EVENT-DRIVEN FEEDS

QUANTITATIVE RESEARCH

Embedded Value in Bloomberg News & Social Sentiment Data

Prepared by: Xin Cui, Daniel Lam, Arun Verma, Bloomberg L.P.

Abstract

To explore the value embedded in Bloomberg News & Social Sentiment data, we have built three types of equity trading strategies based on sentiment data and shown that these strategies significantly outperform the corresponding benchmark indexes.

Disclaimer: This white paper offers research for idea-generation purposes and should not be used as direct advice on trading or investment. Although best practices have been followed in the research process, Bloomberg L.P. assumes no liability or responsibility of any kind.

Bloomberg
FOR ENTERPRISE

Contents

1. Introduction

When rational arbitrageurs have limited risk-bearing capacity and time horizons, the actions of irrational noise traders can affect asset prices (De Long, Shleifer, Summers, & Waldmann, 1990a). Such actions can be interpreted as being driven by fluctuating investor sentiment, which creates the possibility of trading profitably on the basis of investor sentiment—most obviously by being a contrarian. Under some circumstances, however, it may be rational to "jump on the bandwagon" and bet with rather than against noise traders (De Long, Shleifer, Summers, & Waldmann, 1990b). Various proxies for investor sentiment have been proposed (Baker & Wurgler, 2006), but perhaps the most direct way to measure sentiment in the stock market is to analyze the words of people commenting on stocks. One traditional source of such words is stories in the news media (Tetlock, 2007); more recently, Google searches and Twitter feeds have been used (Mao, Counts, & Bollen, 2015).

The high volume and time sensitivity of news and social media stories necessitate automated processing to extract actionable information, while the unstructured nature of textual information presents challenges that are well-suited to machine-learning techniques. Bloomberg has applied such techniques to identify a news story or tweet as being relevant for an individual stock ticker and to assign a sentiment score to each story or tweet in the feed. In this paper, we examine these scores, focusing on whether and how using News and Social Sentiment information in trading strategies can achieve good risk-adjusted returns.

2. Bloomberg News and Social Sentiment data

Supervised statistical machine-learning techniques are used to construct News & Social Sentiment from the stories text at Bloomberg. Bloomberg News and Social Sentiment classification engines are trained to mimic a human expert in processing textual information. First, a human expert manually assigns a positive, negative or neutral score to each news story or tweets. The labeling is based on the question: "If an investor having a long position in the security mentioned were to read this news or tweet, is he/she bullish, bearish or neutral on his/her holdings?" Then the annotated data is fed into machine learning models, for example, a support vector machine. Once the model has been trained, when new information comes in, the model will automatically assign probability of being positive, negative or neutral to each news story or tweet.

Bloomberg provides two types of sentiment analytics: story-level sentiment and company-level sentiment.

- Story-level sentiment is generated in real time upon the arrival of news or tweets. It consists of two parts: score and confidence. Score is a categorical value, i.e., 1, -1 and 0, which indicate positive, negative and neutral sentiment, respectively. Confidence is a numerical value ranging from 0 to 100, which can be interpreted as the probability of being positive, negative or neutral.

- Company-level sentiment is the confidence-weighted average of story-level sentiment. It only delivers one score as a numerical value ranging from -1 to 1, with -1 being the most negative sentiment and 1 being the most positive sentiment.

  For company-level intraday sentiment, the computation covers feeds with a rolling window. The news intraday sentiment score is recomputed every two minutes with an eight-hour rolling window; while the Twitter intraday sentiment score is recomputed every minute with a 30-minute rolling window. Company-level daily sentiment scores are a confidence-weighted average of the past 24 hours' story-level sentiments for both news and Twitter and are published every morning about 10 minutes before market open.

3. Trading on Sentiment

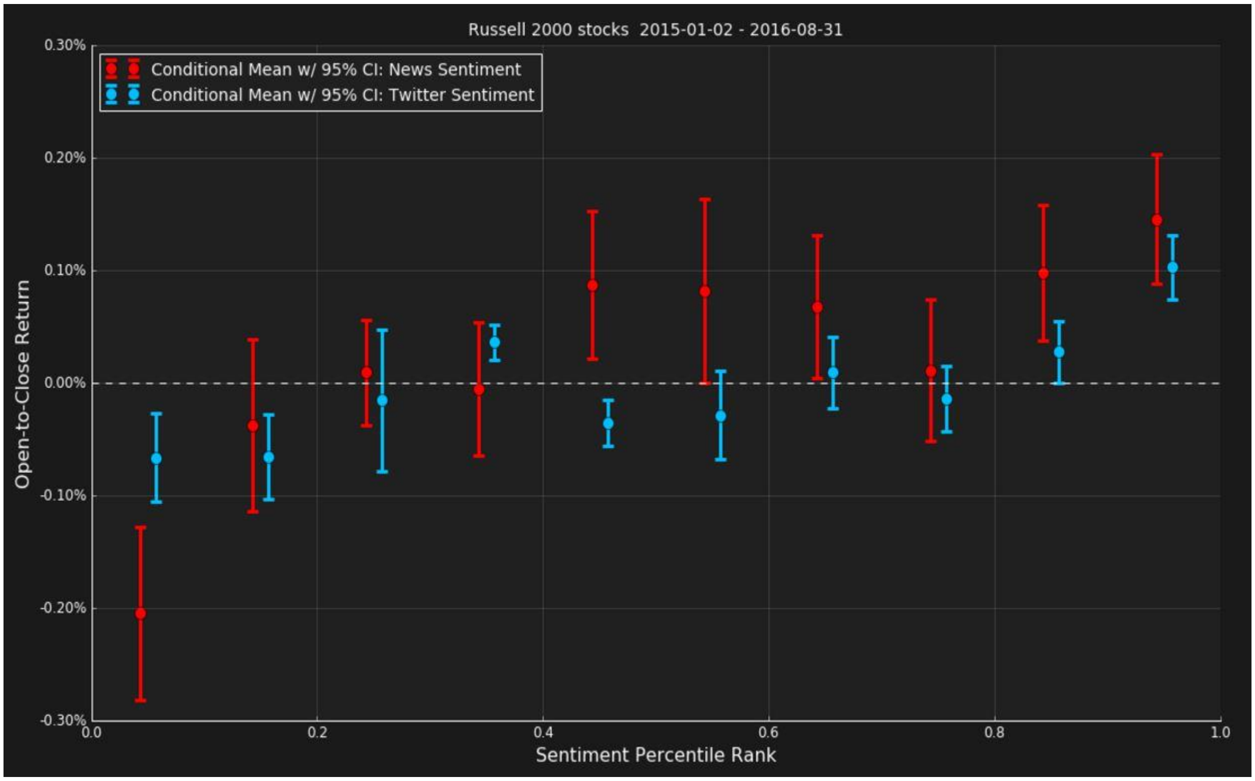
Sentiment can be used as a directional signal for trading purposes. Intuitively, if there is positive information about a particular company, we expect that company's stock price to increase; whereas if there is negative information, we expect the stock price to decrease.

To show the predictive power of Bloomberg's sentiment data, we constructed three different types of systematic equity long/short trading strategies: daily sentiment-driven strategy, daily earnings event–driven strategy and intraday sentiment-driven strategy—in all cases, we assume no transaction costs and no risk management.

3.1 Strategy I: Daily sentiment-driven strategy

The company-level daily sentiments are published every day before market open, thus a natural idea is to test if these sentiments have any predictive power for the open-to-close returns on the same day. To begin, we first compute daily percentile rank of each stock based on cross-sectional sentiment scores. Then we estimate the average return conditioned on the percentile rank across all the stocks and over a 20-month period.

In the following graph, we show the conditional average open-to-close return with 95% confidence interval for each sentiment percentile rank bucket of Russell 2000 stocks. The y-axis is the open-to-close return; the x-axis is the percentile rank of the sentiment score. Each percentile rank bucket covers roughly 10% of the data points. Because of sentiment clustering, especially for Twitter sentiment, the exact number of data points differs from bucket to bucket.



As clearly shown in the graph, conditional returns for both news and Twitter sentiments indicate momentum trading opportunity. Stocks ranked in the top quantiles have significant positive average returns, while stocks ranked in the bottom quantiles have significant negative average returns.

Our daily sentiment-driven strategy builds a daily-rebalanced long/short portfolio based on Bloomberg sentiment daily average scores. The daily sentiment is a lagging indicator as it is an aggregation of the past 24 hours' story-level sentiment. Our strategy based on daily sentiment is actually exploring if the market has efficiently priced in sentiment information.

Our strategy uses Twitter sentiment as an example and is described as follows:

- Each day before market open, rank all stocks in the given stock universe by their sentiment daily average scores.

