Deutsche Bank

19 July 2010

Signal Processing

Beyond the headlines

Key summary
In this research we show that quantified news sentiment can offer a relatively uncorrelated new alpha source for quantitative investors. We show how to use advanced learning-type models to extract this alpha.

Using news flow to predict stock returns

News sentiment as an alpha signal
In this report we show that news sentiment is a promising alpha signal for quantitative investors. Advances in Natural Language Processing (NLP) techniques make it possible to use statistical algorithms to quantify the "sentiment" of a news story in real-time, based on the language and grammatical features of the article. This is extremely useful for quantitative investors because it opens up a whole new data source – textual data – that was previously impossible to use in a systematic fashion.

Using non-linear models to harness the alpha
We show that the alpha in news sentiment is best extracted using non-linear "learning" type models. These models capture the often complex interactions between sentiment and market-related variables like price and volume. Sentiment in absolute terms is less important than sentiment relative to market expectations; therefore non-linear models that allow conditional relationships between news and the market reaction to that news are useful.

An orthogonal data source
We find that the alpha generated from news sentiment is relatively orthogonal to more traditional quant factors, even though many traditional quant signals have their roots in how the market interprets news – for example price momentum. We also find that news sentiment is a high turnover signal, but it can still add value even in a lower turnover strategy.

Team contacts

Rochester Cahan, CFA
Strategist
(+1) 212 250-8983
rochester.cahan@db.com

Yin Luo, CFA
Strategist
(+1) 212 250-8983
yin.luo@db.com

Javed Jussa
Strategist
(+1) 212 250-4117
javed.jussa@db.com

Miguel-A Alvarez
Strategist
(+1) 212 250-8983
miguel-a.alvarez@db.com

Table of Contents

A letter to our readers

Harnessing the information in textual data

A recurring theme in our research is that the traditional "bread and butter" quant factors are no longer enough

One of the strongest criticisms of quantitative investing in recent years has been the assertion that many, if not most, of the common quant signals have been arbitraged out of existence, particularly in highly developed markets like the U.S. In our recent research we have documented at length how many of the "bread and butter" quant factors like PE, earnings revisions, and accruals have morphed from alpha factors to risk factors over the past few years.

Faced with this challenge, quantitative investors have two options: become better at timing the transient value-add of existing factors, or find new factors that are less widely used and hopefully can still deliver relatively stable alpha through time. The latter is of course easier said then done, and often requires the use of new, less scrutinized data sets. Indeed, finding such data sources has been a focus of our *Signal Processing* research series for the past few months. Last month we looked at how investors can use industry-specific datasets to augment existing factors and boost alpha [1], and the month before we showed how information from the options market can be used to pick stocks[2].

In this report we explore new factors constructed from an innovative dataset: news sentiment

This month we tackle another new dataset: news sentiment. Regular readers of our research will know that this is a topic we find particularly interesting, and one that we have already done a lot of work in. In this particular report, we take what we think is an innovative approach to studying the predictive power of news sentiment; instead of using standard linear models, we focus on three non-linear, "learning" type models: classification and regression trees, forests of classification and regression trees, and multivariate adaptive regression splines. All three of these models are unique in that they allow us to take a data-centric approach to our analysis. Instead of predefining a hypothetical relationship and then testing it, we allow the data to determine the form of the model. This allows us to better understand which variables within our dataset are most important in determining post-event abnormal returns. It also allows us to model complex non-linear relationships that may not be apparent at first glance.

We find that non-linear, "learning" type models are useful for extracting alpha from textual data

Overall we find that news sentiment, in conjunction with non-linear models, can generate alpha. Even better, we find this alpha is relatively uncorrelated with the more traditional quant factors. Of course, there is also a downside. The predictive ability of news sentiment is short-lived; the best results are obtained when forecasting only the next five days. Therefore, for some quantitative investors, the signal on its own may have too much turnover to be viable. Nonetheless, we do show that there are ways for even lower-frequency investors to use news sentiment data to enhance their stock-selection process.

As always, if you have any questions or comments, or would like to discuss any of these ideas in more detail you can reach us on +1 212 250 8983 or DBEQS.Americas.db.com.

Regards,

Yin, Rochester, Miguel, and Javed
Deutsche Bank North American Equity Quantitative Strategy

---

[1] Luo, Y., R. Cahan, J. Jussa, and M. Alvarez, 2010, "Signal Processing: Industry-specific factors", *Deutsche Bank Quantitative Strategy*, 8 June 2010

[2] Cahan, R., Y. Luo, J. Jussa, and M. Alvarez, 2010, "Signal Processing: The options issue", *Deutsche Bank Quantitative Strategy*, 12 May 2010

The data

Natural language processing: The next frontier?

Natural Language Processing is a cutting edge field where computer algorithms are used to "read" and interpret textual information

This month we tackle a completely new dataset: textual information. Recent advances in the field of natural language processing have led to a broad array of computer algorithms that can be used to extract quantitative metrics from textual data. These algorithms work by analyzing the words, phrases, and linguistic patterns in a block of text relative to historical patterns or a predetermined set of criteria. The ultimate aim is the use the algorithm to try to replicate (or indeed improve) on how a human reader might interpret the text. For example, one dimension of interest to investors is the *sentiment* of a news story; is the story positive or negative in tone, and is that positive or negative tone directed at a particular company mentioned within the news story?

Why quantify sentiment?

The ability to classify sentiment algorithmically is useful for two reasons:

- First, it allows a "hard" number to be assigned to a "soft" concept. As quants, we are always seeking numerical representations of abstract characteristics. Using a sentiment score to measure the sentiment of a news story is helpful in the same way that using a PE ratio is helpful in measuring the nebulous characteristic of "value". Both allow us to take a systematic approach to picking stocks that best represent that characteristic.
- Second, it allows us to process an almost unlimited number of news stories, essentially in real-time. Of course, a big part of the traditional fundamental analyst's job has been to read and interpret new information as it is released to the market. But with the explosion in information sources, this is becoming increasingly difficult. Along with 10-Ks and company press releases, one also needs to keep on top of internet stories and blogs, not to mention social networks and Twitter. A computer is much more efficient at processing such a huge volume of textual information in a timely manner, albeit arguably at a lower degree of accuracy. Essentially we are taking the view that the loss in skill is more than compensated for by the increased breadth.

An orthogonal data source?

But many traditional quant factors – like momentum – are intrinsically tied to news events, so the question of whether this data is truly orthogonal is an important one

In our analysis we investigate whether news sentiment data contains useful information for picking stocks. In particular, we are interested in whether there is any information in news sentiment data that is not already captured in traditional quant signals. This point is an important one. Many traditional factors are intrinsically linked to the release of news. For example, Hong and Stein [1999] argue that the reason price momentum works is because investors are slow to incorporate new information into their decisions. Likewise, the earnings revisions signal is also tied to news releases, because sell-side analysts typically revise their forecasts when important new information is released to the market. So there is the possibility that investing based on news sentiment is just another way to articulate the exposures that most quant investors already have. Therefore, we have to be careful to ensure that any predictability we find is truly orthogonal to existing quant signals.

The Reuters NewsScope database

We use the Reuters NewsScope database for this study

In this study we use the Reuters NewsScope database as our source for news data and for news sentiment scoring. The database is constructed by applying natural language processing (NLP) algorithms to news wire stories. The data is available from 2003 onwards, and includes U.S. as well as international companies. The underlying content used for sentiment processing comes exclusively from the Reuters news service, and includes a mixture of content types including alerts (typically used for breaking news) along with more in-depth news stories. Figure 1 shows a snapshot of the database format. The exact definition of each variable is specified in Figure 12 on page 11.

Figure 1: Snapshot of Reuters NewsScope database

RIC	ESTTIME	RLVN	SENT	SPOS	SNEU	SNEG	CNT1	CNT2	CNT3	CNT4	CNT5	ITYP	GENR	SEQN
IBM.N	1/23/2003 11:20:22	0.14	1	0.49	0.39	0.12	1	1	1	1	1	ARTICLE	NOT DEFINED	1
IBM.N	1/23/2003 10:51:01	0.71	-1	0.12	0.25	0.63	0	0	0	0	0	ALERT	NOT DEFINED	1
IBM.N	7/21/2009 9:46:42	1.00	-1	0.19	0.11	0.69	0	0	0	0	0	ARTICLE	NOT DEFINED	1
IBM.N	7/31/2009 12:30:01	1.00	-1	0.16	0.29	0.55	0	0	0	0	0	ALERT	NOT DEFINED	1
IBM.N	7/31/2009 12:31:50	1.00	1	0.41	0.40	0.20	1	1	1	1	1	ARTICLE	NOT DEFINED	1
IBM.N	7/22/2009 10:31:49	1.00	1	0.77	0.15	0.08	3	3	3	3	3	ARTICLE	NOT DEFINED	1
IBM.N	7/22/2009 10:30:13	1.00	-1	0.06	0.13	0.82	2	2	2	2	2	ALERT	NOT DEFINED	3
IBM.N	7/22/2009 10:30:04	1.00	1	0.85	0.12	0.03	1	1	1	1	1	ALERT	NOT DEFINED	2
IBM.N	7/22/2009 10:30:01	1.00	1	0.81	0.17	0.02	0	0	1	1	1	ALERT	NOT DEFINED	1
IBM.N	7/21/2009 7:20:24	1.00	-1	0.06	0.13	0.81	0	0	0	0	0	ARTICLE	NOT DEFINED	1
IBM.N	7/20/2009 22:11:44	0.20	0	0.20	0.79	0.01	0	0	0	0	0	ARTICLE	NOT DEFINED	1
IBM.N	7/20/2009 22:11:15	1.00	-1	0.06	0.13	0.82	1	1	1	1	1	ALERT	NOT DEFINED	2
IBM.N	7/20/2009 22:10:54	1.00	1	0.86	0.12	0.03	0	0	0	0	0	ALERT	NOT DEFINED	1
IBM.N	10/30/2009 2:30:20	1.00	1	0.74	0.20	0.05	1	1	1	1	1	ALERT	NOT DEFINED	5

Source: Reuters NewsScope, Deutsche Bank

Essentially, the NLP algorithms are designed to classify every news story along three broad dimensions:

News stories are scored across three dimensions: relevance, sentiment, and novelty

- **Relevance** assesses how targeted a news story is to a particular company. For example, a news story about IBM's 3Q earnings will have a high relevance for IBM, whereas a broad market commentary that mentions in passing that IBM was down 2% today will likely be less relevant to IBM.
- **Sentiment** measures how positive or negative the tone of the article is, with respect to a given company. For example and article that starts with "IBM's 3Q earnings disappoint" is likely to be tagged with a negative sentiment score for IBM. Sentiment scores depend not just on the words used in a story, but also the grammatical structure.
- **Novelty** quantifies how unique a given news story is. News stories tend to develop over time. For example, when a company releases earnings, the first news flash will usually just contain a headline with the EPS number reported. This will be followed up over the remained of the day with longer articles that analyze the reported results in more detail and discuss the market reaction. Novelty scoring is useful because it allows one to determine the first instance of each news event.

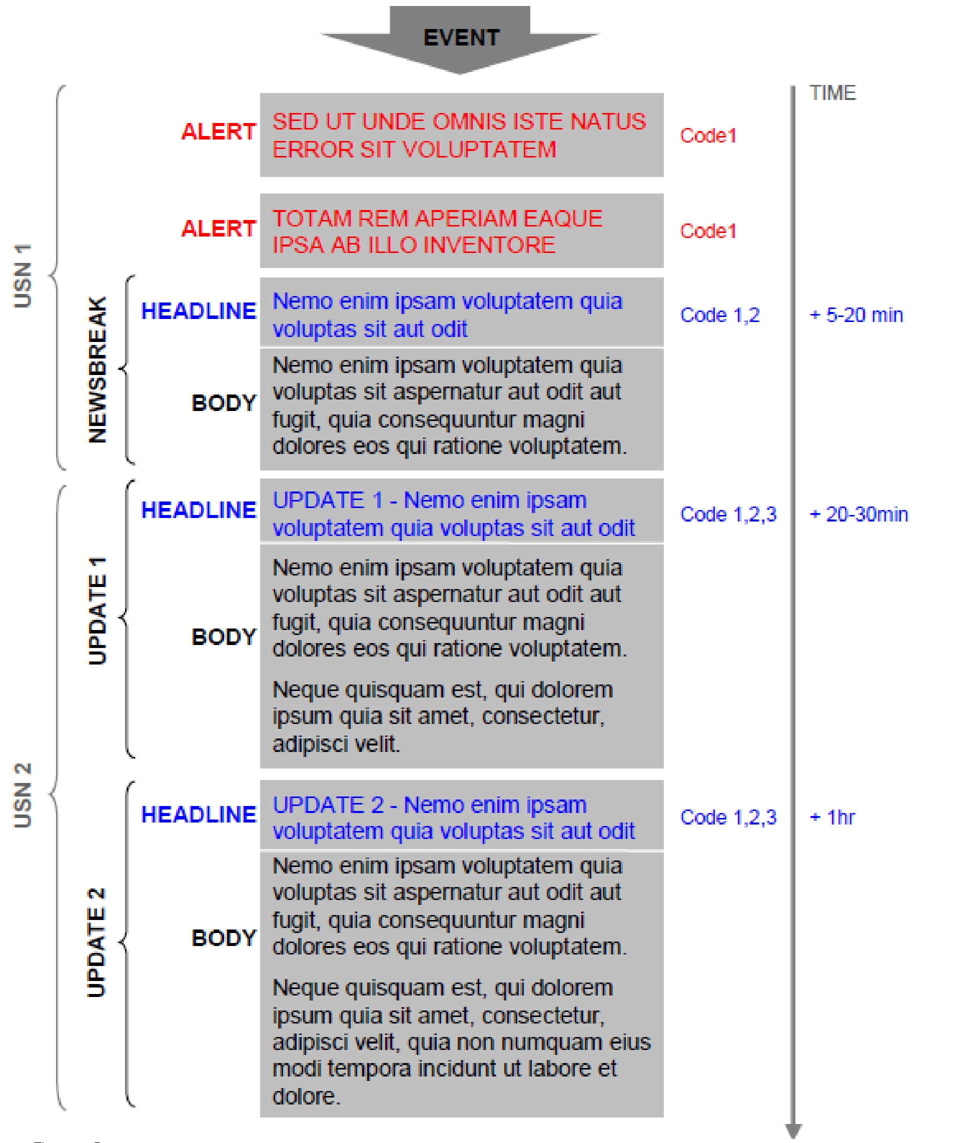
Evolution of a news story

A Reuters news story evolves in a number of steps: an alert is released first, followed by a succession of more in-depth stories

Further to the last bullet point, Figure 2 shows the typical evolution of a Reuters news story. When news breaks, typically the first one or two news items related to the event are *alerts*, i.e. brief one sentence headlines with just the essential information. These are then followed up in the next 5-20 minutes by a *newsbreak*, which consist of a headline plus two to four paragraphs of text giving more details. Finally, the story may be extended by an *update*, typically 20-30 minutes after the newsbreak. The update will contain more detailed information, context, and analysis around the event.

An important feature of the database is the fact that the evolution of a story is chained together by a unique ID. This allows one to assess how information about an event flows into the market over a period of time.

Figure 2: Evolution of a Reuters news story



Source: Thomson Reuters

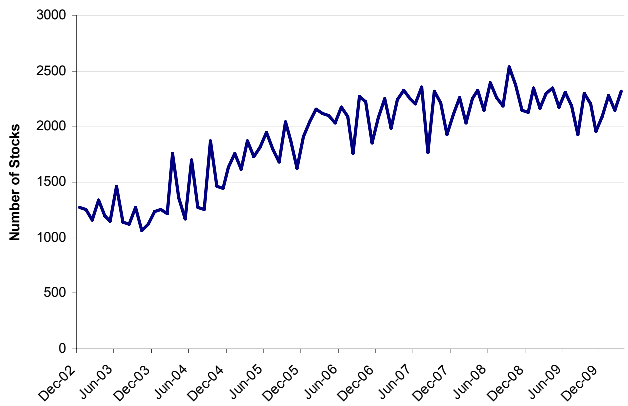
Summary statistics

Coverage is reasonable: typically around two thirds of the universe will have at least one news story in a given month

What does this news data actually look like? Figure 3 shows the number or stocks from the Russell 3000 index that have at least one news story in a given month. Coverage is reasonable, starting at around 1,250 stocks and rising to over 2,000 in recent years. The increase in coverage over time is driven by the fact that the number of news stories released on a given day has been steadily increasing over time.

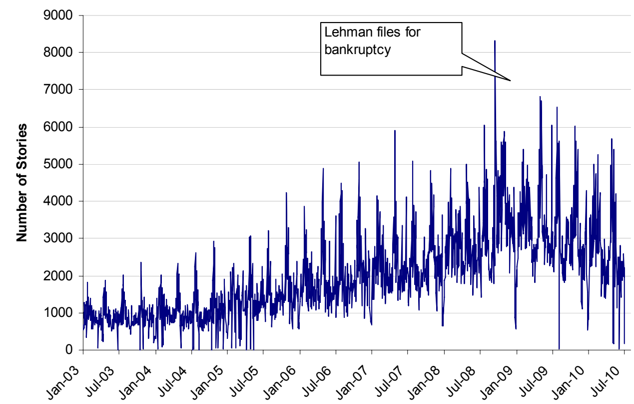
Another feature of news is that it tends to exhibit high seasonality (Figure 4). During reporting season there are of course many more news stories than at other times in the year. Abnormal events tend to also generate large spikes in news volume. September 15th, 2008 is a case in point – the day Lehman Brothers filed for bankruptcy.

Figure 3: Number of stocks with at least one news story in a given month, Russell 3000 universe



Source: Reuters NewsScope, Deutsche Bank

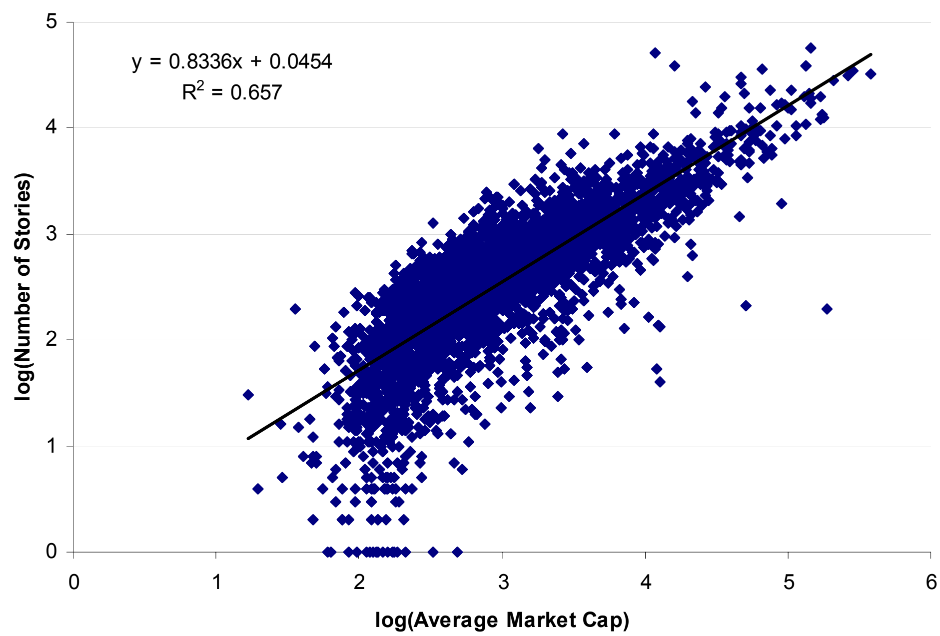
Figure 4: Number of news stories by day, Russell 3000 universe



Source: Reuters NewsScope, Deutsche Bank

As we might expect, news volume is also highly related to size; large companies tend to have more news stories than small companies (Figure 5). Thus it is important to control for a size bias in any investment strategy based on news sentiment.

