Evaluation of News-Based Trading Strategies

Stefan Feuerriegel[(✉)] and Dirk Neumann

Information Systems Research, University of Freiburg, Platz der Alten Synagoge,
79098 Freiburg, Germany
stefan.feuerriegel@is.uni-freiburg.de

Abstract. The marvel of markets lies in the fact that dispersed information is instantaneously processed by adjusting the price of goods, services and assets. Financial markets are particularly efficient when it comes to processing information; such information is typically embedded in textual news that is then interpreted by investors. Quite recently, researchers have started to automatically determine news sentiment in order to explain stock price movements. Interestingly, this so-called news sentiment works fairly well in explaining stock returns. In this paper, we attempt to design trading strategies that are built on textual news in order to obtain higher profits than benchmark strategies achieve. Essentially, we succeed by showing evidence that a news-based trading strategy indeed outperforms our benchmarks by a 9.06-fold performance.

Keywords: Financial news · Decision support · Trading strategies · Text mining · Sentiment analysis

1 Introduction

Market efficiency relies, to a large extent, upon the availability of information. Nowadays, market information can be accessed easily as it comes naïvely with the prevalence of electronic markets and, because of the straightforward access, decision makers can use such information (e. g. [18]) to make purchases and sales more beneficial. In the same context, the so-called (weak) efficient market hypothesis [11] asserts that financial markets are *informationally efficient*, in the sense that financial stock prices accurately reflect all public information at all times; price adjustments occur when previously unknown information enters the market, as is the case with *news*.

Several publications (e. g. [5,6,43]) study the market reception of news announcements, finding a causal and clearly measurable relationship between financial disclosures and stock market reaction. Market reception is not only triggered by the quantitative facts embedded in financial disclosures, but, more importantly, qualitative information drives stock market reactions to financial disclosures, since news is typically embodied in text messages. In order to extract tone from a textual content, one frequently measures the polarity of news by measuring the so-called *news sentiment*.

While previous research [2, 21, 47, 48] succeeded in establishing a link between news tone and stock market prices, it is not clear how the extracted sentiment signals can then be utilized to facilitate investment decisions. To close this gap, this paper studies how news sentiment, as an emergent trend of IT in the finance industry, can enrich news-based trading. News trading combines real-time market data and natural language processing to detect suitable news announcements in order to trigger transactions. Its mechanisms are often part of an algorithmic trading system, while many regard it as an enabling Decision Support System (DSS) for use in banking and financial markets [16].

Consequently, this paper investigates how a Decision Support System can utilize news sentiment to perform stock trading in practice. Overall, our contribution is as follows: first, we propose different rule-based trading strategies. Second, we find quantitative evidence that our news trading system can outperform benchmark scenarios. In addition, news-based trading profits from incorporating other external variables, such as price momentum, to achieve better estimates of possible market reaction in that specific economic cycle.

The remainder of this paper is structured as follows. In Sect. 2, we review related research on the sentiment analysis of financial disclosures and news trading, in which we focus particularly on how both can be tied within a trading system. Next, Sect. 3 describes the data sources, as well as the news corpus, that is integrated into the sentiment analysis to extract the subjective tone of financial disclosures. The calculated sentiment values are then integrated (Sect. 4) into various news trading strategies and, finally, Sect. 5 evaluates these strategies in terms of their financial performance.

2   Related Work

In this section, we present related literature grouped into two categories: first, we compare algorithmic approaches that measure news sentiment in financial disclosures. Second, we review previous works from both IT and finance that distill news into trading actions as part of a decision support system for investments.

2.1   News Sentiment in the Financial Domain

Methods that use the textual representation of documents to measure the positivity and negativity of the content are referred to as opinion mining or *sentiment analysis*. Sentiment analysis can be utilized to extract subjective information from text sources, as well as to measure how market participants perceive and react to financial materials. In this case, one uses the observed price reactions following financial text to validate the accuracy of the news sentiment. Based upon sentiment measures, one can study the relationship between financial documents and their effect on markets. Empirical evidence, for example, shows that a discernible relationship between news content and stock market reaction exists, see e.g. [2, 47] for some of the first analyses.

As sentiment analysis is applied to a broad variety of domains and textual sources, research has devised various approaches to measure sentiment. A recent literature overview [37] provides a comprehensive domain-independent survey and, within the domain of finance, a number of surveys [32,35] compare studies aimed at stock market prediction. For example, *dictionary-based approaches* are very frequently used in recent financial text mining research [8,21,23,29,48]. These methods count the frequency of pre-defined positive and negative words from a given dictionary, producing results that are straightforward and reliable. In comparison, *machine learning approaches* [2,28,34,41, for example] offer a broad range of methods, but may suffer from overfitting [42].

