, such as Reuters sentiment, can better explain stock returns than fundamental factors, such as the Conference Board Index. This finding is manifested in the results of out-of-sample forecasts that were constructed for the year 2010. Furthermore, we find that negative sentiment has much higher explanatory and predictive power than positive sentiment in Reuters news. This is also in line with Tetlock (2007), whereas this study goes further by extending the analysis to positive sentiment and greater accuracy as well as to a higher annual return of the trading strategy, which is achieved through a novel sentiment algorithm developed by Thomson Reuters.

Appendix

A.1

In a VECM, the a vector of endogenous variables is denoted by y_t. If the process y_t is stationary, it has a Wold moving average (MA) representation

$$y_t = \Phi_0 u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \cdots ,$$

where $\Phi_0 = I_K$ and the Φ_s can be computed recursively as

$$\Phi_s = \sum_{j=1}^{s} \Phi_{s-j} A_j, \ s = 1, 2, \ldots ,$$

with $\Phi_0 = I_K$ and $A_j = 0$ for $j > p$. The coefficients of this representation may be interpreted as reflecting the responses to impulses hitting the system. The (i, j)th elements of the matrices Φ_s, regarded as a function of s, trace out the expected response of $y_{i,t+s}$ to a unit change in y_{jt} holding constant all past values of y_t. The elements of Φ_s represent the impulse responses of the components of y_t with respect to the u_t innovations.

Because the underlying shocks are not likely to occur in isolation if the components of u_t are not instantaneously uncorrelated, that is, if $\sum_u$ is not diagonal, in many applications the innovations of the VAR/VECM are orthogonalized using a Cholesky decomposition of the covariance matrix $\sum_u$. Denoting by P a lower triangular matrix such that $\sum_u = PP'$, the orthogonalized shocks are given by $\varepsilon_t = P^{-1} u_t$. Thus, we obtain

$$y_t = \Psi_0 \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \cdots ,$$

where $\Psi_i = \Phi_i P \, (i = 0, 1, 2, \ldots)$. Here $\Psi_0 = P$ is lower triangular so that an ε shock in the first variable may have an instantaneous effect on all the variables, whereas a shock in the second variable cannot have an instantaneous impact on y_{1t} but only on the other variables and so on.

It is important to notice that if a different ordering of the variables in the vector y_t is chosen this may produce different impulse responses. Hence, the effects of a shock may depend on the way the variables are arranged in the vector of y_t. Breitung *et al* (2004) discuss this issue in detail.

For the impulse responses that are computed from the estimated Structural Vector Error Correction Model (SVEC) coefficients, the confidence intervals

(CIs) are contrsucted with the bootstrap method according to Efron and Tibshirani (1993). The standard percentile interval is determined as

$$CI_s = \left[s^*_{\gamma/2}, s^*_{(1-\gamma/2)} \right],$$

where $s^*_{\gamma/2}$ and $s^*_{(1-\gamma/2)}$ are the $\gamma/2-$ and $(1-\gamma/2)-$quantiles, respectively, of the bootstrap distribution of the corresponding bootstrap estimator of the impulse response coefficient $\hat{\Phi}^*$.

A.2

The matrix B is defined such that $u_t = B\varepsilon_t$ in (1) and the matrix Ξ of long-run effects of the u_t residuals is

$$\Xi = \beta_\perp \left(\alpha_\perp^{'} \left(I_K - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_\perp \right)^{-1} \alpha_\perp^{'}. \tag{2}$$

Hence, the long-run effects of ε shocks are given by ΞB. $rk(\Xi) = K - r$ and, hence, ΞB has rank $K - r$. Thus, the matrix ΞB can have at most r columns of zeros. Therefore, there can be at most r shocks with transitory effects (zero long-run impact) and at least $k* = K - r$ shocks have permanent effects.

A.3

The SVEC Forecast Error Variance Decomposition (FEVD) separates the variation in an endogenous variable into the component shocks to the Structural VAR (SVAR), or, in this case, the SVEC. The FEVD provides information about the relative importance of each random innovation in affecting the variables in the SVEC. Denoting the ij-th element of the orthogonalized impulse response coefficient matrix ψ_n, the variance of the forecast error $y_{k,T+h} - y_{k,T+h|\mathrm{T}}$ is

$$\sigma_k^2(h) = \sum_{n=0}^{h-1} \left(\psi_{k1,n}^2 + \cdots + \psi_{kK,n}^2 \right) = \sum_{j=1}^{K} \left(\psi_{kj,0}^2 + \cdots + \psi_{kj,h-1}^2 \right).$$

A.4

The corresponding forecast errors for the forecasts are

$$y_{T+h} - y_{T+h|\mathrm{T}} = u_{T+h} + \phi_1 u_{T+h-1} + \cdots + \phi_{h-1} u_{T+1},$$

where $\phi_s = \sum_{j=1}^{s} \phi_{s-j} A_j$, $s = 1, 2, \ldots$, $with\, \phi_0 = I_K\, and\, A_j = 0\, for\, j > p$. Thus, the forecast errors have zero mean and, hence, the forecasts are unbiased.

A.5

The Sharpe ratio is calculated according to Sharpe (1994):

$$\frac{R_p - R_f}{\sigma_p},$$

where R_p is the annualized return of the portfolio, R_f the annualized rate of a risk-free asset (in this study we use the 1-month Treasury Bill rate), and σ_p is the annualized standard deviation of the portfolio returns.