Figure 5: Number of Stories versus Average Market Cap for stocks in Russell 3000 universe, 2003-present



Source: Reuters NewsScope, Deutsche Bank

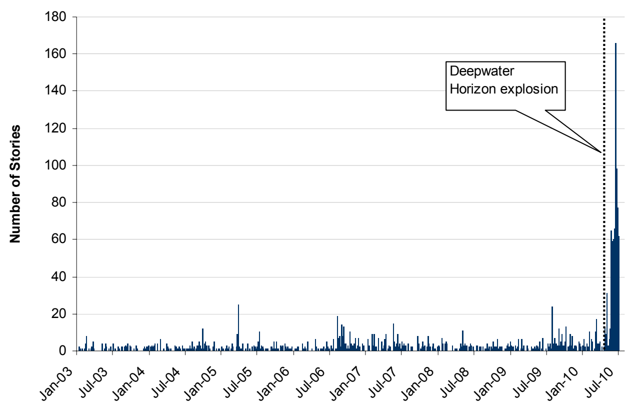
A real-life example

We use BP as a topical case study to illustrate what sentiment actually looks like

To get a better picture of how sentiment evolves through time, it is useful to examine a single stock. We pick BP as a particularly topical case study. Figure 6 shows the number of news stories about BP since 2003. It comes as no surprise that in the months after the Deepwater Horizon rig explosion, the volume of stories about the company has jumped by an order of magnitude.

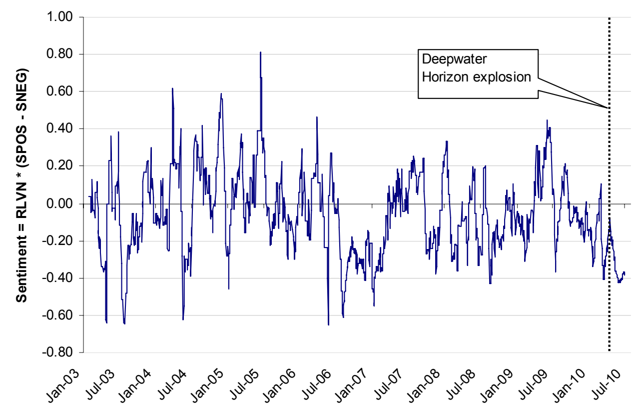
Figure 7 shows the news sentiment for BP through time. Clearly sentiment fell sharply after the disaster, and is currently tracking near the bottom of its long-term range. However, it is interesting to note that sentiment is currently no lower than it was though the 2006-07 period, a time when BP was dealing with another environmental issue (an oil spill in Prudhoe Bay, Alaska).

Figure 6: Number of stories for BP.N



Source: Reuters NewsScope, Deutsche Bank

Figure 7: Sentiment for BP.N (21D average)

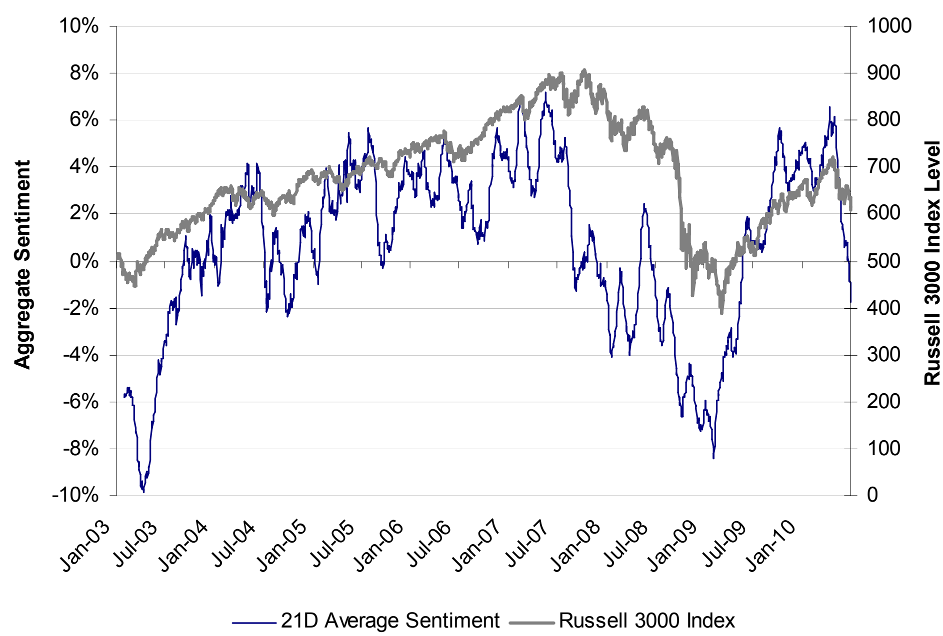


Source: Reuters NewsScope, Deutsche Bank

We can also aggregate sentiment up to a market or sector level, which is an interesting area for future research

It is also interesting to track aggregate sentiment for the market as a whole. Figure 8 shows the aggregate sentiment for Russell 3000 stocks, overlaid with the level of the Russell 3000 index. Clearly sentiment and the market tend to move in step, which is not particularly surprising given one could argue the market is the ultimate measure of sentiment. In this report we focus on picking individual stocks using sentiment, but using aggregate sentiment as a market or sector timing tool is an interesting area for future research.

Figure 8: Aggregate market sentiment



Source: Reuters NewsScope, Deutsche Bank

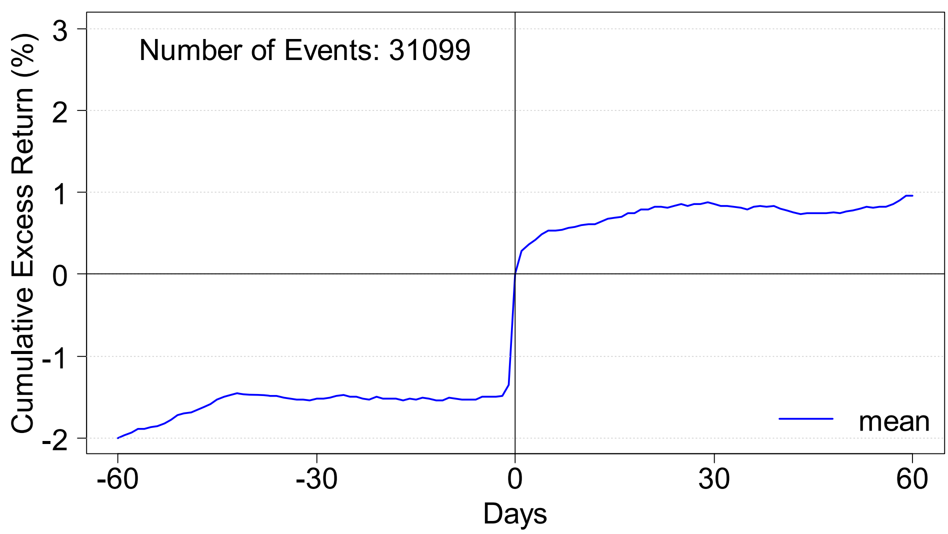
Event studies

Positive news events

We are particularly interested in whether there is any meaningful post-event drift after certain types of news events

Another way to visualize the dataset is to use an event study approach, to examine the typical reaction of stocks to certain types of news events. To do this, we start with a naïve approach where we predefine what we think constitutes positive, negative, or neutral event. For example, suppose we look at the average market-excess returns for stocks that experience a "positive" news event, where positive is defined to mean a news story where *RLVN=1*, *SENT*=1, *SPOS*>0.75, and *SEQN*=1 (these variables are defined in more detail on page 11). In other words, we are considering only news stories that are highly relevant to a particular stock, have a high probability of having positive sentiment, and are the first instance of a particular story. It seems reasonable to expect that such stories might lead to a positive market reaction. Figure 9 shows the results.

Figure 9: Average cumulative returns around positive news events, Russell 3000 universe, 2003-present



Source: Reuters NewsScope, Deutsche Bank

Even using simple definitions of positive events, we do find a moderate positive drift in the next 30 days

As mid-frequency investors, we are not trying to profit from the intraday price reaction in the seconds (or indeed milliseconds) after a news event. Rather, we are focused on whether there is any post-event drift that we might be able to profit from, should we take a long position following a positive news event. As the chart shows, there is moderate evidence of some post-event drift. In the 30 trading days after the event date, there is a market-adjusted positive drift of around 50bps on average. While not spectacular, the advantage of quant investing is that we may be able to profit from even a relatively small alpha opportunity so long as we can take the position many times – which with news stories we can.

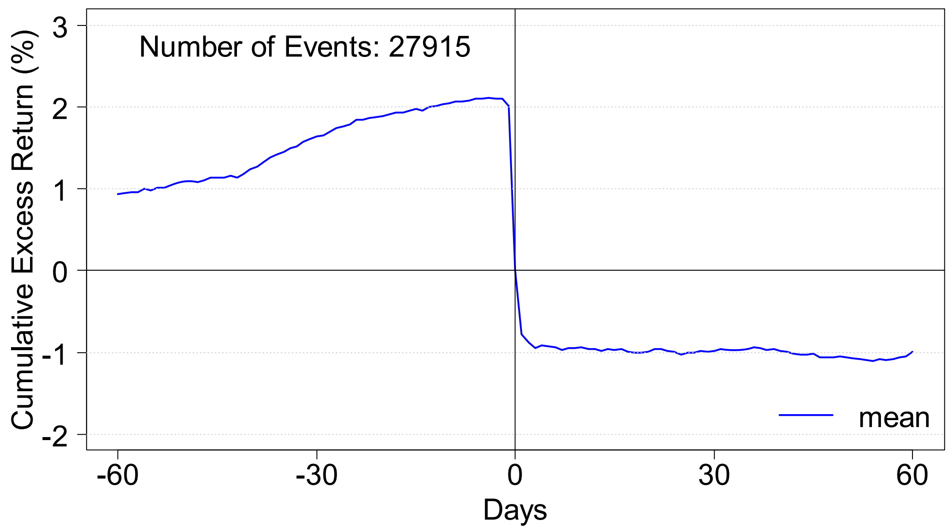
Negative news events

Figure 10 shows the same chart for negative news events. Again, we use a simple definition: *RLVN=1*, *SENT*=-1, *SNEG*>0.75, and *SEQN*=1. Interestingly, for negative news events there is very little drift in the days after the story is released. This is consistent with our past research, where we have found a "sell first and ask questions later" mentality when it comes to bad news, which tends to cause bad news to be priced in a lot quicker than good news.

For negative events we find less post-event drift, which is consistent with our past research

Also interesting is the fact stocks tend to run-up on average before bad news. This is puzzling because traditionally researchers tend to find some negative drift before bad news announcements in the days leading up to an event, due to information leakage. However, keep in mind that our study begins in 2003, which is after the adoption of Reg FD. This may well have reduced the extent to which investors act on negative news before it becomes public information.

Figure 10: Average cumulative returns around negative news events, Russell 3000 universe, 2003-present



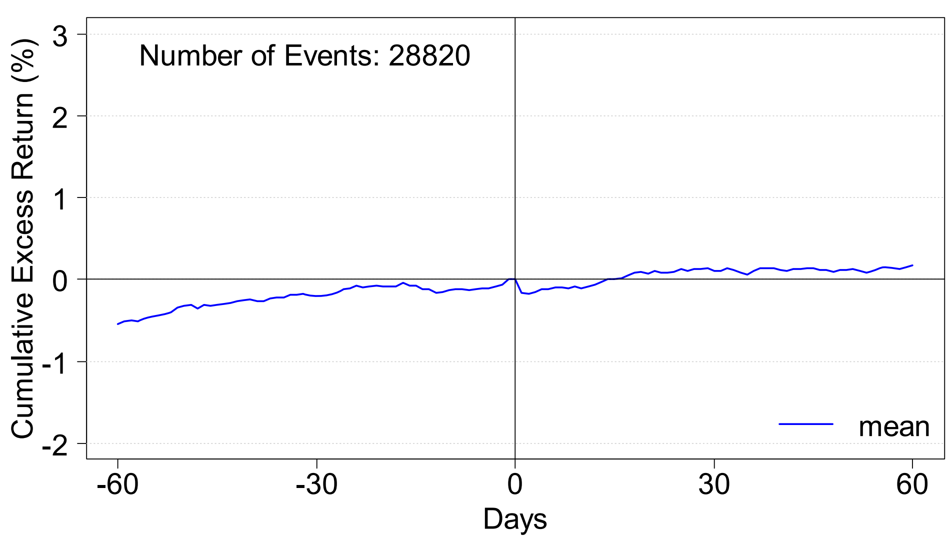
Source: Reuters NewsScope, Deutsche Bank

Neutral news events

Neutral events are unsurprisingly not very interesting!

As one could probably guess, neutral news events are uninteresting, but for those who want all the details, Figure 11 shows the results.

Figure 11: Average cumulative returns around neutral news events, Russell 3000 universe, 2003-present



Source: Reuters NewsScope, Deutsche Bank

Setting up our analysis

Our key research question is whether we can find a set of variables that can better help us predict which events will lead to large post-event drift

Based on the event studies above, the obvious question to ask is whether there is a better way to define positive and negative events, such that we get more pronounced post-event drift. To answer this question, we start by collecting a set of variables that may be useful in making this determination. Figure 12 shows the variables we selected. We call these variables the independent variables, because we are interested in using them to try to predict post-event abnormal returns. Our research question is simple: can we use news sentiment variables, along with more traditional market variables, to determine which news events are more likely to lead to post-event abnormal returns?

The independent variables are selected to capture the three dimensions mentioned above – relevance, sentiment, and novelty. For example *RLVN* is a continuous score from 0 to 1 that measures how relevant that particular news story is to a given stock. A score of 1 means the story is highly relevant. *SENT* is a discrete variable that takes on a value of -1 for negative sentiment, 0 for neutral sentiment and 1 for positive sentiment. Related to this score are the variables *SPOS*, *SNEU*, and *SNEG*, which measure the probability that the sentiment is positive, neutral, or negative respectively.

We select a number of news-based variables, plus market variables like event-day volume and return

Novelty is captures by the *CNT1* to *CNT5* variables. Essentially these variables measure how many related news stories have appeared in the past 12 hours through to seven days. Clearly a story that has a value of zero for all these variables is more novel than one that has high counts. A related variable is *SEQN*, which measures the sequence number of a story (1 for the first iteration, and incremented thereafter).

In addition to the news-related variables, we also want to include market-related variables in our analysis. Sentiment by itself may be useful for assessing the impact of an event, but clearly the market reaction to news events also tells us something important about the characteristics of the event. With this in mind, we also include volume and turnover metrics for the day before each event, and also on the day the event is released.

Figure 12: Independent variables

Variable	Source	Description	Type
RLVN	Reuters NewsScope	Relevance score (0 = no relevance, 1 = high relevance)	Continuous
SENT	Reuters NewsScope	Sentiment score (-1 = negative, 0 = neutral, 1 = positive)	Categorical
SPOS	Reuters NewsScope	Probability that sentiment score is positive	Continuous
SNEU	Reuters NewsScope	Probability that sentiment score is neutral	Continuous
SNEG	Reuters NewsScope	Probability that sentiment score is negative	Continuous
CNT1	Reuters NewsScope	Number of related stories in past 12 hours	Discrete
CNT2	Reuters NewsScope	Number of related stories in past 24 hours	Discrete
CNT3	Reuters NewsScope	Number of related stories in past 3 days	Discrete
CNT4	Reuters NewsScope	Number of related stories in past 5 days	Discrete
CNT5	Reuters NewsScope	Number of related stories in past 7 days	Discrete
ITYP	Reuters NewsScope	Type of story (ALERT, ARTICLE, APPEND, OVERWRITE)	Categorical
GENRE	Reuters NewsScope	Was the story an automatically generated share imbalance alert?	Categorical
SEQN	Reuters NewsScope	Sequence number for the story (1 for first iteration of a story, and incremented thereafter)	Discrete
TCODE	Reuters NewsScope	Topic code for story	Categorical
RTN1D	Deutsche Bank	Excess return in day before news story (vs. Russell 3000 total return)	Continuous
RTN0D	Deutsche Bank	Excess return in day of news story (vs. Russell 3000 total return)	Continuous
VOLUME1D	Deutsche Bank	Turnover (in percent of float shares outstanding) in day before news story release	Continuous
VOLUME0D	Deutsche Bank	Turnover (in percent of float shares outstanding) in day of news story	Continuous
INTRADAY	Deutsche Bank	Flag if news was released during trading hours (1 = during trading, 0 = outside of trading)	Categorical

Source: Reuters NewsScope, Deutsche Bank

Our goal is to use the independent variables in Figure 12 to predict in turn each of the dependent variables listed in Figure 13. The dependent variables are simple; they are just the market excess returns following a news event. We consider a number of horizons from one day to 252 trading days after the event.

Figure 13: Dependent variables (only one selected at a time)

Variable	Source	Description	Type
FRTN1D	Deutsche Bank	Excess return from close on day of news event to close 1D later (vs. Russell 3000 total return)	Continuous
FRTN5D	Deutsche Bank	Excess return from close on day of news event to close 5D later (vs. Russell 3000 total return)	Continuous
FRTN10D	Deutsche Bank	Excess return from close on day of news event to close 10D later (vs. Russell 3000 total return)	Continuous
FRTN21D	Deutsche Bank	Excess return from close on day of news event to close 1M later (vs. Russell 3000 total return)	Continuous
FRTN42D	Deutsche Bank	Excess return from close on day of news event to close 2M later (vs. Russell 3000 total return)	Continuous
FRTN64D	Deutsche Bank	Excess return from close on day of news event to close 3M later (vs. Russell 3000 total return)	Continuous
FRTN252D	Deutsche Bank	Excess return from close on day of news event to close 12M later (vs. Russell 3000 total return)	Continuous

Source: Deutsche Bank

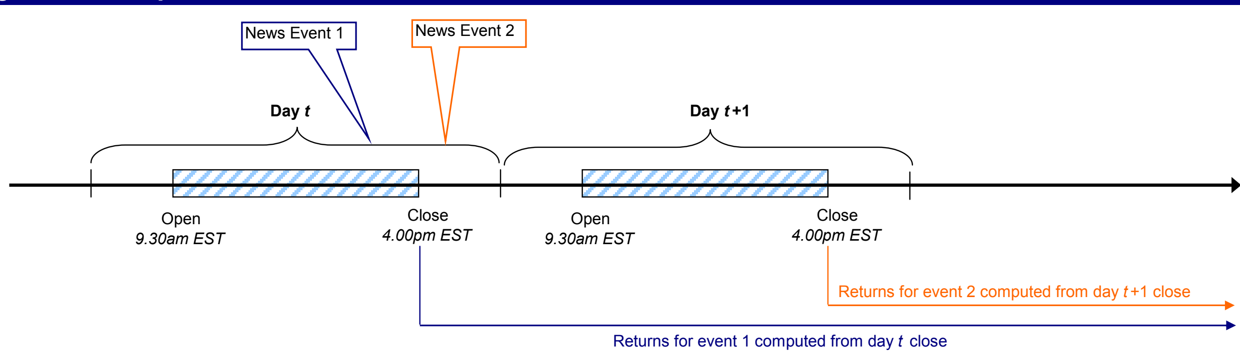
When using news sentiment data it is important not to introduce a look-ahead bias

Real-world considerations

In our study, we are primarily interested in whether there is any mid-term predictive power in news sentiment. By mid-term we mainly refer to abnormal returns from one to 21 days after a news event. We are not trying to forecast intraday returns in the seconds after the news is released; rather we are interested in whether a mid-term investor can buy or sell a news event after the fact, and still profit from any post-event drift. With this in mind, we make the conservative assumption that an investor will buy or sell a news event *at the next close price*. In practice this means that if a news event occurs during trading hours on a give day, then the investor will buy or sell the stock at that day's close price. However, if the news comes out after trading hours, then the transaction will not take place until the close on the following day.

