Global

Quantitative Strategy

Signal Processing

Date
6 July 2017

Social Media Sentiment

George Zhao
Quantitative Strategist
+1-212-250-4219

Ronnie Shah, PhD
Quantitative Strategist
+1-212-250-1173

Spyros Mesomeris, PhD
Quantitative Strategist
+44-20-754-71684

Andy Moniz, PhD
Chief Data Scientist
+44-20-754-71684

David Elledge
Quantitative Strategist
+1-212-250-8925

Part I

Natural language processing is in vogue

In this research report, we explore how natural language processing a.k.a. text mining techniques can be used to process qualitative data to form equity return forecasts. Specifically, we analyze Social Media Analytics (SMA) Indicators provided by IHS Markit which estimate stock-level sentiment from twitter data for US equities.

Twitter data as a measure for market sentiment?

We aggregate stock-level twitter data to form a measure of market sentiment. As we show, our measure is highly positively correlated with contemporaneous market returns and negatively correlated with VIX in a way that is consistent with social media reacting to fear associated with certain macroeconomic events.

Social media sentiment for stock selection

Using various measures, we find twitter social media sentiment explains the cross-section of stock returns yielding annualized Sharpe ratios that range from 2.0 to 2.5. However, we caution that the return forecast decays very quickly and the turnover of the signal is also high. We suggest certain alternative measures to improve risk-adjusted returns while reducing turnover and examine whether social media sentiment can improve post-event return predictability.

A Letter to our readers

Big data is here, and unlikely to go anywhere. The advent of the information age initially spawned computerized techniques to largely glean the value of an investment from financial statement disclosures and past stock return prices and volumes. The second coming of applications in quantitative finance involves leveraging vast improvements in micro-processing power to analyze and identify trends in large, often un-structured datasets. Examples include RFID data for retail transactions, satellite imagery to track business conditions, Google search trends [1] and smart phone data. Indeed, a simple Google search for "Big Data" yields 247 Million results. [2]

Natural language processing or text mining is largely an application of big data. These techniques can be loosely described as a translation of text-based data into quantitative data through the identification of keywords that convey sentiment.

One such application of big data is the use of natural language processing (NLP) to harness qualitative, non-numeric information. We recognize that "big data" has become a catch-all phrase for non-standard quantitative strategies, and we have previously written on this topic to explain certain applications more broadly in Rohal et al (2016). As is the case of this report, we caution clients that the signal-to-noise ratio on certain big data strategies can be quite low (although better, more refined techniques will improve predictions over time) and the time history that we can construct our back-tests can often be quite short due to lack of data availability. Despite these criticisms, the mounting evidence that big data can be an important source of alpha going forward is very compelling.

This piece leverages output from text mining / NLP algorithms, and suggests novel market and stock selection strategies based on twitter data to gauge social media attitudes towards various US companies. The "tweets" that users make are unrestricted, and potentially reflect opinions of goods produced and services rendered of covered companies. As we show in this report, twitter data can be useful for security selection and the factor performance is largely orthogonal to more traditional, quantitative factors.

1 For more information, see our previous research on Google trends (Luo et al [2014]).
2 Google search results as of April 3, 2017.

Translating Words to Numbers

Using Natural language processing to create return forecasts

Machines have replaced human in many aspects of everyday life. Natural language processing (NLP) is a technique that uses computer algorithms to extract meaningful conclusions from a vast quantity of unstructured, qualitative, text-based data. NLP usually identifies sentiment by examining text for certain linguistic combinations or grammatical patterns, similar to how a human might react when reading text. We then use this information to construct equity forecasts.

We use Twitter feed data to form views on social media sentiment on different US stocks.

The role of media in influencing stock market index prices has been previously examined by academic literature. Tetlock [2007] finds that high media pessimism forecasts negative market returns and unusually high or low pessimism is associated with high market trading volume. In related work, Bollen, Mao and Zeng [2011], analyze the relation between collective mood from large-scale Twitter feeds and the performance of the Dow Jones index. They develop a strategy using Twitter-based sentiment which predicts the daily direction (up and down) of DJIA.

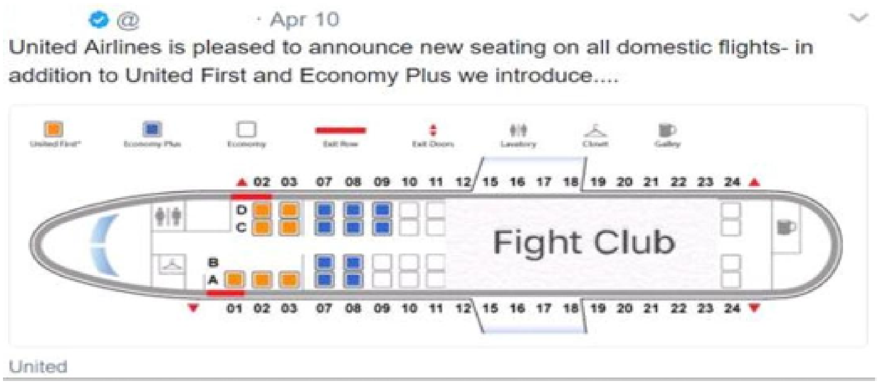
There is also evidence that text-based data predicts individual equity returns. Zhou, Wang, Fan and Wang (2015) and Jung, Naughton, Tahoun and Wang (2016) find that close to 50% of S&P Capital IQ 1500 firms use social media (Twitter or Facebook) to communicate information regarding new products or business initiatives. Tetlock, Saar-Tsechansky and Mackassy [2008] find that firms that have new articles with negative sentiment have low subsequent earnings and poor next-day stock returns. Giannini, Irvine and Shu [2017] examine twitter posts from stocktwits.com, a social media website that twitter members can use to comment on financial investments to analyze whether the sentiment from twitter tweets is consistent with the sentiment in publically available news before earnings announcements. They find that divergence (convergence) of opinion is associated with higher (lower) earnings announcement returns. In this report, we extend such research by exploring whether SMA social media indicator measures, which aggregate data from Twitter users can be used to forecast future returns.

A Twitter case study: United Airlines (UAL)

Social media often provides timely information on market events that forecast future profitability and returns. One example involves a passenger being forcibly removed from a United Airlines flight on Sunday April 9, 2017. Soon after the incident, fellow passengers posted videos of the encounter on the internet. The tweets associated with United over that period were very negative. For example, one user tweeted “@united @FoxNews @CNN not a good way to treat a Doctor trying to get to work because they are overbooked” [Social Market Analytics SMA, Twitter: April 9, 2017].

Another user posted a more colorful depiction:

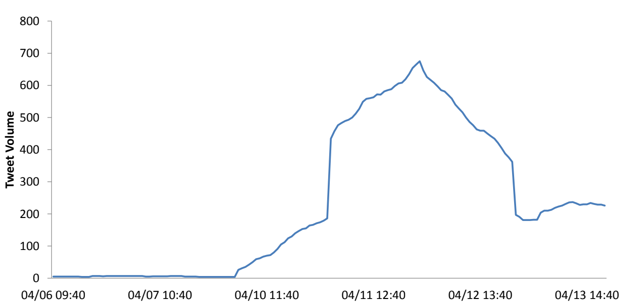
Figure 1: Tweet Example for United Airlines on April 10, 2017



Source: Twitter; Social Market Analytics SMA

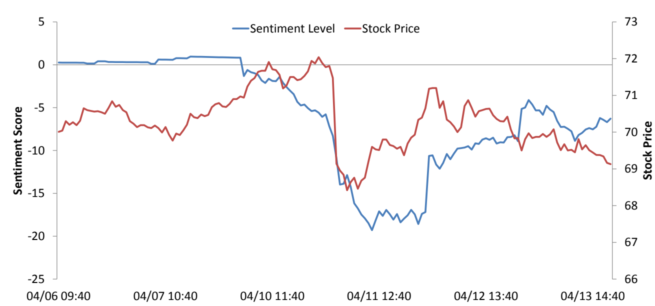
As we show, social media sentiment contains opinions regarding company-specific events that can potentially forecast future stock returns. Figure 2 reports the tweet volume for United Airlines around the event. The graph shows large increases in tweet volume regarding United Airlines after April 9th, which subsequently correlates with a drop in the stock price of United Airlines.

Figure 2: United Airlines (UAL) tweet volume



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 3: SMA social media sentiment score and United Airlines (UAL) stock price performance



Source:Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 3 illustrates social media sentiment and stock price performance for United Airlines for the five days surrounding the passenger incident on April 9th, 2017. We report the sentiment score provide by Social Market Analytics SMA [3] (based on past 24 hour trailing sentiment scores from Twitter tweets). While the sentiment is close to zero on April 6th, the social media sentiment declines substantially during the morning of April 10th, the day after the incident, but ahead of the drop in the stock price. There is also weaker evidence that increases in sentiment lead to improvements in the stock price after the passenger removal incident.

3 Social Market Analytics (SMA)'s Social Media Indicators provided through IHS Markit.

Social media analytics (SMA) indicator construction

Social Market Analytics (SMA) is a Chicago based data provider founded in early 2012 by an experienced team of financial market professionals. SMA patented data technology represents the aggregated intentions of investors as expressed on Twitter and StockTwits. Their data delivers real time aggregations of conversations on Twitter and StockTwits to provide directional and volatility indications. SMA has over five years of out-of-sample data with no survivorship or look-ahead bias.

SMA's social media indicators are provided by IHS Markit. [4]IHS Markit (Ticker: INFO) delivers next-generation information, analytics and solutions to customers in business, finance and government. IHS Markit has more than 50,000 business and government customers - including 85 percent of the Fortune Global 500 and the world's leading financial institutions. The data we analyze starts in December 2011 and sentiment scores are available in real-time. For our analysis, we obtain information from SMA as of 3:25 PM, or approximately thirty-five minutes before the market close each day.

Figure 4: Sentiment score generation process

Source: Social Market Analytics, Deutsche Bank

Social Market Analytics SMA filters out 90% of irrelevant twitter posts to improve the signal-to-noise ratio.

Figure 4 describes how SMA applies their text analysis algorithms. First, they collect market-related tweets by identifying keywords in the body of the text that link to specific stocks, including tickers, company names, names of company executives and a company's consumer products. The tweets come from both stocktwits and twitter, with special attention to those anagrams that start with either "@", "#" or "$".

Next, they filter the tweets to only include "sentimental indicative" tweets, i.e. those tweets which contain cognitive sentiment relevant to the particular company. More than 90% of the tweet feeds collected are filtered out during this step. Certain tweets are filtered out by the number of user followers, the time the twitter account has been active, whether the tweet is a re-tweet or a duplicate, whether the user deletes posts and then re-posts tweets later.

The algorithm then matches elements of the tweet text to a numerical measure using a proprietary word-to-sentiment dictionary. The dictionary contains 1-6 word phrases, each of which indicates a different level of sentiment. Last, numerical scores for certain identified keywords from tweets regarding a company for a 24 hour period are then summed up into a daily sentiment score. The final score is then calculated by averaging the sentiment in the tweets from

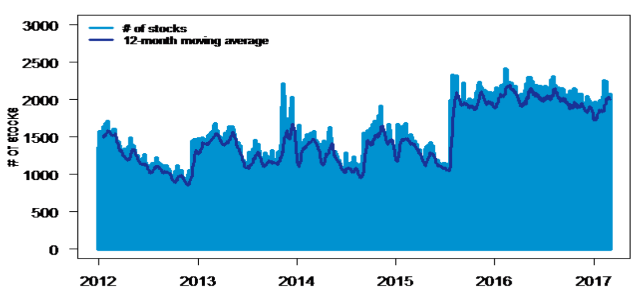
4 For more information on SMA please see www.socialmarketanalytics.com.

the previous 24 hours. The entire process takes around 3/10th of a second, and company sentiment scores are updated on a real-time basis.

SMA coverage for the US universe

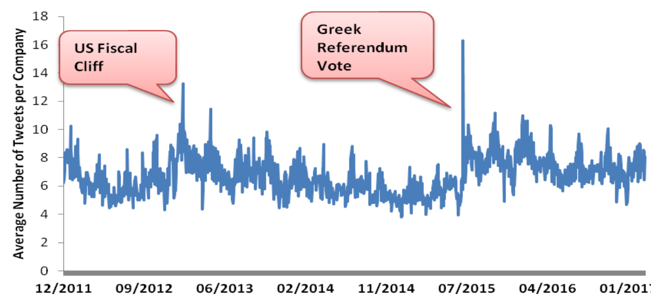
Figure 5 reports the number of companies listed in the Russell 3000 universe (red line) that have at least one reported tweet during the day over the period December 1, 2011 – March 31, 2017. The coverage varies over time – on average, there are over 20,000 tweets per day made by twitter users mentioning over 1,500 companies in the Russell 3000 universe. In recent periods, the number of companies covered in the index is close to two-thirds.