- Construct portfolio holdings in three different variations:
  1) High-Minus-Low portfolio (HML 1/3): long (short) the top (bottom) 1/3 of stocks ranked by sentiment scores. Stocks in long and short portfolios are equally weighted.
  2) High-Minus-Low portfolio (HML 5%): long (short) the top (bottom) 5% of stocks ranked by sentiment scores. Stocks in long and short portfolios are equally weighted.
  3) Proportional portfolio (Prop): long (short) stocks with positions proportional to the difference of the sentiment score from its cross-sectional mean. If the sentiment score is above the mean, take a long position; if it is below, take a short position. The further away from the mean, the greater the position.
- Positions are created at market open and closed out at market close.

The portfolio daily return can be computed as:

$$Ret_j = \sum_{i \in Long_j} w_{ij}^{Long} \left(\frac{P_{ij}^{close}}{P_{ij}^{open}} - 1 \right) - \sum_{i \in Short_j} w_{ij}^{Short} \left(\frac{P_{ij}^{close}}{P_{ij}^{open}} - 1 \right)$$

Where

Ret_j is the portfolio return on day j;

P_{ij}^{close} is the close price of stock i on day j, P_{ij}^{open} is the open price of stock i on day j;

$Long_j$ is the basket of stocks to long on day j, w_{ij}^{Long} is the weight of stock i in $Long_j$;

$Short_j$ is the basket of stocks to short on day j, w_{ij}^{Short} is the weight of stock i in $Short_j$;

- For HML portfolio, $w_{ij}^{Long} = \frac{1}{\#\, of\ Stocks\ in\ Long_j}$, $w_{ij}^{Short} = \frac{1}{\#\, of\ Stocks\ in\ Short_j}$

- For proportional portfolio, $w_{ij}^{Long} = \frac{SS_{ij}^{Long} - \mu_j}{\sum_{i \in Long_j} (SS_{ij}^{Long} - \mu_j)}$, $w_{ij}^{Short} = \frac{\mu_j - SS_{ij}^{Short}}{\sum_{i \in Short_j} (\mu_j - SS_{ij}^{Short})}$

  μ_j is the cross-sectional mean of the company-level sentiment on day j;
  SS_{ij}^{Long} is the sentiment score of stock i in $Long_j$ on day j;
  SS_{ij}^{Short} is the sentiment score of stock i in $Short_j$ on day j.

In the backtesting, we tried four different holding periods:

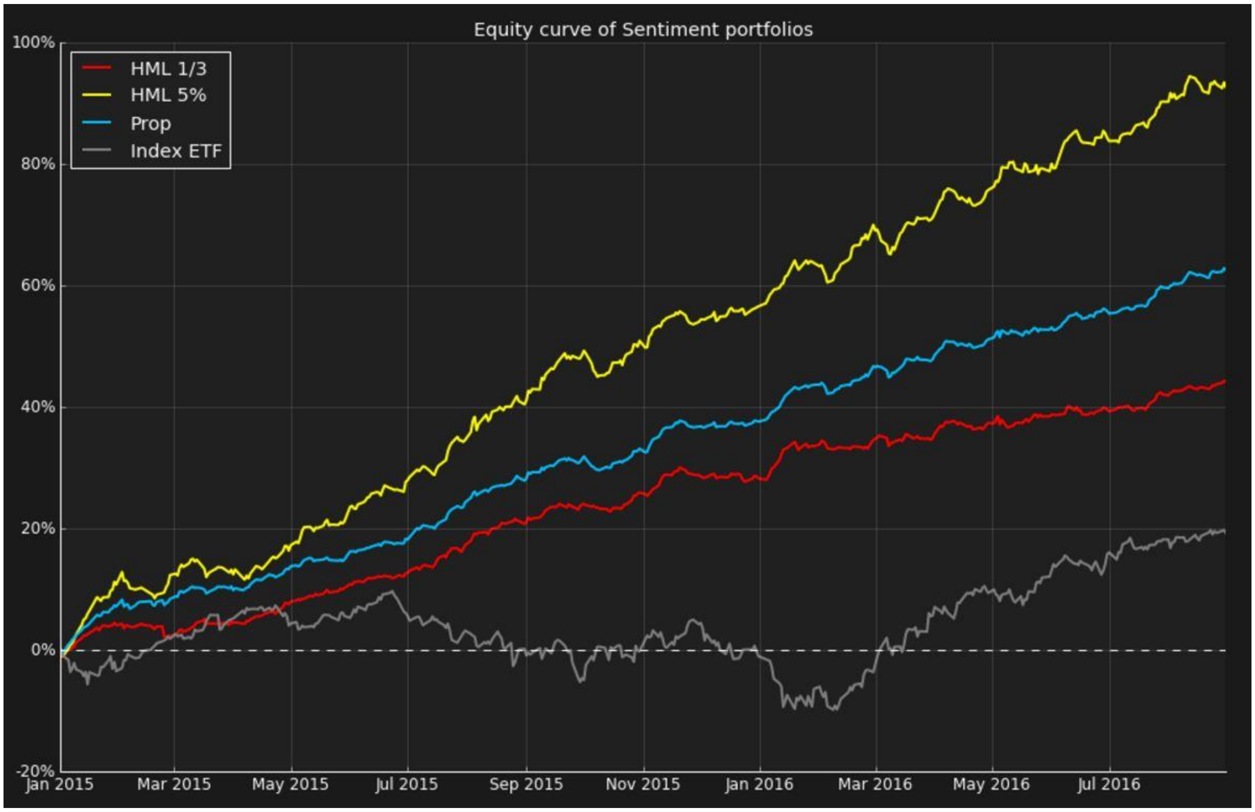
- “Same O2O” is the open-to-open return from the previous day to the same day as the sentiment score. Since the score is computed about 10 minutes before the market open, this return is not attainable by any feasible strategy using the sentiment score. Instead, it may be viewed as a rough proxy for the contemporaneous impact of news or tweets.
- “Same O2C” is the open-to-close return on the same day as the sentiment score. This is theoretically attainable but requires quick action, trading a few minutes after observing the score.
- “Next C2C” is the close-to-close return from the same day to the next day as the sentiment score. In other words, after observing the sentiment score at 9:20 AM, you wait until 4:00 PM to trade.
- “SO2NC” is just the sum of “Same O2C” and “Next C2C”. It stands for “Same-day open to next-day close.”

We backtested this strategy for S&P 500 stocks, Russell 3000 stocks and Russell 2000 stocks in the order of descending average market cap. The backtesting period was from January 2, 2015, to August 31, 2016.

The most statistically significant returns are for “Same O2O”—this is consistent across all stock universes. This also is in line with the intuition that news should drive contemporaneous return, but this return is not exploitable. The “Same O2C” also shows some statistical significance for low- and mid-cap universes, e.g., Russell 2000 and Russell 3000 stocks; it results in an exploitable trading strategy.

Equity curves of the strategies for Russell 2000 stocks are shown below with a performance statistics table. Benchmark index ETF (IWM) equity curve is calculated with open-to-close returns in order to be consistent with the sentiment strategy. The results for other portfolios are included in Appendix 5.1.

Illustration for Russell 2000 Stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
HML 1/3	-0.05	26%	6%	4.25	232	386
HML 5%	-0.09	56%	12%	4.80	58	59
Proportional	-0.04	38%	6%	5.87	796	381
Index ETF (IWM)	1.00	12%	14%	0.80	NaN	NaN

The performance statistics show that all sentiment portfolios outperform the benchmark index ETF (IWM) significantly and that the realized betas are all very close to zero, which means very small market exposure. Also, based on our backtesting results, the more diversified portfolio, i.e., the Proportional portfolio, substantially improves the risk-adjusted return.

3.2 Strategy II: Daily earnings event–driven strategy

U.S. publicly traded companies are required to file their earnings reports with the SEC every quarter. An earnings release shows the profitability of the company in the past quarter and also gives guidance for the future potential performance of a company. Based on the actual earnings results, stocks are likely to move significantly on earnings reporting days. With the earnings event–driven strategy, we show that Bloomberg News and Twitter sentiment data prior to the market open of the earnings reporting day can be used to explore the earnings day's return.

The earnings strategy goes long or short on stocks on their earnings days based on Bloomberg News/Twitter sentiment daily average scores. Our strategy is the following:

> On each day, if companies are scheduled to release earnings between current market open and the next market open:

- Long those with positive sentiment and short those with negative sentiment.
  Since the portfolio by construction is not necessarily long/short balanced, in order to reduce the market exposure, if a portfolio has only long positions, we short market ETF; if a portfolio has only short positions, we long market ETF. Stocks in long and short portfolios are equally weighted.
- Positions are created at current market open and close out at the next market open.