For example, in Figure 14, News Event 1 happens during trading hours on day *t*. Therefore, we assume investors can enter or exit the stock at the close on day *t*. News Event 2 happens after the close on day *t*. Therefore, we assume any transaction occurs at the close on day *t*+1. As shown in the graphic, this is relatively conservative; a more aggressive approach would be to assume trades based on News Event 2 happen at the open on day *t*+1.

Figure 14: Example of forward return calculations



Source: Deutsche Bank

Non-linear models

Let the data speak

Traditionally quants specify a model, and then use data to either accept or reject that model…

The next step is to take the variables we selected in the previous section, and devise a model to help determine which news events are worthy of taking investment positions on, i.e. which news events are most likely to signal longer-term abnormal returns. The traditional way to approach this problem would be to follow the tried and true scientific method: first devise a hypothesis, and then build a model to either confirm or reject the theory.

…however in recent years the rise in computing power has given rise to a new data-centric approach to model building

However, the explosion in computing power and data volumes in recent years has led to an alternative approach: start with the data and see where it leads. We might call this new approach the Google approach. Instead of predetermining a model and then using the available data to test it, we start with the raw data and see where it leads. Such an approach is well suited to tackling new datasets, particularly datasets as rich as news sentiment. With so many possible variables to consider, it makes sense to start with a clean slate and see what relationships we can tease out of the data.

There is a plethora of potential techniques we might use for this task. Here we focus on three non-linear models that we think show promise: classification and regression trees, forests of classification and regression trees, and multivariate adaptive regression splines. We will call these *TREE*, *FOREST*, and *PLANET* models respectively.[3] Below we describe each type of model in more detail.

In this report we use non-linear models to hopefully better capture the complex dependencies between variables in our dataset

The reason we favor these models is because we believe news sentiment lends itself to non-linear analysis since sentiment variables rarely act linearly in predicting future returns. For example, sentiment – as determined purely from natural language processing – tells us nothing about market expectations. Suppose a negative news story comes out on BP; this will be a surprise to no one and the odds are it will lead to little market reaction. However, if a negative story comes out on a media darling – let's say Apple – then the chances of a negative price reaction are likely much higher. In other words, the impact of sentiment is conditional on other variables – perhaps abnormal volume for example. A purely linear model may struggle to capture such complex relationships between independent variables.

Classification and regression trees (*TREE*)

TREE models are like a flow-chart that allow us to classify out-of-sample data points

Classification and regression trees (*TREE*) models are the most common of the three models we consider. Indeed, *TREE* models were originally in used in stock selection models as far back as the early 2000s.[4] As the name suggests, *TREE* models are constructed by "growing" a (upside-down) tree-like classification structure. Think of it as something of a flow-chart, that can then be used to classify out-of-sample data points.

The model starts with the whole dataset and looks for the explanatory variable in the data that best splits the dependent variable into two groups. Within the groups, the dependent variable is homogenous as possible, and between groups it is as different as possible. The process is then repeated for each of the two new groups. Again, we look for dependent variables that will best subdivide the data into smaller groups.

---

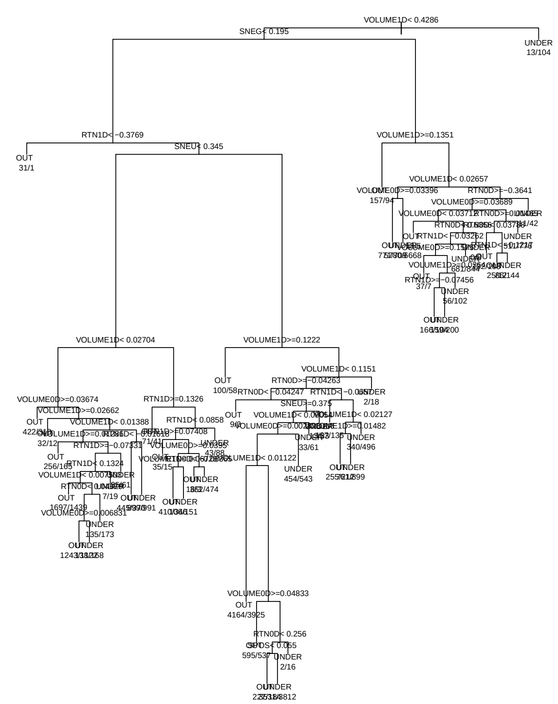
[3] These models are often referred to as CART, Random Forest, and MARS respectively. However, all three of these names are registered trademarks of Salford Systems. Therefore we follow the standard practice in the literature and use alternative names for each model.

[4] For example, Sorensen, Miller, and Ooi [2000]

TREE models can either be categorical or regression based

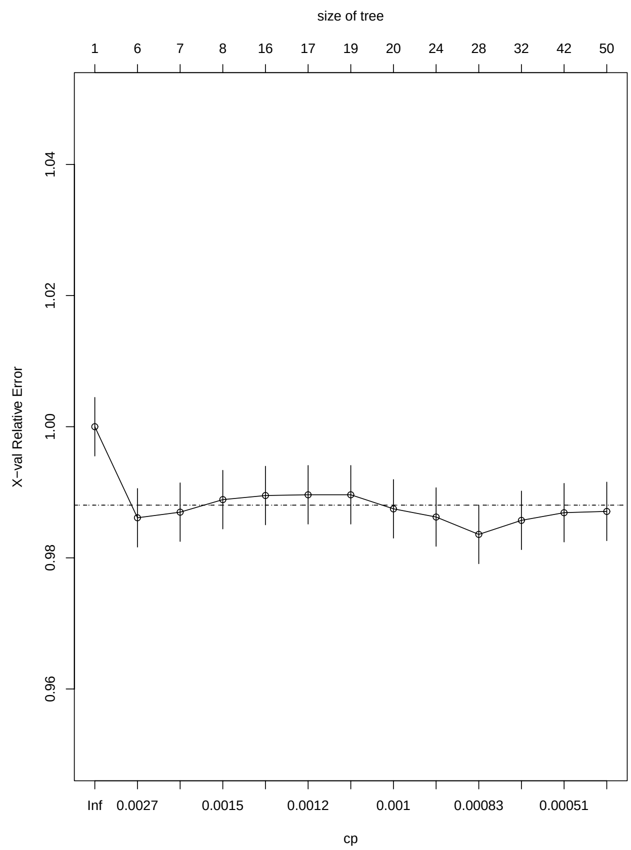
For example, let's take the simple example of a classification tree on news sentiment. In this case the dependent variable (i.e. what we want to predict) is post-event abnormal returns. *TREE* models come in two flavors: classification trees or regression trees. Classification trees are used when the dependent variable is categorical, whereas regression trees are used for a continuous dependent variable. To keep things simple, we focus in this example on a classification tree; this requires us to transform continuous returns into a categorical variable. This is easily done. We just define two categories: outperform (forward returns > 0) and underperform (forward returns < 0).[5]

Figure 15: Example of overfit *TREE*



Source: Reuters NewsScope, Deutsche Bank

Figure 16: 1-SE cutoff rule



Source: Reuters NewsScope, Deutsche Bank

Figure 15 shows the results of fitting the *TREE* model to explain the forward returns. One of the most useful features of the output is that is gives a good sense of the hierarchy of importance for explanatory variables. In this example, *VOLUME1D*, which is the turnover of the stock on the day before the news event, is the single most important variable in determining whether a stock is likely to outperform or underperform. We will come back to this point in the next section.

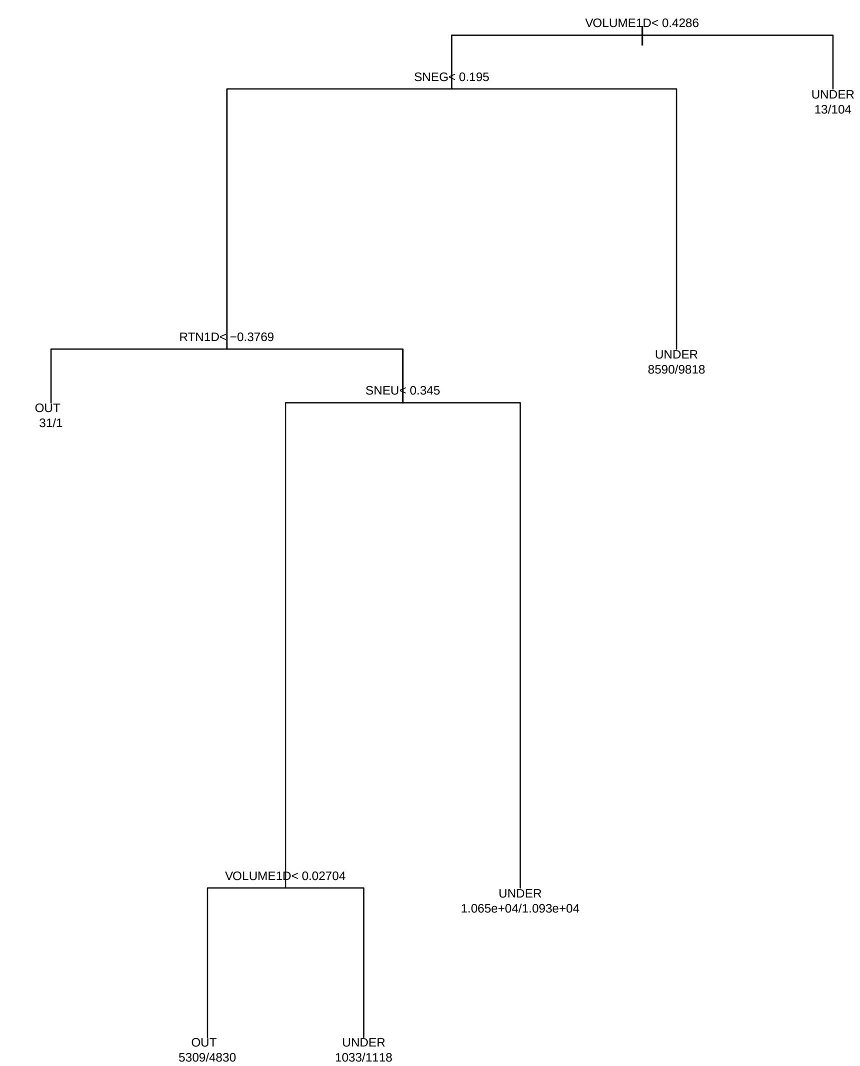
The biggest weakness of TREE models is their tendency to overfit

The tree in Figure 15 also illustrates one of the main weaknesses with *TREE* models – the tendency to overfit. Left unchecked, there is nothing to stop the tree from continuing to grow until it fits the data perfectly, i.e. the end of each branch (or node) represents a single stock. Clearly such a tree will work beautifully in-sample, but will fail completely when used to

[5] In this example, we focus on five day forward returns. Our analysis in the following sections will show that news sentiment tends to be most useful at horizons of around one week. In the following section we explore the sensitivity of our analysis to the prediction horizon.

forecast out-of-sample news events. For this reason we would never use the tree as shown above; instead we apply a technique called (naturally enough) pruning. This is a process where we cut back the tree based on how well it works out-of-sample. Figure 16 shows how we can do this with a technique called cross-validation. Essentially in cross-validation we divide the dataset up into *n* subsets, and then build a *TREE* using data from *n*-1 subsets and test it on the remaining subset. The chart shows the error (essentially a measure of how well the tree works out of sample) as a function of tree size. Initially making the tree bigger helps reduce error, since using more variables better describes the data. But the improvement quickly reaches a plateau, i.e. at this point the loss in out-of-sample predictive ability from overfitting outweighs the gain from adding more explanatory variables.

Figure 17: Example of pruned *TREE*



Source: Reuters NewsScope, Deutsche Bank

A technique called pruning can be used to reduce the chance of overfitting

Figure 17 shows the pruned tree in the previous example, after applying 10-fold cross-validation. With this simplified tree, we can start to see some interesting relationships. For example, stocks that have high turnover on the day before a news event tend to underperform on average (see the first node on the right). This is interesting because it reconciles with our previous finding using options data that stocks with heavy options trading tend to underperform on average.[6]

Forest models (*FOREST*)

A FOREST is just a large collection of TREE models

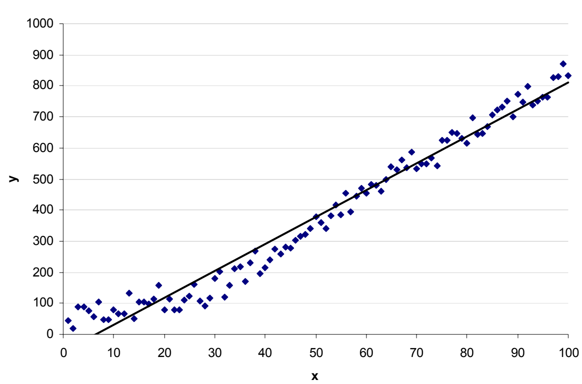
A *FOREST* model takes (or rather grows) the concept of a *TREE* to the next level. Instead of one growing one *TREE*, we grow a whole forest! The idea is to overcome the overfitting problem, which even with pruning is still the main drawback of *TREE* models. The *FOREST* model was first proposed by Breiman [2001], but as far as we know has never been used extensively in quantitative investing.

At a high level the concept is simple. Each *TREE* in a *FOREST* has a vote for how a particular stock should be classified (i.e. as a post-event outperformer or underperformer). Each *TREE* is constructed by taking a random subset of the dataset, and building a *TREE* using only that data.[7] The *TREE* is then tested out-of-sample on the remaining data, and the final vote that *TREE* gets is proportional to how well it works on the out-of-sample data. Thus a *FOREST* of, say, 1,000 *TREE*s is hopefully a more robust forecasting tool than a single lonely *TREE* in isolation.

Multivariate adaptive regression splines (*PLANET*)

The *PLANET* model is easiest to describe with an example (Figure 18 and Figure 19).

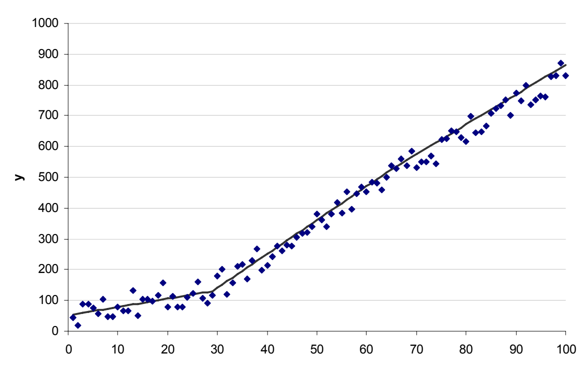
Figure 18: Hypothetical linear model



$$\hat{y} = 8.67x - 57.18$$

Source: Deutsche Bank

Figure 19: Hypothetical *PLANET* model



$$\begin{aligned}\hat{y} = 130.0 \\ &+ 11.0 \max(0, x - 29) \\ &- 2.7 \max(0, 29 - x) \\ &- 1.3 \max(0, x - 65)\end{aligned}$$

Source: Deutsche Bank

[6] Cahan, R., Y. Luo, J. Jussa, and M. Alvarez, 2010, "Signal Processing: The options issue", *Deutsche Bank Quantitative Strategy*, 12 May 2010

[7] This description is a little simplified. The construction of each *TREE* in the *FOREST* is not quite the same as an individual *TREE* model. At each split, instead of considering all explanatory variables as split candidates, only a random subset are selected for testing. This speeds up performance without significant loss of accuracy.

A PLANET model is designed to fit non-linear relationships by breaking up the problem into smaller linear sections

Suppose we have a simple two variable dataset as shown in the carts. Clearly the relationship is close to but not quite linear; there is a kink in the pattern around *x*=30. Fitting a linear model is one option, as shown in Figure 18. But can we explicitly account for the non-linearity in the data with a *PLANET* model? The answer is of course yes. A *PLANET* model breaks the variable space up into rectangular sections, and fits linear functions within those spaces. This allows the model to "kink" to approximate non-linear relationships, as shown in Figure 19.

More technically, a *PLANET* model takes the form

$$\hat{f}(x) = \sum_{i=1}^{k} c_i B_i(x)$$

where c_i is a constant and $B_i(x)$ is a so-called basis function. The basis function can either be a simple constant, or what is called a hinge function. A hinge function takes the form $\max(0, x-c)$ or $\max(0, c-x)$ where c is a constant. If we allow higher-order relationships, then a basis function could also be the product of two or more hinge functions.

To fit the model, the process is very similar to the *TREE* model. First we use a forward pass which adds pairs of basis functions. The idea is to add a pair of functions that will most reduce the residual error in the data. Like a *TREE* the forward pass leads to a seriously overfit model. The second step, or backwards pass, then goes back and deletes hinge functions based on cross-validation, similar to the *TREE* model.

What does a real *PLANET* model look like? Figure 20 shows the results for our news sentiment data, at a point in time.

Figure 20: Example of *PLANET* model

$$
\begin{aligned}
\text{UNDER} = {} & 0.73 \\
& -0.03 \max(0,\ \text{SPOS} - 0.18) \\
& -0.12 \max(0,\ 0.18 - \text{SPOS}) \\
& -0.12 \max(0,\ 0.34 - \text{SNEG}) \\
& +0.12 \max(0,\ \text{RTN1D} - 0.52) \\
& +0.76 \max(0,\ \text{-}0.52 - \text{RTN1D}) \\
& -0.21 \max(0,\ \text{RTN1D} - 0.12) \\
& -2.20 \max(0,\ \text{VOLUME1D} - 0.07) \\
& +3.4 \max(0,\ \text{VOLUME1D} - 0.28) \\
& -1.2 \max(0,\ 0.28 - \text{VOLUME1D}) \\
& +0.43 \max(0,\ 0.19 - \text{VOLUME0D}) \\
& +2.60 \max(0,\ \text{VOLUME0D} - 0.48) \\
& -11.01 \max(0,\ \text{VOLUME0D} - 0.68) \\
& +0.84 \max(0,\ \text{VOLUME0D} - 0.75)
\end{aligned}
$$

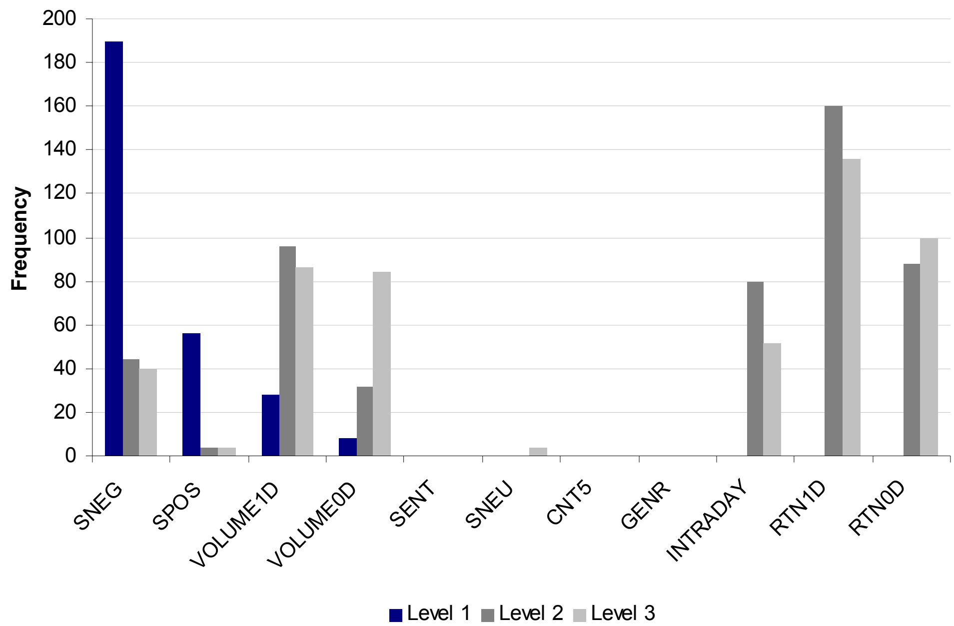
Source: Reuters NewsScope, Deutsche Bank

What really matters?