References

[1] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* AC-19: 716-723.

[2] Antweiler, W. and M.Z. Frank (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance* 59:1259-1293.

[3] Baker, M. and J. Wurgler (2007). Investor sentiment and the cross-section of stock returns. *Journal of Finance* 61: 1645-1680.

[4] Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of Financial Economics* 49: 307-343.

[5] Breitung, J., R. Brüggemann, and H. Lütkepohl (2004). Structural vector autoregressive modelling and impulse responses. *In* H. Lütkepohl and M. Krätzig (eds.), *Applied Time Series Econometrics*. Cambridge: Cambridge University Press.

[6] Cao, M. and J. Wei (2005). Stock market returns: A note on temperature anomaly. *Journal of Banking and Finance* 29: 1559-1573.

[7] Chan, W.S. (2003). Stock Price Reaction to News and No-news: Drift and Reversal after Headlines. *Journal of Financial Economics* 70: 223-260.

[8] Cutler, D.M., J.M. Poterba, and L.H. Summers (1989). What moves stock prices? *Journal of Portfolio Management* 15: 4-12.

[9] De Bondt, W.F.M. and R. Thaler (1985). Does the Stock Market Overreact? *Journal of Finance* 40: 793-805.

[10] DeLong, J.B., A. Shleifer, L.H. Summers and R. J. Waldmann (1990). Noise trader risk in financial markets. *Journal of Political Economy* 98: 703- 738.

[11] Edmans, A., D. Garcia, and O. Norli (2007) Sports Sentiment and Stock Returns. *Journal of Finance* 62, 1967 - 1998.

[12] Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman & Hall.

[13] Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25: 383-417.

[14] Fang, L., and J. Peress (2009). Media Coverage and the Cross-Section of Stock Returns. *Journal of Finance* 64: 2023-2052.

[15] Hirshleifer, D. (2001). Investor psychology and asset pricing. *Journal of Finance* 91: 342-346.

[16] Hirshleifer, D. and T. Shumway (2003). Good day sunshine: Stock returns and the weather. *Journal of Finance* 58: 1009-1032.

[17] Hong, H. and J.C. Stein (1999). A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets. *Journal of Finance* 54: 2143-2184.

[18] Huberman, G. and T. Regev (2001). Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. *Journal of Finance* 56: 387-396.

[19] Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica* 59: 1551-1580.

[20] Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

[21] Kahneman, D. and A. Tversky (1981). The Framing of Decisions and the Psychology of Choice. *Science* 211: 453-458.

[22] Kamstra, M., L. Kramer, and M. Levi (2003). Winter blues: a SAD stock market cycle. *American Economic Review* 93: 324-343.

[23] Klibanoff, P., O. Lamont, and T.A. Wizman (1998). Investor Reaction to salient news in closed-end country funds. *Journal of Finance* 53: 673-699.

[24] Livnat, J. and C. Petrovits (2009). Investor Sentiment, Post-Earnings Announcement Drift, and Accruals. *SSRN Working Paper* available at http://ssrn.com/abstract=1262757.

[25] Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer Verlag.

[26] Maier, S.R. (2005). Accuracy Matters: A Cross-Market Assessment of Newspaper Error and Credibility. *Journalism & Mass Communication Quarterly* 82: 533-551.

[27] Meschke, F. and Y.H. Kim (2011). CEO Interviews on CNBC. *SSRN Working Paper* available at http://ssrn.com/abstract=1627683.

[28] Mullainathan, S. and A. Shleifer (2005). The Market for News. *American Economic Review* 95: 1031–1053.

[29] Sharpe, W.F. (1994). The Sharpe Ratio. *Journal of Portfolio Management* 21: 49-58.

[30] Shiller, R. J. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review* 71: 421-436.

[31] Summers, L.H. (1986). Does the Stock Market Rationally Reflect Fundamental Values? *Journal of Finance* 41: 591-601.

[32] Tetlock, P.C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *Journal of Finance* 62: 1139-1168.

[33] Tetlock, P.C. (2011). All the news that's fit to reprint: Do investors react to stale information? *Review of Financial Studies*, forthcoming.

[34] Tetlock, P.C., M. Saar-Tsechansky, and S. Macskassy (2008). More than words: Quantifying language to measure firm's fundamentals. *Journal of Finance* 63:1437-1467.

[35] Yuan, K. L. Zheng, and Q. Zhu (2006). Are investors moonstruck? Lunar phases and stock returns. *Journal of Empirical Finance* 13: 1-23.

Figure 1: Time-Series Charts of the Dow Jones Stock Index (dj_log) and Volume (dj_vol_log), The Conference Board Index (Conf_B_log), and Reuters sentiment - all values (tr_ns_u_eq_sel), Reuters negative sentiment (tr_ns_eq_sel_neg), and Reuters positive sentiment (tr_ns_eq_sel_pos).