In the backtesting, we compared three sentiment sources: News only, Twitter only and News/Twitter combined. For the combined version, we go long on stocks only if both News and Twitter sentiment scores are positive; we go short on stocks only if both measures are negative.

The portfolio daily return can be computed as:

$$
\begin{aligned}
Ret_j = & \sum_{i \in Long_j} \frac{1}{N_j^{Long}} \left(\frac{P_{i(j+1)}^{open}}{P_{ij}^{open}} - 1 \right) \mathbb{I}\left(N_j^{Long} > 0\right) + \left(\frac{M_{(j+1)}^{open}}{M_j^{open}} - 1 \right) \mathbb{I}\left(N_j^{Long} = 0\right) \\
& - \sum_{i \in Short_j} \frac{1}{N_j^{Short}} \left(\frac{P_{i(j+1)}^{open}}{P_{ij}^{open}} - 1 \right) \mathbb{I}\left(N_j^{Short} > 0\right) - \left(\frac{M_{(j+1)}^{open}}{M_j^{open}} - 1 \right) \mathbb{I}\left(N_j^{Short} = 0\right)
\end{aligned}
$$

Where

Ret_j is the portfolio return on day j;

P_{ij}^{open} is the open price of stock i on day j;

M_j^{open} is the open price of benchmark market ETF on day j;

$Long_j$ is the basket of stocks to long on day j, N_j^{Long} is the number of stocks in $Long_j$;

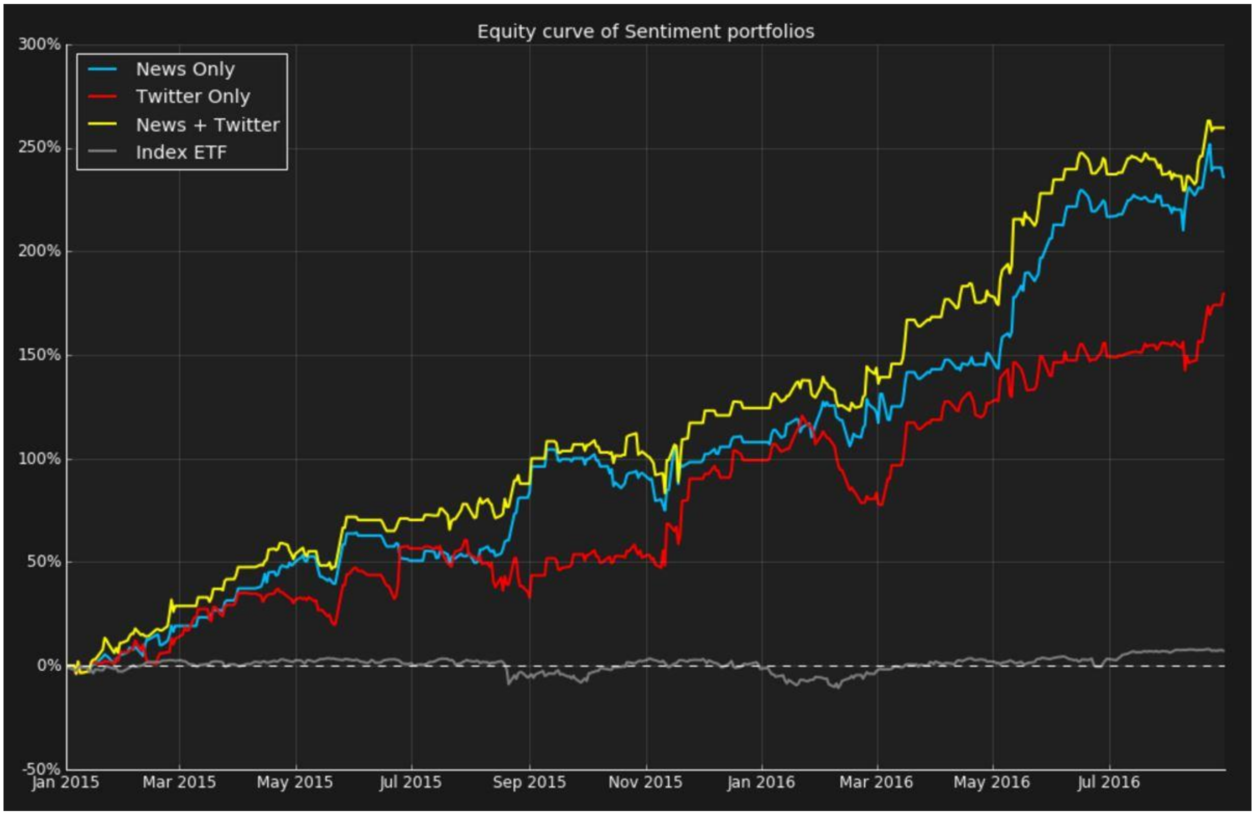
$Short_j$ is the basket of stocks to short on day j, N_j^{Short} is the number of stocks in $Short_j$.

We backtested the strategy for S&P 500 stocks, Russell 1000 stocks and Russell 2000 stocks in the order of descending average market cap. The backtesting period was from January 2, 2015, to August 31, 2016. Based on our results, this strategy works better for S&P 500 stocks. We believe that is because S&P 500 companies attract more attention and analyst coverage, so their sentiments before earnings are more likely to contain earnings-related information.

However, even for S&P 500 stocks, the signal-to-noise ratio is very low at stock level—and improves at portfolio level. One way to refine this strategy is to use topic codes to filter out irrelevant information. For example, if we are interested only in earnings-related information, we can just aggregate earnings stories or tweets to get earnings sentiment scores, which may further enhance the signal. We will test this in our future work.

Equity curves of the strategy for S&P 500 stocks are shown below with a performance statistics table. The benchmark index ETF (SPY) equity curve is calculated with open-to-open returns in order to be consistent with the sentiment strategy. The results for other portfolios are included in Appendix 5.1.

Illustration for S&P 500 Stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
News Only	0.08	142%	59%	2.39	2	7
Twitter Only	-0.14	108%	61%	1.76	4	7
News + Twitter	0.12	156%	58%	2.68	2	5
Index ETF (SPY)	1.00	4%	15%	0.29	NaN	NaN

The performance statistics show that our strategies based on News only, Twitter only and News/Twitter combined all significantly outperform the benchmark index ETF (SPY). The realized betas are all relatively small by design. More interestingly, combining News and Twitter sentiment further boosts the performance in terms of the risk-adjusted return, i.e., Sharpe ratio.

Because of the long-run sentiment bias, we also notice that taking long positions is more common than taking short positions for this strategy. For news sentiment, long positions appeared 70% of time and short positions appeared 40% of the time in the backtesting period; the average number of stocks in long positions was 7 and the average number of stocks in short positions was 2. For Twitter sentiment, long positions appeared 70% of time and short positions appeared 50% of time; the average number of stocks in long positions was 7 and the average number of stocks in short positions was 4.

3.3 Strategy III: Intraday sentiment-driven strategy

Bloomberg's story-level sentiment tracks news and tweets in real time, which enables subscribers to respond quickly to the new information in the market. Our intraday strategy longs or shorts stocks based on Bloomberg News intraday story-level sentiment. Since news stories of different companies arrive asynchronously, intraday trading in practice usually involves rebalancing positions according to these asynchronous signals. For simplicity, in order to convey the main idea, our strategy trades at evenly-spaced intervals instead. We assume we only trade during market trading hours, e.g., from 9:30 AM to 4:00 PM EST for U.S. exchange-listed stocks. The trading strategy is as follows:

- Divide each day's trading hours into N-minute intervals;
- Within each N-minute interval, if there are multiple stories on the same company, we take the average of the scores and confidence;
- At the end of each N-minute interval, long stocks with perfect positive sentiment (average score equals 1 and average confidence equals 100) and short stocks with perfect negative sentiment (average score equals -1 and average confidence equals 100). Stocks in long and short portfolios are equally weighted;
- Close out positions after N minutes.