The three models we pick allow one to determine which variables are most important in explaining the dependent variable

One of the nice features of all three non-linear models is that they give a sense of which variables are most important in determining post-event outperformance or underperformance. For example, in a *TREE* model, assessing the hierarchy of explanatory power is easy; variables that appear near the top of the tree are most important in discriminating between future outperformance and underperformance. Figure 21 shows the average variable importance for the *TREE* model through time.[8]

Figure 21: *TREE* average variable importance



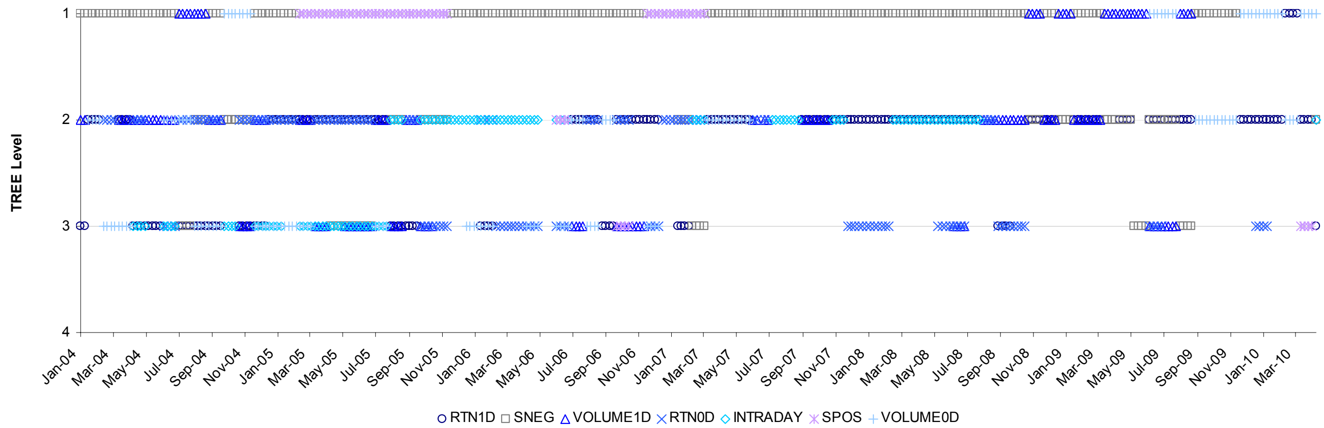
Source: Reuters NewsScope, Deutsche Bank

The sentiment variables are most important based on the TREE model

Interestingly, *SNEG*, the probability that the sentiment of the story is negative, appears most often at the first split. In other words, this is the variable that is most frequently selected to make the first split in the data. This is interesting because it suggests that using natural language processing to determine sentiment is useful in determining future returns, and indeed is more useful than the pure market variables like volume and price action. Figure 22 expands on the previous chart by showing the time-series evolution of *TREE* variable importance.

[8] We construct a new model every month from 2003-present using one year of trailing data. Thus the shape of the tree and the hierarchy of variable importance changes over time. See the backtesting section for further details.

Figure 22: *TREE* variable importance through time

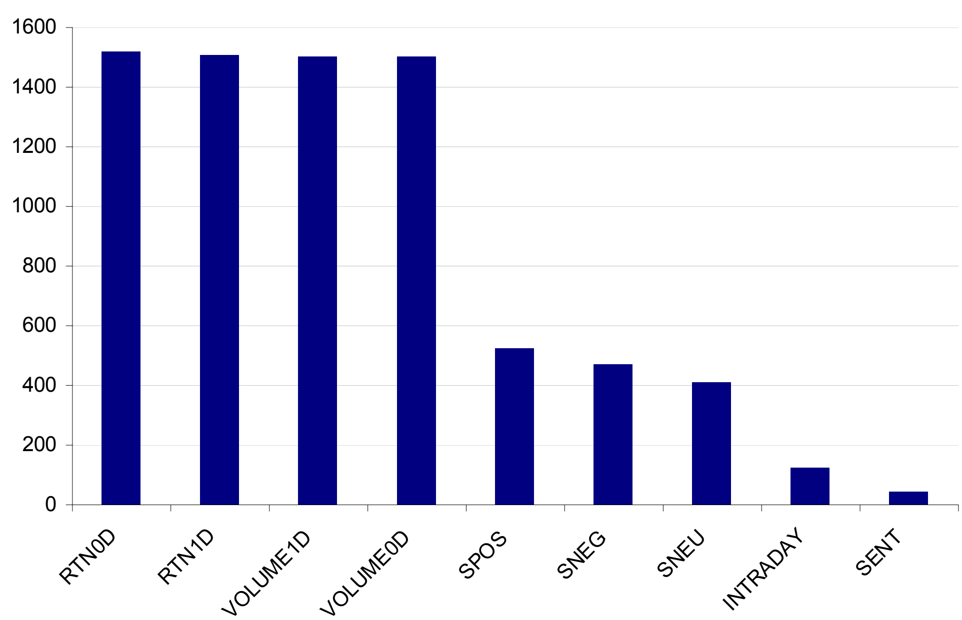


Source: Reuters NewsScope, Deutsche Bank

For the FOREST models, the market variables are most important

We can carry out the same exercise for the *FOREST* model. For this class of model, the concept of variable importance is a little more complicated. Essentially, variable importance in a *FOREST* is measured by looking at how much the error in the model increases by when a single variable is permuted while leaving all others constant. Figure 23 shows the average variable importance through time for the *FOREST* model. Interestingly, with this model the market-related variables are much more important. However, note that if one "added" together the importance for *SPOS*, *SNEG*, and *SNEU* than it would be roughly equivalent to the market variables in importance.

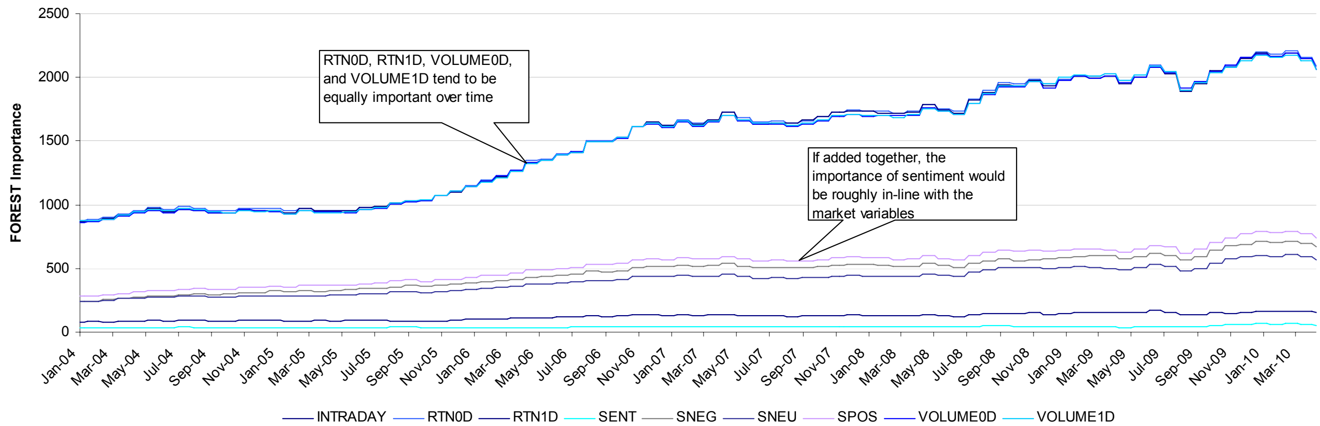
Figure 23: *FOREST* average variable importance



Source: Reuters NewsScope, Deutsche Bank

Figure 24 shows the time evolution of the previous chart. Because the *FOREST* model is more stable through time, the variable importance is also quite consistent over time.

Figure 24: *FOREST* variable importance through time

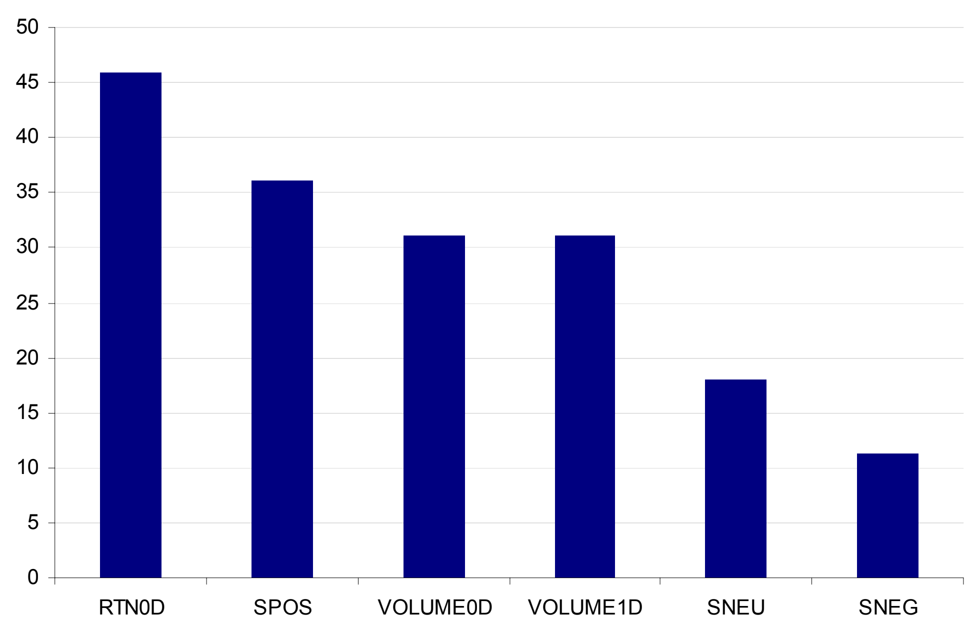


Source: Reuters NewsScope, Deutsche Bank

The PLANET models show that a mixture of market and sentiment variables is important

For *PLANET* models, there is also an importance metric that is analogous to those in the previous models. Once again we see that on average market variables and sentiment variables are both considered useful by the model in determining future performance (Figure 25).

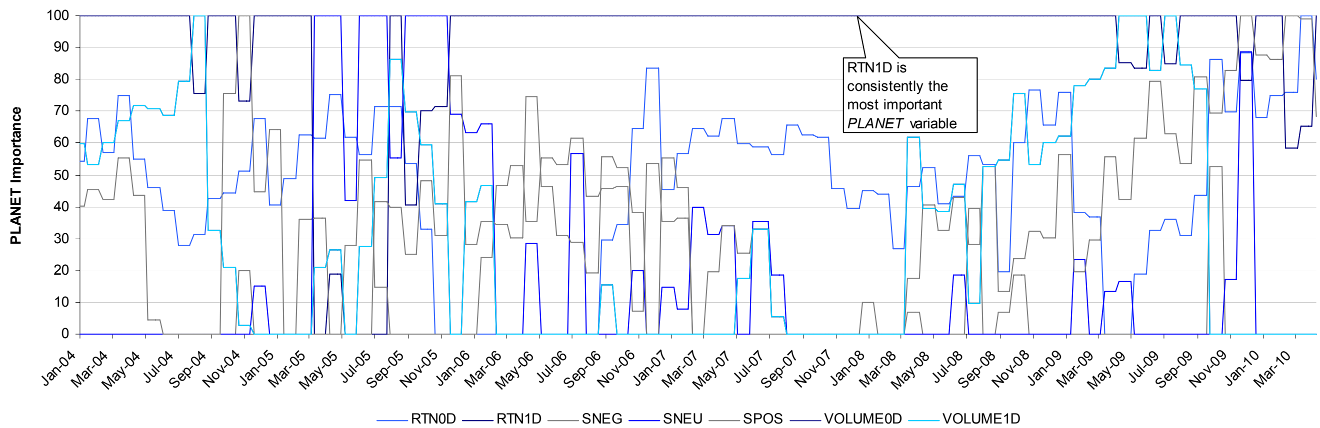
Figure 25: *PLANET* average variable importance



Source: Reuters NewsScope, Deutsche Bank

As before, we also show the time-series evolution of the *FOREST* variable importance. In the time-series, the *PLANET* results are more volatile than the *FOREST* results. This may or may not be a good thing, depending on whether the model is evolving to matching changing market conditions, or just reacting to noise. The way we answer this is to backtest strategies based on each model; see the following section.

Figure 26: *PLANET* variable importance through time



Source: Reuters NewsScope, Deutsche Bank

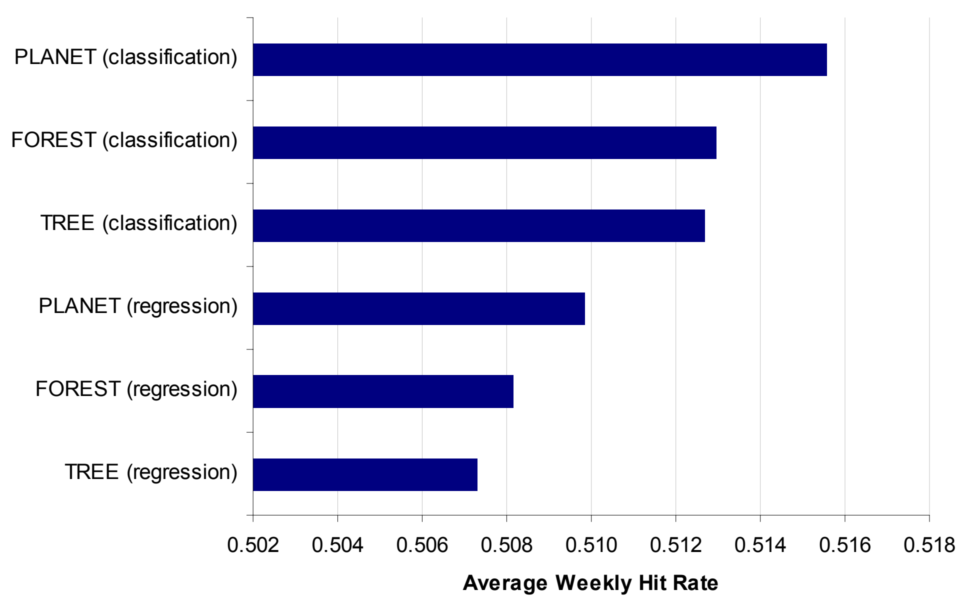
Assessing the accuracy

Assessing the hit rate of each model is important for getting a sense of how well they work out-of-sample

Of course, knowing the variable importance is irrelevant if the models have no predictive power. What we are primarily interested in is how well the models work out-of-sample in predicting which news events will generate positive or negative post-event drift.

To test this somewhat crudely, we look at a simple hit rate. Each week we construct a model for each of our model types, and then use it to try to predict which events in the following week will generate positive or negative post-event returns. Figure 27 shows the results averaged over time.

Figure 27: Average weekly hit rate



Source: Reuters NewsScope, Deutsche Bank

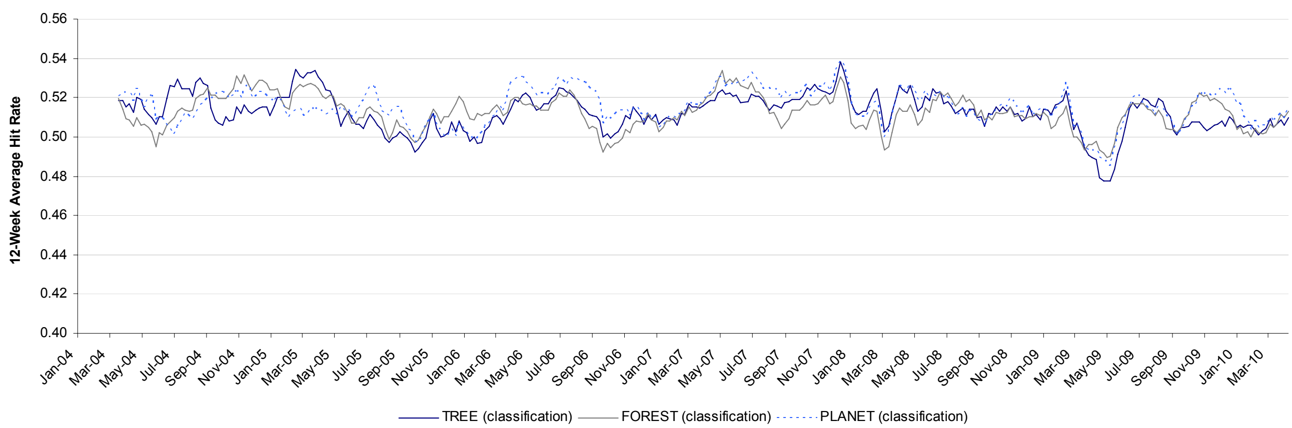
In general, hit rates are low but consistency through time is good

The hit rates are low at best, ranging from 0.505 to 0.515 depending on the model. This tells us that forecasting post-event drift is hard, even with the sophisticated models we are employing. However, as quants we know that event a slightly positive hit rate can be useful so long as we can invest in the strategy enough times, in the same way that flipping a slightly biased coin is on average almost be guaranteed to be in our favor if we flip it enough times.

In the next section we apply a more relevant statistic for assessing the efficacy of news sentiment-based investing: the backtest.

Figure 28 shows that the hit rates shown above are reasonably stable through time.

Figure 28: Hit rate for classification models through time



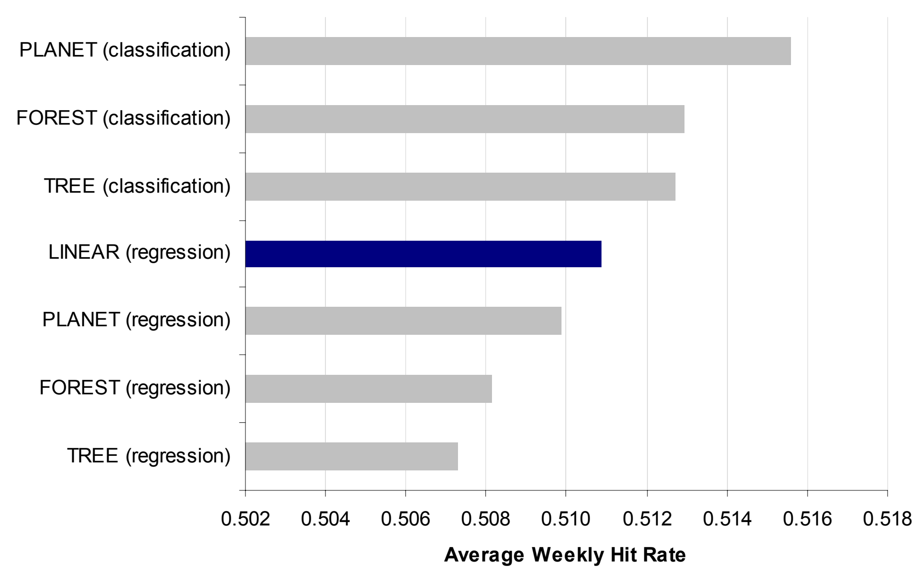
Source: Reuters NewsScope, Deutsche Bank

Could we just use a simple linear model instead?