In this paper, we want to address only the trading simulation itself and so utilize a dictionary-based approach to allow for easier verification of our results. Furthermore, dictionary approaches seem to be the more widespread technique nowadays in finance literature.

2.2   News Trading

This section provides a brief overview of components that are necessary for a news trading system. A more detailed review and taxonomy can be found in [16]. In behavioral finance, news trading has long been associated with both noise and arbitrage trading [3]. Noise traders chase profit from overreactions to momentary events, whereas arbitrageurs exploit mispricing, possibly indicated by news stories [44,45]. As a consequence, both a lack of information and misunderstanding may contribute to gains [1,31] from news trading. Most likely, an even larger advantage would result from automated transactions that a Decision Support System could trigger, before human information processing triggered a buy/sell decision, following news releases.

As a first key ingredient for a realistic study of news trading, we need to account for incurred *transaction fees*. Several previous research papers perform trading simulations, but many neglect the importance of transaction fees.

– For example, a support vector machine using financial news yields an accuracy of 71 % in predicting the direction of asset returns [41]. Overall, this gives an excess return of 2.88 % compared to the S&P 500 index between October 25, 2005 and November 28, 2005. A different approach uses decision rules with risk words [27] which yields an average annual excess of 20 % compared to U. S. Treasury bills. Similarly, Tetlock [47] incorporates pessimistic words and yields a 7.3 % plus, when compared to the Dow Jones, with the applied trading strategy. However, all the aforementioned papers share a lack of consideration for transaction fees, which, in fact, can be substantial [20].
– One of the few papers that considers transaction fees also utilizes German ad hoc announcements and yields an accuracy of 65 % when predicting the direction of returns. Along with transaction fees of 0.1 %, the average return per transaction accounts for 1.1 %. However, this paper [20] relies on one basic strategy (similar to our simple news-based strategy) and does not compare other trading strategies.

As a second addition to our Decision Support System, we need to test our trading strategies using *benchmark scenarios*. A common approach is to utilize a benchmark stock index for comparison. The author of [49] predicts the direction of stock price movements via sparse matrix factorization utilizing news articles from the Wall Street Journal. The results show an accuracy rate of 55.7 %, higher than when compared to a reference index. As an alternative, related research also integrates simple buy-and-hold strategies of stocks with the highest historic returns.

The third component of a news trading system involves *trading strategies*. Simple buy-and-hold strategies are common when testing news trading in historic portfolio simulation. For example, the authors of [25] hypothesize that a sentiment-based selection strategy outperforms a classical buy-and-hold benchmark strategy (holding all stocks over the whole test period). This approach is what the authors call a portfolio selection test. In a different paper [33], the authors build a news categorization and trading system to predict stock price trends. The system is combined with a trading engine that generates trading recommendations in the form of *"buy stock X and hold it until the stock prices hit the +d % barrier"*. In a similar fashion, a trading strategy can be built around the Google query volume [39] for search terms related to finance. This variable is integrated into a simple buy-and-hold strategy (without transaction costs) to buy the Dow Jones index at the beginning and sell it at the end of the hold period, in which the authors tested various lengths of holding strategies. This strategy yields a 16 % profit, equal to the overall increase in value of the Dow Jones index in the time period from January 2004 until February 2011.

3   Background

This section introduces background knowledge for both datasets and the sentiment analysis. First, we describe the construction of the news corpus that is used throughout this paper. We then transform this running text into machine-readable tokens to measure news sentiment.

3.1   Data Sources

Our news corpus originates from regulated ad hoc announcements[1]. These announcements must be published for all listed firms in Germany in English. We choose this data source primarily because companies are bound to disclose these ad hoc announcements as soon as possible through standardized channels, thereby enabling us to study the short-term effect of news disclosures on stock prices. In research, ad hoc announcements are a frequent choice [35] when it comes to evaluating and comparing methods for sentiment analysis. In addition, this type of news corpus shows several advantages: ad hoc announcements must be authorized by company executives, the content is quality-checked by the

---

[1] Kindly provided by Deutsche Gesellschaft für Ad-Hoc-Publizität (DGAP).

Federal Financial Supervisory Authority[2] and several publications analyze their relevance to the stock market – finding a direct relationship (e. g. [14,15,36]).

Our collected announcements date from the beginning of January 2004 until the end of June 2011. We investigate such a long time period to avoid the possibility of only analyzing news driven predominantly by a single market event, for example, the financial crisis. In addition, we apply the following *filter rules*. First, each announcement must have at least 50 words. Second, we focus only on ad hoc press releases from German companies which are written in the English language. Our final corpus consists of 14,463 ad hoc announcements. To study stock market reaction, we use the daily stock market returns of the corresponding company, originating from Thomson Reuters Datastream. We only include business days and, due to data availability, we yield a total of 2108 observations. In addition, we adjust the publication days of ad hoc announcements according to the opening times of the stock exchange. This is achieved by counting all disclosures after 8 p.m. to the next day.