The portfolio daily return can be computed as: $Ret_j = \sum_{k=2}^{M} Ret_{jk}^{Nmin}$

Where

Ret_j is the portfolio return on day j;

M is the number of non-overlapping N-minute intervals in a trading day;

Ret_{jk}^{Nmin} is the portfolio return of the k-th N-minute on day j and is computed as

$$Ret_{jk}^{Nmin} = \sum_{i \in Long_{jk}} \frac{1}{N_{jk}^{Long}} \left(\frac{P_{ijk}^{close}}{P_{ijk}^{open}} - 1 \right) \mathbb{1}\left(N_{jk}^{Long} > 0\right) - \sum_{i \in Short_{jk}} \frac{1}{N_{jk}^{Short}} \left(\frac{P_{ijk}^{close}}{P_{ijk}^{open}} - 1 \right) \mathbb{1}\left(N_{jk}^{Short} > 0\right)$$

P_{ijk}^{open} is the stock i's open price of the k-th N-minute on day j;

P_{ijk}^{close} is the stock i's close price of the k-th N-minute on day j;

$Long_{jk}$ is the basket of stocks to long at the beginning of the k-th N-minute on day j;

N_{jk}^{Long} is the number of stocks in $Long_{jk}$;

$Short_{jk}$ is the basket of stocks to short at the beginning of the k-th N-minute on day j;

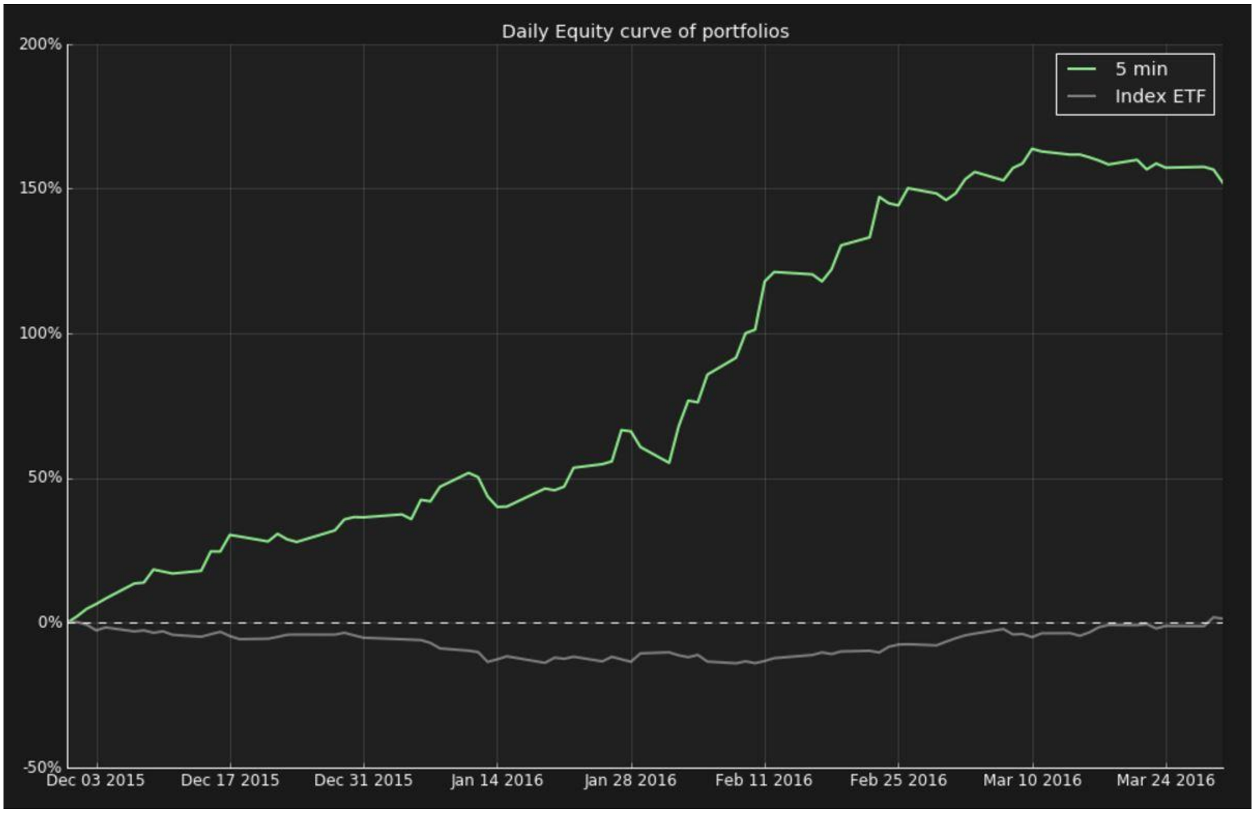
N_{jk}^{Short} is the number of stocks in $Short_{jk}$.

We used 5-min interval as an example and backtested this strategy for S&P 500 stocks, Russell 2000 stocks and NASDAQ biotechnology stocks. The backtesting period was from December 1, 2015, to March 30, 2016, with intraday 1-minute trade price bars.[1] If bars are missing for certain 5-min periods, we drop that 5-min return. Although this doesn't introduce any upward or downward bias in expectation, to be more realistic, one may use bid/ask intraday bars instead of trade bars. Also, in real trading, the close bar is not really exploitable. One may use the open bar of the next 5 minutes instead in the backtesting; however, in this case, one needs to deal with the overlapping period with the new positions of the next 5 minutes.

Based on our results, this strategy works better for small-cap stocks, e.g., Russell 2000 stocks. The equity curve of the strategy for Russell 2000 stocks is shown below with a performance statistics table. The benchmark index ETF (IWM) equity curve is calculated with open-to-close returns in order to be consistent with the sentiment strategy. Results for other portfolios are included in Appendix 5.1.

[1] The point-in-time index members and their corresponding 1-min price bars were retrieved separately on different days. If any stock changed its ticker during this period, that stock's price bars may be missing. However, this should only affect a few stocks at most.

Illustration for Russell 2000 Stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe
News	-0.94	461%	69%	6.69
Index ETF (IWM)	1.00	4%	19%	0.22

The performance statistics show that the news story–based strategy outperforms the benchmark index ETF (IWM) significantly. Since this is an intraday strategy and thus more subject to the transaction costs, a proper transaction cost model should be included for real trading purposes. We also backtested this strategy with different interval lengths, e.g., 1 minute, 10 minutes, 30 minutes and 60 minutes. For most of the cases, the best Sharpe ratio is achieved between 5 minutes and 30 minutes, depending on the stock universe. Beyond 30 minutes, we see almost no meaningful return, which reflects market efficiency in incorporating new information.

4. Conclusion

In this paper, we discuss and give illustrations of three different types of trading strategies based on Bloomberg News & Social Sentiment data. According to our backtesting results, the sentiment strategies outperform the corresponding benchmark index ETFs significantly, thus strongly demonstrating the value embedded in Bloomberg News & Social Sentiment data.

5. APPENDIX

5.1 Additional backtesting results

In this section, we show more backtesting results for different stock universes.

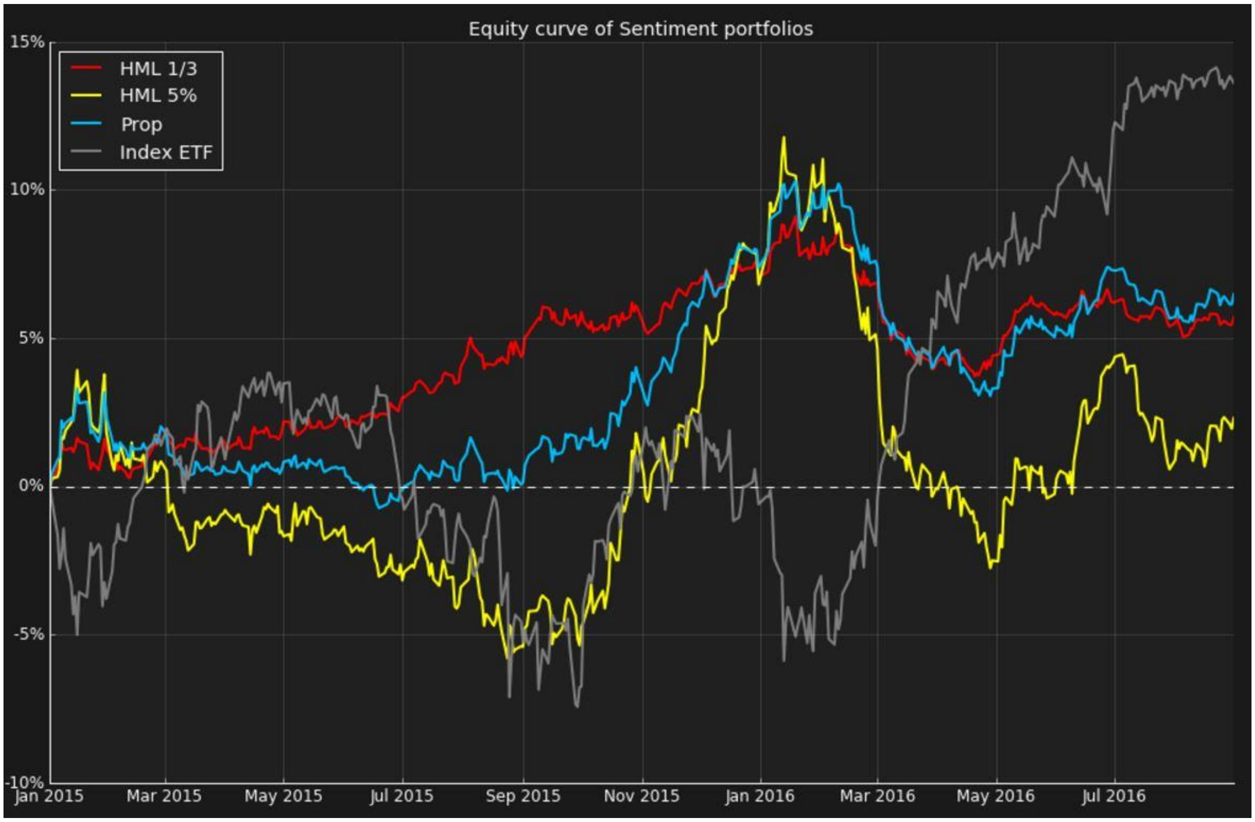
Strategy I: Daily sentiment-driven strategy

The backtesting period was from January 2, 2015, to August 31, 2016.