Comparison with linear model

Before tackling the backtesting, one important question needs to be addressed: are the complex non-linear models we use actually worth the effort? To test this, we add a simple multivariate linear regression model to the mix. Figure 29 shows the hit rates again, with the linear model added in.

Figure 29: Average weekly hit rate – Linear model versus non-linear models



Source: Reuters NewsScope, Deutsche Bank

Non-linear models do seem to do a little better than a simple linear model

Interestingly the linear model finishes in the middle of the pack. It tends to do better than the non-linear regression models but worse than the classification-based non-linear models. The question of whether the roughly 1% difference in hit rate between the best non-linear model and the linear model is significant or not is best addressed by backtesting, since the accuracy statistic we most care about is the risk-adjusted returns to an actual investment strategy based on these models.

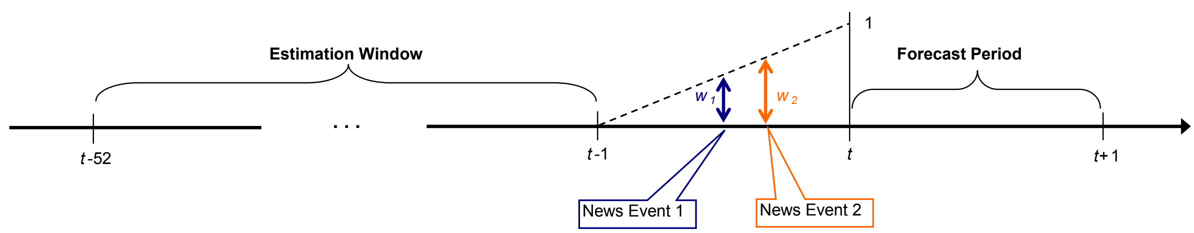
Backtesting results

Methodology

Our backtesting methodology is complicated by the asynchronous nature of news events

Our basic backtesting methodology is straightforward. At the end of each rebalancing period we build a model based on one year of trailing data at that point in time. We then use that model to forecast the likelihood that recent news events will lead to positive or negative market-adjusted returns in the coming period. Figure 30 shows more details of the exact backtesting setup.

Figure 30: Backtesting scheme



Source: Deutsche Bank

We focus mainly on a weekly rebalancing frequency because in general we find the predictive power of news sentiment to be relatively short-lived. We come back to some sensitivity analysis on this point later in this report.

We use a time-weighting scheme to put more emphasis on recent news events

Suppose we are at week *t*. We want to construct a model to estimate the expected returns of events that have happened recently, say in the last week (i.e. between *t*-1 and *t*). Because events in this window are still relatively recent, we are interested in forecasting their post-event drift. However, events that happened almost a week ago will have already realized most of their post-event returns; therefore we apply a simple linear time-weighting scheme to put more weight on recent news events. The recent news events are the ones where we have more time to benefit from any predictability in post announcement drift. For example, in the scenario above, News Event 2 is more recent and so is given a larger weight, w_2, than News event 1.

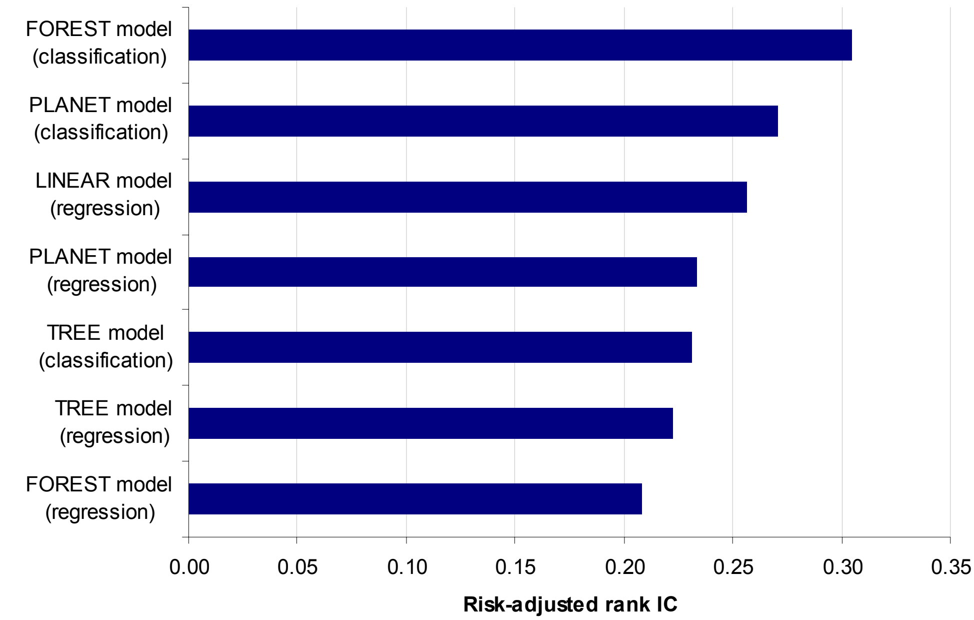
In summary, at each point in time *t*, we estimate our model in the trailing year *t*-52 to *t*-1. Then we use that model to forecast expected excess returns for events occurring in the last week. These forecasts are time-weighted to put more emphasis on the very latest events.

The bottom line

Our backtests show that there is predictive power in news sentiment

Overall, we find that our news sentiment models do add value. Figure 31 shows the average weekly risk-adjusted rank information coefficients (IC) for the various models. This is simply measured as the average weekly rank IC divided by the time-series standard deviation of IC.

Figure 31: Comparison of risk-adjusted rank ICs for various models

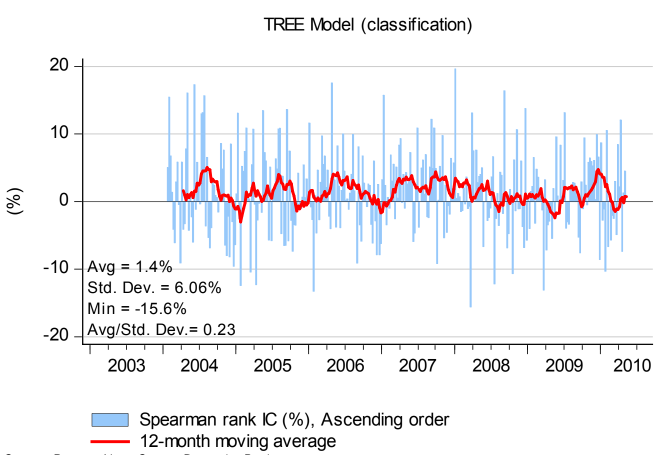


Source: Reuters NewsScope, Deutsche Bank

The best models are quite consistent over time

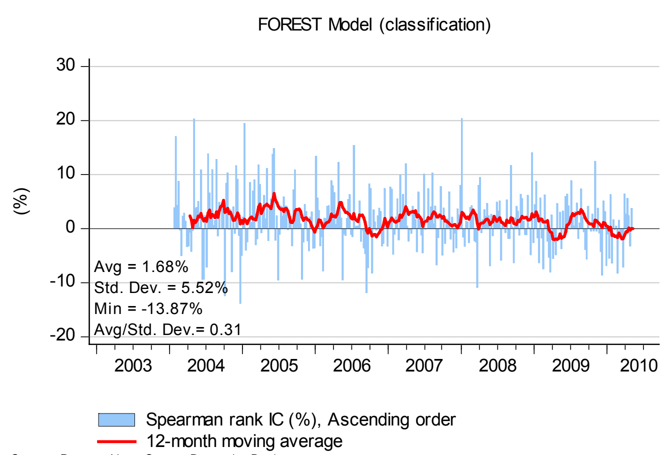
The next set of charts shows the time-series characteristics of each backtest. Again, we are looking at weekly rank ICs. A pleasing feature of most the models, but particularly the *FOREST* and *PLANET* models, is the consistency of performance through time. The 12-week rolling average tends to stay above zero for almost the entire backtest, which indicates the factor is adding consistent value over time.

Figure 32: Rank IC – *TREE* model (classification)



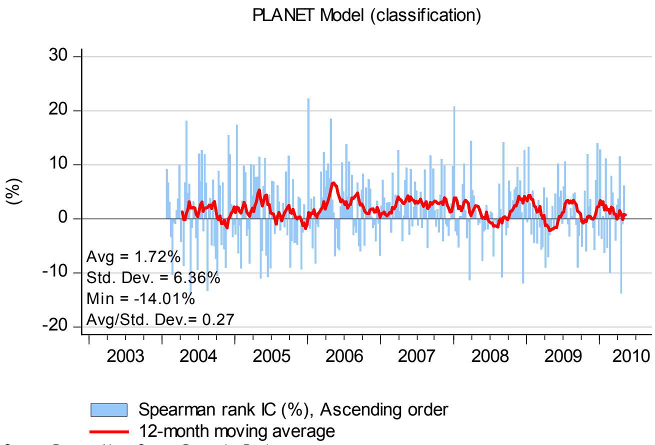
Source: Reuters NewsScope, Deutsche Bank

Figure 33: Rank IC – *FOREST* model (classification)



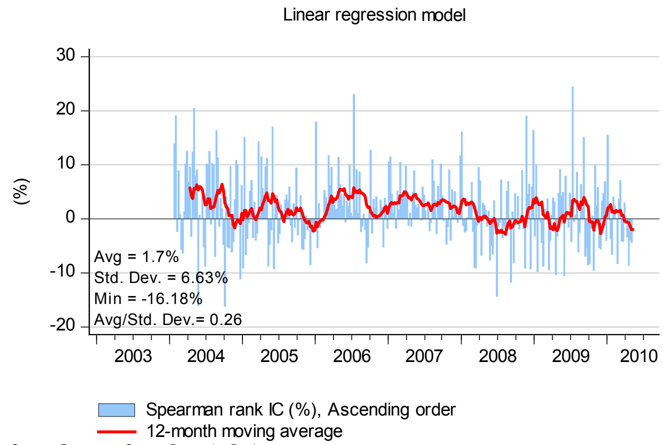
Source: Reuters NewsScope, Deutsche Bank

Figure 34: Rank IC – *PLANET* model (classification)



Source: Reuters NewsScope, Deutsche Bank

Figure 35: Rank IC – *LINEAR* model (regression)



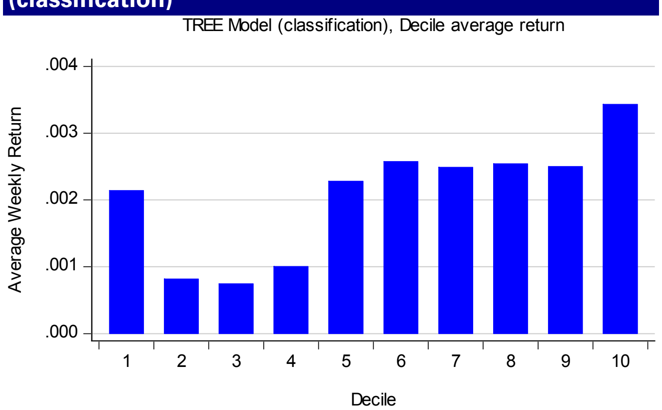
Source: ReuterewsScope, Deutsche Bank

The simpler models – TREE and linear – don't do as

Looking at return space instead of IC space, the charts below show the average decile returns for each model over the backtesting period. This is where we see a clear advantage for the more complex non-linear models. Both the *TREE* and linear model do a poor job at picking stocks for the extreme short (or underweight) decile. In both cases the Decile 1 stocks tend to outperform rather than underperform as we would like.

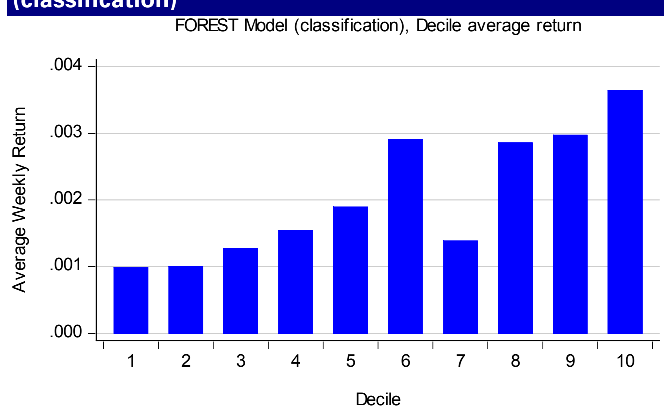
In contrast, the *FOREST* and *PLANET* models do a better job at predicting returns over the whole cross-section of the universe. In both cases, the average returns are reasonably monotonic from Decile 1 through to Decile 10.

Figure 36: Average decile returns – *TREE* model (classification)



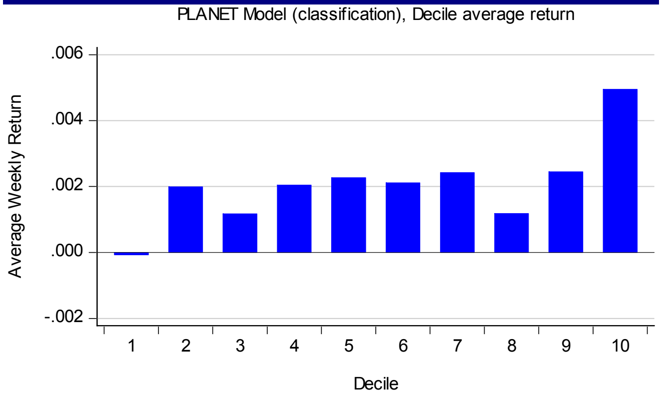
Source: Reuters NewsScope, Deutsche Bank

Figure 37: Average decile returns – *FOREST* model (classification)



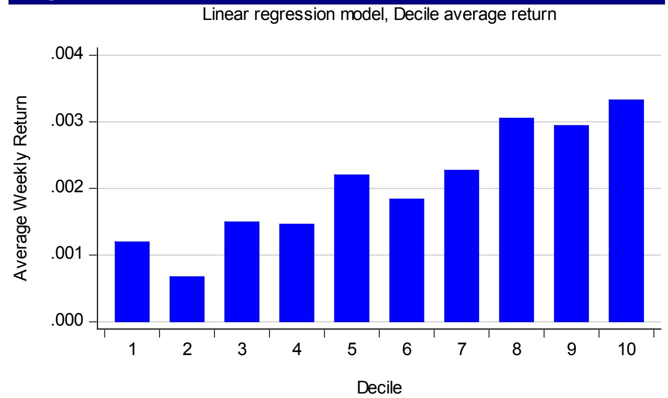
Source: Reuters NewsScope, Deutsche Bank

Figure 38: Average decile returns – *PLANET* model (classification)



Source: Reuters NewsScope, Deutsche Bank

Figure 39: Average decile returns – *LINEAR* model (regression)

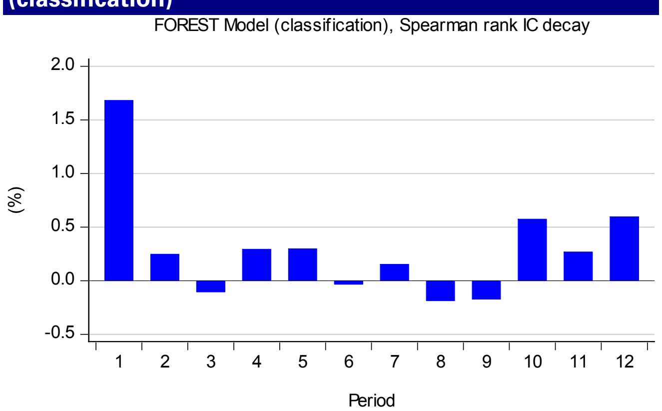


Source: Reuters NewsScope, Deutsche Bank

Turnover for the news sentiment based factors is relatively high – most information decays away in the first week

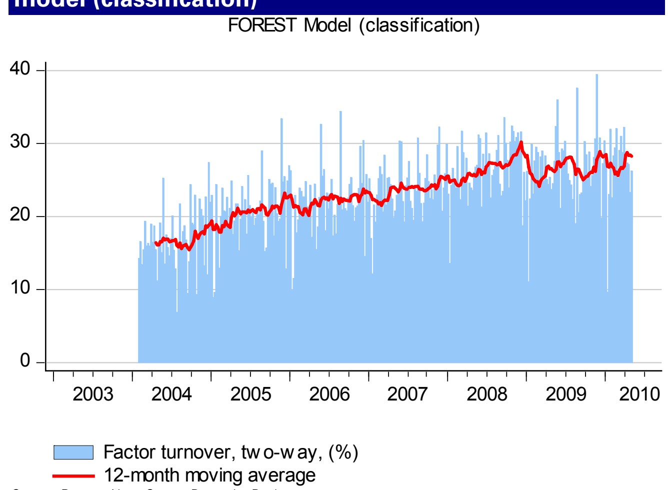
As mentioned previously, for many investors the downside of news sentiment will be relatively high turnover of the strategies. Figure 40 shows the weekly information decay for the *FOREST* model. The chart shows that most of the predictive power in the news sentiment signal has vanished after the first week. This of course leads to high turnover. Figure 41 shows the weekly two-way turnover for the *FOREST* model, which averages around 20-30% per week, or around 1500% per annum. Clearly this is too high for many investors to use this factor in its own right. However, a more important question is the incremental turnover that is added to an alpha model if we include this factor. We will address this in more detail in the last section of this report.

Figure 40: Rank IC decay – *FOREST* model (classification)



Source: Reuters NewsScope, Deutsche Bank

Figure 41: Weekly decile portfolio turnover – *FOREST* model (classification)

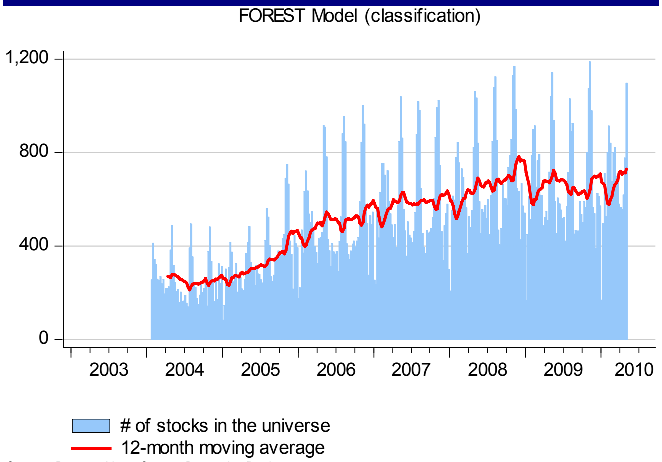


Source: Reuters NewsScope, Deutsche Bank

What happens if we include all stocks in the backtesting universe, instead of just those with news? The results are still reasonably strong

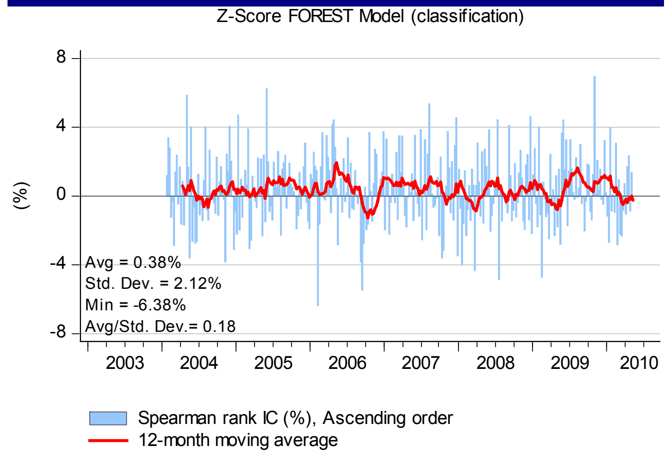
Another important point to note is that our results to date have focused on the universe of stocks that have news data. Clearly this is only a subset of the universe, as shown in Figure 42. On any given week, only around one third of the universe (Russell 3000) has at least one news story. What happens if we run the backtesting over the whole universe, and just set stocks with no news to have a score of zero (i.e. fall in the middle of cross-sectional score distribution)? Figure 43 gives the answer. As we would expect, the average IC is reduced somewhat, but the model does still add value, and still appears to have reasonably consistent positive performance through time.