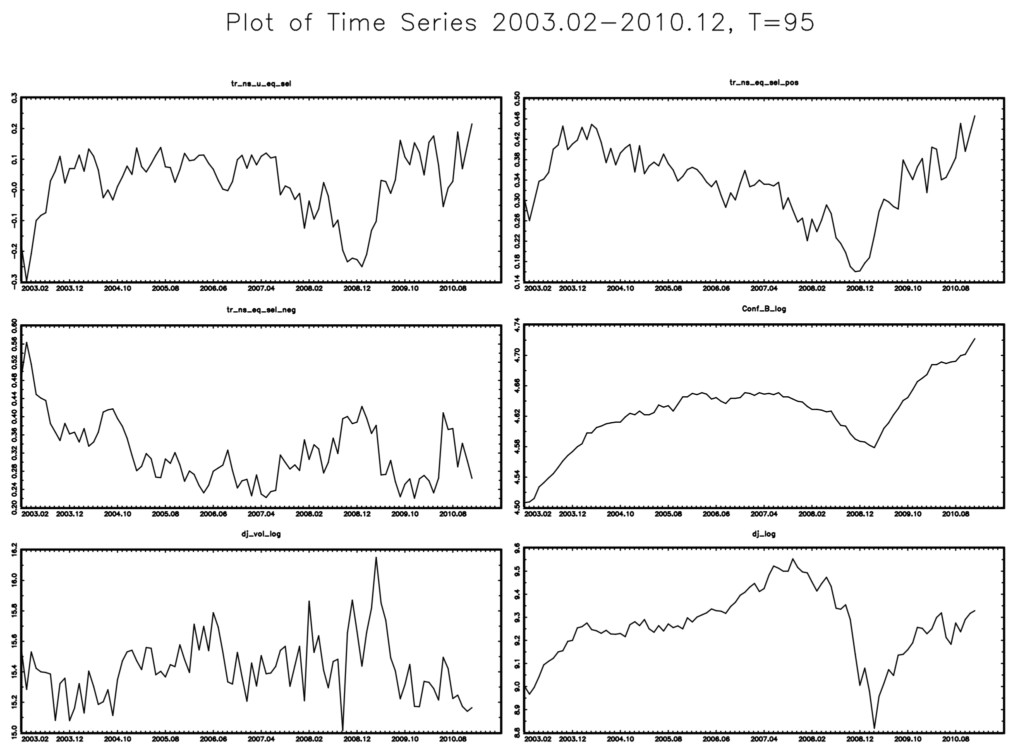

Figure 2: Cross Correlations of the Dow Jones Stock Index and the Conference Board Index, Dow Jones Volume and Reuters Equities Sentiment (all values)

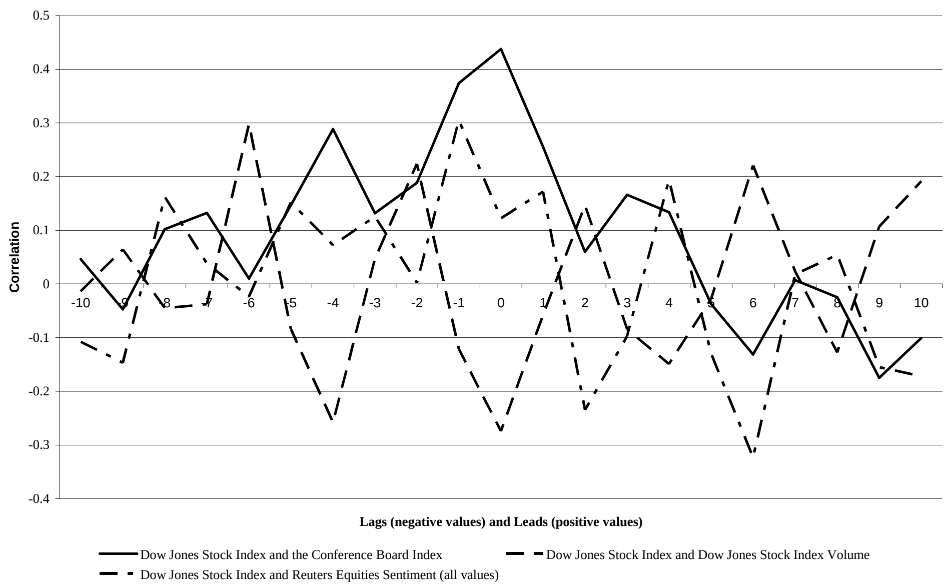

Figure 3: Cross Correlations of the Dow Jones Stock Index with negative and positive Reuters Equities Sentiment

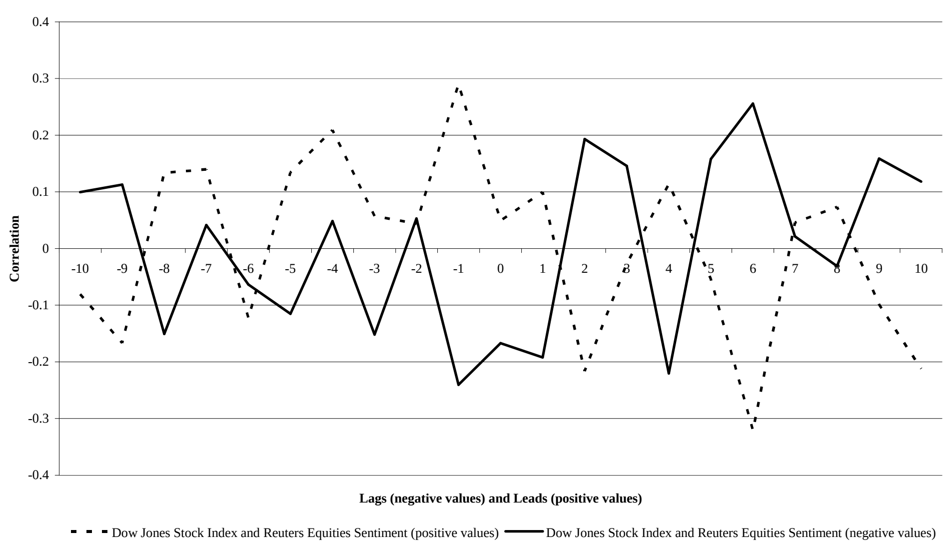

Figure 4: Impulse Responses from the Structural Vector Error Correction Model (SVEC) with 95% Bootstrap Confidence Intervals according to Efron and Tibshirani (1993). Abbreviations used denote the following: logarithmized Dow Jones Industrials Stock Index (dj_log_d1), logarithmized Dow Jones Industrials Stock Index Volume (dj_vol_log_d1), the Conference Board Index (Conf_B_log_d1), and Reuters sentiment from the equities section (tr_ns_U_eq_sel_d1).