We later also integrate a stock market index into our analysis as follows: in our analysis, the so-called CDAX index works as a benchmark which the trading strategies need to combat in terms of performance. The CDAX (see Fig. 1) is a German stock market index calculated by Deutsche Börse. It is a composite index of all stocks traded on the Frankfurt Stock Exchange that are listed in the General Standard or Prime Standard market segments, giving a total of 485 stocks. It does not contain foreign firms or foreign stocks, thus serving as a suitable match to our ad hoc news corpus.

Fig. 1. The CDAX as a German stock market index representing our benchmark.

---

[2] Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin).

3.2   Sentiment Analysis

Methods that use the textual representation of documents to measure positive and negative content are referred to as opinion mining or *sentiment analysis* [37]. In fact, sentiment analysis can be utilized [32,35] to extract subjective information from text sources, as well as to measure how market participants perceive and react to news. One uses the observed stock price reactions following a news announcement to validate the accuracy of the sentiment analysis routines. Thus, sentiment analysis provides an effective tool chain to study the relationship between news content and its market reception [2,12,47].

Before performing the actual sentiment analysis, there are several preprocessing steps as follows:

1. **Tokenization.** Corpus entries are split into single words named *tokens* [19].
2. **Negations.** Negations invert the meaning of words and sentences [7,40]. When encountering the word *no*, each of the subsequent three words (i.e the object) are counted as words from the opposite dictionary. When other negating terms are encountered (*rather, hardly, couldn't, wasn't, didn't, wouldn't, shouldn't, weren't, don't, doesn't, haven't, hasn't, won't, hadn't, never*), the meaning of all succeeding words is inverted.
3. **Stop Word Removal.** Words without a deeper meaning, such as *the, is, of,* etc. are named *stop words* [30] and can be removed. We use a list of 571 stop words proposed in [26].
4. **Synonym Merging.** Synonyms, though spelled differently, convey the same meaning. In order to group synonyms by their meaning, we follow a method that is referred to as pseudoword generation [30]. Approximately 150 frequent synonyms or phrases from the finance domain are aggregated according to their meanings.
5. **Stemming.** Stemming refers to the process of reducing inflected words to their stem [30]. Here, we use the so-called Porter stemming algorithm [38].

Having completed the preprocessing, we can continue to analyze news sentiment. As shown in a recent study [13] on the robustness of sentiment analysis, the correlation between news sentiment and stock market returns varies across different sentiment metrics. A sentiment approach that results in a reliable correlation is the Net-Optimism metric [8]. Net-Optimism works well in coordination with Henry's Finance-Specific Dictionary [21]. Consequently, we rely upon this approach in the following evaluation. The metric is calculated as the difference between the number of positive $W_{\text{pos}}(A)$ and negative $W_{\text{neg}}(A)$ words divided by the total number of words $W_{\text{tot}}(A)$ in an announcement A. Thus, Net-Optimism sentiment $S(A)$ is defined by

$$S(A) = \frac{W_{\text{pos}}(A) - W_{\text{neg}}(A)}{W_{\text{tot}}(A)} \in [-1, +1]. \tag{1}$$

4   Trading Strategies

This section introduces all the trading strategies that serve as a foundation for our analysis. Consistent with the existing literature, we start by present-

ing our benchmark, namely, a momentum trading approach. This strategy derives purchase decisions solely from the historic returns of assets by maximizing the so-called rate-of-change. In addition, we propose news-based trading strategies in which investment decisions are triggered by news sentiment signals. Then, we combine both methods and develop a strategy that utilizes both historic prices and news sentiment.

In the subsequent algorithms, we use the following notation: let $p_{i,t}$ denote the closing price of a stock i at time t. Furthermore, the variable $S(A)$ gives the news sentiment of an announcement A corresponding to stock i as defined above.

When trading, we exclude all so-called penny stocks (i.e stocks below €5) from our evaluation [4,9]. The reason behind this is that these penny stocks tend to react more unsystematically to trends and news announcements and, consequently, may introduce a larger noise component to our data.

4.1   Benchmark: Momentum Trading

Past stock returns can be a predictor of future firm performance. This is what we define as *momentum*, in which historic stock prices continue moving in their previous direction. The (partly) predictable connection between past and future return has been proven in the finance literature, such as in [22]. Nevertheless, finance academics have trouble with the finding that a simple strategy of buying winners and selling losers can apparently be profitable, since this contradicts the theory of efficient markets [11], where markets quickly absorb new information and adjust asset prices accordingly. Momentum is, consequently, also named a *"premier anomaly"* in stock returns [10]. By extrapolating historic stock trends, we motivate the following *momentum trading*, which picks up the subtle patterns in returns. Developing a successful momentum trading strategy is primarily a merit from the manual efforts of finance academics and practitioners to hand-engineer features from historical prices [46].