Figure 5: Russell 3000 signal coverage



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 6: Average number of tweets per company



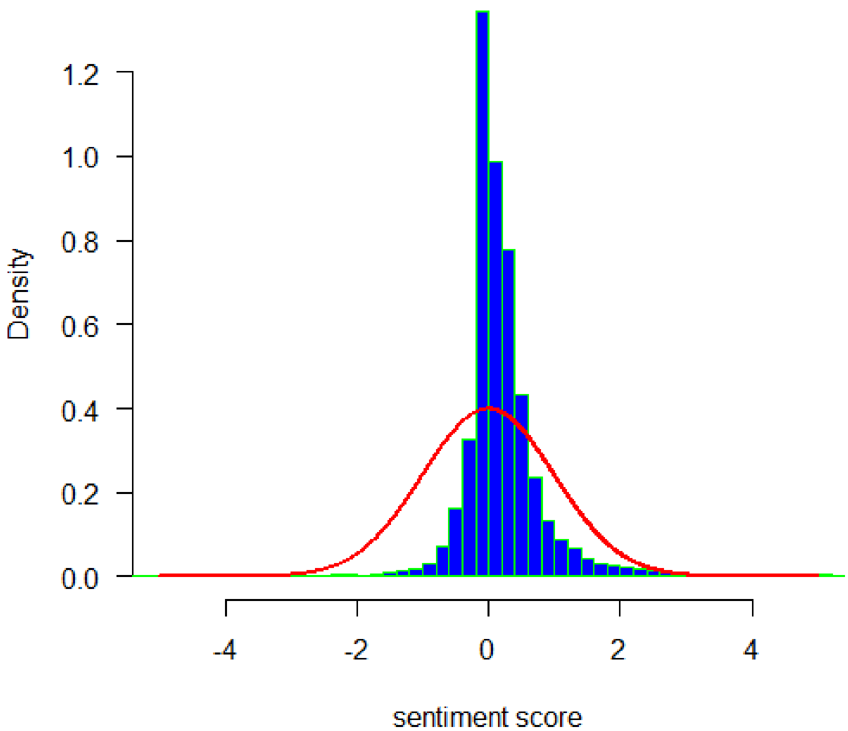
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 6 graphs the average number of tweets per company over time. During our sample period, there are on average seven tweets per company each day. There are certain periods in that the volume of relevant tweets increases, particularly around periods of macro uncertainty. For example, during the US Fiscal cliff debate which occurred in December 2012 or 'Grexit' referendum that happened in July 2015, the average number of tweets was more than double the average number. Despite those outliers caused by specific macro events, the number of tweets per company is remarkably stable over our sample period.

SMA sentiment score distribution

Figure 7: SMA provides a histogram of the sentiment scores which is estimated from 2,237,181 firm-day sentiment scores. We also super-impose a normal distribution illustrated by the red curve. As we show, the mean of the distribution is close to 0. There is some evidence of left skewness, or a few observations that have extreme positive sentiment. Relative to a normal distribution, there is much more clustering around the mean.

Figure 7: SMA sentiment score S distribution

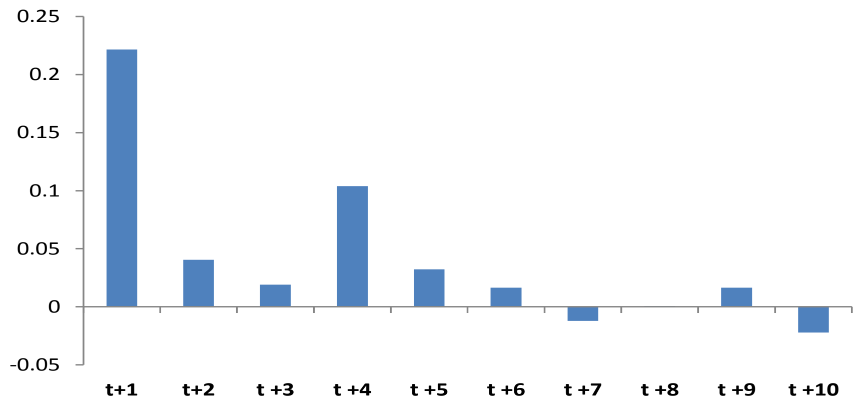


Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank.

Social media sentiment scores are slightly left skewed, and have less kurtosis when compared to a standard normal distribution.

Figure 8 reports the average partial autocorrelation (PACF) in factor scores to examine how quickly the sentiment information changes over time. To form this chart, we estimate partial autocorrelation for all stocks (N=3,504) that have more than 100 days of trading data during our sample period. The PACF provides a correlation of the variable with its lagged values after controlling for shorter lags. We then take the average of each lag coefficient and display that value in Figure 8. As we show, the innovations in correlations are relatively small after first day indicating that a quantitative strategy based on the raw sentiment information will tend to generate high turnover.

Figure 8: Partial autocorrelation function for SMA sentiment

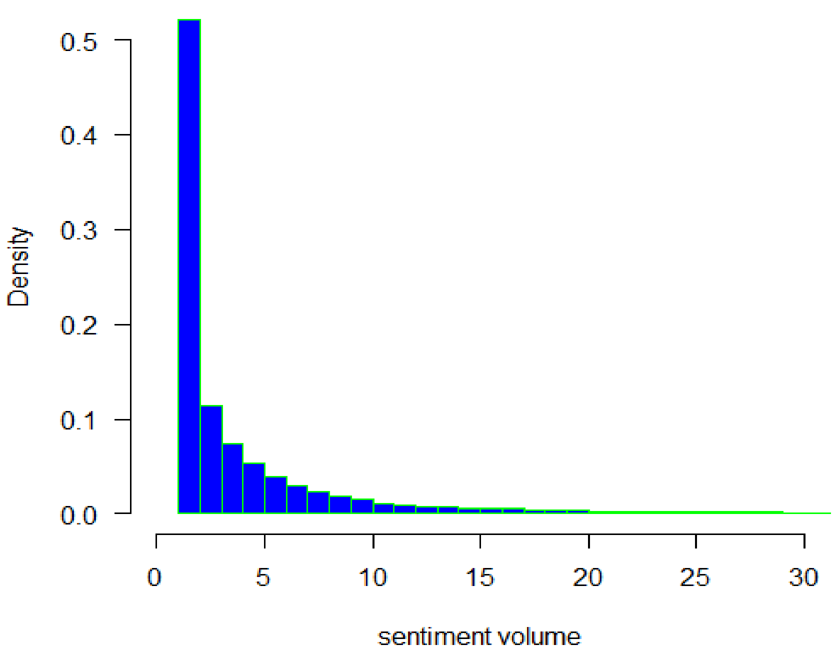


Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

We also examine the distribution of the number of tweets (a.k.a SMA sentiment volume) for individual companies each day. In our sample, more than half of the observations have only one relevant tweet per day. The overall distribution of SMA sentiment volume conforms the Benford's law. [5]

Social media sentiment changes rapidly.

Figure 9: SMA sentiment volume distribution



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank.

The below chart provides statistics for the sentiment score. The three rows report the sentiment score mean, standard deviation, 25th and 75th percentiles for sentiment score by sentiment volume (SVOL). Reading down the table, as the number of tweets increases, the sentiment scores mean and standard deviation

More than half of company-day observation have only one relevant tweet that make up the sentiment score.

5 For applications of Benford's law on Forensic Accounting, see our previous research (Jussa et al [2015]).

increase substantially. These results indicate that on average, tweets provide more positive than negative sentiment and the dispersion in sentiment increases when there are more tweets.

Figure 10: Sentiment score (S) distribution on sentiment volume

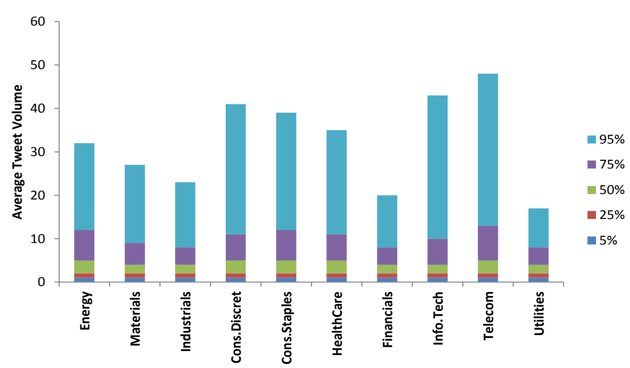
	Mean	STD	p25	p75
SVOL=1	0.048	0.195	0.000	0.157
SVOL=2	0.101	0.289	-0.010	0.270
SVOL>2	0.467	1.811	-0.025	0.695

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank.

Sector effects

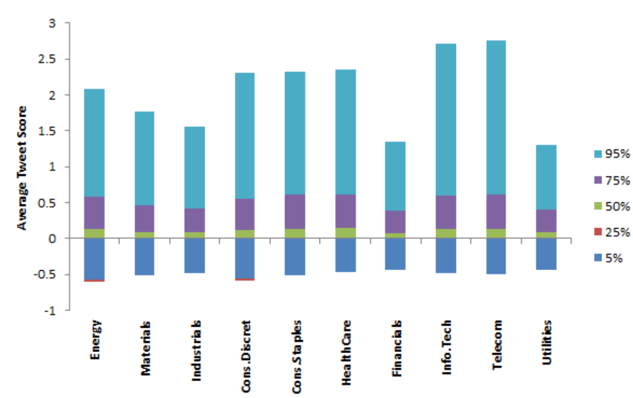
Figure 11 displays the equal-weighted average daily company tweet volume (number of tweets per company each day) by GICs Sector. We report the tweet volume distribution (5th, 25th, median or 50th, 75th and 95th percentiles) across the entire sample by industry. Stocks from the telecommunications and technology sectors, (which include Twitter, the firm that provides data used in the SMA indicators) have the highest median average daily tweet volume when compared to other industries. The next industry that has high visibility among social media is the Consumer Discretionary sector which includes durable, non-essential goods including apparel and automobiles. Utility and Financial companies are among those that have the least mentions by twitter users. Generally, cyclical industries tend to have higher twitter volumes, when compared to defensive industries.

Figure 11: Tweet volume across sectors



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 12: Tweet sentiment score across sectors



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 12 shows the equal-weighted average tweet sentiment level by GICs sector using the daily S scores. There are some commonalities with the previous chart – the top three industries that have some of the highest (95th percentile) tweet volumes (Telecommunications, Information Technology and Consumer Discretionary) also have the highest (95th percentile) sentiment scores, while the bottom three industries (Industrials, Financial and Utilities) have the lowest average sentiment scores.

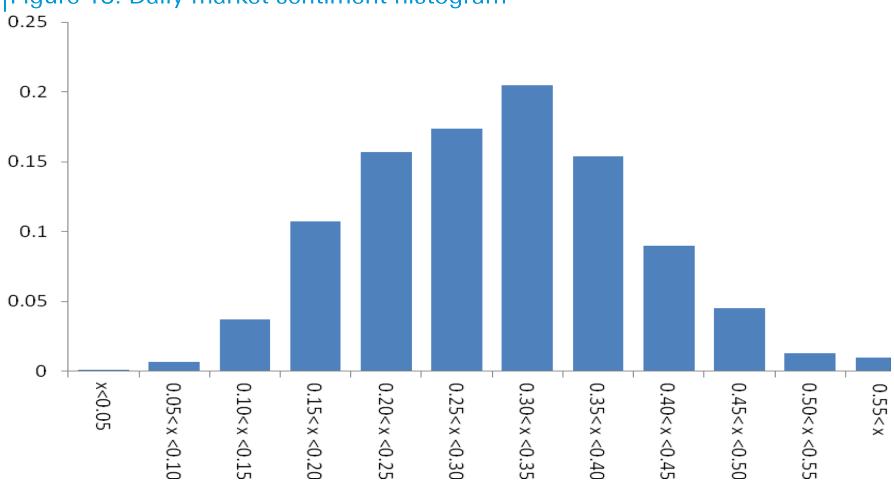
Social Media Market Sentiment Index

Social media sentiment often captures reactions to market events, like news reported by the financial press. In additional to the standard methodology presented in the previous section, SMA reports scores based on exponentially-weighted sums of tweet sentiment captured over the previous 24-hour period that gives greater weight to more recent twitter posts. We find that both measures are highly correlated, and for much of the analysis including the analysis presented in this section, we use the time-decayed metrics for social media sentiment.

We aggregate individual sentiment scores to form an index that captures social media market sentiment.

We estimate a market-based measure of social media sentiment formed by taking the mean of the exponentially-weighted time-decayed sentiment scores across Russell 3000 stocks. Figure 13 describes the distribution of our social media market sentiment index estimated using 1,298 daily index values.

Figure 13: Daily market sentiment histogram



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

The recommendations in the table above may or may not reflect those of Deutsche Bank's fundamental analysts, given the different criteria used in evaluating the stocks.