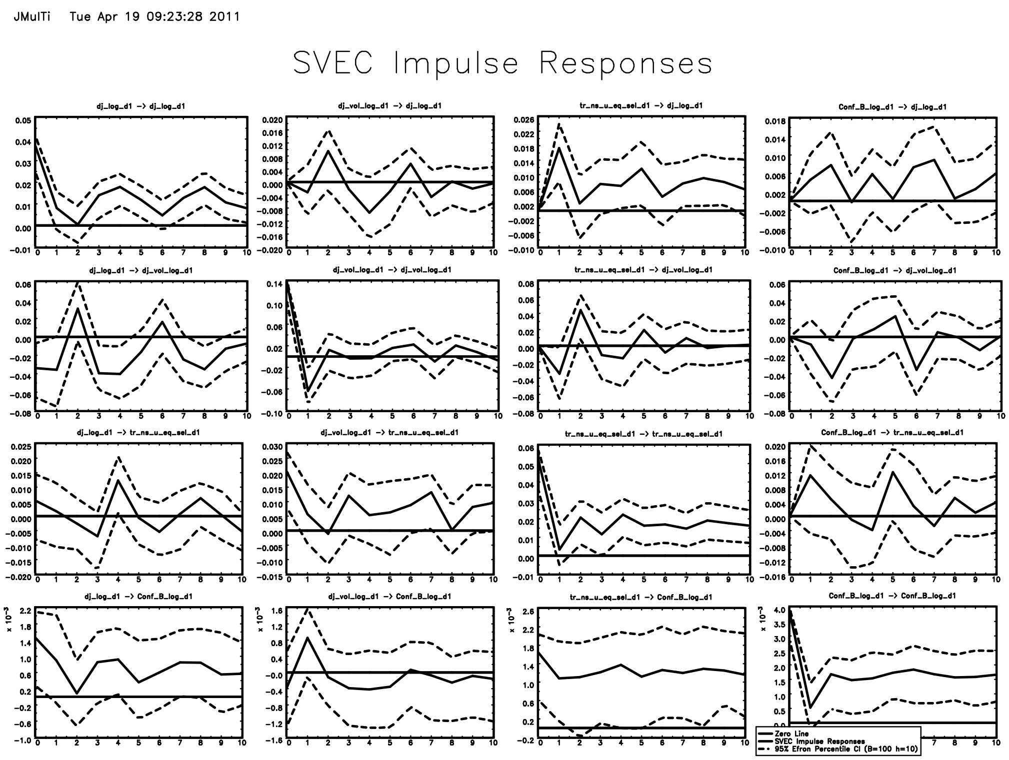

Figure 5: Impulse Responses from the Structural Vector Error Correction Model (SVEC) with 95% Bootstrap Confidence Intervals according to Efron and Tibshirani (1993). Abbreviations used denote the following: logarithmized Dow Jones Industrials Stock Index (dj_log_d1), logarithmized Dow Jones Industrials Stock Index Volume (dj_vol_log_d1), the Conference Board Index (Conf_B_log_d1), and Reuters negative sentiment from the equities section (tr_ns_eq_sel_neg_d1).