Let us define both the terms momentum and rate-of-change respectively. The so-called *momentum $Mom_{i,t}$* is the absolute difference in stock i defined by

$$Mom_{i,t} = p_{i,t} - p_{i,t-\delta} \tag{2}$$

with a time span of δ days. In short, momentum denotes the difference between today's closing price and the closing price N days ago, thus referring to prices continuing to trend. In comparison, the *rate-of-change RoC_i* represents the relative change as a fraction, i.e

$$RoC_{i,t} = \frac{p_{i,t} - p_{i,t-\delta}}{p_{i,t-\delta}} = \frac{Mom_{i,t}}{p_{i,t-\delta}} \ . \tag{3}$$

Both the momentum and rate-of-change indicators indicate trend by remaining positive during an uptrend or negative during a downtrend.

Altogether, this results in the momentum trading strategy [24], formally introduced by the following pseudocode. In short, the key idea is to always

choose the stock that has the highest rate-of-change. Step 1 initializes the variable s which stores the stock that our Decision Support System currently holds. The subsequent for-loop iterates through all time steps of our simulation horizon T. In each iteration, Step 3 updates the rate-of-change scales for all stocks, excluding the current business day. If the previously held stock was empty, then Step 5 invests in the stock with the highest absolute value of all historic rate-of-change values. However, if the rate-of-change of the currently held stock drops below a threshold θ_{RoC}, then we trigger transactions to sell the previous stock (Step 7) and buy (or short-sell) new stock with the highest rate-of-change in Step 8.

As free parameters, we can vary the time span δ calculating the rate-of-change and the threshold θ_{RoC}. For the former, we find good results with δ set to 200 business days. This value serves as a good trade-off in between the range of 20 days to 12 months proposed in the literature [22]. We choose the latter variable θ_{RoC} by testing different values, and decide for $\theta_{\mathrm{RoC}} = 50\,\%$.

1: Initialize stock $s \leftarrow \bot$.
2: **for** t **in** T **do**
3:    Compute $RoC_{i,t-1} \leftarrow \dfrac{p_{i,t-1} - p_{i,t-1-\delta}}{p_{i,t-1-\delta}}$ for all stocks i.
4:    **if** $s = \bot$ **then**
5:       Buy or short-sell stock $s \leftarrow \arg\max_i |RoC_i|$.
6:    **else if** $|RoC_{s,t-1}| < \theta_{\mathrm{RoC}}$ **then**
7:       Remove investment in stock s.
8:       Buy or short-sell stock $s \leftarrow \arg\max_i |RoC_i|$.
9:    **end if**
10: **end for**

4.2   News Trading

While the previous strategy only utilizes historic stock prices, we now instead focus on news sentiment in order to enable news-based purchase decisions. In order to react to news sentiment signals, our Decision Support System needs to continuously scan the news stream and compute the sentiment once a new financial disclosure is released. When the news sentiment associated with this press release is either extremely positive or negative, this implies a strong likelihood of a subsequent stock market reaction in the same direction. We benefit from the stock market reaction if an automated transaction is triggered beforehand.

To achieve this goal, we specify the so-called simple *news trading strategy*, given in the pseudocode below. Steps 2 and 3 trigger buy and short-sell decisions, whenever the absolute value of the news sentiment metric of an incoming announcement exceeds a certain positive or negative threshold. This decision is given by the if-statement in Step 1, i.e the condition that $S(A)$ is smaller than a negative threshold θ_S^- or larger than a positive θ_S^+ must be fulfilled. We choose suitable threshold values for both θ_S^- and θ_S^+ as part of our evaluation in Sect. 5.

Input: Released announcement A that corresponds to stock i.
1: **if** $S(A) > \theta_S^+$ **or** $S(A) < \theta_S^-$ **then**
2:    Remove investment in previous stock s.
3:    Buy or short-sell stock $s \leftarrow i$.
4: **end if**

4.3   Combined Strategy with News and Momentum Trading

The subsequent trading strategy combines the above approaches by utilizing both news sentiment and historic prices in the form of momentum. We develop this trading strategy around the idea that we want to invest in assets with both (1) a news disclosure with a high polarity and (2) a previous momentum in the same direction. The combined pseudocode is given below. Only if both the news release and historic prices give an indication of a development in the same direction, Steps 3 and 4 trigger a corresponding trading decision. Thus, this strategy expects the same direction in terms of the return-of-change and sentiment metric as tested in Step 2.