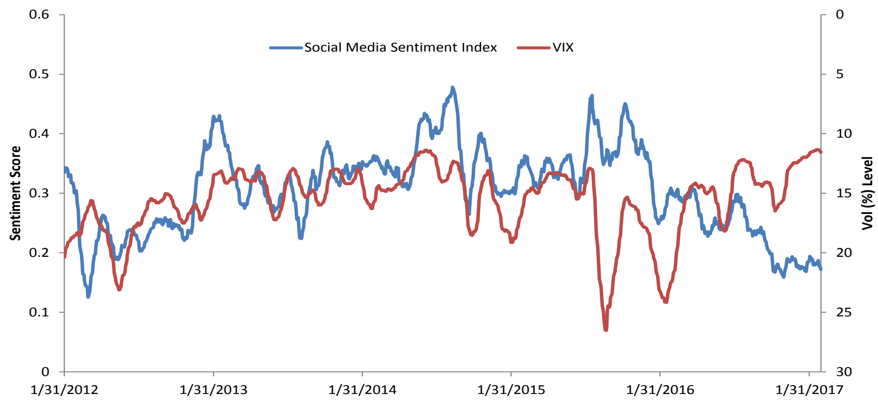
The median daily market sentiment score is 0.30, indicating that SMA's text mining algorithm identify more positive than negative words from twitter tweets. The distribution is fairly symmetric, but there are a few days that the sentiment is quite high (x>0.55). The standard deviation of the sentiment scores is 0.10.

Does social media sentiment proxy for investor sentiment?

Figure 15 compares the social media market sentiment index (blue line) to the VIX index (red line), which measures the market's expectation for 30-day volatility of the S&P Capital IQ 500 index. We report trailing 21-day moving averages for both series. The VIX index has been coined by the financial press as a way to measure investor fear versus greed and tends to spike during adverse macroeconomic events. We invert the VIX index, so high values reflect lower expectations of future volatility. As we show, the correlation between both indices is negative (-21.3%)

suggesting that during times of economic turmoil social media sentiment is low and VIX is high.

Figure 14: Relation between social media market sentiment and VIX

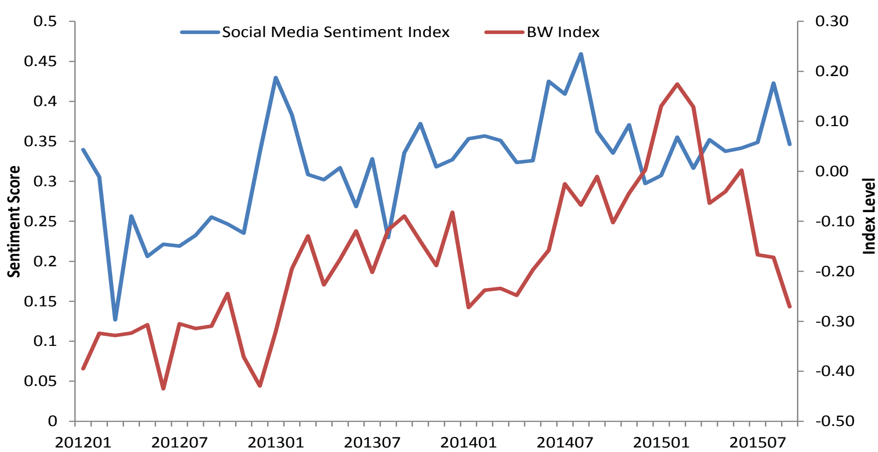


Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Social media market sentiment is highly correlated with investor sentiment.

Over the last two years, there is some evidence that the different sentiment indicators have diverged – VIX has increased substantially since July 2015, while Twitter market sentiment has fallen. Figure 15 compares the social media sentiment index to a more academic measure of investor sentiment first expressed in Baker and Wurgler (2006). Their measure is the first principal component from five sentiment proxies: value-weighted dividend premium, first-day IPO returns, IPO volume, closed-end fund discount and equity share in new issues.

Figure 15: Relation between social media market sentiment and investor sentiment



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

The correlation between both monthly time-series is 36.9%, suggesting that high retail-oriented social media sentiment is associated with higher values of investor sentiment. Again, starting in early 2015 we observe a divergence between the different measures of market sentiment.

Timing equity markets using social media sentiment

We also examine whether social media sentiment index can be useful for market timing. Our experiment can be best described as taking a long position in the market portfolio when social media sentiment is high relative to other time periods, and taking a short position in the market portfolio when social media sentiment is low. We also include simulations that invest in cash when market sentiment is neither high nor low. We then compare our market-timed portfolio to the market index that continuously invests in the Russell 3000.

Figure 16 presents the results from our market timing analysis. This table is organized as follows. The first row presents the annualized compound returns, standard deviations, annualized Sharpe ratios and Hit Ratio (defined as the percentage of daily returns that are greater than 0) for the Russell 3000 index from January 1, 2012 to January 31, 2017. The hit ratio suggests that close to 55% of daily returns during our sample period were positive.

Social media sentiment is unlikely to improve market selection models except over the very short run (<1 day).

Figure 16: Using social media sentiment for market timing (Simulation 1)

	Compound Return	Standard Deviation	Sharpe Ratio	Hit Ratio
Index				
R3000	12.5%	13.3%	0.95	53.5%
Perfect Foresight				
Ret_t	52.1%	13.0%	3.29	59.5%
Predictive Simulation				
$Ret_{Close(t)\text{-to-}Open(t+1)}$	3.6%	4.8%	0.76	49.1%
$Ret_{Open(t+1)\text{-to-}Close(t+1)}$	0.5%	11.7%	0.10	50.6%
Ret_{t+2}	3.9%	13.3%	0.36	50.9%
Ret_{t+3}	-2.7%	13.3%	-0.14	50.6%
Ret_{t+4}	-5.2%	13.3%	-0.34	50.6%

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

The simulations then invest in the market if the social media sentiment index is greater than the 45th percentile (x>28.8%) and take a short position otherwise. We chose the 45th percentile as our cut-off as close to 45% of the days in our sample have negative market returns. The market sentiment information is available at 3:25 EST each day, about thirty minutes before market close. The first two columns report contemporaneous returns, or the return to deploying a market timing strategy with foresight of the level of market sentiment. This is not an investible strategy. As we show, the average annualized return for this strategy is 52.1% and the Sharpe ratio is 3.29. [6]

The last five rows reflect the market timing strategy applied to predict future returns. We break the first day into the overnight return (Close(t)-to-Open(t +1)) and the next day return (Open(t+1)-to-Close(t+1)). We find that the market

6 The Sharpe ratio, a measure of risk-adjusted returns is equal to the square root of 250 multiplied by the average daily return less the risk free rate, all divided by the standard deviation of daily returns.

sentiment index has some predictive power (Sharpe ratio = 0.76) associated with timing the over-night return, but the next-day return during market hours is very close to 0. We also analyze our market timing model on returns after the next day, listed in the last three rows of the table. There is weak evidence that market sentiment predicts returns two days following observing the sentiment data, and the 3rd and 4th day future returns is actually negative. The hit rates of all our forecasting simulation models are close to 50%, which is a bit disappointing.

Figure 17: Using social media sentiment for market timing (Simulation 2)

	Compound Return	Standard Deviation	Sharpe Ratio	Hit Ratio
Index				
R3000	12.5%	13.3%	0.95	53.5%
Perfect Foresight				
Ret_t	45.8%	11.0%	3.48	61.8%
Predictive Simulation				
$Ret_{Close(t)\text{-to-}Open(t+1)}$	1.6%	4.2%	0.39	46.8%
$Ret_{Open(t+1)\text{-to-}Close(t+1)}$	-3.1%	9.6%	-0.27	50.4%
Ret_{t+2}	-1.6%	10.8%	-0.10	50.4%
Ret_{t+3}	-6.1%	10.8%	-0.53	50.2%
Ret_{t+4}	-5.7%	11.3%	-0.46	49.8%

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

As a refinement, we also consider another market timing model presented in Figure 17 that invests in cash when the sentiment is neither high (above the 70th percentile) or low (below the 30th percentile). The results for the predictive simulation listed in the last five rows of the table are actually weaker than those reported in Figure 16. As a generalization, our market sentiment index is highly correlated with contemporaneous market returns, but is a poor predictor of future market returns except over the very short-run.

Social Media Sentiment as a Stock-Picking Strategy

For the standard finance model, efficient capital markets imply prices reflect aggregate, un-biased forecasts on cash flows and allow no room for emotions to impact prices. Investor sentiment is a belief regarding future cash flows and investment risk that is not justified by observable, fundamental information and that psychological biases can on aggregate impact stock prices. In this article, we extend previous research by examining whether twitter data can be used to explain differences in average stock returns.

Social Market Analytics SMA factor set and signal construction

Does sentiment predict individual stock returns?

In this section, we analyze portfolio construction and performance for various metrics provided by IHS Markit . Figure 7 reports a number of different metrics to capture social mood-related sentiment. We examine six of these indicators, which are listed in Figure 18. The first indicator is S, the average sentiment score across all tweets for a particular company for the previous 24 hours and the time-decayed exponentially-weighted average sentiment score, respectively.

What is the best way to construct the strategy?

Figure 18: IHS Markit sentiment indicator list

factor mnemonic	factor definition
S	Exponentially Weighted Aggregated Sentiment Score
S_SCORE	Normalized Exponentially Weighted Aggregated Sentiment Score
CHG1D_S_SCORE	1-day Change in Normalized Weighted Sentiment Score
CHG5D_S_SCORE	5-day Change in Normalized Weighted Sentiment Score

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 19 displays a snapshot of the data that is provided by SMA and IHS Markit for our research. The first column reports S, which is equal to the sum of sentiment scores from tweets captured over the previous 24-hour window. The second and third columns, S_{Mean} is the 20-day average of the daily S and σ(S) is the standard deviation of the past 20 day S scores. The variable we use in our analysis S_{SCORE} is then calculated as:

$$S_{Score} = \frac{S - S_{Mean}}{\sigma(S)}$$

Figure 19: Snapshot of social media database

Date	Cusip	Ticker	Name	S	S_Mean	S_Volatility	S_Score
1-Jun-16	01858110	ADS	ALLIANCE DATA SYSTEMS CORP	0.67	0.24	0.31	1.38
1-Jun-16	05276910	ADSK	AUTODESK INC	-0.18	0.15	0.55	-0.59
1-Jun-16	00738A10	ADTN	ADTRAN INC	0.31	0.00	0.23	1.32
1-Jun-16	00673910	ADUS	ADDUS HOMECARE CORP	0.35	0.05	0.13	2.29
1-Jun-16	02360810	AEE	AMEREN CORP	0.74	0.02	0.36	2.00
1-Jun-16	00770F10	AEGN	AEGION CORP	0.00	0.01	0.11	-0.08
1-Jun-16	00767E10	AEGR	AEGERION PHARMACEUTICALS INC	0.00	-0.01	0.33	0.04
1-Jun-16	00797310	AEIS	ADVANCED ENERGY INDS	0.54	0.19	0.30	1.18
1-Jun-16	02567620	AEL	AMERICAN EQTY INVT LIFE HLD	0.12	0.21	0.38	-0.22

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

The S_{score} measure compares the current level of sentiment surrounding a stock, S to the historical, past-month's sentiment, S_{Mean} and then scales by the standard deviation of past daily sentiment scores. Since this variable is "change" variable, we should expect the measure to change rapidly over time. The de-meaning process also removes any positive or negative sentiment bias.

Returning to previous figure, The last two factors reported in Figure 18 , Δ1D S_{SCORE} and Δ5D S_{SCORE} are the one-day and five-day changes in S_{SCORE}. These factors are "acceleration" variables that capture changes in z-score variables – which are capturing a change of sentiment (S) relative to a historical mean (last twenty days of S).

There are additional factor constructs that IHS Markit includes that we do not cover in our report. First, they report the number of tweets recorded over the previous 24 hour period for each stock. This information can be used to gauge a stock's popularity with social media participants. Second, IHS Markit reports an equal weighted and a volume-weighted sentiment measures (similar to S). In our results (unreported), we find these measures is economically similar to S.

Signal construction considerations and back-tested performance

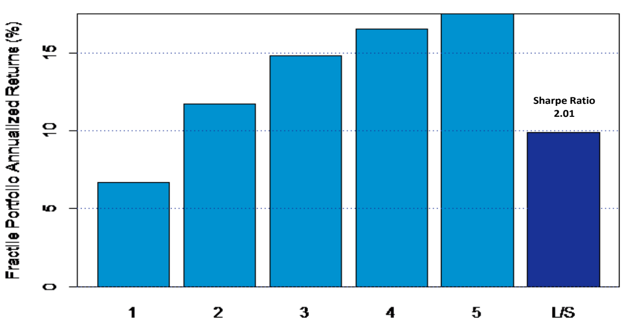
In this section, we analyze different ways to leverage qualitative data from social media for security selection. We begin our analysis in this section by analyzing strategy performance for S, the SMA indicator that is equal to the average time-weighted sentiment over the previous twenty-four hours.

For the sentiment measure S, we construct a long/short portfolio that takes long positions in stocks with scores that are in the top 20% of the investible universe (Russell 3000) and short positions in stocks that are in the bottom 20%. Stocks within a portfolio are equally-weighted. We then rebalance the long/short portfolio each day, and then aggregate returns on a monthly basis.

The performance charts for the strategies we consider in this section are organized as follows. The upper-left graph, Figure 20 displays annualized returns for portfolios formed on S over the period January 2012 to February 2017. Each day, stocks are sorted into five portfolios each with a roughly equal number of stocks. The long/short portfolio return (L/S) formed by taking a long position in those stocks with the highest S (top 20%) each day and a short position in those stocks with the lowest S. The returns illustrated in the graph are monotonic – the stocks with the highest sentiment scores have annualized returns of 17%, while the stocks with the lowest scores have annualized returns that are around 8%.