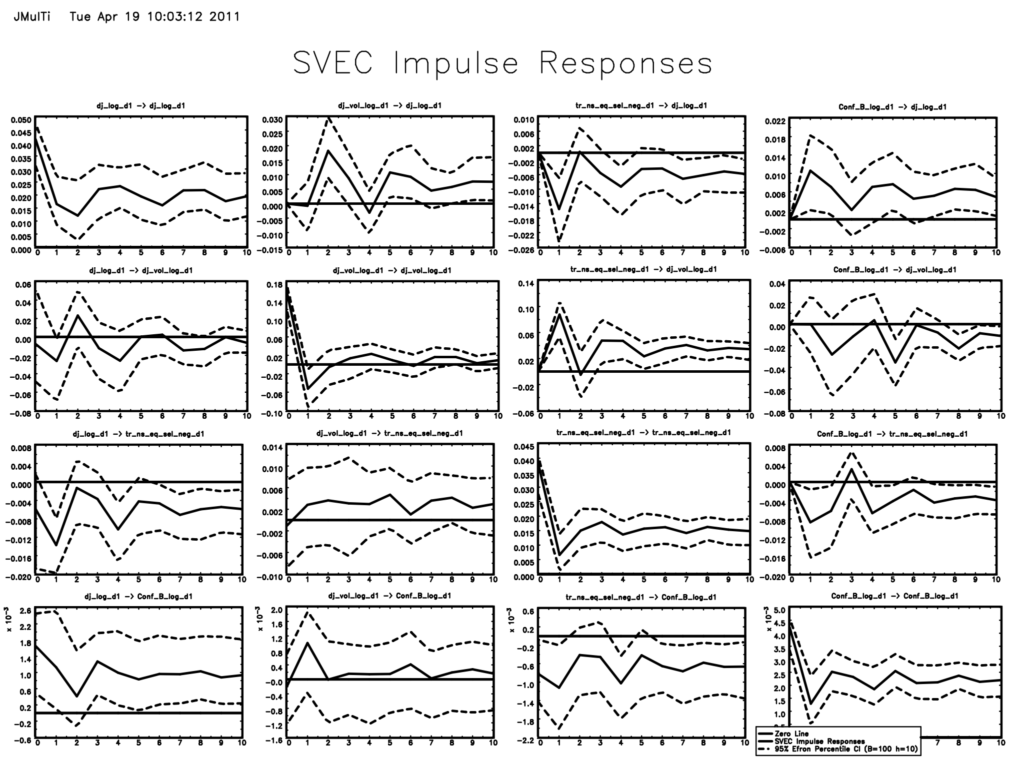

Figure 6: Impulse Responses from the Structural Vector Error Correction Model (SVEC) with 95% Bootstrap Confidence Intervals according to Efron and Tibshirani (1993). Abbreviations used denote the following: logarithmized Dow Jones Industrials Stock Index (dj_log_d1), logarithmized Dow Jones Industrials Stock Index Volume (dj_vol_log_d1), the Conference Board Index (Conf_B_log_d1), and Reuters positive sentiment from the equities section (tr_ns_eq_sel_pos_d1).

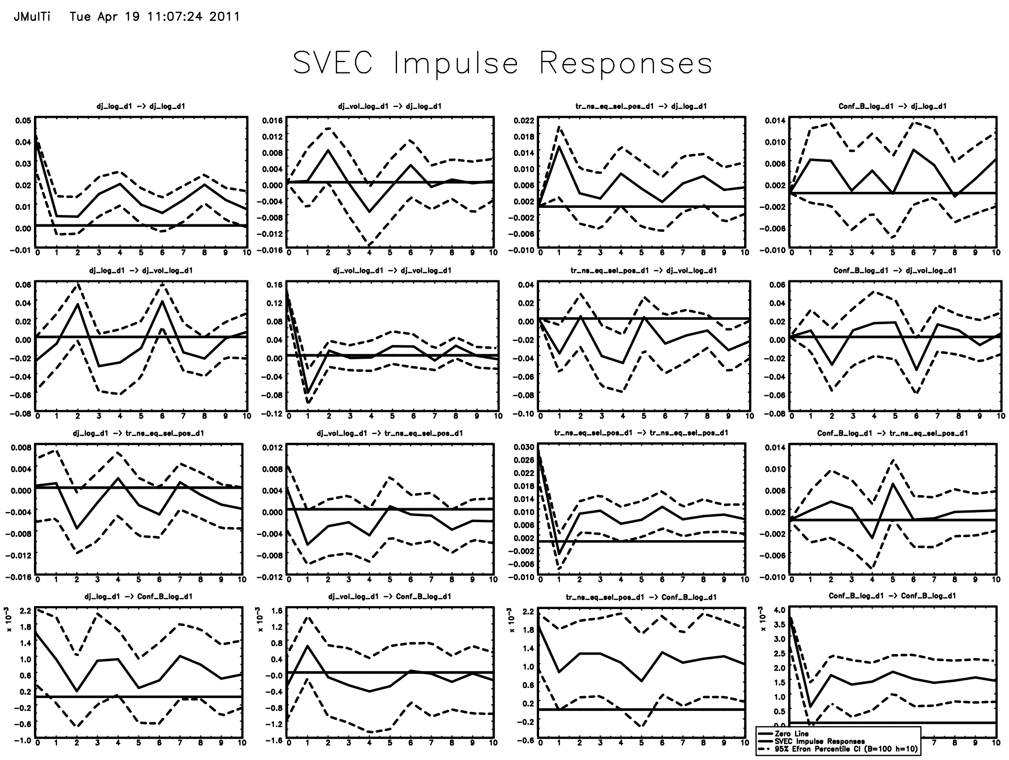

Figure 7: SVEC Forecast Error Variance Decomposition of Dow Jones Industrials Stock Index (differenced logs)

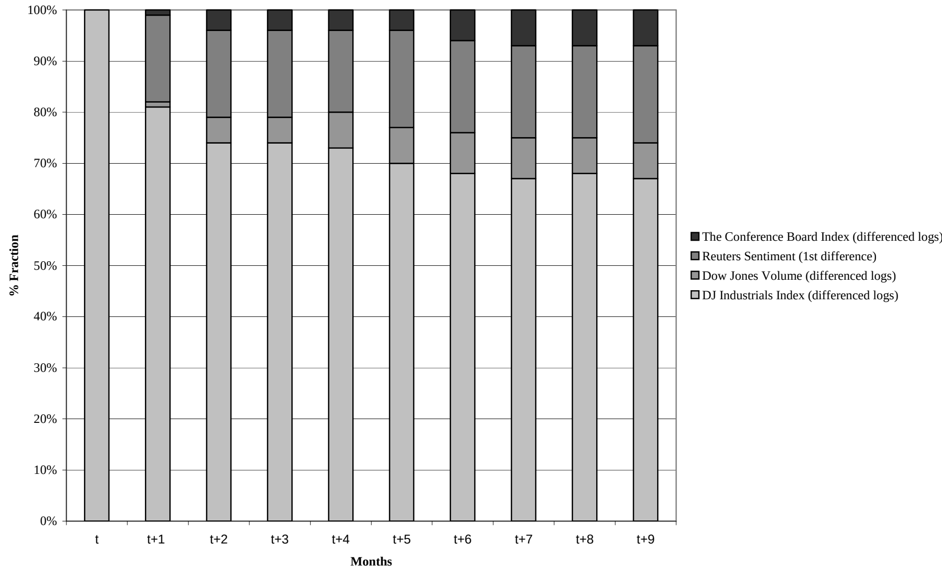

Table I

Sentiment Sources

	Number of News Articles examined for sentiment 2003 - 2010
*Thomson Reuters News Items**	3'636'917
Total	3'636'917