1. Illustration for Russell 3000 stocks

	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
HML 1/3	-0.05	14%	4%	3.29	449	661
HML 5%	-0.10	30%	8%	3.94	99	100
Proportional	-0.06	21%	4%	4.67	1332	664
Index ETF (IWV)	1.00	5%	11%	0.44	NaN	NaN

2. Illustration for S&P 500 stocks

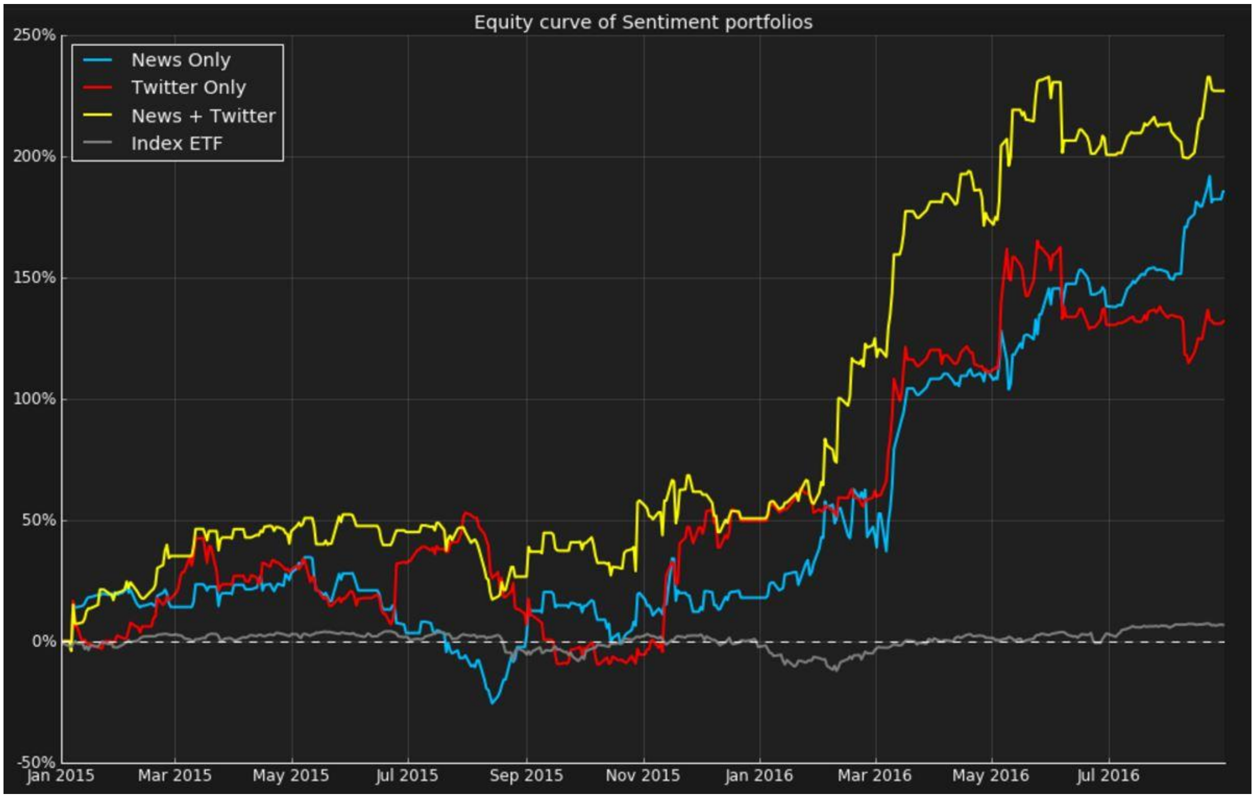


	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
HML 1/3	-0.08	3.42%	3.71%	0.92	130	151
HML 5%	-0.15	1.38%	8.42%	0.16	23	23
Proportional	-0.11	3.89%	4.93%	0.79	291	160
Index ETF (SPY)	1.00	8.17%	10.97%	0.74	NaN	NaN

Strategy II: Daily earnings event–driven strategy

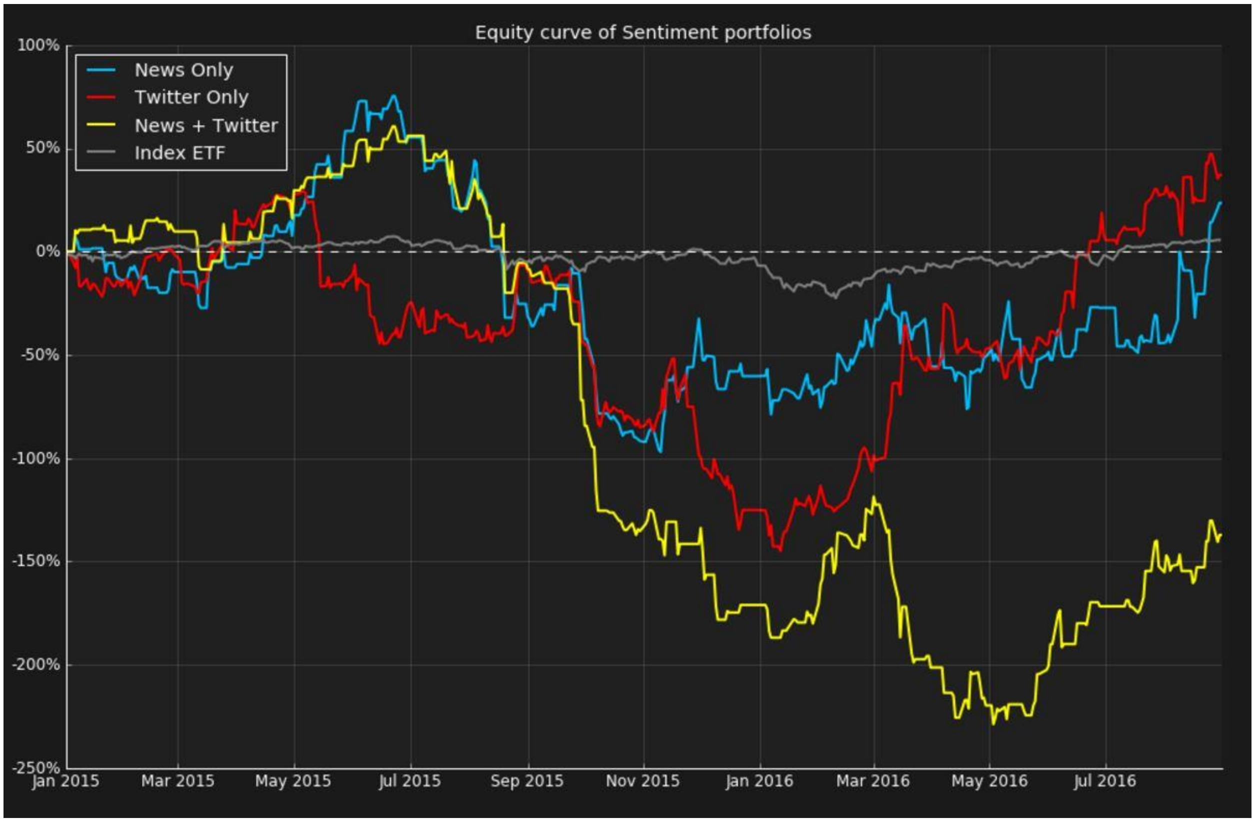
The backtesting period was from January 2, 2015, to August 31, 2016.