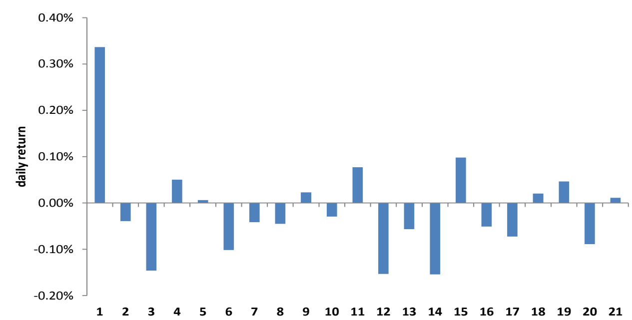
In our analysis, we use the Sharpe ratio as a measure of risk-adjusted returns to compared different portfolios' performance. The Sharpe ratio is defined as the ratio of average long/short portfolio monthly returns scaled by volatility multiplied by the square root of 12.The Sharpe ratio for the Long/Short portfolio is 2.01. In our back-test, we obtain from SMA sentiment data at 3:25 PM EST, and securities are purchased at market-on-close.

Figure 20: Returns for portfolios formed on S



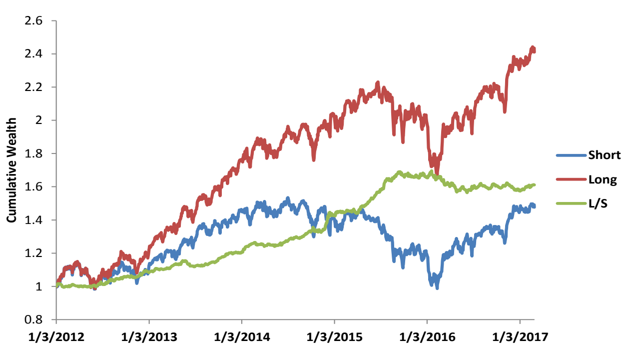
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 21: IC decay for L/S portfolios formed on S



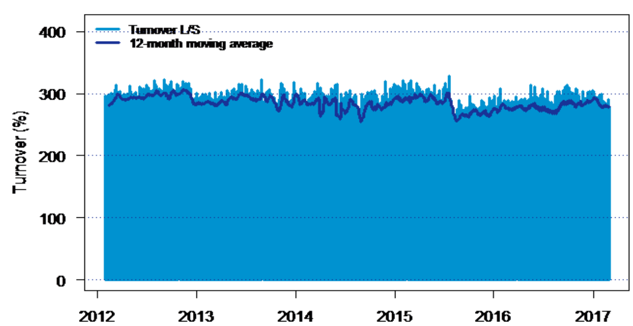
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 22: Growth in wealth for portfolios formed on S



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 23: Long/Short turnover for portfolios formed on S



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

The upper-right chart, Figure 21 reports the IC over time, or the correlation between factor scores and returns measured with over different horizons (x-axis) in days. As we show, implementation on the first day is imperative, as the IC decay shows that the correlation between the factor scores and returns is close to zero on the 2nd day (first lag). We find evidence of a reversal after the first day, as positive sentiment today (weakly) forecasts negative returns for rest of the month.

The lower-left figure, Figure 22 presents the growth in wealth associated with investing $1 in each of the three portfolios (Red: Highest quintile portfolio formed on S, Blue: lowest quintile portfolio formed on S, Green: High minus Low Long/ Short factor portfolio). The long/short portfolio generates a buy-and-hold return of 60% from January 2012 to June 2015. Subsequent performance from July 2015 to February 2017, has not been very robust. The performance deterioration in recent periods is potentially driven by (i) de-coupling of overall social media sentiment with other forms of investor sentiment [See Figure 14 and Figure 15] or (ii) the inclusion of social media sentiment in quantitative investment models reducing the return premia of such strategies. This evidence, however, should be interpreted with caution given the short sample period post-June 2015.

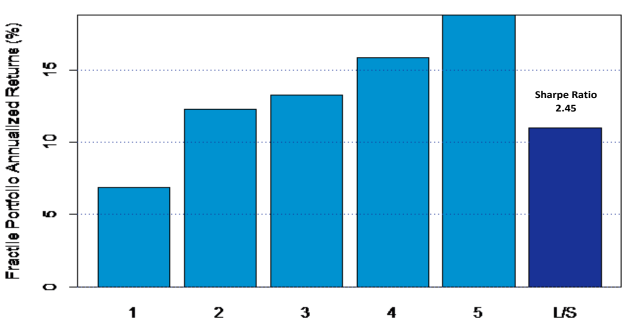
The lower-right image, Figure 25 shows the daily turnover for the Long/Short S portfolio over the sample period. We find turnover of close to 300% per day, indicating that 75% of the stocks in the long and short portfolios that make-up the Long/Short factor are replaced each day. These results stem from the low autocorrelation reported in Figure 8. The high levels of turnover likely preclude an investor from directly profiting from S after accounting for transactions cost.

S_{Score} explains return differences in the short-run (next few days), but has very high turnover.

In unreported results, we find that S has low correlations with conventional quant factors such as Dividend Yield strategy.

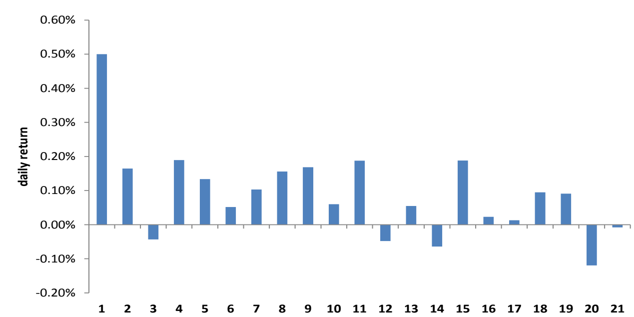
We also examine the performance of S_{Score}, the z-scored version of S. As we show in Figure 24, the long/short portfolio returns for scored version are more robust (Sharpe ratio= 2.45) than S (Sharpe ratio=2.01), indicating that differences in sentiment are a better predictor of returns than sentiment levels. S_{Score} also decays much slower than S – with positive correlations between factor scores (i.e. forecasts) and returns out to ten days.

Figure 24: Returns for portfolios formed on S_{Score}



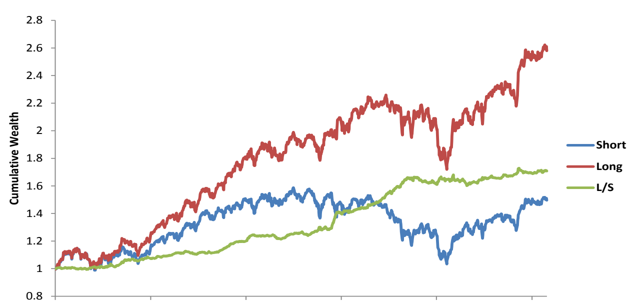
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 25: Daily return decay for L/S portfolios formed on S_{Score}



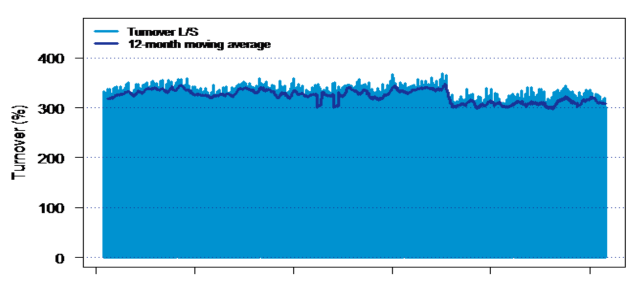
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 26: Growth in wealth for portfolios formed on S_{Score}



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 27: Long/Short turnover for portfolios formed on S_{Score}



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

One shortcoming of the z-scored version of S is increased turnover - Figure 27 reports average daily turnover of 330% for a long/short strategy based on S_{Score} .

Advanced signal construction techniques

As a refinement, we examine ways to improve risk-adjusted returns, while reducing strategy turnover. We leverage our previous research on technical factors[7] to create new constructs to capture sentiment-related information efficiently for stock selection. Our measures are based on technical analysis that leverages past price movements to forecast future equity returns. Figure 28 displays a description of new sentiment technical factors that we evaluate.

Advanced signal processing techniques can preserve the return predictability while lowering turnover.

The first construct, Bollinger Band, measures movements of sentiment around a mean scaled by standard deviation. As standard deviation is a measure of volatility, the bands are self-adjusting: widening during volatile markets and contracting during calmer periods. This measure is a generalized version of S_{SCORE}, which we apply over the past fifteen days.

The second measure, Moving Average Convergence/Divergence (MACD) captures the difference in three exponentially-smoothed moving-average time series – 9 trading days (short), 12 days (medium) and 26 days (long). The MACD is the difference between the medium and long measures. Intuitively, the MACD captures the difference in returns between days 13-26 (as the measure is the long series measured over 26 days the less the medium series measured over 12 days). The final trading signal is calculated as the short series minus the MACD: when the 9-day exponentially-smoothed variable is greater than the MACD the model yields a buy signal.

Figure 28: SMA Sentiment technical Factor list

Technical Indicator	Formula
Bollinger Bands (BB)	BB=(SMA score-MA(SMA score,N))/stdev(SMA score,N)
Moving Average Convergence Divergence (MACD)	DIFF : EMA(SMA score,SHORT) - EMA(SMA score,LONG); DEA : EMA(DIFF,M); MACD : (DIFF-DEA),
Relative Strength Indicator (RSI)	RSI = 100 - (100/1+RS) RS = Average SMA sentiment Gain / Average SMA sentiment Loss
Percentage Price Oscillator (PPO)	PPO = 100* (Fast_EMA - Slow_EMA) / Fast_EMA
Percentage Volume Oscillator (PVO)	PVO = 100* (Fast_EMA - Slow_EMA) / Fast_EMA

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

The recommendations in the table above may or may not reflect those of Deutsche Bank's fundamental analysts, given the different criteria used in evaluating the stocks.

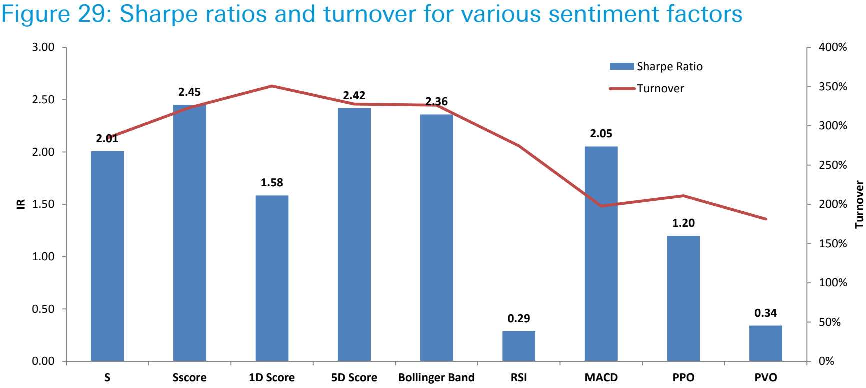
7 See more details in our report by J. Jussa, Z. Chen, Y. Luo, R. Cahan, M. Alvarez 2011, "Canada Quant Technically Savvy Alpha", Deutsche Bank Quantitative Strategy, May 6, 2011

The third measure, Relative Strength Index (RSI) captures the historical strength or weakness in a stock price based on sentiment values over the past 14 days. This measure is a momentum oscillator, measures the velocity and magnitude of price movements. The relative strength measure (RS) is equal to exponentially-smoothed moving average of positive past changes in sentiment divided by the exponentially-smoothed moving average of negative past changes in sentiment. The RSI factor is then equal to 100 minus 100 divided by 1 + RS. By construction, RSI takes values between zero and 100. One characteristic of the RSI is that it moves slower when it reaches increased overbought or oversold conditions, and then snaps back very quickly when the market enters even a mild correction. This brings the RSI back to more neutral levels and indicates that the sentiment trend may be able to resume.

The fourth and fifth measures, Percentage Price Oscillator (PPO) and Percentage Volume Oscillator (PVO) measures percentage change in sentiment and sentiment volume (number of tweets) respectively. Our measure for both variables is one minus the percentage change between the 26-day exponential moving average and the 12-day exponential moving average.

Figure 29 reports Sharpe ratios and daily turnover for long/short portfolios formed on various factors discussed in Figure 18 and Figure 28. The IHS Markit signals (four left-most bars) have strong performance during our back-test period with Sharpe ratios that range from 1.6 to 2.5. There are a few more takeaways that relate to factor construction choices. The “z-score” factor S_{SCORE} compares today's sentiment to the 20-day trailing average sentiment score and divide by the standard deviation out-perform the standard, vanilla factor S. These factors control for the level of sentiment over the past twenty days, and can be construed as the daily change in sentiment scaled by the variation in the sentiment score. The z-scored variables also control for persistent differences in sentiment across types of stocks, i.e. tech stocks having more positive sentiment on average when compared to airline stocks. Δ5D S_{SCORE} represents five-day changes in S_{Score} have slightly worse performance when compared to the S_{Score} factor.