Source: Thomson Reuters NewsAnalytics*

Table II

Optimal Endogenous Lags from Information Criteria

Deterministic Variables: Constant

Endogenous Variables	Sample Range	Optimal number of lags (searched up to 10 lags) according to Akaike Info Criterion
Dow Jones Industrials Stock Index and Volume (differenced logs), The Conference Board Business Cycles Indicator (differenced logs), Reuters Sentiment - all values (first differences)	[2004 M1, 2010 M12], T = 84	4
Dow Jones Industrials Stock Index and Volume (differenced logs), The Conference Board Business Cycles Indicator (differenced logs), Reuters Sentiment - negative values (first differences)	[2004 M1, 2010 M12], T = 84	2
Dow Jones Industrials Stock Index and Volume (differenced logs), The Conference Board Business Cycles Indicator (differenced logs), Reuters Sentiment - positive values (first differences)	[2004 M1, 2010 M12], T = 84	4

Table III

Vector Error Correction Model Coefficient Estimates *(monthly values)*

Endogenous Variables	Dow Jones Industrials Stock Index (log_dj), Dow Jones Industrials Stock Inde Volume (log_dj_vol), Reuters Sentiment - all values (r_s), The Conference Board Index (log_Conf_B)
Exogenous Variables	*none*
Deterministic Variables	Constant (CONST)
Endogenous Lags (Differences - in months)	4
Exogenous Lags	*0*
Sample Range	[2003 M7, 2010 M12], T = 90
Estimation Procedure	One stage Johansen approach according to Johansen (1995)

Lagged endogenous term [coefficient, standard deviation, p-values in {}-parentheses]

	d(log_dj)	d(log_dj_vol)	d(r_s)	d(log_Conf_B)
d(log_dj) (t-1)	-0.684	1.733	-0.112	0.033
	-0.156	-0.605	-0.239	-0.019
	{0.000}	{0.004}	{0.640}	{0.074}
d(log_dj_vol) (t-1)	0.103	0.574	-0.009	0.013
	-0.077	-0.3	-0.119	-0.009
	{0.182}	{0.056}	{0.936}	{0.168}
d(r_s) (t-1)	0.172	-2.031	-0.983	0.01
	-0.092	-0.358	-0.142	-0.011
	{0.062}	{0.000}	{0.000}	{0.364}
d(log_Conf_B) (t-1)	1.203	-2.318	2.979	-0.862
	-1.019	-3.95	-1.562	-0.122
	{0.238}	{0.557}	{0.056}	{0.000}
d(log_dj) (t-2)	-0.776	2.42	-0.335	0.017
	-0.164	-0.637	-0.252	-0.02
	{0.000}	{0.000}	{0.184}	{0.382}
d(log_dj_vol) (t-2)	0.141	0.276	-0.059	0.014
	-0.064	-0.247	-0.098	-0.008
	{0.027}	{0.264}	{0.549}	{0.077}
d(r_s) (t-2)	0.078	-0.684	-0.636	0.012
	-0.112	-0.433	-0.171	-0.013
	{0.488}	{0.114}	{0.000}	{0.371}
d(log_Conf_B) (t-2)	2.004	-11.403	4.023	-0.502
	-1.399	-5.422	-2.144	-0.168
	{0.152}	{0.035}	{0.061}	{0.003}
d(log_dj) (t-3)	-0.264	1.433	-0.565	0.017
	-0.154	-0.595	-0.235	-0.018
	{0.086}	{0.016}	{0.016}	{0.365}
d(log_dj_vol) (t-3)	0.119	0.181	-0.053	0.008
	-0.046	-0.179	-0.071	-0.006
	{0.010}	{0.313}	{0.454}	{0.164}
d(r_s) (t-3)	0.22	-0.416	-0.452	0.011
	-0.102	-0.396	-0.157	-0.012
	{0.031}	{0.294}	{0.004}	{0.370}
d(log_Conf_B) (t-3)	0.518	-16.541	2.118	-0.223
	-1.396	-5.41	-2.14	-0.167
	{0.711}	{0.002}	{0.322}	{0.182}
d(log_dj) (t-4)	-0.038	0.746	-0.025	0.013
	-0.11	-0.426	-0.168	-0.013
	{0.727}	{0.079}	{0.883}	{0.311}
d(log_dj_vol) (t-4)	0.063	0.089	-0.056	0.003
	-0.025	-0.099	-0.039	-0.003
	{0.014}	{0.366}	{0.151}	{0.411}
d(r_s) (t-4)	0.043	-0.244	0.024	0.005
	-0.086	-0.334	-0.132	-0.01
	{0.616}	{0.466}	{0.855}	{0.656}
d(log_Conf_B) (t-4)	1.148	-15.484	-1.039	-0.032
	-1.024	-3.969	-1.569	-0.123
	{0.262}	{0.000}	{0.508}	{0.792}