Input: Released announcement A that corresponds to stock i at day t.
1: Compute $RoC_{i,t-1} \leftarrow \dfrac{p_{i,t-1} - p_{i,t-1-\delta}}{p_{i,t-1-\delta}}$ for stock i.
2: **if** $|S(A)| > \theta_S$ **and** $\operatorname{sign} S(A) = \operatorname{sign} RoC_{i,t-1}$ **then**
3:    Remove investment in stock s.
4:   Buy or short-sell stock $s \leftarrow i$.
5: **end if**

5   Evaluation

The above sections have presented a number of trading strategies that differ in the way in which operations are derived; this section evaluates these trading strategies in terms of their achieved performance. We first focus on our benchmark strategies, i.e momentum trading and German composite stock market index. We then determine suitable values for all free parameters inside the news trading algorithms and analyze their performance.

5.1   Benchmarks: Stock Market Index and Momentum Trading

As our first benchmark, we choose the so-called CDAX, a German stock market index calculated by Deutsche Börse. It is a composite index of all stocks traded on the Frankfurt Stock Exchange that are listed in the General Standard or Prime Standard market segments. Figure 1 shows the overall development of the stock index from the year 2004 until mid-2011. During that period, the index increased by 50.99 %, which correspond to 5.65 % at an annualized rate. The number of days with positive returns outweigh the negative days by 1092 to 864.

Simple momentum trading acts as our second benchmark. This strategy works with no data input other than historic stock prices. When historic prices continue their trend, we can invest in the specific stock to profit from this development. While news trading yields high returns at first, the profits later plummet, resulting in a negative cumulative return of -65.33%. Nevertheless, the average daily returns remain positive at 0.0464% and are even higher than that of the CDAX index (0.0298%). The number of days with positive returns outweighs the negative by 1154 to 767.

The results of both benchmarks, namely, the CDAX stock market index and momentum trading, are presented in Fig. 2 where we see different performance patterns.
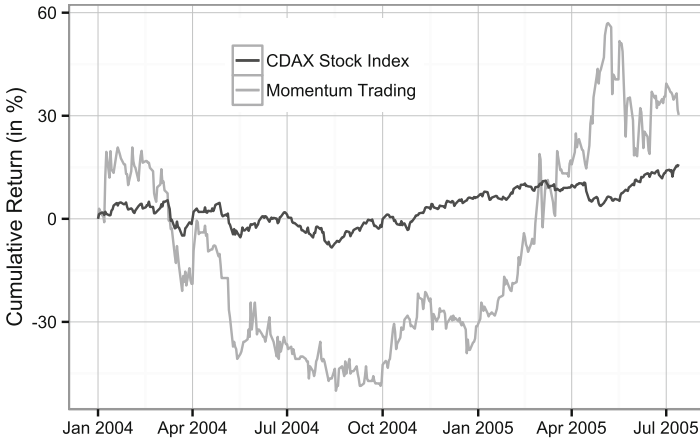
Fig. 2. Cumulative returns of both the CDAX index and the momentum trading strategy compared across the first 400 business days.

5.2 News Trading

This section evaluates different variants of news trading, starting with a simple trading strategy. This strategy triggers transactions whenever a very positive or negative ad hoc announcement is disclosed. Here, we take into account only a single announcement per day and business days, giving a total corpus of 1894 disclosures.

What remains unanswered thus far is av value for the threshold θ_S above which our news-based trading strategies perform a purchase decision. In order to find the optimal parameter, Fig. 3 compares the thresholds θ_S^+ and θ_S^- against the average returns. For reasons of simplicity, we measure these thresholds in terms of quantiles of the news sentiment distribution. We see that a threshold value of

around 10 % appears in a cluster of high daily returns and yields good results. Thus, we decided to set θ_S^+ to the 90 % quantile (and θ_S^- to the 10 % quantile) of the sentiment values $S(A)$ in order to make this variable exogenously given. However, it is important to stress that there large variations in performance depending on the threshold.

Evaluating the above strategies with historic data reveals the following findings:

– With the threshold set to the 10 % quantile, we gain an overall cumulative return of 178.53 % at a volatility of 0.078. The average daily return accounts for 0.421 %.
– In addition, we include a combination of news and momentum trading. This strategy leads to a lower performance with a cumulative return of 361.60 % and average daily returns of 0.1180 %. However, this strategy simultaneously reveals a reduced risk component in the form of less volatility, which accounts for 0.028.