Figure 29: Sharpe ratios and turnover for various sentiment factors



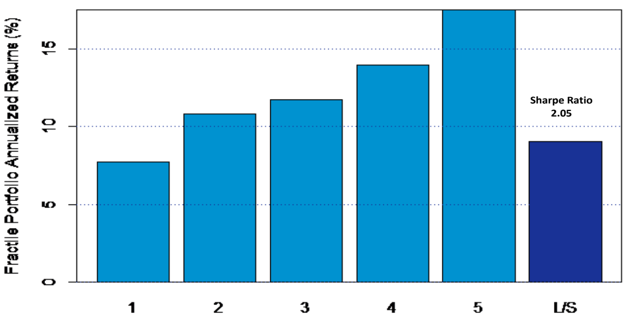
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

There is much more variation in performance for our proprietary technical factors which are reported in the five right-most bars. For example, RSI, PPO and PVO have Sharpe ratios of 0.29, 1.20 and 0.34, respectively. In contrast, the Bollinger

Band strategy that closely resembles the S_{Score} factor has a Sharpe ratio of 2.36. One of the challenges with the higher performing signals (S, S_{Score}, Δ5D S_{Score}, Bollinger Band) is the high turnover which ranges from 300% to 350%. As we show, the MACD strategy performs relatively well (Sharpe ratio=2.05) and generates much less turnover (200%) than the other four signals we considered.

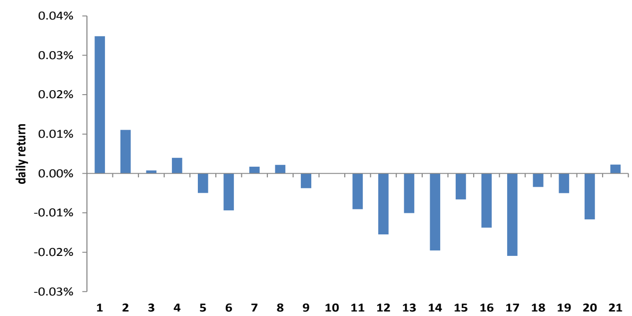
Performance charts providing details on returns, Sharpe ratios, IC decay, growth in wealth and turnover can be found below in Figure 30, Figure 31, Figure 32 and Figure 33.

Figure 30: Quintile return of the backtest portfolio on MACD sentiment



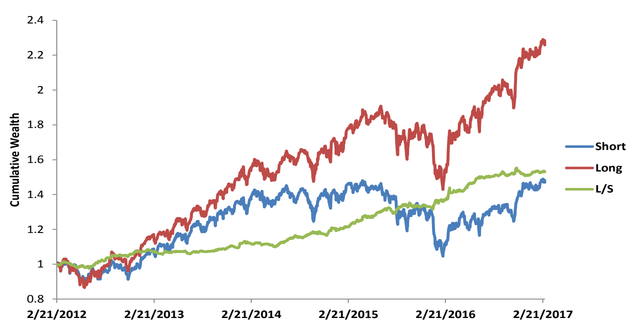
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 31: MACD sentiment daily return decay of the backtest portfolio



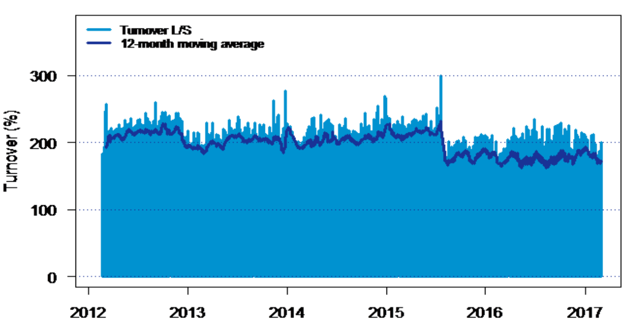
Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 32: Wealth curve of the backtest portfolio on MACD sentiment



Source: oomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Figure 33: MACD sentiment turnover of the backtest portfolio



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

As we show in the upper right-hand graph, the IC decay for MACD is slower than S and S_{SCORE}. Recall that the MACD strategy involves forming three exponentially-smoothed moving averages over 9, 12 and 26 days. This strategy is in some sense taking the difference between sentiment over days 1-9 and days 13-26. These results largely follow the S IC decay reported in Figure 21, which shows positive correlations between forecast and returns for the first 10 days, and negative correlations after 10 days.

Figure 34 illustrates the correlations between the nine different factors. This table is organized as follows: the lower-left diagonal contains the spearman rank correlation between the different factor scores; the upper-right diagonal contains the correlation between different factor long/short portfolios.

Figure 34: Factor score and long/short portfolio return correlation

	S	Sscore	1D Score	5D Score	Bollinger I	RSI	MACD	PPO	PVO
S		0.71	0.58	0.56	0.61	0.33	0.31	-0.21	0.05
Sscore	0.85		0.77	0.78	0.79	0.36	0.45	-0.01	0.08
1D Score	0.68	0.80		0.64	0.67	0.34	0.29	0.00	0.02
5D Score	0.69	0.83	0.68		0.69	0.35	0.33	0.00	0.08
Bollinger I	0.76	0.89	0.72	0.74		0.45	0.52	0.11	0.12
RSI	0.36	0.39	0.34	0.33	0.46		0.25	0.08	0.09
MACD	0.35	0.41	0.22	0.28	0.47	0.22		0.26	0.22
PPO	-0.01	0.06	0.03	0.03	0.12	0.05	0.23		0.32
PVO	0.07	0.05	0.02	0.02	0.07	0.02	0.17	0.26	

Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

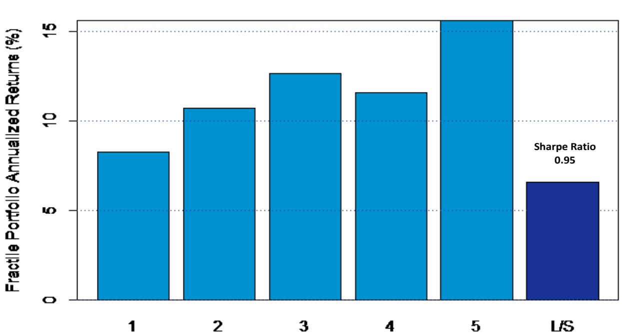
The correlations are highest between S and S_{SCORE}, and S_{SCORE} and Bollinger Band which is not surprising given the similarity in signal constructions. Interestingly, MACD has low correlations with most other factor constructs. We also find very low correlations for RSI, PPO and PVO with other factors.

Longer-term systematic social media sentiment strategies

We propose a longer-term quantitative strategy that uses the range of sentiment scores over the past 60 days. A long/short strategy based on this construct generated a Sharpe ratio of 0.95.

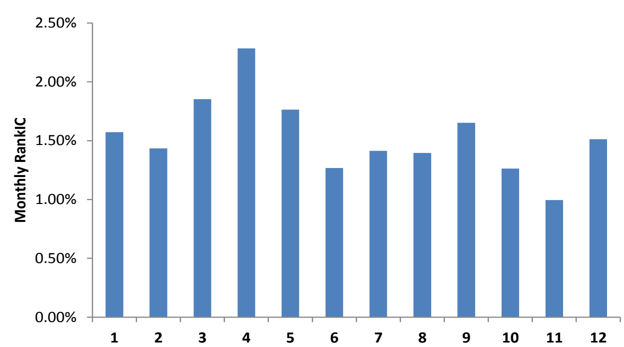
One of the challenges associated with previously examined strategies that use sentiment data is high turnover. In this section, we suggest an alternative formulation that generates lower turnover and also has higher capacity. Specifically, we examine a range-based strategy that involves taking the *minimum* sentiment less the *maximum* sentiment over the past sixty days. The performance of strategy is displayed below and generates an long (top quintile) / short (bottom quintile) Sharpe Ratio of 0.95 with an information coefficient of 1.5% which is relatively stable over the past 12 months.

Figure 35: Returns for portfolios formed on Sentiment Range



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

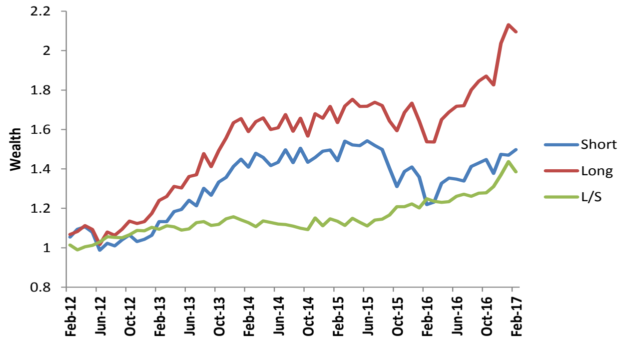
Figure 36: IC decay for L/S portfolios formed on Sentiment Range



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

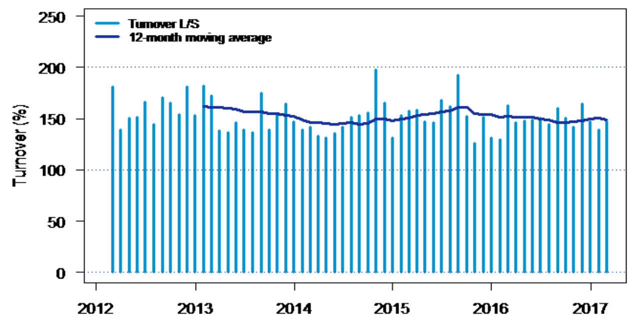
Previously, we explained that sentiment volume and the range of sentiment tends to be larger around certain news-worthy events that cause individuals to post their opinions on social media. The sentiment range strategy captures the dispersion of opinions surrounding different events that have occurred in the recent past. The reason why sentiment range explains differences in average returns is similar to the story behind analyst dispersion. While retail investors buy following positive sentiment events, they are much less likely to short sell following negative sentiment events. Thus, positive sentiment information is incorporated into stock prices while negative sentiment information is not - higher sentiment ranges (our measure is calculated as minimum sentiment less maximum sentiment) that predicts negative future stock returns.

Figure 37: Wealth curve of the backtest portfolio on Sentiment Range



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

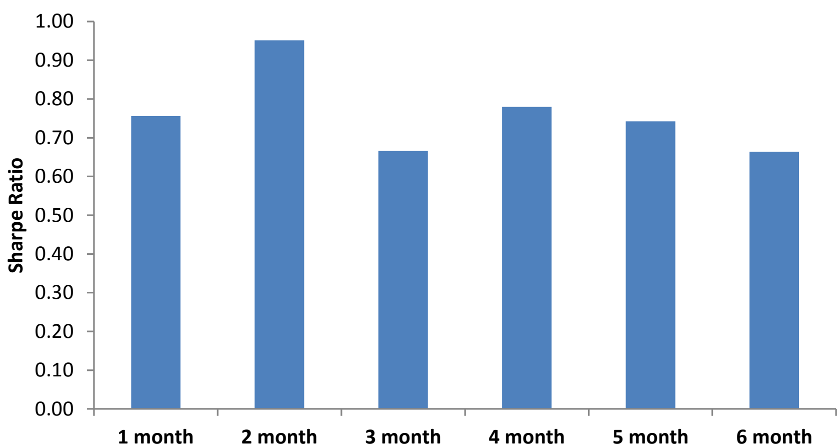
Figure 38: Sentiment Range turnover of the backtest portfolio



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

As a robust check, we also calculated the sentiment range factors using different time windows to estimate the historical sentiment range, which we report on Figure 39. While the 2 month window has the best performance, all of the other versions that we try have Sharpe ratio that are greater than 0.60. The two month window is long enough to allow social sentiment to change but not so long a window to keep the social sentiment from becoming outdated.

Figure 39: Sentiment Range strategy Sharpe Ratio under different calculation windows

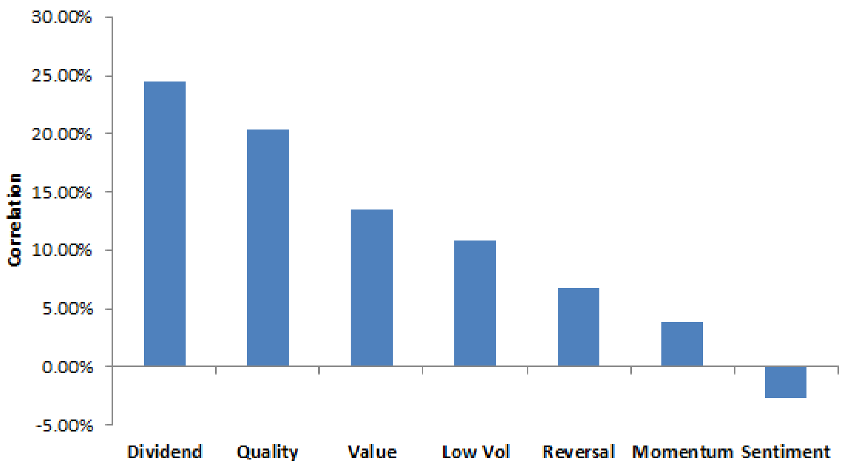


Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Performance attribution

Figure 40 presents correlations between Long/Short portfolios formed on Sentiment Range and different standard long/short quantitative strategies: Momentum, Analyst Revision, Low Volatility, Value, Quality, Reversal and Dividend-to-Price. We find relatively low correlations with most, major traditional quantitative factors. As we show, the range factor has the highest correlations with Dividend-to-Price and Quality. In contrast, we find the range factor has low correlations with Analyst Revisions and Momentum.