Loading coefficients

	d(log_dj)	d(log_dj_vol)	d(r_s)	d(log_Conf_B)
ec1(t-1)	-0.246	-2.893	-0.013	0.093
	-0.13	-0.503	-0.016	-0.199
	{0.058}	{0.000}	{0.410}	{0.642}

Estimated cointegration relation

	ec1(t-1)
log_dj(t-1)	1
	0
	{0.000}
log_dj_vol(t-1)	0.676
	-0.092
	{0.000}
r_s(t-1)	-0.496
	-0.248
	{0.046}
log_Conf_B(t-1)	-0.053
	-1.338
	{0.968}
CONST	0.001
	-0.005
	{0.872}

Table IV

Vector Error Correction Model Coefficient Estimates *(monthly values)*

Endogenous Variables	Dow Jones Industrials Stock Index (log_dj), Dow Jones Industrials Stock Index Volume (log_dj_vol), Reuters Sentiment negative values (r_s_neg), The Conference Board Index (log_Conf_B)
Exogenous Variables	*none*
Deterministic Variables	Constant (CONST)
Endogenous Lags (Differences - in months)	2
Exogenous Lags	*0*
Sample Range	[2003 M5, 2010 M12], T = 92
Estimation Procedure	One stage Johansen approach according to Johansen (1995)

Lagged endogenous term [coefficient, standard deviation, p-values in { }-parentheses]

	d(log_dj)	d(log_dj_vol)	d(r_s_neg)	d(log_Conf_B)
d(log_dj) (t-1)	-0.673	-1.055	-0.173	0.016
	-0.086	-0.325	-0.079	-0.01
	{0.000}	{0.001}	{0.029}	{0.098}
d(log_dj_vol) (t-1)	-0.173	0.169	-0.1	0.001
	-0.042	-0.159	-0.039	-0.005
	{0.000}	{0.290}	{0.010}	{0.907}
d(r_s_neg) (t-1)	0.042	-1.17	-0.599	-0.009
	-0.14	-0.531	-0.129	-0.016
	{0.762}	{0.028}	{0.000}	{0.573}
d(log_Conf_B) (t-1)	1.982	4.625	-2.439	-0.712
	-0.951	-3.611	-0.88	-0.106
	{0.037}	{0.200}	{0.006}	{0.000}
d(log_dj) (t-2)	-0.596	0.225	0.031	-0.008
	-0.085	-0.323	-0.079	-0.009
	{0.000}	{0.487}	{0.698}	{0.416}
d(log_dj_vol) (t-2)	-0.066	-0.016	-0.06	0.002
	-0.025	-0.093	-0.023	-0.003
	{0.007}	{0.861}	{0.008}	{0.518}
d(r_s_neg) (t-2)	0.288	-1.14	-0.418	-0.007
	-0.109	-0.413	-0.101	-0.012
	{0.008}	{0.006}	{0.000}	{0.544}
d(log_Conf_B) (t-2)	1.532	3.621	-2.453	-0.288
	-0.95	-3.607	-0.879	-0.106
	{0.107}	{0.315}	{0.005}	{0.006}

Loading coefficients

	d(log_dj)	d(log_dj_vol)	d(r_s_neg)	d(log_Conf_B)
ec1(t-1)	-0.079	0.693	-0.003	-0.054
	-0.025	-0.096	-0.003	-0.023
	{0.002}	{0.000}	{0.314}	{0.020}

Estimated cointegration relation

	ec1(t-1)
log_dj(t-1)	1
	0
	{0.000}
log_dj_vol(t-1)	-2.146
	-0.316
	{0.000}
r_s_neg(t-1)	5.016
	-1.381
	{0.000}
log_Conf_B(t-1)	-6.754
	-5.788
	{0.243}
CONST	0.017
	-0.024
	{0.474}

Table V

Vector Error Correction Model Coefficient Estimates *(monthly values)*

Endogenous Variables	Dow Jones Industrials Stock Index (log_dj), Dow Jones Industrials Stock Index Volume (log_dj_vol), Reuters Sentiment - positive values (r_s_pos), The Conference Board Index (log_Conf_B)
Exogenous Variables	*none*
Deterministic Variables	Constant (CONST)
Endogenous Lags (Differences - in months)	4
Exogenous Lags	*0*
Sample Range	[2003 M7, 2010 M12], T = 90
Estimation Procedure	One stage Johansen approach according to Johansen (1995)

Lagged endogenous term [coefficient, standard deviation, p-values in {}-parentheses]