1. Illustration for Russell 1000 stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
News Only	0.02	112%	67%	1.67	3	9
Twitter Only	0.11	79%	71%	1.12	6	11
News + Twitter	0.23	136%	76%	1.80	2	7
Index ETF (IWB)	1.00	4%	14%	0.28	NaN	NaN

2. Illustration for Russell 2000 stocks

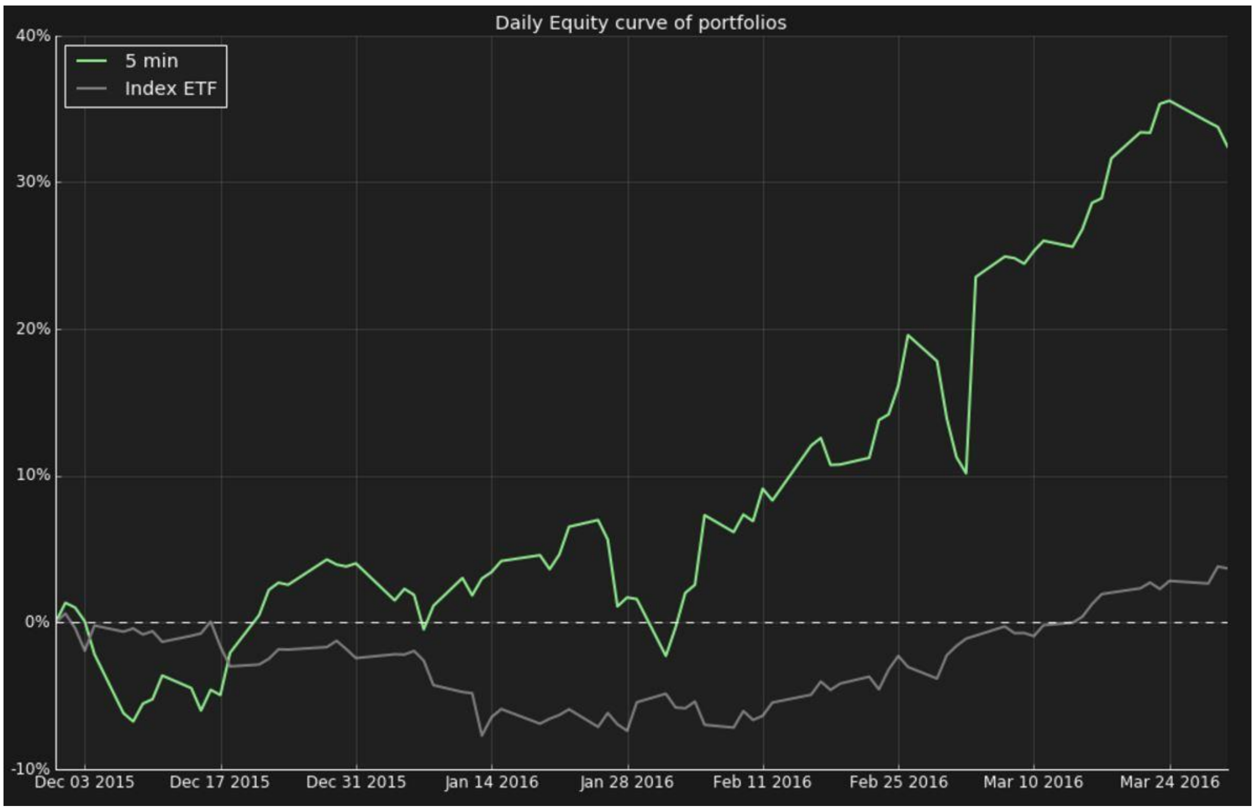


	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
News Only	0.42	14%	96%	0.15	3	5
Twitter Only	-0.11	22%	95%	0.23	6	12
News + Twitter	0.24	-83%	92%	-0.90	2	3
Index ETF (IWM)	1.00	3%	17%	0.19	NaN	NaN

Strategy III: Intraday sentiment-driven strategy

1. Illustration for S&P 500 stocks

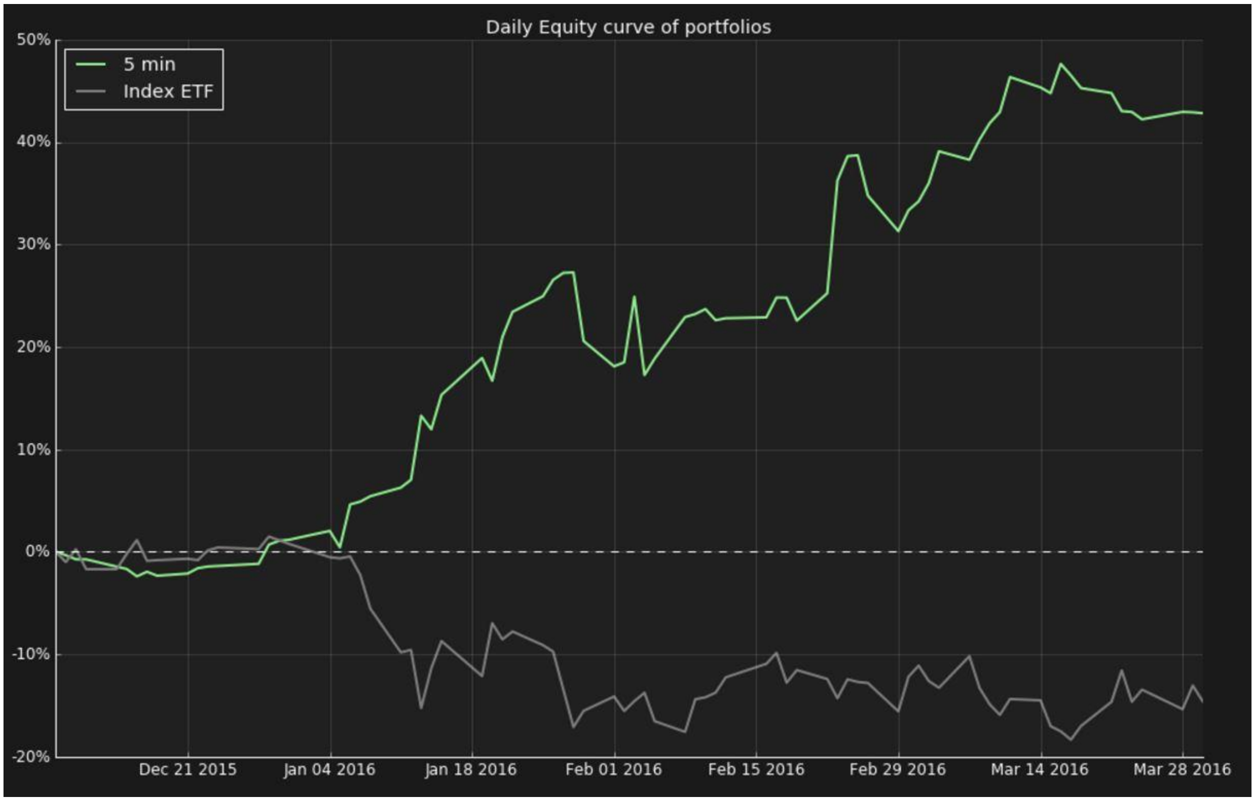
The backtesting period was from December 1, 2015, to March 30, 2016, with 1-minute price bars.



	Beta	Annualized Ret	Annualized Vol	Sharpe
News	-0.29	98%	36%	2.73
Index ETF (SPY)	1.00	11%	13%	0.84

2. Illustration for NASDAQ biotechnology stocks

The backtesting period was from December 9, 2015, to March 30, 2016, with 1-minute price bars.



	Beta	Annualized Ret	Annualized Vol	Sharpe
News	-0.34	144%	41%	3.49
Index ETF (IBB)	1.00	-48%	32%	-1.47

5.2 Transaction costs

A proper estimation of transaction costs is an essential component in a successful trading strategy. Without properly taking transaction costs into account, real trading could underperform the backtesting result substantially, even leading to a loss. In general, transaction costs include bid/ask spreads, slippage costs and market impact. However, properly estimating all components depends on the details of trade implementation and order execution, which is out of the scope of this paper. In this section, we will just discuss the impact of transaction costs on the aforementioned strategies.

Strategy I: daily sentiment-driven strategy trades small-cap stocks. Daily open-to-close rebalancing will incur a lot of transaction costs, which can wipe out the profit completely. However, we could reduce the transaction costs by holding the portfolio from market open to the next market open and only trading stocks needing to be rebalanced. For the Russell 2000 stocks, we see about 30% reductions in the transaction costs for HML 1/3 portfolio.

Strategy II: daily earnings event–driven strategy trades large-cap stocks, which are less affected by the transaction costs. With reasonable transaction costs assumed, the strategy still shows a reasonable Sharpe ratio.

Strategy III: intraday sentiment-driven strategy trades small-cap stocks–and very frequently. Therefore, this strategy is more subject to transaction costs. A trader definitely needs to incorporate a proper transaction costs model into the backtesting. When taking transaction costs into account, we may see that the length of the optimal holding period increases, for example, from 5 minutes to 30 minutes, since trading less frequently reduces transaction costs.

5.3 Alternative aggregation methodology

In this section, we propose a new methodology for company-level sentiment aggregation. Instead of just one score representing the average sentiment, the new methodology produces two components: average sentiment and dispersion.