Figure 42: Number of stocks with data – *FOREST* model (classification)



Source: Reuters NewsScope, Deutsche Bank

Figure 43: Rank IC – *FOREST* model (classification) when all stocks are included in backtest



Source: Reuters NewsScope, Deutsche Bank

Importance of quantifying sentiment

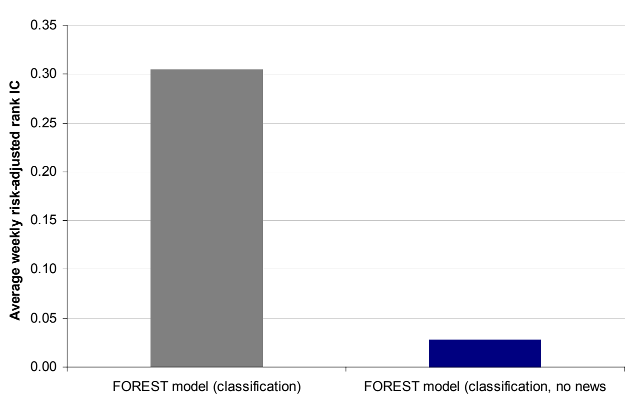
We test whether including sentiment is useful, or whether just knowing a new event occurred is enough

Another question we can legitimately ask is whether the news sentiment data is adding any value above and beyond the market variables like volume and stock returns. In other words, would it be enough just to know that a news event occurred, rather than going through the additional trouble of quantifying sentiment using complex natural language processing algorithms? After all, when we looked at the variable importance in the previous section, we saw that market variables are as important, if not more important, than the sentiment variables.

It turns out that quantifying sentiment is very important – without it there is little alpha

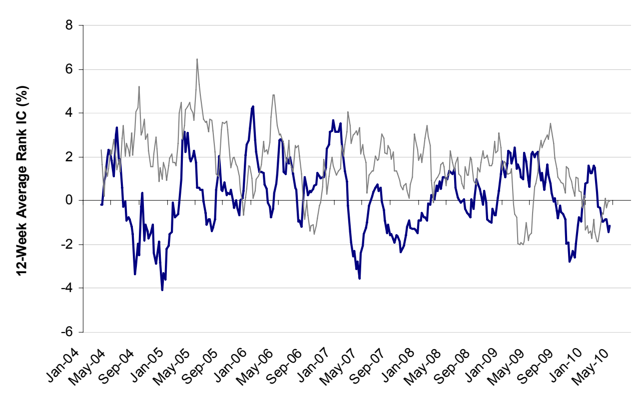
Figure 44 shows the answer is a resounding no. To generate this chart, we create another *FOREST* model, only this time we only use market variables in our set of explanatory variables. The average rank IC of this new model is close to zero, which shows that the market variables in isolation are not enough to even come close to predicting future post-event performance. It is only when adding exogenous information – in this case quantified textual data – that we can produce a reasonable forecast. Figure 46 shows the time-series performance of the model with and without news sentiment variables.

Figure 44: Comparison of risk-adjusted IC, with and without news sentiment variables



Source: Reuters NewsScope, Deutsche Bank

Figure 45: Rank IC time-series for *FOREST* model with and without news sentiment variables



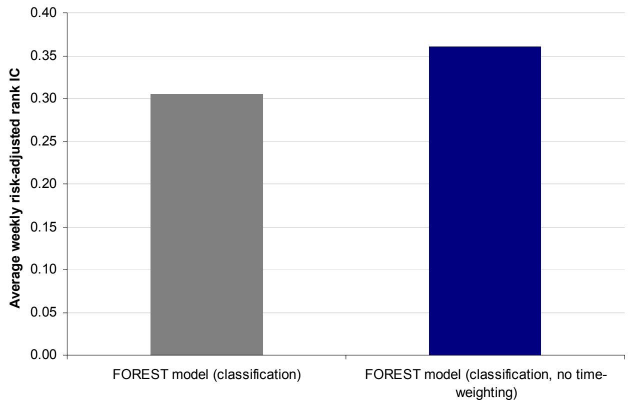
Source: Reuters NewsScope, Deutsche Bank

Sensitivity to time-weighting

It turn out the news factors are not sensitive to the time-weighting scheme used

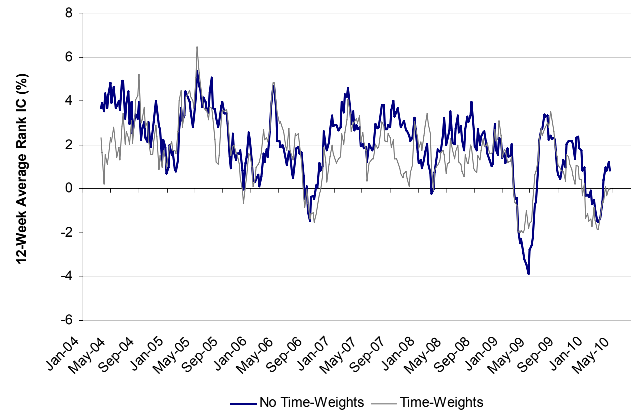
Another important question is whether our models are robust to the time-weighting scheme we use. In Figure 46 we show the average risk-adjusted rank IC for the *FOREST* model with and without the time-weighting component. Interestingly the model actually does slightly better without the time-weighting scheme. In terms of time-series performance though, the two models track tightly together (Figure 47).

Figure 46: Comparison of risk-adjusted IC, with and without time-weighting



Source: Reuters NewsScope, Deutsche Bank

Figure 47: Rank IC time-series for *FOREST* model with and without time weighting



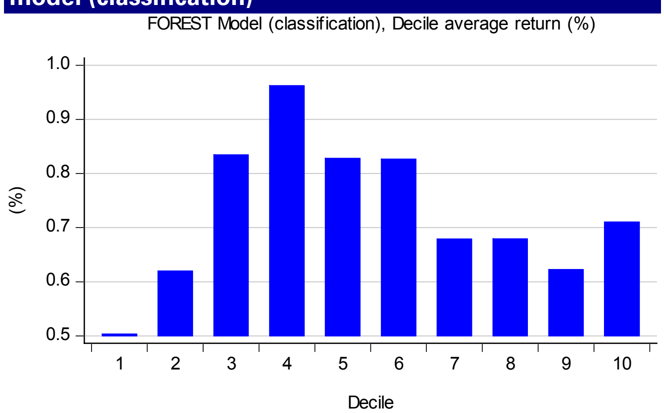
Source: Reuters NewsScope, Deutsche Bank

Sensitivity to rebalancing frequency

We find that using news sentiment at a monthly rebalancing frequency adds little value

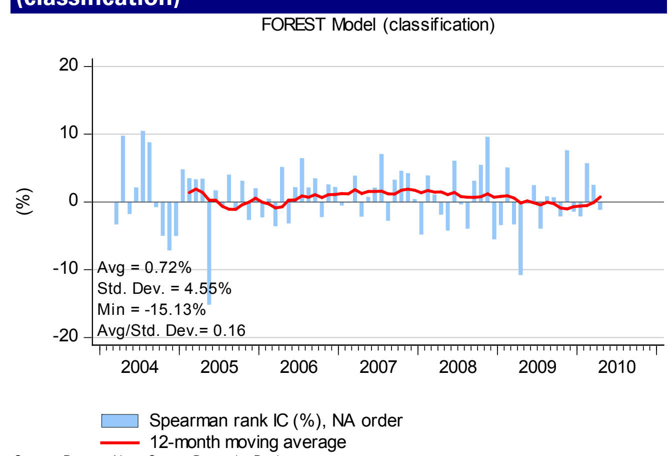
As we have seen, the information decay of our news sentiment models is fairly rapid. Here we test whether there is any alpha to be had at a monthly rebalancing frequency. This does not appear to be the case. Figure 48 shows the average decile returns, which do not show any attractive monotonicity. Similarly, the rank IC through time is not very compelling (Figure 49).

Figure 48: Average monthly decile returns – *FOREST* model (classification)



Source: Reuters NewsScope, Deutsche Bank

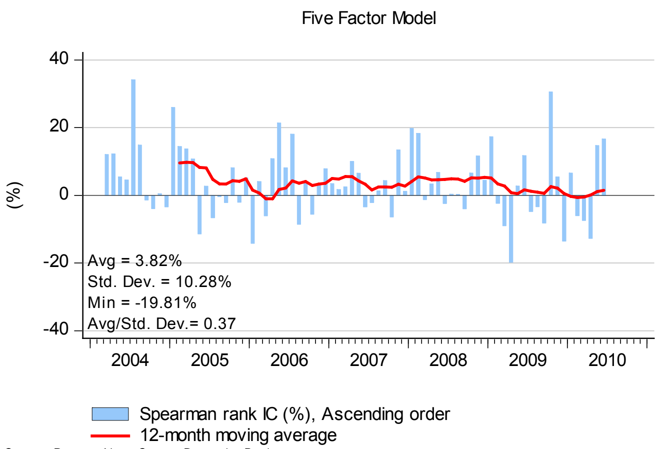
Figure 49: Monthly rank IC – *FOREST* model (classification)



Source: Reuters NewsScope, Deutsche Bank

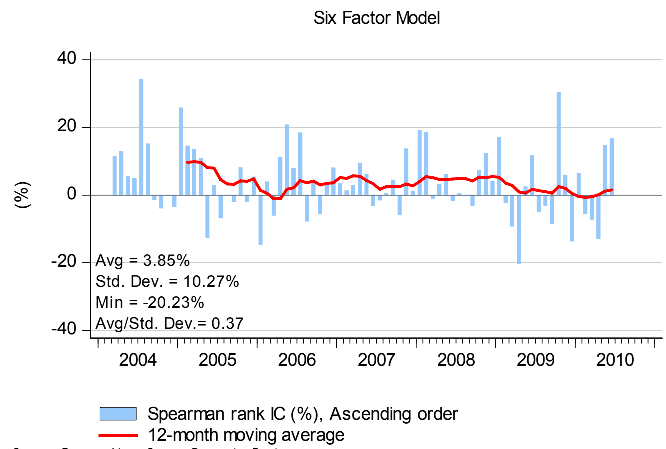
As a further test, we take a generic five factor alpha model, and add the *FOREST* model as an additional factor to that model. We rebalance both models monthly. The charts below show that adding news sentiment to the model adds zero value – the risk-adjusted information coefficient remain exactly the same (0.37) with and without news sentiment.

Figure 50: Rank IC – Five factor alpha model



Source: Reuters NewsScope, Deutsche Bank

Figure 51: Rank IC – Six factor alpha model



Source: Reuters NewsScope, Deutsche Bank

Factor correlations

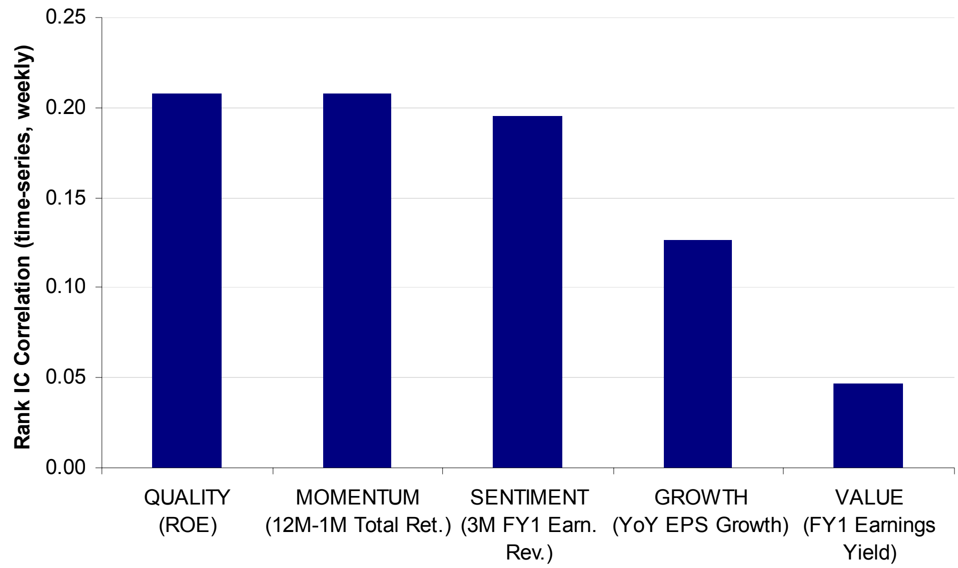
One of the most important questions is whether the news factor is correlated with existing quant factors

Returning to our weekly rebalancing, perhaps the most important question to ask for any new factor is whether it is a genuinely new alpha source, or whether it is just another way to express an existing factor. One way to test this is to look at the correlation with existing factors. Figure 52 shows the correlation of the weekly time-series of ICs for the *FOREST* model with a number of common quant factors which were selected to broadly represent a range of common quant styles.

We find correlations with common factors to be very low

The results are very good. Overall the news sentiment model has a very moderate correlation with the existing factors – typically under 0.2. This of course is a great result because it suggests we can potentially add news sentiment to a quant alpha model and get some diversification benefits.

Figure 52: Time-series factor performance correlation



Source: Reuters NewsScope, Deutsche Bank

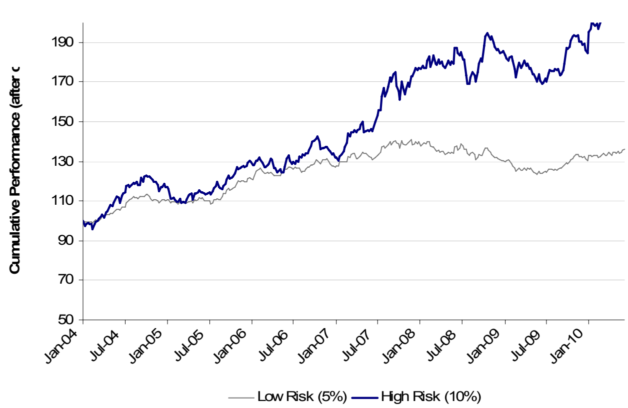
Real-world simulation

Our final test is to run a real-world portfolio simulation with realistic constraints, transaction costs, and risk controls

All our results so far have been pre-transaction costs. But clearly, given this is a high turnover signal, transaction costs are a significant consideration. To test the efficacy of the strategy after transaction costs, we perform a more real-world portfolio simulation. To do this we use the Axioma portfolio optimizer and construct an optimized portfolio based on the news sentiment signal. We rebalance this portfolio weekly, and include constraints to neutralize for beta, size, and sector exposures. We test two versions of our strategy, a low-risk/low-turnover strategy where we target an active risk of 5% p.a. and 700% two-way turnover p.a., and a high-risk/high-turnover strategy targeting 10% risk and 1500% turnover. For all our backtests, we assume transaction costs of 20bps one-way (i.e. this is charged twice per rebalance).

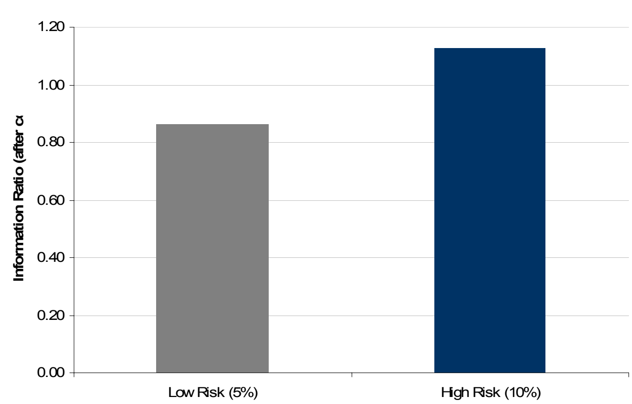
Figure 53 shows the cumulative performance (after costs) for each strategy, and Figure 54 shows the information ratio (after costs) over the backtest period.

Figure 53: Cumulative performance (after costs) for optimized news sentiment backtests



Source: Axioma, Compustat, IBES, S&P, Russell, Reuters NewsScope, Deutsche Bank

Figure 54: Information ratio (annualized, after costs) for optimized news sentiment backtests



Source: Axioma, Compustat, IBES, S&P, Russell, Reuters NewsScope, Deutsche Bank

We find the factor can add value after costs, even at a moderate turnover

For both the low-risk and high-risk strategy, the signal does appear to add value above and beyond transaction costs. Another important point is that for the low-risk strategy, the turnover of 786% p.a. two-way (Figure 55) is not outside the realm of many quant portfolios, which gives some comfort that there is scope to harvest some of the alpha in textual information even for lower-frequency quant investors.

Figure 55: Performance statistics for optimized news sentiment backtests

Model	Return (annualized, after costs)	Volatility (annualized)	Information Ratio (annualized, after costs)	Turnover (annualized, two-way)
Low Risk (5%)	5.0%	5.8%	0.86	786%
High Risk (10%)	12.4%	11.0%	1.13	1560%

Source: Axioma, Compustat, IBES, S&P, Russell, Reuters NewsScope, Deutsche Bank

References

Akbras, F., E. Kocatulum, and S. Sorescu, 2008, "Mispricing following public news: Overreaction for losers, underreaction for winners", SSRN, available at http://ssrn.com/abstract=1107690

Barber, B. and T. Odean, 2007, "All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors", *Review of Financial Studies*, Volume 21, Number 2

Breiman, L., 2001, "Random forests", *Machine Learning*, Volume 45, Number 1

Cahan, R., Y. Luo, J. Jussa, and M. Alvarez, 2010, "Signal Processing: The options issue", *Deutsche Bank Quantitative Strategy*, 12 May 2010

Da, Z., J. Engelberg, and P. Gao, 2009, "In search of attention", SSRN, available at http://ssrn.com/abstract=1364209

Fang, L. and J. Peress, 2009, "Media coverage and the cross-section of stock returns", *Journal of Finance*, Volume 63 Number 1

Friedman, J. H., 1991, "Multivariate adaptive regression splines", *Annals of Statistics*, Volume 19, Issue 1

Hong, H. and J. Stein, 1999, "A unified theory of underreaction, momentum trading and overreaction in asset markets", *Journal of Finance*, Volume 54, Number 6

Hou, K., L. Peng, and P. Xiong, 2009, "A tale of two anomalies: The implications of investor attention for price and earnings momentum", SSRN, available at http://ssrn.com/abstract=976394

Luo, Y., R. Cahan, J. Jussa, and M. Alvarez, 2010, "Signal Processing: Industry-specific factors", *Deutsche Bank Quantitative Strategy*, 8 June 2010

Sorensen, R., K. Miller, and C. Ooi, 2000, "The decision tree approach to stock selection: An evolving tree model performs the best", *Journal of Portfolio Management*, Volume 27, Number 1

Tetlock, P., 2009, "Does public news resolve asymmetric information?", SSRN, available at http://ssrn.com/abstract=1303612

Factor performance review

In this section we present our standard factor performance tables

Every month, we review the performance of about 80 factors from our factor library. Please note that this is only a tiny fraction of our factor library, which includes over 1,200 factors for the US market. We choose these 80 factors to provide a balanced view for each broad factor category.