Both strategies, namely, simple news trading and the combined version, are further evaluated in the following diagrams. Figure 4 depicts how the cumulative returns develop during the first 1500 business days, showing clearly the superiority of simple news trading.
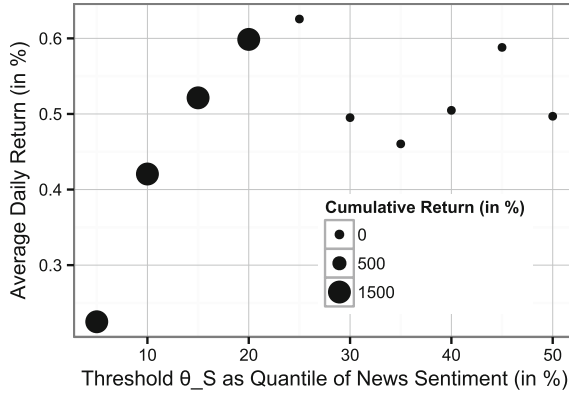
Fig. 3. Comparison of thresholds θ_S^+ and θ_S^- against average daily returns. The threshold is measured as quantiles from both ends of the average news sentiment in the corpus. In addition, the point size indicates the total cumulative return.

5.3   Comparison

The simulation horizon spans January 2014 until the end of June 2011, giving a total of 1956 business days. All results of our trading simulation are provided

Fig. 4. Cumulative returns of both news trading and the combination of news and momentum trading compared across the first 1500 business days.

in Table 1. Here, we evaluate how well the investment decisions of each strategy play together with market feedback. We focus mainly on average daily return, since cumulative returns can be misleading. The reason is as follows: one wrong trade can make performance plummet, while a high average daily return indicates a continuously high benefit. In addition, we want to direct attention to the volatility column. These values serve as an indicator of the level of risk associated with each strategy. Even though simple news trading achieves higher returns, it is linked with higher volatility and higher risks. Thus, it may be beneficial for practitioners to follow a combined strategy of news and momentum trading that results in smaller returns, while also decreasing the associated risk.

We now compare our benchmarks to news trading. The benchmarks feature mean returns of 0.0298 % for the CDAX and 0.0464 % for momentum trading. In comparison, news trading reaches 0.4206 %. This is the 9.06-fold value, but linked with a considerably higher volatility.

While we put an emphasis on raw returns, we also provide a performance measure that incorporates transaction costs. Thus, the last column of Table 1 reveals how an initial invest of €1000 would evolve over time. Consistent with [20, 48][3], we simulate the portfolio with a transaction fee for each buy/sell operation of 0.1 %, equivalent to 10 bps.[4]

---

[3] Using a proportional transaction fee is common in financial research. For example, other papers [17] mostly vary transaction costs mostly in the range of 0.1 % to 0.3 % or assume a fixed transaction fee [33] of U. S. $ 10 for buying and selling stocks respectively.

[4] A frequent unit in finance is basis point (bps). Here, one unit is equal to 1/100th of 1 %, i.e 1 % = 100 bps.

Table 1. Comparison of benchmarks and trading strategies across several key performance characteristics.

Trading strategy	Cumulative return	Returns: median	Returns: mean	Annualized return	#Positive returns	#Negative returns	Volatility	ΔTrades	Portfolio outcome (€)
Benchmarks									
CDAX Index	50.99 %	0.0703 %	0.0298 %	5.6480 %	1092	864	0.01319	—	1506.92
Momentum Trading ($\theta = 50\,\%$, $\delta = 200d$)	−65.33 %	0.0000 %	0.0464 %	−13.1702 %	1129	827	0.04496	102.95d	322.94
News Trading									
Simple News Trading ($\theta_S^+ = 90\,\%$, $\theta_S^- = 10\,\%$)	17853.31 %	0.0000 %	0.4206 %	99.7803 %	1136	820	0.07780	10.57d	967448.60
Combined: News &Momentum ($\theta_S^+ = 90\,\%$, $\theta_S^- = 10\,\%$, $\delta = 200d$)	361.60 %	0.0000 %	0.1180 %	22.6222 %	1113	843	0.02839	19.96d	9513.92

6   Conclusion and Outlook

Although it is a well-known fact that financial markets are very sensitive to the release of financial disclosures, the way in which this information is received is far from being studied sufficiently. Not until recently have researchers started to look at the content of news stories using very simple techniques to determine news sentiment. Typically, these research papers concentrate on finding a link between the qualitative content and the subsequent stock market reaction. To harness this relationship in practice, news trading combines real-time market data and sentiment analysis in order to trigger investment decisions. Interestingly, what previous approaches all have in common is that they rarely study and compare trading strategies.