Figure 40: Long/Short portfolio return correlation between range factor and conventional factors



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, IHS Markit, Deutsche Bank

Social Media Sentiment in Corporate Events

Event study based on sentiment

In this section, we analyze whether social media sentiment can be used to improve stock return predictability on and around corporate events.

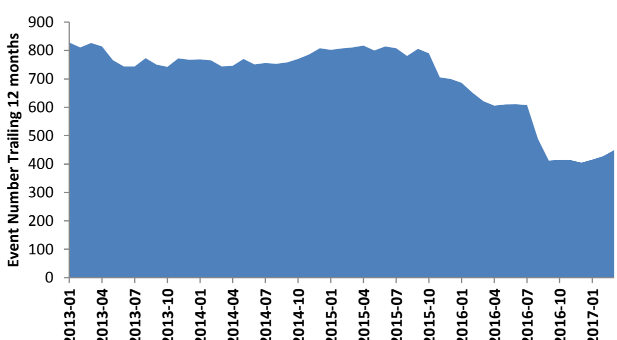
Quantifying language can be used to estimate impact on a firm's earnings prospects and stock price for a variety of different corporate events.

Specifically, the events that we examine in this report include:

- Raise/Lowering of earnings guidance
- Increase/Decrease in dividends
- Merger and Acquisition announcements
- M&A rumors

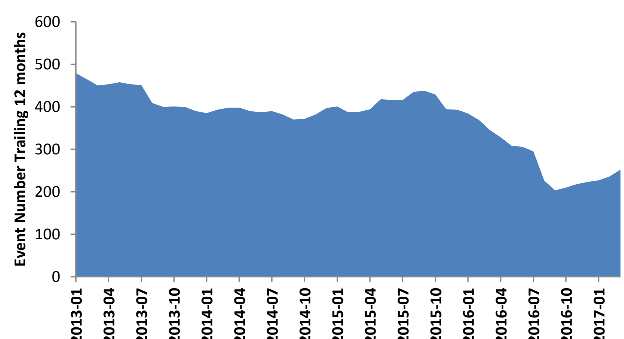
As part of our analysis, we evaluate whether sentiment around events can predict post-event drift. We obtain information on events from S&P Capital IQ's key event dataset [8]. Figure 41 through Figure 46 report the number of corporate event (based on trailing 12-months) from January 2012 to February 2017.

Figure 41: Earnings Guidance Raised



Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

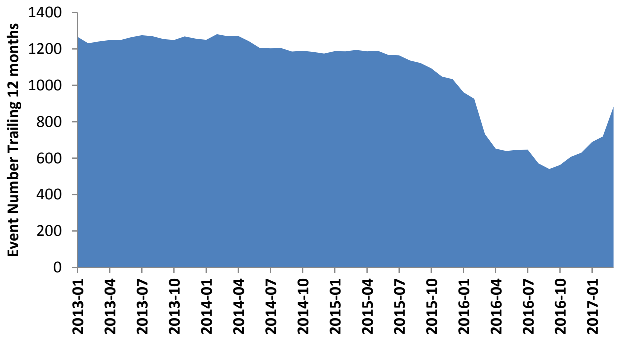
Figure 42: Earnings Guidance Lowered



Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

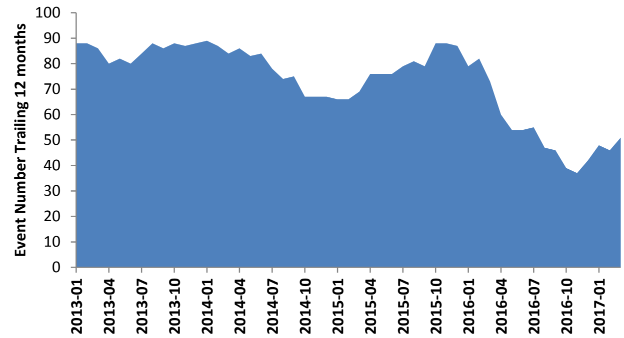
8 Please refer to Luo Y et al [2013] for detailed description of the dataset.

Figure 43: Dividend Increase



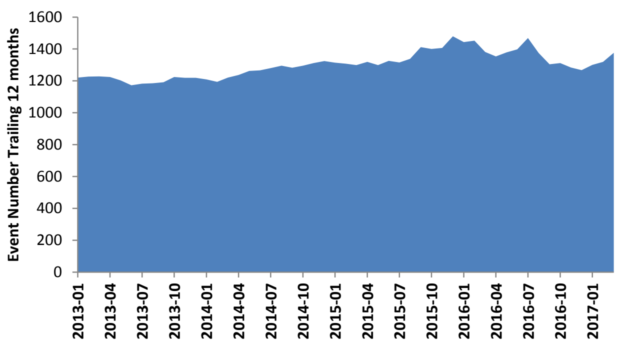
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 44: Dividend Decrease



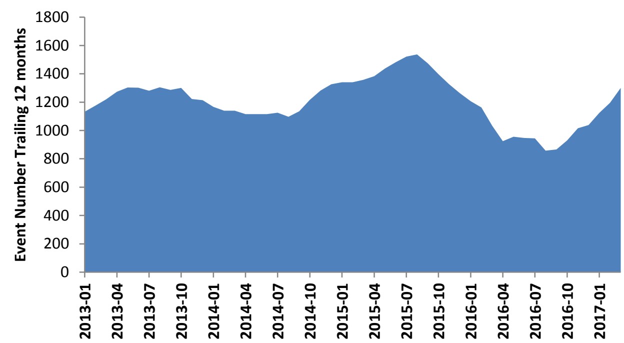
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 45: M&A announcement (target company)



Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 46: M&A rumor (target company)

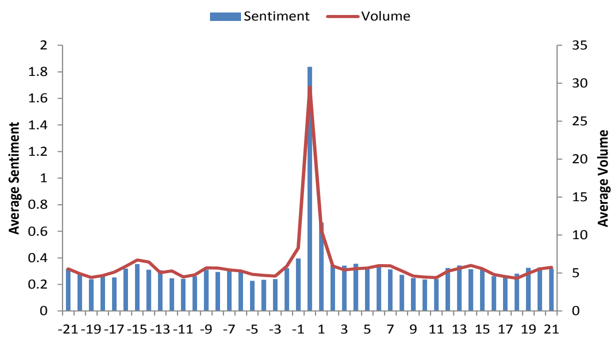


Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

As we show, close to twice the number of firms raised guidance compared to those that lowered guidance. Ten times the number of firms raised dividends when compared to those that lowered dividends.

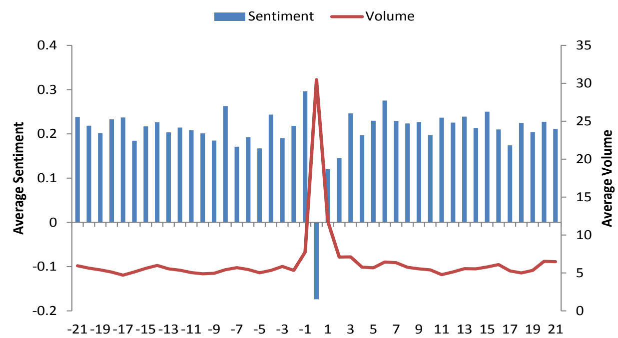
As a first step, we report the level of social media sentiment and volume of tweets for various corporate events Figure 47 to Figure 52 in event time (x-axis on each chart displays the number of trading days till or after the corporate event). Generally, our measure of sentiment has a positive mean as the number of positive words exceeds the number of negative words as reported in Figure 7. Most of the results show large spikes in volume surrounding corporate events. We also find sentiment is correlated in a way that is consistent with the expected impact of the corporate event on the stock price: guidance and dividend increases are associated with positive sentiment; while guidance and dividend decreases are associated with negative sentiment. The sentiment estimated for days after and before the event is generally not that different suggesting that sentiment associated with corporate events is short-lived.

Figure 47: Earnings Guidance Raised



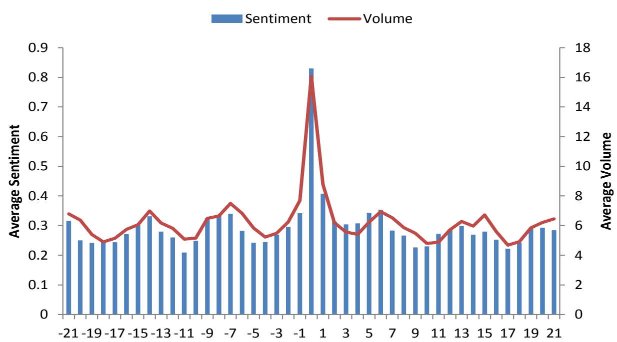
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 48: Earnings Guidance Lowered



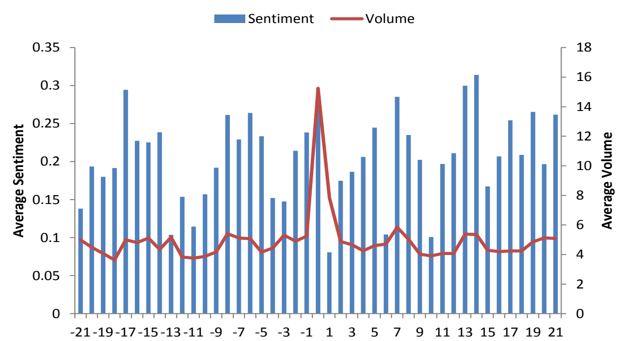
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 49: Dividend Increase



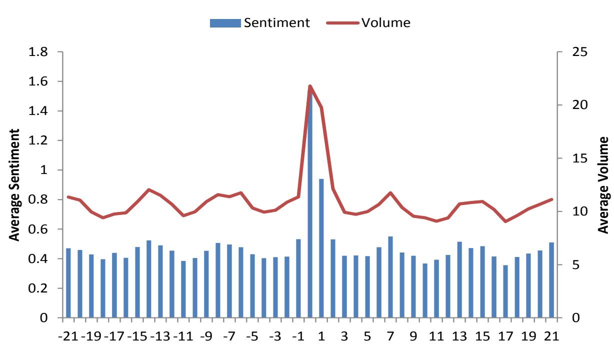
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 50: Dividend Decrease



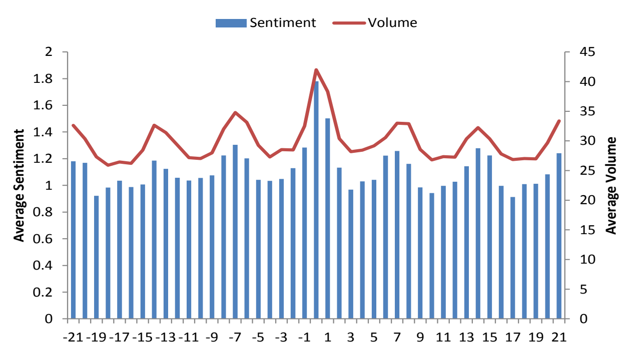
Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 51: M&A Announcement (target company)



Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 52: M&A Rumor (target company)

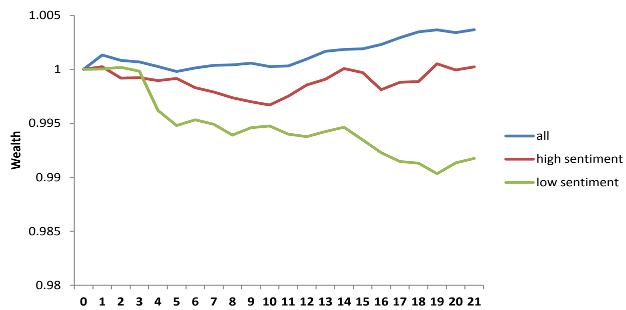


Source:, Russell, S&P Capital IQ Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

We extend the analysis by evaluating whether sentiment predicts post-event return drift. Our methodology involves sorting stocks based on sentiment score (S_{score}) on event announcement date (T=0) and examine average market-adjusted performance over the following month (day T+1 to T+21). The next eight charts are organized as follows. For each event, we report the average cumulative buy-and-hold return associated with all firms, firms with the highest (top 10% across all corporate events) sentiment score S_{score} and firms with the lowest (bottom 10%

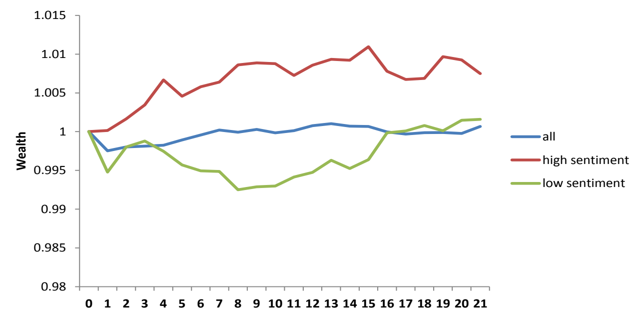
across all corporate events) sentiment score. We then evaluate whether social media sentiment can help explain post-event drift.