We measure factor performance in five standard analyses: long/short hedged portfolio, Pearson information coefficient, Spearman rank IC, sector-neutral IC, and risk-adjusted IC. For simplicity, we only present Spearman rank IC in this report.

Due to space limitation, we will only present the results for the broad investable universe, two size (Russell 1000 and Russell 2000), two style (Russell 3000 Value and Russell 3000 Growth) universes, and 10 GICS sectors on a monthly basis. However, we perform factor backtesting for more sub-universes on a daily basis, e.g., S&P index family, GICS industry groups, etc. We plan to publish a factor performance review on our research website in the next few months. In the meantime, please contact us for customized factor backtesting.

The tables in the next few pages reveal a large amount of useful information. Not only do we summarize recent factor performance (last month, last year, last three years), but also provide factor distribution (standard deviation, skewness, kurtosis, percent positive, performance in up and down markets), along with factor turnover (factor score serial correlation).

Figure 56: US factor performance, Spearman rank IC

Factor Name	Direction[1]	Current # of Stocks	Average (%) Last M	12M Avg	3Y Avg	Since Inception Avg	Std Dev	Max	Min	p-value[2]	# of Months	Avg # of Stocks	%Positive	Avg in Up Mkt (%)	Avg in Dn Mkt (%)	Serial Corr (%)[3]
1. Value																
1 Dividend yield, trailing 12M	Ascending	2,932	10.60	3.03	0.80	2.83	14.90	42.71	(32.80)	0.00	270	2,805	54.07	(2.39)	11.85	99.31
2 Expected dividend yield	Ascending	2,932	11.00	2.38	1.07	3.08	15.25	44.00	(33.21)	0.00	270	2,805	54.07	(2.36)	12.48	99.31
3 Price-to-operating EPS, trailing 12M, Basic	Descending	2,209	7.61	2.62	0.76	3.15	11.05	31.00	(30.93)	0.00	185	2,384	58.92	1.05	6.77	95.05
4 Operating earnings yield, trailing 12M, Basic	Ascending	2,917	11.41	1.48	1.88	4.88	13.85	46.30	(33.54)	0.00	185	2,884	58.92	(0.28)	13.77	96.17
5 Earnings yield, forecast FY1 mean	Ascending	2,805	8.05	0.99	0.52	4.43	12.78	47.86	(34.28)	0.00	270	2,511	61.48	0.88	10.56	94.94
6 Earnings yield, forecast FY2 mean	Ascending	2,789	4.24	2.56	0.49	4.04	12.27	45.81	(33.97)	0.00	270	2,403	64.07	1.60	8.26	94.16
7 Earnings yield x IBES 5Y growth	Ascending	1,718	(0.75)	0.70	(0.41)	1.87	10.44	40.84	(27.11)	0.02	185	1,975	61.08	4.08	(1.94)	93.47
8 Sector-rel Operating earnings yield, trailing 12M, Basic	Ascending	2,917	10.98	1.15	1.79	4.11	8.75	28.31	(15.71)	0.00	185	2,884	66.49	1.18	9.16	95.66
9 Hist-rel Operating earnings yield, trailing 12M, Basic	Ascending	2,511	10.96	(1.06)	0.65	1.17	7.40	18.91	(17.10)	0.05	150	2,359	54.67	0.71	1.85	93.20
10 Operating cash flow yield (income stmt def)	Ascending	2,919	5.49	0.77	0.57	0.95	10.43	39.73	(31.83)	0.14	270	2,758	49.63	(2.27)	6.51	96.10
11 Cash flow yield, FY1 mean	Ascending	1,496	0.28	2.81	0.38	1.58	13.97	35.25	(46.90)	0.12	192	817	54.69	0.77	2.99	96.30
12 Free cash flow yield	Ascending	2,897	(11.66)	2.35	1.35	4.36	8.12	32.21	(19.15)	0.00	233	2,503	71.67	1.81	8.87	94.16
13 Price-to-sales, trailing 12M	Ascending	2,892	22.15	(4.49)	0.54	1.43	10.29	27.13	(40.36)	0.02	270	2,729	58.15	1.60	1.14	99.23
14 Price-to-book	Ascending	2,838	3.25	(3.16)	1.07	2.96	9.84	25.21	(34.91)	0.00	270	2,710	66.30	3.48	2.06	98.20
15 EBITDA/EV	Ascending	2,600	5.66	0.65	0.68	0.92	9.63	35.45	(27.96)	0.12	270	2,411	50.37	(1.86)	5.72	95.72
16 Price-to-book adj for ROE, sector adj	Ascending	2,717	8.09	(2.53)	0.49	2.00	8.40	21.68	(32.67)	0.00	270	2,430	62.96	1.63	2.63	96.05
2. Growth																
17 Hist 5Y operating EPS growth	Descending	2,867	18.66	(0.82)	1.25	0.47	7.31	20.14	(20.92)	0.40	178	2,640	51.69	(1.20)	3.15	97.10
18 Hist 5Y operating EPS acceleration	Ascending	2,867	4.44	(2.64)	0.75	1.35	6.28	13.86	(17.11)	0.00	178	2,640	62.36	0.54	2.66	94.34
19 IBES 5Y EPS growth	Ascending	1,913	20.94	(1.35)	0.78	0.58	8.77	22.81	(30.66)	0.28	270	1,881	53.70	2.10	(2.05)	98.20
20 IBES 5Y EPS growth/stability	Ascending	1,913	21.38	(1.86)	1.38	1.07	8.00	21.38	(20.62)	0.03	270	1,881	55.93	1.04	1.10	98.57
21 IBES LTG EPS mean	Descending	2,107	(2.57)	2.00	(0.16)	2.00	16.49	52.36	(37.15)	0.05	270	2,156	48.89	(3.70)	11.84	98.10
22 IBES FY2 mean DPS growth	Ascending	1,856	12.25	(1.07)	1.39	0.74	8.75	23.32	(20.96)	0.41	97	1,318	51.55	(3.05)	7.17	87.67
23 IBES FY1 mean EPS growth	Ascending	2,093	(9.57)	(3.70)	(0.99)	0.75	8.43	20.47	(29.33)	0.14	270	2,140	58.52	2.07	(1.53)	88.65
24 Year-over-year quarterly EPS growth	Ascending	2,916	(6.96)	(2.15)	0.59	2.28	7.29	24.49	(21.36)	0.00	185	2,894	67.03	2.14	2.52	81.22
25 IBES FY1 mean CFPS growth	Descending	923	(7.76)	0.93	1.22	0.39	11.04	41.21	(26.95)	0.67	142	479	51.41	(0.22)	1.45	92.62
26 IBES SUE, amortized	Ascending	2,571	0.07	(1.44)	(0.10)	1.65	6.36	18.82	(15.09)	0.00	230	2,176	60.87	2.51	0.17	73.74
3. Price momentum and reversal																
27 Total return, 1D	Descending	2,931	(7.12)	2.10	2.09	5.12	7.12	34.11	(15.21)	0.00	270	2,760	78.15	5.08	5.19	1.63
28 Total return, 21D (1M)	Descending	2,931	(18.94)	2.49	(0.26)	2.16	10.75	41.99	(27.58)	0.00	270	2,754	60.74	3.76	(0.60)	0.10
29 Maximum daily return in last 1M (lottery factor)	Descending	2,929	22.21	2.20	4.32	5.13	14.94	55.54	(38.44)	0.00	270	2,636	64.81	(0.84)	15.44	52.14
30 21D volatility of volume/price	Descending	2,929	(6.30)	2.18	0.88	0.25	6.93	17.82	(24.82)	0.55	270	2,636	50.74	1.16	(1.31)	56.28
31 Total return, 252D (12M)	Ascending	2,662	(12.78)	(3.93)	(0.26)	2.40	14.02	38.89	(54.01)	0.01	270	2,466	59.26	0.97	4.88	89.29
32 12M-1M total return	Ascending	2,662	(16.79)	(3.12)	0.24	3.37	13.15	37.39	(47.61)	0.00	270	2,466	64.81	2.42	5.01	87.70
33 Price-to-52 week high	Ascending	2,888	12.27	(5.01)	0.76	3.14	16.42	48.67	(58.45)	0.00	270	2,694	61.85	(2.23)	12.41	82.93
34 Total return, 1260D (60M)	Ascending	2,082	25.29	(4.27)	0.44	0.20	11.62	26.31	(36.15)	0.81	198	1,354	52.02	(0.54)	1.45	97.21

Note

1 Direction indicates how the factor scores are sorted. Ascending order means higher factor scores are likely to be associated with higher subsequent stock returns, and vice versa for descending order.
2 P-value indicates the statistical significance of a factor's performance. A smaller p-value suggests that it is more likely the factor's performance is different from zero.
3 This is the autocorrelation of a factor's scores over time. Higher serial correlation is likely to have lower portfolio turnover based on the factor.

Source: Compustat, IBES, Russell, S&P, Thomson Reuters, and Deutsche Bank

Source: Compustat, IBES, S&P, Russell, Deutsche Bank

Figure 56: US factor performance, Spearman rank IC (cont'd)

Factor Name	Direction[1]	Current # of Stocks	Average (%) Last M	Average (%) 12M Avg	Average (%) 3Y Avg	Since Inception Avg	Std Dev	Max	Min	p-value[2]	# of Months	Avg # of Stocks	%Positive	Avg in Up Mkt (%)	Avg in Dn Mkt (%)	Serial Corr (%)[3]
4. Sentiment																
35 IBES LTG Mean EPS Revision, 3M	Ascending	1,926	(4.51)	(0.38)	(0.61)	0.81	3.95	11.83	(12.37)	0.00	270	2,068	61.11	0.54	1.29	59.36
36 IBES FY1 Mean EPS Revision, 3M	Ascending	2,751	(8.73)	(1.01)	0.19	2.83	8.76	24.73	(32.99)	0.00	270	2,453	64.81	2.50	3.41	76.24
37 IBES FY1 EPS up/down ratio, 3M	Ascending	2,729	(5.93)	(0.98)	0.70	3.08	8.06	22.71	(24.98)	0.00	270	2,306	65.56	3.42	2.50	79.80
38 Expectation gap, short-term - long-term	Ascending	2,093	(5.66)	(3.05)	(0.67)	1.27	5.23	15.88	(23.30)	0.00	270	2,140	64.81	1.26	1.29	87.22
39 IBES FY1 Mean CFPS Revision, 3M	Ascending	1,338	(0.91)	(0.37)	0.17	0.77	10.54	29.88	(38.74)	0.33	183	759	60.66	(0.18)	2.41	65.29
40 IBES FY1 Mean SAL Revision, 3M	Ascending	2,709	(3.90)	(0.24)	0.66	0.91	8.09	27.43	(24.67)	0.15	168	2,092	58.93	0.31	1.84	71.13
42 IBES FY1 Mean DPS Revision, 3M	Ascending	1,002	(4.46)	(1.77)	0.02	0.35	5.58	15.02	(16.93)	0.55	94	953	53.19	0.25	0.53	61.25
43 IBES FY1 Mean ROE Revision, 3M	Ascending	1,871	(9.01)	(0.53)	1.42	0.51	7.05	20.79	(21.68)	0.49	94	1,644	56.38	(0.29)	1.92	65.96
44 Recommendation, mean	Descending	2,106	6.74	(1.92)	1.00	0.68	8.54	21.95	(23.25)	0.26	199	2,282	55.78	2.47	(2.40)	94.17
45 Mean recommendation revision, 3M	Descending	1,936	(4.61)	1.11	0.62	1.35	4.37	12.25	(20.10)	0.00	196	2,192	62.24	1.11	1.78	60.05
46 Target price implied return	Ascending	2,097	(11.13)	3.71	0.53	0.70	16.48	61.38	(38.06)	0.62	135	2,088	53.33	8.44	(9.88)	78.38
47 Mean target price revision, 3M	Ascending	1,921	(16.68)	0.59	0.64	2.01	13.93	31.67	(43.96)	0.10	132	1,990	62.12	(0.78)	5.80	75.54
5. Quality																
48 ROE, trailing 12M	Ascending	2,823	12.26	0.75	2.83	6.01	11.76	38.09	(31.57)	0.00	185	2,814	68.65	2.15	12.66	98.24
49 Return on invested capital (ROIC)	Ascending	2,902	12.45	0.21	3.36	5.65	10.85	34.26	(30.33)	0.00	185	2,875	70.27	2.14	11.69	98.35
50 Sales to total assets (asset turnover)	Ascending	2,917	(13.92)	2.60	2.25	1.39	8.92	22.83	(22.06)	0.01	270	2,754	56.67	2.15	0.09	99.16
51 Operating profit margin	Ascending	2,875	15.46	0.71	1.51	1.11	5.42	16.21	(14.26)	0.00	270	2,584	58.89	0.79	1.66	97.86
52 Current ratio	Descending	2,310	(1.72)	1.13	0.06	2.17	10.69	38.98	(31.17)	0.00	270	2,204	54.81	(0.73)	7.20	97.84
53 Long-term debt/equity	Ascending	2,807	(7.17)	2.84	(1.61)	0.70	9.85	35.27	(28.03)	0.25	270	2,691	48.15	(1.22)	4.00	98.47
54 Altman's z-score	Ascending	2,284	7.41	(2.03)	1.28	2.25	8.62	30.28	(21.13)	0.00	270	2,136	58.89	2.64	1.57	98.37
55 Merton's distance to default	Ascending	2,166	23.34	(1.22)	4.07	4.60	11.63	30.81	(42.61)	0.00	270	2,119	70.00	0.95	10.89	95.39
56 Ohlson default model	Descending	2,280	6.70	(1.13)	1.43	1.06	6.04	18.57	(14.17)	0.01	233	2,078	59.66	0.60	1.88	97.97
57 Campbell, Hilscher, and Szilagyi model	Descending	2,683	17.83	(0.70)	2.93	1.98	11.70	26.74	(39.01)	0.02	186	2,592	55.38	(1.98)	8.86	96.43
58 Accruals (Sloan 1996 def)	Descending	1,673	(2.44)	1.08	0.63	0.53	4.55	12.70	(11.78)	0.05	270	1,394	55.93	0.49	0.60	89.46
59 Firm-specific discretionary accruals	Descending	1,542	3.79	1.86	0.80	0.28	4.29	16.97	(12.49)	0.39	177	1,275	51.98	(0.04)	0.79	98.75
60 Hist 5Y operating EPS stability, coef of determination	Ascending	2,867	(1.61)	(0.80)	0.68	0.49	5.13	13.12	(13.13)	0.20	178	2,640	52.81	0.42	0.61	96.74
62 IBES FY1 EPS dispersion	Descending	2,557	14.77	(0.33)	2.36	2.21	10.35	26.02	(36.04)	0.00	270	2,290	59.63	(0.58)	7.04	84.72
63 Payout on trailing operating EPS	Ascending	2,114	8.10	1.51	0.45	0.82	13.85	38.83	(30.80)	0.33	270	2,193	50.37	(4.00)	9.16	99.21
64 YoY change in # of shares outstanding	Descending	2,894	2.74	0.91	1.29	2.55	9.17	45.82	(18.74)	0.00	270	2,709	58.52	(0.93)	8.57	93.83
65 YoY change in debt outstanding	Descending	2,241	(8.00)	0.29	0.13	0.31	3.73	11.22	(11.24)	0.18	270	2,133	55.56	0.72	(0.41)	89.75
66 Net external financing/net operating assets	Ascending	2,883	(5.13)	3.25	0.95	2.91	10.41	47.89	(27.14)	0.00	270	2,438	57.78	(0.30)	8.45	94.62
67 Piotroski's F-score	Ascending	2,204	(1.88)	(1.96)	(0.57)	1.86	10.77	35.31	(27.28)	0.02	185	2,137	56.76	(2.50)	9.36	90.22
68 Mohanram's G-score	Ascending	596	11.33	0.84	2.28	1.54	8.50	23.73	(26.98)	0.01	185	438	57.30	(0.63)	5.28	93.64
6. Technicals																
70 CAPM beta, 5Y monthly	Descending	2,535	35.07	(3.54)	(0.24)	0.85	16.08	45.71	(47.10)	0.43	227	2,139	48.46	(6.64)	13.84	98.70
71 CAPM idosyncratic vol, 1Y daily	Descending	2,860	20.21	(0.31)	3.30	4.79	17.98	57.80	(40.02)	0.00	270	2,649	60.74	(1.81)	16.18	99.15
72 Realized vol, 1Y daily	Descending	2,888	23.67	(1.19)	3.24	4.73	18.53	59.14	(40.26)	0.00	270	2,649	60.00	(2.40)	17.04	99.08
73 Skewness, 1Y daily	Descending	2,888	6.08	(0.57)	0.49	1.14	5.46	20.25	(14.21)	0.00	270	2,649	57.04	0.57	2.12	89.65
74 Kurtosis, 1Y daily	Descending	2,888	(1.27)	0.86	1.02	1.31	5.75	16.85	(15.19)	0.00	270	2,649	62.22	0.91	1.99	91.28
75 Idiosyncratic vol surprise	Descending	2,855	(4.14)	(1.92)	1.82	2.84	7.66	26.36	(26.55)	0.00	269	2,635	65.80	1.05	5.90	86.16
76 Normalized abnormal volume	Ascending	2,932	(3.64)	3.31	1.44	0.94	6.96	20.03	(21.03)	0.03	270	2,804	58.15	2.48	(1.73)	81.70
77 Float turnover, 12M	Descending	2,932	18.19	(3.11)	1.01	2.23	16.03	54.02	(36.07)	0.02	270	2,814	51.48	(4.88)	14.52	99.34
78 Moving average crossover, 15W-36W	Ascending	2,859	(14.91)	0.50	0.98	2.15	13.37	44.46	(52.88)	0.01	270	2,372	58.15	0.76	4.55	90.87
79 Log float-adj capitalization	Ascending	2,932	(6.45)	2.47	2.07	2.84	10.93	26.57	(38.74)	0.00	270	2,805	60.74	2.69	3.10	99.34
80 # of month in the database	Ascending	2,932	(15.06)	3.95	2.06	2.49	8.76	38.37	(22.43)	0.00	269	2,814	58.74	0.38	6.11	98.42

Note

1 Direction indicates how the factor scores are sorted. Ascending order means higher factor scores are likely to be associated with higher subsequent stock returns, and vice versa for descending order.
2 P-value indicates the statistical significance of a factor's performance. A smaller p-value suggests that it is more likely the factor's performance is different from zero.
3 This is the autocorrelation of a factor's scores over time. Higher serial correlation is likely to have lower portfolio turnover based on the factor.