5.3.1 Alternative sentiment score and new dispersion indicator

Each news story/tweet is scored with "confidences" $C_+, C_-, C_n > 0$ for positive, negative and neutral sentiment, respectively. These can be interpreted as "probabilities" and the following identity holds $C_+ + C_- + C_n = 1$.

Background: These probabilities are formed from output of three SVMs—each of which helps with a binary classification of a story/tweet as *Positive Vs Neutral, Positive Vs Negative* & *Negative Vs Neutral,* respectively—using various features that are constructed from the text analysis of the story/tweet.

The probabilities C_+, C_-, C_n are essentially a function of the output of the three SVMs; we derive final label $L \in \{Positive, Negative, Neutral\}$ for each story/tweet from the probability values $C_+, C_- \& C_n$. This essentially amounts to assigning the class with highest probability as the label.

Sentiment Average: For each story i, we define story-specific sentiment polarity score $S^i \in [-1,1]$ simply as $C_+^i - C_-^i$ (this passes the smell test—S values that are highly positive should correspond to positive stories or stories with high "positive probability," negative values to negative stories and values around 0 to neutral stories, but may be not be fully consistent with labeling that is currently used).

For average sentiment calculation, we simply propose an average of the sentiment polarity scores from each story that is part of the set, i.e.:

$$\text{Average Sentiment} = \mu = \frac{\sum_{i=1}^{N} S^i}{N} = \frac{\sum_{i=1}^{N}(C_+^i - C_-^i)}{N} = \overline{C_+} - \overline{C_-}$$

Here $\overline{C_+}$, $\overline{C_-}$ are the average positive and negative sentiments.

Sentiment Dispersion: To calculate the overall dispersion metric to go with the average sentiment, we need to track two components. One is the variance of the average sentiment across different stories, the second is the specific variances of the sentiment per story.

Sentiment dispersion = Inter-story variance + story-specific dispersion, which simplifies to:

$$\text{Dispersion} = \overline{C_+} + \overline{C_-} - \mu^2$$

5.3.2 Backtesting on alternative aggregation method

We backtested strategy I: Daily sentiment-driven strategy for Russell 2000 stocks using the new methodology sentiment score and dispersion indicator. The backtesting period was from January 2, 2015, to March 31, 2016.[2]

5.3.2.1 Effect of dispersion

In this part, we demonstrate how the dispersion indicator can improve the Sharpe ratio based on Twitter sentiment. We constructed portfolio holdings in four different variations:

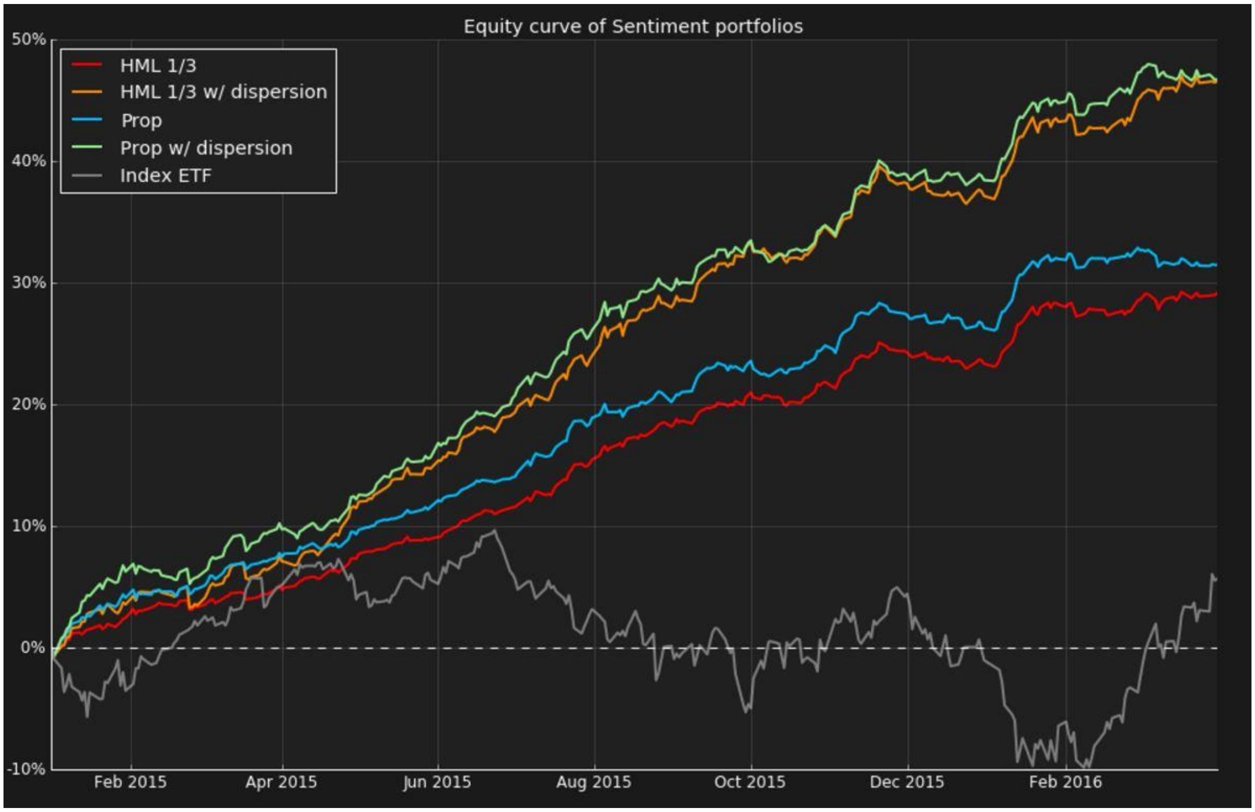
1) High-Minus-Low portfolio (HML 1/3): long (short) the top (bottom) 1/3 of stocks ranked by sentiment scores. Stocks in long and short portfolios are equally weighted.
2) High-Minus-Low with dispersion portfolio (HML 1/3 w/dispersion): long (short) the top (bottom) 1/3 of stocks ranked by sentiment scores and filtered with dispersion

[2] The story-level sentiment data provided covers the period from Jan. 1, 2015, to Mar. 31, 2016.

indicator above its cross-sectional median. In other words, it's the same as HML 1/3 except those stocks with dispersion below median are removed. Stocks in long and short portfolios are equally weighted.

3) Proportional portfolio (Prop): long (short) stocks with positions proportional to the difference of the sentiment score from the sentiment cross-sectional mean. If the sentiment score is above the mean, take a long position; if it is below, take a short position.
4) Proportional dispersion portfolio (Prop w/dispersion): long (short) stocks with positions proportional to the difference of the sentiment score from the sentiment cross-sectional mean and filtered with dispersion indicator above its cross-sectional median. In other words, it's the same as Proportional portfolio except those stocks with dispersion below median are removed.

Illustration for Russell 2000 stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
HML 1/3	-0.04	23%	4%	5.37	426	425
HML 1/3 + dispersion	-0.08	38%	7%	5.48	197	325

Prop	-0.06	25%	5%	5.06	752	529
Prop + dispersion	-0.06	38%	7%	5.39	265	375
Index ETF (IWM)	1.00	5%	15%	0.30	NaN	NaN