As a remedy, this paper evaluates algorithmic trading strategies within a Decision Support System for news trading. We propose and compare different rule-based strategies for news-based trading. As a result, our Decision Support System outperforms all benchmark scenarios by relying upon news-based investment decisions. Further performance improvements can be achieved by including external variables that, for example, describe the economic environment or lagged prices respectively. Altogether, we contribute to the understanding of information processing in electronic markets and show how to enable decision support in financial markets.

This paper opens avenues for further research into two directions. First, a multi-asset strategy could be beneficial to spread risks. To model these, intriguing approaches include Value-at-Risk (VaR) measures, as well as techniques from portfolio optimization. Second, it is worthwhile to improve the forecast of asset returns by including a broader set of exogenous predictors. As such, possible external variables range from stock market indices, fundamentals describing the economy and additional lagged variables. Further enhancements would also result from embedding innovative news sources, such as social media. Altogether, the accuracy of predicting stock return directions and trigger trading signals would be greatly improved.

References

1. Alfano, S.J., Feuerriegel, S., Neumann, D.: Is news sentiment more than just noise? In: 23rd European Conference on Information Systems (ECIS 2015), Münster, Germany, 26-29 May 2015
2. Antweiler, W., Frank, M.Z.: Is all that talk just noise? The information content of internet stock message boards. J. Finance **59**(3), 1259–1294 (2004)
3. Barberis, N., Thaler, R.: A Survey of Behavioral Finance. Financial Markets and Asset Pricing. Handbook of the Economics of Finance, pp. 1053–1128. Elsevier, Amsterdam (2003)
4. Beatty, R., Kadiyala, P.: Impact of the penny stock reform act of 1990 on the Initial public offering market. J. Law Econ. **46**(2), 517–541 (2003)
5. Cenesizoglu,T.: The reaction of stock returns to news about fundamentals. Management Science (2014)
6. Cutler, D.M., Poterba, J.M., Summers, L.H.: What moves stock prices? J. Portfolio Manag. **15**(3), 4–12 (1989)

7. Dadvar, M., Hauff, C., de Jong, F.: Scope of negation detection in sentiment analysis. In: Proceedings of the Dutch-Belgian Information Retrieval Workshop (DIR 2011), Amsterdam and Netherlands, pp. 16–20 (2011)
8. Demers, E.A., Vega, C.: Soft Information in Earnings Announcements: News or Noise? INSEAD Working Paper No. 2010/33/AC, SSRN Electronic Journal (2010)
9. Fama, E.F., French, K.R.: Multifactor explanations of asset pricing anomalies. J. Finance **51**(1), 55–84 (1996)
10. Fama, E.F., French, K.R.: Dissecting anomalies. J. Finance **63**(4), 1653–1678 (2008)
11. Fama, F.E.: The behavior of stock-market prices. J. Bus. **38**(1), 34–105 (1965)
12. Feuerriegel, S., Heitzmann, S.F., Neumann, D.: Do investors read too much into news? How news sentiment causes price formation. In: 48th Hawaii International Conference on System Sciences (HICSS) (2015)
13. Feuerriegel, S., Neumann, D.: News or noise? How news drives commodity prices. In: Proceedings of the International Conference on Information Systems (ICIS 2013), Association for Information Systems (2013)
14. Ratku, A., Feuerriegel, S., Rabhi, F., Neumann, D.: Finding evidence of irrational exuberance in the oil market. In: Workshop on Enterprise Applications, Markets and Services in the Finance Industry, FinanceCom 2014, Sydney, Australia, December 12, 2014, to be published in Springer's LNBIP (2014)
15. Feuerriegel, S., Ratku, A., Neumann, D.: Analysis of how underlying topics in financial news affect stock prices using latent dirichlet allocation. In: Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS), Kauai, January 5–8, 2016, IEEE Computer Society (2016)
16. Gagnon, S.: Rules-based integration of news-trading algorithms. J. Trading **8**(1), 15–27 (2013)
17. Graf, F.: Mechanically Extracted Company Signals and their Impact on Stock and Credit Markets (2011)
18. Granados, N., Gupta, A., Kauffman, R.J.: Research commentary-information transparency in business-to-consumer markets: concepts, framework, and research agenda. Inf. Syst. Res. **21**(2), 207–226 (2010)
19. Grefenstette, G., Tapanainen, P.: What is a word, what is a sentence? Problems of Tokenization (1994)
20. Hagenau, M., Liebmann, M., Hedwig, M., Neumann, D.: Automated news reading: stock price prediction based on financial news using context-specific features. In: 45th Hawaii International Conference on System Sciences (HICSS), pp. 1040–1049 (2012)
21. Henry, E.: Are investors influenced by how earnings press releases are written? J. Bus. Commun. **45**(4), 363–407 (2008)
22. Jegadeesh, N., Titman, S.: Returns to buying winners and selling losers: implications for stock market efficiency. J. Finance **48**(1), 65–91 (1993)
23. Jegadeesh, N., Di, W.: Word power: a new approach for content analysis. J. Financ. Econ. **110**(3), 712–729 (2013)
24. Kim, K.: Financial time series forecasting using support vector machines. Neurocomputing **55**(1–2), 307–319 (2003)
25. Klein, A., Altuntas, O., Hausser, T., Kessler, W.: Extracting investor sentiment from weblog texts: a knowledge-based approach. In: IEEE 13th Conference on Commerce and Enterprise Computing (CEC), pp. 1–9 (2011)
26. Lewis, D.D., Yang, Y., Rose, T.G., Li, F.: RCV1: A new benchmark collection for text categorization research. J. Machine Learning Res. **5**, 361–397 (2004)