Figure 53: Post-Event Return Drift: Earnings Guidance Raised



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

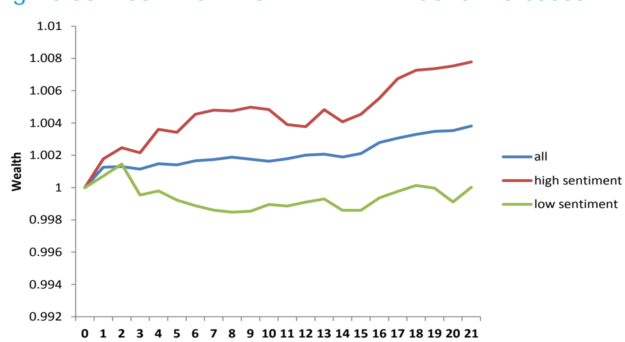
Figure 54: Post-Event Return Drift:Earnings Guidance Lowered



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

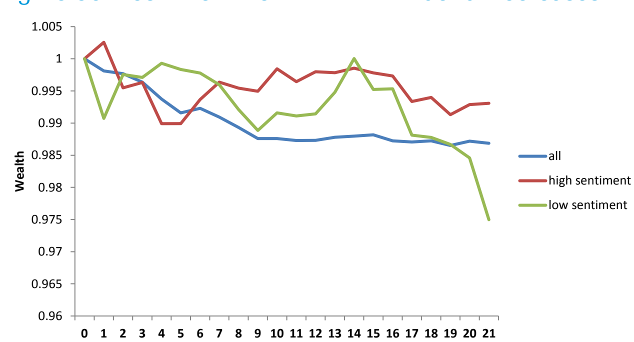
As we show in Figure 53 and Figure 54, sentiment generates post-event earning guidance return predictability. In the left-most graph, the low sentiment firms underperform the average firm that increases their earnings guidance by 1.2% over the next month. High sentiment firms have slightly lower returns when compared to the average firm that receives an earnings increase. The right-most graph shows that high social media sentiment firms that generate close to 0.75% in cumulative abnormal returns in the month following the guidance raise compared to close to 0% for the low social media sentiment firms.

Figure 55: Post-Event Return Drift:Dividend Increases



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 56: Post-Event Return Drift:Dividend Decreases



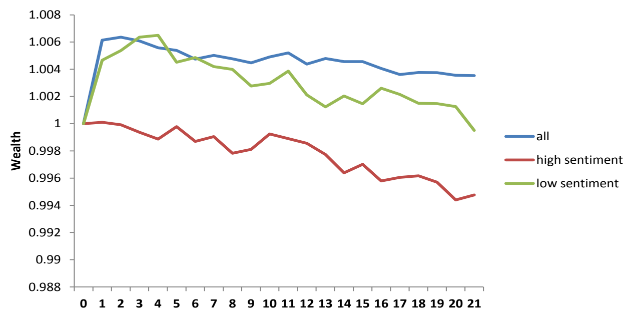
Source:Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

A dividend increase can signal that future business prospects are strong and the company can commit to payments to shareholders. A dividend decrease could signal that the business is in trouble and the company cannot maintain its current dividend payout, or that the management could not find profitable investment opportunities.

Following dividend increases, there are small average increases in stock prices (0.30%) as shown by the blue line in Figure 55. However, for firms that have high sentiment on the announcement day, we find an increase of 0.80% over the next month. We also find that low sentiment firms have flat performance over the next month.

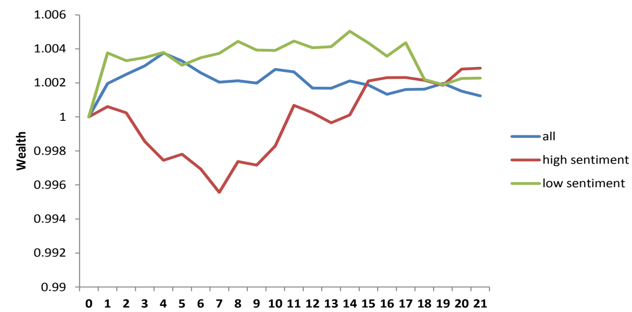
The results are more striking for dividend decreases reported in Figure 56. On average, dividend decreasing companies experience negative returns (-1.25%) in the month following the announcement day. Low sentiment firms decline by 2.5% over the next month, while high sentiment firms only decline by 0.6%.

Figure 57: Post-Event Return Drift:M&A Announcements (target company)



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

Figure 58: Post-Event Return Drift: M&A Rumors (target company)



Source: Bloomberg Finance LP, Compustat, IBES, MSCI, Russell, S&P Capital IQ, Thomson Reuters, Social Market Analytics, Deutsche Bank

We also analyze the performance of target companies during merger and acquisition (M&A) announcements (Figure 57) and takeover rumors (Figure 58). As we show, sentiment is much less useful for explaining post-event drift for M&A announcements and takeovers. High sentiment (calculated on the announcement day) firms actually have poorer returns than low sentiment firms for actual M&A announcements and have similar returns to low sentiment firms for M&A rumors. We generally find weak evidence for social media sentiment to be able to explain differences in post-event drift for M&A events.

Overall the results that we present in this section are mixed. For earnings guidance and dividend changes we find that sentiment can predict post-event drift. For M&A announcements and rumors social media sentiment does not explain differences in post-event next-month returns.

Concluding Remarks

While natural language processing of social media sentiment information shows great promise, applications towards stock selection tend forecast returns over very short horizons (few days) and generate high turnover. The big challenge involves improving the signal-to-noise ratio through better NLP techniques and subsequent signal construction. This is an aspect of text mining / NLP research which we hope to address in future work.

Bibliography

J.Blaschak, A. Blinov, et al. [2015], "Systems and Methods of Detecting, Measuring, and Extracting Signatures of Signals Embedded in Social Media Data Streams", U.S. Patent No 9,104,734, August 2015.

G. Brown, M. Cliff [2004], "Investor sentiment and the near-term stock market", Journal of Empirical Finance, May 13, 2002.

R. Giannini, P. Irvine, T. Shu, [2012], "The Convergence and Divergence of Investors' Opinions around Earnings News: Evidence from a Social Network", Asian Finance Association (AsFA) 2013 Conference, March 8, 2012.

IHS Markit Research Signals [2014], "Alpha: Extracting Market Sentiment From 140 Characters", March 3, 2014.

J. Bollen, H. Mao, X. Zeng [2011], "Twitter mood predicts the stock market", Journal of computational science, 2011.

M. Jung, J. Naughton, et al. [2014], "Corporate use of social media" Unpublished paper, Northwestern University, 2014.

J. Jussa, Z. Chen, Y. Luo, R. Cahan, M. Alvarez [2011], "Canada Quant: Technically Savvy Alpha", Deutsche Bank Quantitative Strategy, May 6, 2012.

M. Baker and J. Wurgler [2006], "Investor sentiment and the cross-section of stock returns", Journal of Finance, August, 2006.

M. Alvarez, Y. Luo, R. Cahan, J. Jussa, Z. Chen, S. Wang [2012[, "Portfolios Under Construction: Correlation & Consequences", Deutsche Bank Quantitative Strategy, January 24, 2012.

P. Tetlock [2007], "Giving Content to Investor Sentiment: The Role of Media in the Stock Market", Journal of Finance, June 2007.

S. Wang, Y. Luo, R. Cahan, M. Alvarez, J. Jussa, Z Chen [2012], "Signal Processing: The Rise of the Machines", Deutsche Bank Quantitative Strategy, 5 June 2012.

Y. Choueifaty, T. Froidure, and J. Reynier [2011], "Properties of the Most Diversified Portfolio", Journal of Investment Strategies, July 6, 2011.

Y. Luo, R. Cahan, M. Alvarez, J. Jussa, Z. Chen [2011], "Signal Processing: Quant Tactical Asset Allocation", Deutsche Bank Quantitative Strategy, 19 September 2011.

Y. Luo, S. Wang, R. Cahan, M. Alvarez, J. Jussa, Z, Chen [2012], "Signal Processing: New Insights in Country Rotation", Deutsche Bank Quantitative Strategy, 9 February 2012.

Y. Luo, R. Cahan, J. Jussa, M. Alvarez [2012], "QCD Model: DB Quant Handbook", Deutsche Bank Quantitative Strategy, 22 July 2010.

Y. Luo, S. Wang, R. Cahan, J. Jussa, Z. Chen, M. Alvarez [2013], "DB Handbook of Portfolio Construction, Part I", Deutsche Bank Quantitative Strategy, May 30, 2013.

Y. Luo, S. Wang, M. Alvarez, J. Jussa, A. Wang, G. Rohal [2014], "Signal Processing: Macro Uncertainty, Investor Sentiment, and Asset Returns", Deutsche Bank Quantitative Strategy, September 15, 2014.

M. Zhou, L. Lei, J. Wang, W. Fan, A. Wang [2015], "Social media adoption and corporate disclosure", Journal of Information Systems, 2015.

Appendix 1

Important Disclosures

*Other information available upon request

*Prices are current as of the end of the previous trading session unless otherwise indicated and are sourced from local exchanges via Reuters, Bloomberg, and other vendors. Other information is sourced from Deutsche Bank, subject companies, and other sources. For disclosures pertaining to recommendations or estimates made on securities other than the primary subject of this research, please see the most recently published company report or visit our global disclosure look-up page on our website at http://gm.db.com/ger/disclosure/DisclosureDirectory.eqsr. Aside from within this report, important conflict disclosures can also be found at https://gm/db.com/equities under the "Disclosures Lookup" and "Legal" tabs. Investors are strongly encouraged to review this information before investing.

Analyst Certification

The views expressed in this report accurately reflect the personal views of the undersigned lead analyst(s). In addition, the undersigned lead analyst(s) has not and will not receive any compensation for providing a specific recommendation or view in this report. George Zhao, Ronnie Shah, Spyros Mesomeris, Andy Moniz, David Elledge

Hypothetical Disclaimer

Backtested, hypothetical or simulated performance results have inherent limitations. Unlike an actual performance record based on trading actual client portfolios, simulated results are achieved by means of the retroactive application of a backtested model itself designed with the benefit of hindsight. Taking into account historical events the backtesting of performance also differs from actual account performance because an actual investment strategy may be adjusted any time, for any reason, including a response to material, economic or market factors. The backtested performance includes hypothetical results that do not reflect the reinvestment of dividends and other earnings or the deduction of advisory fees, brokerage or other commissions, and any other expenses that a client would have paid or actually paid. No representation is made that any trading strategy or account will or is likely to achieve profits or losses similar to those shown. Alternative modeling techniques or assumptions might produce significantly different results and prove to be more appropriate. Past hypothetical backtest results are neither an indicator nor guarantee of future returns. Actual results will vary, perhaps materially, from the analysis.

Additional Information

The information and opinions in this report were prepared by Deutsche Bank AG or one of its affiliates (collectively "Deutsche Bank"). Though the information herein is believed to be reliable and has been obtained from public sources believed to be reliable, Deutsche Bank makes no representation as to its accuracy or completeness. Hyperlinks to third-party websites in this report are provided for reader convenience only. Deutsche Bank neither endorses the content nor is responsible for the accuracy or security controls of these websites.

If you use the services of Deutsche Bank in connection with a purchase or sale of a security that is discussed in this report, or is included or discussed in another communication (oral or written) from a Deutsche Bank analyst, Deutsche Bank may act as principal for its own account or as agent for another person.

Deutsche Bank may consider this report in deciding to trade as principal. It may also engage in transactions, for its own account or with customers, in a manner inconsistent with the views taken in this research report. Others within Deutsche Bank, including strategists, sales staff and other analysts, may take views that are inconsistent with those taken in this research report. Deutsche Bank issues a variety of research products, including fundamental analysis, equity-linked analysis, quantitative analysis and trade ideas. Recommendations contained in one type of communication may differ from recommendations contained in others, whether as a result of differing time horizons, methodologies or otherwise. Deutsche Bank and/or its affiliates may also be holding debt or equity securities of the issuers it writes on. Analysts are paid in part based on the profitability of Deutsche Bank AG and its affiliates, which includes investment banking, trading and principal trading revenues.