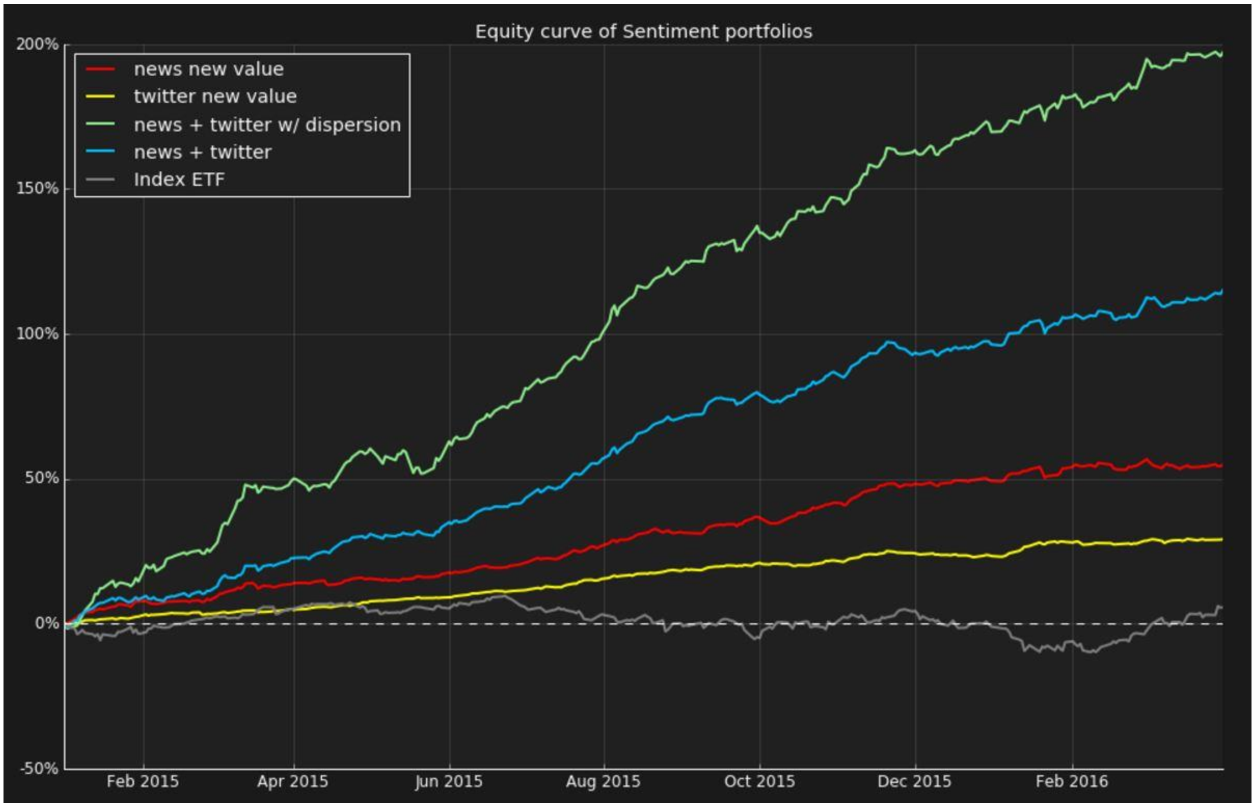
According to the backtesting results, adding the dispersion indicator improves the strategy performance for both HML 1/3 and Proportional portfolios. We believe that the dispersion indicator captures the conflicting market sentiments for a company. When dispersion is small, it means that market views are quite in line with one another and the market is more efficient in pricing in the converging views; when dispersion is large, it means that market sentiments are quite different and the market is less efficient in pricing in diverging views. Our backtesting results suggest long/short stocks with diverging market views provide better risk-adjusted returns.

5.3.2.2 Effect of combining news and Twitter sentiment

In this part, we show that combining news and Twitter sentiment can further boost the performance of the sentiment portfolio in terms of risk-adjusted return. We construct HML 1/3 portfolio based on three types of sentiment:

1) News new value: long (short) the top (bottom) 1/3 of stocks ranked by news sentiment scores. Stocks in long and short portfolios are equally weighted.
2) Twitter new value: long (short) the top (bottom) 1/3 of stocks ranked by Twitter sentiment scores. Stocks in long and short portfolios are equally weighted.
3) News + Twitter: long (short) the top (bottom) 1/3 of stocks ranked by both news and Twitter sentiment scores. In other words, stocks in the long leg have to be in the top 1/3 of both news and Twitter sentiment; stocks in the short leg have to be in the bottom 1/3 of both news and Twitter sentiment. Stocks in long and short portfolios are equally weighted.
4) News + Twitter w/dispersion: long (short) the top (bottom) 1/3 of stocks ranked by both news and Twitter sentiment scores and filtered with dispersion indicator above their cross-sectional median, respectively. This can be seen as combining HML 1/3 w/dispersion portfolio of news and Twitter. Stocks in long and short portfolios are equally weighted.

Illustration for Russell 2000 stocks



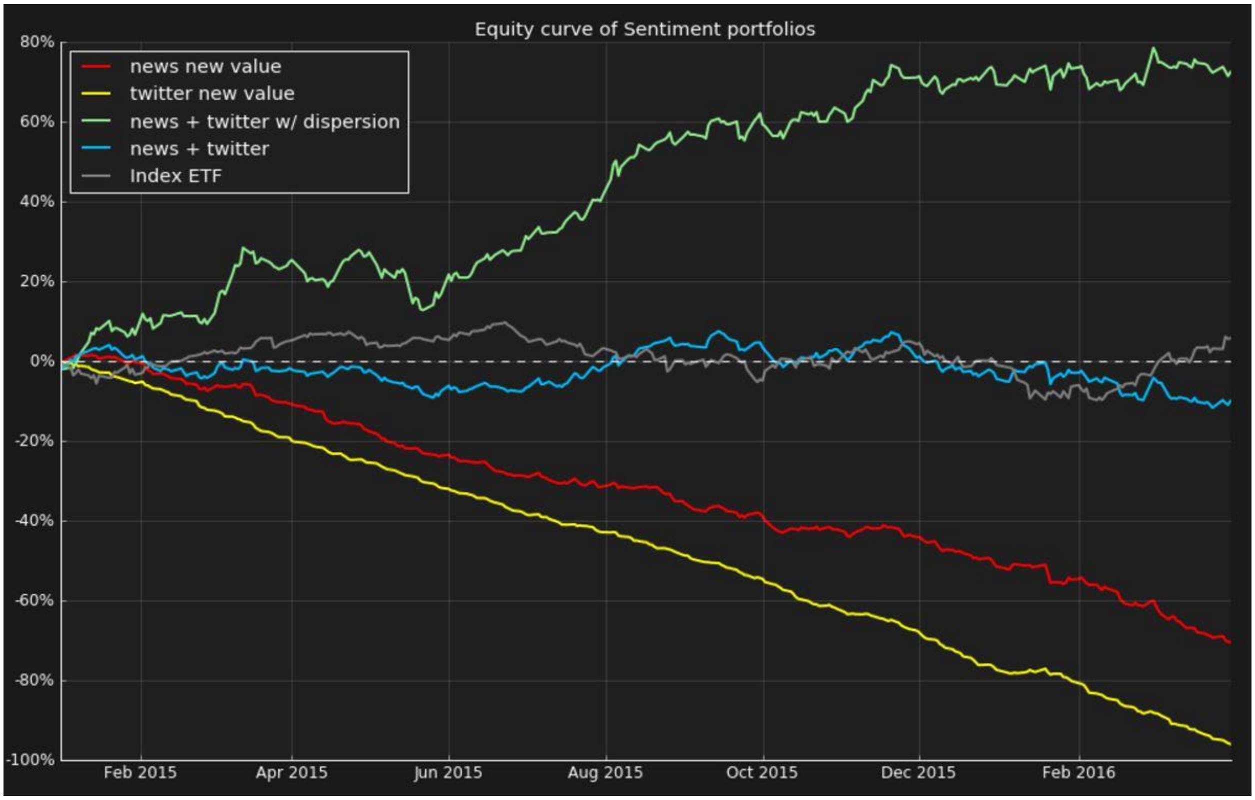
	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
News new value	-0.15	44%	9%	5.10	147	345
Twitter new value	-0.04	23%	4%	5.37	426	425
News + Twitter	-0.12	93%	15%	6.27	55	116
News + Twitter w/ dispersion	-0.17	159%	26%	6.10	24	46
Index ETF (IWM)	1.00	5%	15%	0.30	NaN	NaN

Based on our backtesting results, combining news and Twitter sentiment generates a stronger signal, which improves the strategy performance significantly.

5.3.2.3 Effect of transaction costs

In the following graph, we show the equity curves of the same daily sentiment-driven strategy as in section 5.3.2.2 but with transaction costs. We assume 20 bps average roundtrip costs for both long and short legs. The results below show that the News + Twitter w/dispersion portfolio still has a reasonably good Sharpe ratio.

Illustration for Russell 2000 stocks



	Beta	Annualized Ret	Annualized Vol	Sharpe	Average # of Short	Average # of Long
News Only	-0.15	-57%	9%	-6.58	147	345
Twitter Only	-0.04	-77%	4%	-17.72	426	425
News + Twitter	-0.12	-8%	15%	-0.55	55	116
News + Twitter w/ dispersion	-0.17	58%	26%	2.24	24	46
Index ETF (IWM)	1.00	5%	15%	0.30	NaN	NaN

References

Baker, M., & Wurgler, J. (2006, Aug). "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance* 61(4), 1645-1680.

Bloomberg. (2016). *EDF Research Note—Social Sentiment, DOCS— 2080750.*

De Long, J. B., Shleifer, A., Summers, L. H. & Waldmann, R. J. (1990a, Aug). "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98(4), 703-738.

De Long, J. B., Shleifer, A., Summers, L. H. & Waldmann, R. J. (1990b, Jun). "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance* 45(2), 379-395.

Mao, H., Counts, S. & Bollen, J. (2015). *Quantifying the Effects of Online Bullishness on International Financial Markets.* Statistics Paper Series, European Central Bank.

Tetlock, P. C. (2007, Jun). "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *Journal of Finance* 62(3), 1139-1168.