27. Li, F.: Do stock market investors understand the risk sentiment of corporate annual reports? SSRN Electronic Journal (2006)
28. Li, F.: The information content of forward-looking statements in corporate filings: a Naïve Bayesian Machine learning approach. J. Account. Res. **48**(5), 1049–1102 (2010)
29. Loughran, T., McDonald, B.: When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. J. Finance **66**(1), 35–65 (2011)
30. Manning, C.D., Schütze, H.: Foundations of Statistical Natural Language Processing. MIT Press, Cambridge (1999)
31. Mendel, B., Shleifer, A.: Chasing Noise. J. Financ. Econ. **104**(2), 303–320 (2012)
32. Minev, M., Schommer, C., Theoharry G.: News And Stock Markets: A Survey On Abnormal Returns and Prediction Models (2012)
33. Mittermayer, M.A.: Forecasting intraday stock price trends with text mining techniques. In: Sprague, R.H. (ed.) the 37th Annual Hawaii International Conference on System Sciences, (2004) IEEE Computer Society Proceedings of Los Alamitos, California
34. Mittermayer, M.A., Knolmayer, G.F.: NewsCATS: A news categorization and trading system. In: Sixth International Conference on Data Mining (ICDM'06), pp. 1002–1007 (2006)
35. Mittermayer, M.A., Knolmayer, G.F.: Text Mining Systems for Market Response to News: A Survey (2006)
36. Muntermann, J., Guettler, A.: Intraday stock price effects of ad hoc disclosures: the German case. J. Int. Finan. Markets. Institutions and Money **17**(1), 1–24 (2007)
37. Pang, B., Lee, L.: Opinion mining and sentiment analysis. FNT Inf. Retrieval **2**(1–2), 1–135 (2008)
38. Porter, M.F.: An algorithm for suffix stripping. Program **14**(3), 130–137 (1980)
39. Preis, T., Moat, H.S., Stanley, H.E.: Quantifying trading behavior in financial markets using Google Trends. Sci. Rep. **3**, 1684 (2013)
40. Pröllochs, N., Feuerriegel, S., Neumann, D.: Enhancing sentiment analysis of financial news by detecting negation scopes. In: 48th Hawaii International Conference on System Sciences (HICSS), pp. 959–968 (2015, Forthcoming)
41. Schumaker, R.P., Chen, H.: Textual analysis of stock market prediction using breaking financial news. ACM Trans. Inf. Syst. **27**(2), 1–19 (2009)
42. Sharma, A., Dey. S.: A comparative Study of Feature selection and machine learning techniques for sentiment analysis. In: Cho, Y. (ed.) and ACM Special Interest Group on Applied Computing: Proceedings of the 2012 Research in Applied Computation Symposium (RACS 2012), pp. 1–7 (2012). ACM, New York
43. Shiller, R.J.: Irrational Exuberance. Princeton University Press, Princeton (2005)
44. Shleifer, A., Summers, L.H.: The noise trader approach to finance. J. Econ. Perspect. **4**(2), 19–33 (1990)
45. Shleifer, A., Vishny, R.W.: The limits of arbitrage. J. Finance **52**(1), 35–55 (1997)
46. Takeuchi, L., Lee, Y.Y.: Applying Deep Learning to Enhance Momentum Trading Strategies in Stocks (2013)
47. Tetlock, P.C.: Giving content to investor sentiment: the role of media in the stock market. J. Finance **62**(3), 1139–1168 (2007)
48. Tetlock, P.C., Saar-Tsechansky, M., Macskassy, S.: More than words: quantifying language to measure firms' fundamentals. J. Finance **63**(3), 1437–1467 (2008)
49. Wong, F.M.F., Liu, Z., Chiang, M.: Stock market prediction from WSJ: text mining via sparse matrix factorization. In: IEEE International Conference on Data Mining (ICDM), pp. 430–439. IEEE Computer Society, Shenzhen, China (2014)

![Springer logo]