Opinions, estimates and projections constitute the current judgment of the author as of the date of this report. They do not necessarily reflect the opinions of Deutsche Bank and are subject to change without notice. Deutsche Bank provides liquidity for buyers and sellers of securities issued by the companies it covers. Deutsche Bank research analysts sometimes have shorter-term trade ideas that are consistent or inconsistent with Deutsche Bank's existing longer term ratings. Trade ideas for equities can be found at the SOLAR link at http://gm.db.com. A SOLAR idea represents a high conviction belief by an analyst that a stock will outperform or underperform the market and/or sector delineated over a time frame of no less than two weeks. In addition to SOLAR ideas, the analysts named in this report may from time to time discuss with our clients, Deutsche Bank salespersons and Deutsche Bank traders, trading strategies or ideas that reference catalysts or events that may have a near-term or medium-term impact on the market price of the securities discussed in this report, which impact may be directionally counter to the analysts' current 12-month view of total return or investment return as described herein. Deutsche Bank has no obligation to update, modify or amend this report or to otherwise notify a recipient thereof if any opinion, forecast or estimate contained herein changes or subsequently becomes inaccurate. Coverage and the frequency of changes in market conditions and in both general and company specific economic prospects make it difficult to update research at defined intervals. Updates are at the sole discretion of the coverage analyst concerned or of the Research Department Management and as such the majority of reports are published at irregular intervals. This report is provided for informational purposes only and does not take into account the particular investment objectives, financial situations, or needs of individual clients. It is not an offer or a solicitation of an offer to buy or sell any financial instruments or to participate in any particular trading strategy. Target prices are inherently imprecise and a product of the analyst's judgment. The financial instruments discussed in this report may not be suitable for all investors and investors must make their own informed investment decisions. Prices and availability of financial instruments are subject to change without notice and investment transactions can lead to losses as a result of price fluctuations and other factors. If a financial instrument is denominated in a currency other than an investor's currency, a change in exchange rates may adversely affect the investment. Past performance is not necessarily indicative of future results. Unless otherwise indicated, prices are current as of the end of the previous trading session, and are sourced from local exchanges via Reuters, Bloomberg and other vendors. Data is sourced from Deutsche Bank, subject companies, and in some cases, other parties.

The Deutsche Bank Research Department is independent of other business areas divisions of the Bank. Details regarding our organizational arrangements and information barriers we have to prevent and avoid conflicts of interest with respect to our research is available on our website under Disclaimer found on the Legal tab.

Macroeconomic fluctuations often account for most of the risks associated with exposures to instruments that promise to pay fixed or variable interest rates. For an investor who is long fixed rate instruments (thus receiving these cash flows), increases in interest rates naturally lift the discount factors applied to the expected cash flows and thus cause a loss.

The longer the maturity of a certain cash flow and the higher the move in the discount factor, the higher will be the loss. Upside surprises in inflation, fiscal funding needs, and FX depreciation rates are among the most common adverse macroeconomic shocks to receivers. But counterparty exposure, issuer creditworthiness, client segmentation, regulation (including changes in assets holding limits for different types of investors), changes in tax policies, currency convertibility (which may constrain currency conversion, repatriation of profits and/or the liquidation of positions), and settlement issues related to local clearing houses are also important risk factors to be considered. The sensitivity of fixed income instruments to macroeconomic shocks may be mitigated by indexing the contracted cash flows to inflation, to FX depreciation, or to specified interest rates – these are common in emerging markets. It is important to note that the index fixings may -- by construction -- lag or mis-measure the actual move in the underlying variables they are intended to track. The choice of the proper fixing (or metric) is particularly important in swaps markets, where floating coupon rates (i.e., coupons indexed to a typically short-dated interest rate reference index) are exchanged for fixed coupons. It is also important to acknowledge that funding in a currency that differs from the currency in which coupons are denominated carries FX risk. Naturally, options on swaps (swaptions) also bear the risks typical to options in addition to the risks related to rates movements.

Derivative transactions involve numerous risks including, among others, market, counterparty default and illiquidity risk. The appropriateness or otherwise of these products for use by investors is dependent on the investors' own circumstances including their tax position, their regulatory environment and the nature of their other assets and liabilities, and as such, investors should take expert legal and financial advice before entering into any transaction similar to or inspired by the contents of this publication. The risk of loss in futures trading and options, foreign or domestic, can be substantial. As a result of the high degree of leverage obtainable in futures and options trading, losses may be incurred that are greater than the amount of funds initially deposited. Trading in options involves risk and is not suitable for all investors. Prior to buying or selling an option investors must review the "Characteristics and Risks of Standardized Options", at http://www.optionsclearing.com/about/publications/character-risks.jsp. If you are unable to access the website please contact your Deutsche Bank representative for a copy of this important document.

Participants in foreign exchange transactions may incur risks arising from several factors, including the following: (i) exchange rates can be volatile and are subject to large fluctuations; (ii) the value of currencies may be affected by numerous market factors, including world and national economic, political and regulatory events, events in equity and debt markets and changes in interest rates; and (iii) currencies may be subject to devaluation or government imposed exchange controls which could affect the value of the currency. Investors in securities such as ADRs, whose values are affected by the currency of an underlying security, effectively assume currency risk.

Unless governing law provides otherwise, all transactions should be executed through the Deutsche Bank entity in the investor's home jurisdiction. Aside from within this report, important conflict disclosures can also be found at https://gm.db.com/equities under the "Disclosures Lookup" and "Legal" tabs. Investors are strongly encouraged to review this information before investing.

Deutsche Bank (which includes Deutsche Bank AG, its branches and all affiliated companies) is not acting as a financial adviser, consultant or fiduciary to you, any of your agents (collectively, "You" or "Your") with respect to any information provided in the materials attached hereto. Deutsche Bank does not provide investment, legal, tax or accounting advice, Deutsche Bank is not acting as Your impartial adviser, and does not express any opinion or recommendation whatsoever as to any strategies, products or any other information presented in the materials. Information contained herein is being provided solely on the basis that the recipient will make an independent assessment of the merits of any investment decision, and it does not constitute a recommendation of, or express an opinion on, any product or service or any trading strategy.

The information presented is general in nature and is not directed to retirement accounts or any specific person or account type, and is therefore provided to You on the express basis that it is not advice, and You may not rely upon it in making Your decision. The information we provide is being directed only to persons we believe to be financially sophisticated, who are capable of evaluating investment risks independently, both in general and with regard to particular transactions and investment strategies, and who understand that Deutsche Bank has financial interests in the offering of its products and services. If this is not the case, or if You are an IRA or other retail investor receiving this directly from us, we ask that you inform us immediately.

United States: Approved and/or distributed by Deutsche Bank Securities Incorporated, a member of FINRA, NFA and SIPC. Analysts located outside of the United States are employed by non-US affiliates that are not subject to FINRA regulations.

Germany: Approved and/or distributed by Deutsche Bank AG, a joint stock corporation with limited liability incorporated in the Federal Republic of Germany with its principal office in Frankfurt am Main. Deutsche Bank AG is authorized under German Banking Law and is subject to supervision by the European Central Bank and by BaFin, Germany's Federal Financial Supervisory Authority.

United Kingdom: Approved and/or distributed by Deutsche Bank AG acting through its London Branch at Winchester House, 1 Great Winchester Street, London EC2N 2DB. Deutsche Bank AG in the United Kingdom is authorised by the Prudential Regulation Authority and is subject to limited regulation by the Prudential Regulation Authority and Financial Conduct Authority. Details about the extent of our authorisation and regulation are available on request.

Hong Kong: Distributed by Deutsche Bank AG, Hong Kong Branch or Deutsche Securities Asia Limited.

India: Prepared by Deutsche Equities India Pvt Ltd, which is registered by the Securities and Exchange Board of India (SEBI) as a stock broker. Research Analyst SEBI Registration Number is INH000001741. DEIPL may have received administrative warnings from the SEBI for breaches of Indian regulations.

Japan: Approved and/or distributed by Deutsche Securities Inc.(DSI). Registration number - Registered as a financial instruments dealer by the Head of the Kanto Local Finance Bureau (Kinsho) No. 117. Member of associations: JSDA, Type II Financial Instruments Firms Association and The Financial Futures Association of Japan. Commissions and risks involved in stock transactions - for stock transactions, we charge stock commissions and consumption tax by multiplying the transaction amount by the commission rate agreed with each customer. Stock transactions can lead to losses as a result of share price fluctuations and other factors. Transactions in foreign stocks can lead to additional losses stemming from foreign exchange fluctuations. We may also charge commissions and fees for certain categories of investment advice, products and services. Recommended investment strategies, products and services carry the risk of losses to principal and other losses as a result of changes in market and/or economic trends, and/or fluctuations in market value. Before deciding on the purchase of financial products and/or services, customers should carefully read the relevant disclosures, prospectuses and other documentation. "Moody's", "Standard & Poor's", and "Fitch" mentioned in this report are not registered credit rating agencies in Japan unless Japan or "Nippon" is specifically designated in the name of the entity. Reports on Japanese listed companies not written by analysts of DSI are written by Deutsche Bank Group's analysts with the coverage companies specified by DSI. Some of the foreign securities stated on this report are not disclosed according to the Financial Instruments and Exchange Law of Japan. Target prices set by Deutsche Bank's equity analysts are based on a 12-month forecast period.

Korea: Distributed by Deutsche Securities Korea Co.

South Africa: Deutsche Bank AG Johannesburg is incorporated in the Federal Republic of Germany (Branch Register Number in South Africa: 1998/003298/10).

Singapore: by Deutsche Bank AG, Singapore Branch or Deutsche Securities Asia Limited, Singapore Branch (One Raffles Quay #18-00 South Tower Singapore 048583, +65 6423 8001), which may be contacted in respect of any matters arising from, or in connection with, this report. Where this report is issued or promulgated in Singapore to a person who is not an accredited investor, expert investor or institutional investor (as defined in the applicable Singapore laws and regulations), they accept legal responsibility to such person for its contents.

Taiwan: Information on securities/investments that trade in Taiwan is for your reference only. Readers should independently evaluate investment risks and are solely responsible for their investment decisions. Deutsche Bank research may not be distributed to the Taiwan public media or quoted or used by the Taiwan public media without written consent. Information on securities/instruments that do not trade in Taiwan is for informational purposes only and is not to be construed as a recommendation to trade in such securities/instruments. Deutsche Securities Asia Limited, Taipei Branch may not execute transactions for clients in these securities/instruments.

Qatar: Deutsche Bank AG in the Qatar Financial Centre (registered no. 00032) is regulated by the Qatar Financial Centre Regulatory Authority. Deutsche Bank AG - QFC Branch may only undertake the financial services activities that fall within the scope of its existing QFCRA license. Principal place of business in the QFC: Qatar Financial Centre, Tower, West Bay, Level 5, PO Box 14928, Doha, Qatar. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available to Business Customers, as defined by the Qatar Financial Centre Regulatory Authority.

Russia: This information, interpretation and opinions submitted herein are not in the context of, and do not constitute, any appraisal or evaluation activity requiring a license in the Russian Federation.

Kingdom of Saudi Arabia: Deutsche Securities Saudi Arabia LLC Company, (registered no. 07073-37) is regulated by the Capital Market Authority. Deutsche Securities Saudi Arabia may only undertake the financial services activities that fall within the scope of its existing CMA license. Principal place of business in Saudi Arabia: King Fahad Road, Al Olaya District, P.O. Box 301809, Faisaliah Tower - 17th Floor, 11372 Riyadh, Saudi Arabia.

United Arab Emirates: Deutsche Bank AG in the Dubai International Financial Centre (registered no. 00045) is regulated by the Dubai Financial Services Authority. Deutsche Bank AG - DIFC Branch may only undertake the financial services activities that fall within the scope of its existing DFSA license. Principal place of business in the DIFC: Dubai International Financial Centre, The Gate Village, Building 5, PO Box 504902, Dubai, U.A.E. This information has been distributed by Deutsche Bank AG. Related financial products or services are only available to Professional Clients, as defined by the Dubai Financial Services Authority.

Australia: Retail clients should obtain a copy of a Product Disclosure Statement (PDS) relating to any financial product referred to in this report and consider the PDS before making any decision about whether to acquire the product. Please refer to Australian specific research disclosures and related information at https://australia.db.com/australia/content/research-information.html

Australia and New Zealand: This research is intended only for "wholesale clients" within the meaning of the Australian Corporations Act and New Zealand Financial Advisors Act respectively.

Additional information relative to securities, other financial products or issuers discussed in this report is available upon request. This report may not be reproduced, distributed or published without Deutsche Bank's prior written consent.

David Folkerts-Landau
Group Chief Economist and Global Head of Research

Raj Hindocha
Global Chief Operating Officer
Research

Michael Spencer
Head of APAC Research
Global Head of Economics

Steve Pollard
Head of Americas Research
Global Head of Equity Research

Anthony Klarman
Global Head of
Debt Research

Paul Reynolds
Head of EMEA
Equity Research

Dave Clark
Head of APAC
Equity Research

Pam Finelli
Global Head of
Equity Derivatives Research

Andreas Neubauer
Head of Research - Germany

Stuart Kirk
Head of Thematic Research

International locations

Deutsche Bank AG
Deutsche Bank Place
Level 16
Corner of Hunter & Phillip Streets
Sydney, NSW 2000
Australia
Tel: (61) 2 8258 1234

Deutsche Bank AG
Große Gallusstraße 10-14
60272 Frankfurt am Main
Germany
Tel: (49) 69 910 00

Deutsche Bank AG
Filiale Hongkong
International Commerce Centre,
1 Austin Road West,Kowloon,
Hong Kong
Tel: (852) 2203 8888

Deutsche Securities Inc.
2-11-1 Nagatacho
Sanno Park Tower
Chiyoda-ku, Tokyo 100-6171
Japan
Tel: (81) 3 5156 6770

Deutsche Bank AG London
1 Great Winchester Street
London EC2N 2EQ
United Kingdom
Tel: (44) 20 7545 8000

Deutsche Bank Securities Inc.
60 Wall Street
New York, NY 10005
United States of America
Tel: (1) 212 250 2500