	d(log_dj)	d(log_dj_vol)	d(r_s_pos)	d(log_Conf_B)
d(log_dj) (t-1)	-1.045	2.493	0.218	0.017
	-0.166	-0.569	-0.115	-0.018
	{0.000}	{0.000}	{0.058}	{0.355}
d(log_dj_vol) (t-1)	-0.07	0.903	0.15	0.002
	-0.086	-0.293	-0.059	-0.009
	{0.415}	{0.002}	{0.012}	{0.822}
d(r_s_pos) (t-1)	0.249	5.511	-0.623	0.012
	-0.272	-0.93	-0.188	-0.03
	{0.359}	{0.000}	{0.001}	{0.678}
d(log_Conf_B) (t-1)	2.046	-11.965	-0.604	-0.831
	-1.182	-4.043	-0.819	-0.131
	{0.084}	{0.003}	{0.460}	{0.000}
d(log_dj) (t-2)	-0.994	3.341	-0.07	0.002
	-0.188	-0.642	-0.13	-0.021
	{0.000}	{0.000}	{0.588}	{0.924}
d(log_dj_vol) (t-2)	0.001	0.549	0.087	0.005
	-0.072	-0.247	-0.05	-0.008
	{0.992}	{0.027}	{0.082}	{0.544}
d(r_s_pos) (t-2)	0.277	5.262	-0.464	0.026
	-0.259	-0.885	-0.179	-0.029
	{0.284}	{0.000}	{0.010}	{0.360}
d(log_Conf_B) (t-2)	3.162	-17.684	0.461	-0.458
	-1.453	-4.97	-1.006	-0.161
	{0.030}	{0.000}	{0.647}	{0.004}
d(log_dj) (t-3)	-0.479	2.74	-0.21	0.003
	-0.17	-0.582	-0.118	-0.019
	{0.005}	{0.000}	{0.075}	{0.868}
d(log_dj_vol) (t-3)	0.044	0.395	0.025	0.002
	-0.053	-0.181	-0.037	-0.006
	{0.407}	{0.029}	{0.500}	{0.711}
d(r_s_pos) (t-3)	0.279	3.518	-0.056	0.042
	-0.238	-0.813	-0.165	-0.026
	{0.241}	{0.000}	{0.732}	{0.110}
d(log_Conf_B) (t-3)	1.619	-18.508	1.057	-0.223
	-1.445	-4.941	-1.001	-0.16
	{0.263}	{0.000}	{0.291}	{0.164}
d(log_dj) (t-4)	-0.088	1.598	-0.092	0.008
	-0.118	-0.403	-0.082	-0.013
	{0.454}	{0.000}	{0.259}	{0.557}
d(log_dj_vol) (t-4)	0.03	0.188	-0.02	0
	-0.029	-0.099	-0.02	-0.003
	{0.296}	{0.058}	{0.308}	{0.894}
d(r_s_pos) (t-4)	0.126	1.123	0.191	0.033
	-0.173	-0.593	-0.12	-0.019
	{0.466}	{0.058}	{0.113}	{0.081}
d(log_Conf_B) (t-4)	1.15	-12.215	-0.438	-0.086
	-1.136	-3.886	-0.787	-0.126
	{0.312}	{0.002}	{0.578}	{0.494}

Loading coefficients

	d(log_dj)	d(log_dj_vol)	d(r_s_pos)	d(log_Conf_B)
ec1(t-1)	0.08	-3.104	0.003	-0.243
	-0.12	-0.412	-0.013	-0.083
	{0.504}	{0.000}	{0.805}	{0.004}

Estimated cointegration relation

	ec1(t-1)
log_dj(t-1)	1
	0
	{0.000}
log_dj_vol(t-1)	0.793
	-0.106
	{0.000}
r_s_pos(t-1)	2.266
	-0.619
	{0.000}
log_Conf_B(t-1)	-4.465
	-1.601
	{0.005}
CONST	0.006
	-0.005
	{0.254}

Table VI

Performance of Trading Strategies based on Out-of-sample Prediction Estimates of VECM (Vector Error Correction Models) and VAR (Vector AutoRegression) Models *(monthly values)*

	Type of Model*	Endogenous Lags**	Performance Stock Index Jan 2010 - Dec 2010	Performance Strategy Jan 2010 - Dec 2010	Outperformance Strategy / Stock Index Jan 2010 - Dec 2010	Sharpe Ratio*** Jan 2010 - Dec 2010	Success Rate Jan 2010 - Dec 2010
Endogenous Variables							
Dow Jones Industrials Stock Index			7.94%				
Dow Jones Industrials Stock Index, Dow Jones Industrials Stock Index Volume, the Conference Board Index, and Reuters Sentiment (all values)	VECM	4		3.69%	-4.25%	0.62	58%
Dow Jones Industrials Stock Index, Dow Jones Industrials Stock Index Volume, and the Conference Board Index	VECM	1		-16.56%	-24.50%	-0.51	50%
Dow Jones Industrials Stock Index, Reuters Sentiment (all values)	VAR	1		31.10%	23.16%	1.17	67%
Dow Jones Industrials Stock Index, Dow Jones Industrials Stock Index Volume, the Conference Board Index, and Reuters Sentiment (negative values)	VECM	2		22.13%	14.20%	1.62	75%
Dow Jones Industrials Stock Index, Dow Jones Industrials Stock Index Volume, the Conference Board Index, and Reuters Sentiment (positive values)	VECM	4		19.85%	11.91%	1.49	58%
Dow Jones Industrials Stock Index, Reuters Sentiment (negative values)	VAR	4		47.63%	39.70%	3.60	83%
Dow Jones Industrials Stock Index, Reuters Sentiment (positive values)	VAR	6		15.74%	7.80%	1.26	50%

*The type of model was selected based on the Johansen Test of Cointegration as in Johansen (1991). If no cointegration was found, a VAR model was used. In all VECMs, one cointegrating function was identified.

**The endogenous lag length was selected according to the Akaike Info Criterion (AIC) as in Akaike (1974).

***For the calculation of the Sharpe Ratio, we use an average of the 1-month T-bill yield for the examined period (0.2% p.a.) as risk-free rate. For the specification of the Sharpe Ratio, see Appendix A.5.