Source: Compustat, IBES, Russell, S&P, Thomson Reuters, and Deutsche Bank

Source: Compustat, IBES, S&P, Russell, Deutsche Bank

Figure 57: Factor performance by size and style index

	Last Month					Three-year				
Factor Name	Universe	Russell 1000	Russell 2000	Russell 3K Value	Russell 3K Growth	Universe	Russell 1000	Russell 2000	Russell 3K Value	Russell 3K Growth
1. Value										
1 Dividend yield, trailing 12M	10.60	21.22	6.58	18.26	4.12	0.80	0.75	0.54	1.28	1.47
2 Expected dividend yield	11.00	21.46	7.09	18.35	4.34	1.07	0.61	0.77	1.43	1.43
3 Price-to-operating EPS, trailing 12M, Basic	7.61	16.01	3.31	13.08	(0.57)	0.76	(0.24)	0.97	1.11	0.44
4 Operating earnings yield, trailing 12M, Basic	11.41	21.62	6.47	15.69	1.31	1.88	0.29	1.81	1.77	1.53
5 Earnings yield, forecast FY1 mean	8.05	12.11	5.13	13.41	(1.60)	0.52	(1.88)	0.94	0.72	0.19
6 Earnings yield, forecast FY2 mean	4.24	4.86	2.28	8.91	(4.89)	0.49	(1.93)	1.21	0.97	0.11
7 Earnings yield x IBES 5Y growth	(0.75)	(2.56)	0.42	1.37	(4.13)	(0.41)	(1.24)	0.17	(0.02)	(0.88)
8 Sector-rel Operating earnings yield, trailing 12M, Basic	10.98	16.76	8.85	14.71	3.05	1.79	0.17	2.03	1.45	1.92
9 Hist-rel Operating earnings yield, trailing 12M, Basic	10.96	21.09	5.33	12.62	3.96	0.65	0.13	0.19	0.28	(0.22)
10 Operating cash flow yield (income stmt def)	5.49	(11.50)	2.07	9.99	(3.72)	0.57	0.68	0.52	0.43	0.05
11 Cash flow yield, FY1 mean	0.28	3.66	(3.67)	8.49	(7.01)	0.38	0.66	0.05	1.15	0.08
12 Free cash flow yield	(11.66)	(13.34)	(11.41)	(11.42)	(11.53)	1.35	0.57	1.55	0.60	2.34
13 Price-to-sales, trailing 12M	22.15	20.13	22.21	21.45	21.09	0.54	1.15	(0.04)	0.73	(0.20)
14 Price-to-book	3.25	8.64	1.87	(3.19)	5.11	1.07	2.15	0.06	1.61	0.13
15 EBITDA/EV	5.66	(10.49)	2.22	9.01	(1.29)	0.68	0.35	0.55	0.60	0.15
16 Price-to-book adj for ROE, sector adj	8.09	15.91	5.09	6.77	6.56	0.49	0.95	0.03	0.89	(0.09)
2. Growth										
17 Hist 5Y operating EPS growth	18.66	(25.06)	15.74	21.93	(12.00)	1.25	0.09	1.27	1.63	(0.11)
18 Hist 5Y operating EPS acceleration	4.44	8.58	2.57	1.98	4.38	0.75	1.14	0.49	1.51	(0.72)
19 IBES 5Y EPS growth	20.94	22.85	18.62	19.79	16.28	0.78	0.17	0.80	0.48	(0.55)
20 IBES 5Y EPS growth/stability	21.38	24.69	18.08	19.14	18.60	1.38	1.31	1.13	0.83	0.90
21 IBES LTG EPS mean	(2.57)	11.60	(10.39)	8.25	(5.14)	(0.16)	0.41	(0.66)	0.21	0.86
22 IBES FY2 mean DPS growth	12.25	14.26	7.55	15.94	(10.80)	1.39	0.99	1.48	2.16	(0.59)
23 IBES FY1 mean EPS growth	(9.57)	(18.66)	(4.15)	(10.65)	(7.49)	(0.99)	(1.35)	(0.99)	(0.51)	(2.32)
24 Year-over-year quarterly EPS growth	(6.96)	(11.74)	(4.81)	(7.68)	(6.66)	0.59	0.18	0.71	1.36	(0.60)
25 IBES FY1 mean CFPS growth	(7.76)	(7.67)	6.71	6.20	4.95	1.22	1.39	(2.20)	(1.53)	(1.55)
26 IBES SUE, amortized	0.07	(1.13)	0.21	(2.92)	0.31	(0.10)	(0.58)	(0.08)	(0.23)	(0.47)
3. Price momentum and reversal										
27 Total return, 1D	(7.12)	(23.63)	(1.91)	(9.06)	(7.32)	2.09	1.26	2.63	2.23	2.65
28 Total return, 21D (1M)	(18.94)	(32.62)	(12.99)	(21.55)	(15.21)	(0.26)	(1.75)	0.35	(0.66)	0.79
29 Maximum daily return in last 1M (lottery factor)	22.21	37.25	15.95	25.71	13.85	4.32	3.01	4.51	4.55	4.05
30 21D volatility of volume/price	(6.30)	0.88	9.03	(5.63)	9.56	0.88	(0.92)	(0.06)	0.70	(0.79)
31 Total return, 252D (12M)	(12.78)	(5.65)	(15.77)	(15.40)	(7.44)	(0.26)	0.38	(0.74)	0.41	(1.24)
32 12M-1M total return	(16.79)	(12.80)	(18.44)	(19.57)	(11.02)	0.24	0.30	0.10	1.02	(0.46)
33 Price-to-52 week high	12.27	(32.51)	3.26	16.08	9.27	0.76	(0.34)	0.65	0.98	0.56
34 Total return, 1260D (60M)	25.29	(22.96)	23.04	26.81	(19.96)	0.44	0.13	0.25	0.43	1.14
4. Sentiment										
35 IBES LTG Mean EPS Revision, 3M	(4.51)	(5.58)	(3.78)	(1.31)	(4.49)	(0.61)	(0.47)	(0.62)	(0.45)	(0.44)
36 IBES FY1 Mean EPS Revision, 3M	(8.73)	(18.35)	(5.68)	(10.55)	(6.63)	0.19	(0.47)	0.27	0.67	(0.67)
37 IBES FY1 EPS up/down ratio, 3M	(5.93)	(11.33)	(3.95)	(8.75)	(3.63)	0.70	0.19	0.75	0.76	0.39
38 Expectation gap, short-term - long-term	(5.66)	(9.39)	(3.24)	(5.36)	(4.94)	(0.67)	(0.89)	(0.67)	(0.09)	(1.84)
39 IBES FY1 Mean CFPS Revision, 3M	(0.91)	1.08	(3.08)	(2.98)	(1.14)	0.17	(0.51)	0.60	0.82	(0.80)
40 IBES FY1 Mean SAL Revision, 3M	(3.90)	12.72	(1.79)	(6.58)	(2.59)	0.66	1.00	1.21	0.86	0.54
41 IBES FY1 Mean FFO Revision, 3M	NA	NA	NA	8.27	NA	NA	NA	NA	0.00	NA
42 IBES FY1 Mean DPS Revision, 3M	(4.46)	(8.12)	1.69	(1.40)	10.63	0.02	(0.68)	0.72	0.49	1.55
43 IBES FY1 Mean ROE Revision, 3M	(9.01)	15.90	(3.98)	(10.35)	6.97	1.42	(0.23)	2.11	2.07	(0.46)
44 Recommendation, mean	6.74	3.45	8.15	6.83	4.51	1.00	0.62	1.02	1.21	(0.43)
45 Mean recommendation revision, 3M	(4.61)	(8.22)	(3.08)	(5.67)	(2.54)	0.62	0.12	0.68	0.79	(0.28)
46 Target price implied return	(11.13)	(23.14)	(4.27)	(13.64)	10.72	0.53	0.47	0.88	0.92	(0.19)
47 Mean target price revision, 3M	(16.68)	(18.00)	(16.12)	(20.60)	(9.79)	0.64	1.43	0.23	1.06	0.26
5. Quality										
48 ROE, trailing 12M	12.26	20.81	8.94	16.24	4.58	2.83	2.93	2.15	3.06	2.25
49 Return on invested capital (ROIC)	12.45	20.83	8.51	14.00	5.77	3.36	3.48	2.65	3.62	2.61
50 Sales to total assets (asset turnover)	(13.92)	(11.02)	(15.54)	(18.83)	(13.00)	2.25	1.88	2.30	2.28	1.66
51 Operating profit margin	15.46	15.03	14.67	12.64	14.45	1.51	0.42	1.87	0.53	2.08
52 Current ratio	(1.72)	10.65	(5.74)	5.04	(4.33)	0.06	(0.00)	(0.58)	0.18	(0.21)
53 Long-term debt/equity	(7.17)	(1.02)	(9.48)	2.86	(6.54)	(1.61)	(1.55)	(1.92)	1.85	(0.90)
54 Altman's z-score	7.41	5.17	7.51	2.58	7.65	1.28	1.20	1.40	1.29	1.21
55 Merton's distance to default	23.34	35.82	17.19	22.96	19.05	4.07	3.46	3.72	4.36	3.42
56 Ohlson default model	6.70	(2.19)	6.85	2.80	5.95	1.43	0.22	1.76	0.70	1.88
57 Campbell, Hilscher, and Szilagyi model	17.83	(22.78)	14.80	17.31	14.43	2.93	(2.94)	2.36	3.13	2.09
58 Accruals (Sloan 1996 def)	(2.44)	(8.96)	(0.48)	(3.54)	(2.05)	0.63	0.46	0.64	0.09	0.67
59 Firm-specific discretionary accruals	3.79	(1.56)	4.23	4.36	(2.79)	0.80	(0.56)	0.83	1.21	(0.82)
60 Hist 5Y operating EPS stability, coef of determination	(1.61)	2.78	(2.99)	(5.47)	4.17	0.68	1.20	0.30	0.28	0.85
62 IBES FY1 EPS dispersion	14.77	20.21	11.77	14.54	13.06	2.36	2.55	1.85	2.19	2.33
63 Payout on trailing operating EPS	8.10	(12.07)	7.05	15.00	4.86	0.45	(0.56)	0.13	1.24	0.99
64 YoY change in # of shares outstanding	2.74	12.24	(2.05)	2.81	(2.14)	1.29	1.24	0.78	0.63	1.74
65 YoY change in debt outstanding	(8.00)	(7.49)	(6.19)	(10.49)	(3.80)	0.13	(0.18)	0.63	(0.14)	0.46
66 Net external financing/net operating assets	(5.13)	8.08	(10.15)	(2.88)	(3.67)	0.95	1.30	0.68	0.81	1.51
67 Piotroski's F-score	(1.88)	(6.72)	(5.19)	1.56	(2.39)	(0.57)	0.65	(0.83)	(0.85)	(0.16)
68 Mohanram's G-score	11.33	22.11	7.22	6.99	4.44	2.28	2.82	1.69	2.32	2.15
6. Technicals										
70 CAPM beta, 5Y monthly	35.07	46.44	29.43	39.71	28.24	(0.24)	0.59	(0.51)	0.12	0.52
71 CAPM idosyncratic vol, 1Y daily	20.21	42.32	14.75	26.29	11.24	3.30	2.86	3.11	3.63	3.20
72 Realized vol, 1Y daily	23.67	47.00	16.66	29.47	14.22	3.24	3.16	3.06	3.55	3.21
73 Skewness, 1Y daily	6.08	3.28	7.32	8.46	0.76	0.49	(1.25)	0.97	0.64	0.30
74 Kurtosis, 1Y daily	(1.27)	(3.87)	(0.79)	3.34	(5.70)	1.02	1.48	0.94	1.39	0.34
75 Idiosyncratic vol surprise	(4.14)	2.95	(7.43)	(6.47)	(1.06)	1.82	1.28	1.78	1.48	1.97
76 Normalized abnormal volume	(3.64)	7.50	8.64	(2.15)	(6.44)	1.44	0.30	(0.84)	1.17	1.41
77 Float turnover, 12M	18.19	32.79	13.35	23.53	13.50	1.01	2.45	0.74	1.68	1.17
78 Moving average crossover, 15W-36W	(14.91)	(9.51)	(17.65)	(16.37)	(9.34)	0.98	1.18	0.83	1.57	0.34
79 Log float-adj capitalization	(6.45)	(13.60)	(15.99)	(7.50)	(9.12)	2.07	0.82	1.62	1.77	2.12
80 # of month in the database	(15.06)	0.29	(21.71)	(15.60)	(13.27)	2.06	0.18	2.05	2.08	2.47

Source: Compustat, IBES, Russell, S&P, Thomson Reuters, and Deutsche Bank

Source: Compustat, IBES, S&P, Russell, Deutsche Bank

Figure 58: Factor performance by GICS sector

	Last Month										
Factor Name	Universe	Energy	Materials	Industrial	Consumer Disc	Consumer Staples	Health Care	Financials	Info Tech	Telecom Services	Utilities
1. Value											
1 Dividend yield, trailing 12M	10.60	8.17	4.05	(6.88)	6.82	21.59	7.19	23.36	(0.43)	26.20	28.43
2 Expected dividend yield	11.00	8.39	4.29	(2.94)	6.51	23.90	7.93	22.89	(0.83)	25.48	27.05
3 Price-to-operating EPS, trailing 12M, Basic	7.61	14.94	27.27	(6.34)	3.57	13.06	(12.28)	20.59	0.51	(21.59)	(4.98)
4 Operating earnings yield, trailing 12M, Basic	11.41	(7.13)	25.55	(1.87)	16.47	20.72	1.26	27.54	1.86	24.26	3.79
5 Earnings yield, forecast FY1 mean	8.05	(0.25)	6.45	1.51	5.83	20.44	3.70	26.89	(7.88)	23.83	3.57
6 Earnings yield, forecast FY2 mean	4.24	(7.89)	(2.62)	(2.29)	6.26	18.52	4.11	17.61	(10.73)	23.53	5.32
7 Earnings yield x IBES 5Y growth	(0.75)	3.29	22.87	1.00	6.95	0.41	(9.82)	13.20	4.32	(6.87)	(12.16)
8 Hist-rel Operating earnings yield, trailing 12M,	10.96	(4.49)	22.80	14.20	13.24	(5.30)	10.00	19.13	(1.43)	(13.19)	2.00
9 Operating cash flow yield (income stmt def)	5.49	(3.58)	(12.18)	4.49	(5.81)	(12.34)	1.08	24.46	(6.79)	6.36	(2.49)
10 Cash flow yield, FY1 mean	0.28	5.91	12.53	(2.79)	(9.13)	(16.50)	(6.01)	8.11	(20.30)	(12.77)	(6.31)
11 Free cash flow yield	(11.66)	13.25	0.84	(13.21)	(17.15)	(9.53)	(0.44)	(3.18)	(11.12)	(13.66)	(16.59)
12 Price-to-sales, trailing 12M	22.15	(3.39)	10.63	30.90	27.91	27.14	(10.50)	10.50	21.75	39.52	16.10
13 Price-to-book	3.25	6.77	20.05	17.70	12.68	17.51	(0.17)	(4.17)	9.15	18.89	7.58
14 EBITDA/EV	5.66	(5.14)	(14.43)	2.93	(14.95)	(20.29)	4.71	14.93	(1.51)	(11.85)	(5.49)
2. Growth											
15 Hist 5Y operating EPS growth	18.66	(10.60)	31.05	(12.50)	31.00	(18.33)	5.89	31.60	(18.35)	1.42	(5.27)
16 Hist 5Y operating EPS acceleration	4.44	(8.44)	22.33	15.30	0.77	2.30	(0.89)	5.67	(5.22)	(11.74)	(15.72)
17 IBES 5Y EPS growth	20.94	(7.92)	36.19	16.57	30.66	(12.29)	4.83	25.47	19.89	(19.42)	1.51
18 IBES 5Y EPS growth/stability	21.38	(7.20)	39.77	19.34	31.86	3.34	1.46	19.53	18.87	(16.65)	(4.57)
19 IBES LTG EPS mean	(2.57)	5.99	(2.52)	(14.94)	(18.66)	28.50	(6.55)	5.68	(26.50)	(5.49)	14.77
20 IBES FY2 mean DPS growth	12.25	13.07	(20.63)	1.84	18.76	(13.29)	16.51	(5.33)	(15.18)	(19.73)	15.50
21 IBES FY1 mean EPS growth	(9.57)	(11.87)	27.56	13.83	(23.31)	8.50	4.50	(17.68)	(8.80)	16.58	8.88
22 Year-over-year quarterly EPS growth	(6.96)	(1.01)	(17.91)	15.12	(11.65)	7.28	(2.97)	(1.26)	(17.79)	18.90	(9.85)
23 IBES FY1 mean CFPS growth	(7.76)	6.94	3.27	(14.96)	21.88	(2.85)	(14.89)	(25.94)	8.39	6.42	(29.08)
24 IBES SUE, amortized	0.07	1.71	15.67	10.28	6.37	2.65	6.52	2.00	(4.22)	31.72	(2.61)
3. Price momentum and reversal											
25 Total return, 1D	(7.12)	3.92	(23.74)	(12.86)	0.92	(1.16)	(1.31)	(2.09)	(9.53)	20.15	16.27
26 Total return, 21D (1M)	(18.94)	(9.04)	(17.94)	(17.98)	(25.30)	(22.82)	(8.09)	(18.22)	(19.63)	(22.07)	(11.36)
27 Maximum daily return in last 1M (lottery factor	22.21	7.18	28.99	7.58	29.73	10.44	9.94	22.59	4.12	37.90	38.05
28 21D volatility of volume/price	(6.30)	(9.63)	3.13	9.92	5.93	14.04	11.49	(6.09)	11.86	(0.79)	9.02
29 Total return, 252D (12M)	(12.78)	(0.77)	(12.56)	(10.60)	(19.73)	(3.07)	(2.99)	(10.94)	(6.60)	14.31	(1.31)
30 12M-1M total return	(16.79)	(2.73)	(17.62)	(14.98)	(23.91)	(4.00)	(3.98)	(14.81)	(10.11)	10.28	(0.05)
31 Price-to-52 week high	12.27	10.01	(22.26)	13.14	14.47	(15.74)	3.82	(11.47)	9.17	21.16	(26.33)
32 Total return, 1260D (60M)	25.29	8.71	12.75	(24.95)	39.69	(46.98)	7.02	28.39	(22.79)	(59.04)	(13.77)
4. Sentiment											
33 IBES LTG Mean EPS Revision, 3M	(4.51)	(22.44)	(1.07)	(2.04)	(1.78)	(6.65)	3.29	0.86	1.43	(20.22)	(7.69)
34 IBES FY1 Mean EPS Revision, 3M	(8.73)	0.60	(2.31)	8.26	(22.27)	8.49	0.31	(6.76)	(8.77)	11.89	16.73
35 IBES FY1 EPS up/down ratio, 3M	(5.93)	4.34	8.36	8.19	(5.52)	9.15	4.43	(7.41)	(9.01)	21.03	4.02
36 Expectation gap, short-term - long-term	(5.66)	(14.22)	28.11	12.14	(24.03)	(11.48)	3.17	(17.46)	(12.51)	15.54	12.67
37 IBES FY1 Mean CFPS Revision, 3M	(0.91)	(4.13)	(4.64)	(33.10)	(12.17)	22.14	(6.81)	(1.04)	(5.21)	17.00	19.62
38 IBES FY1 Mean SAL Revision, 3M	(3.90)	4.92	6.03	7.88	(1.53)	18.15	4.63	0.60	5.19	6.98	(1.03)
39 IBES FY1 Mean FFO Revision, 3M	NA	NA	NA	NA	NA	NA	NA	(3.42)	NA	NA	NA
40 IBES FY1 Mean DPS Revision, 3M	(4.46)	(23.59)	(2.49)	(8.26)	(9.06)	11.89	1.98	3.47	(5.94)	7.44	(5.80)
41 IBES FY1 Mean ROE Revision, 3M	(9.01)	2.32	4.27	(6.01)	(10.88)	(17.26)	(3.79)	(4.31)	17.52	(29.77)	15.31
42 Recommendation, mean	6.74	(11.33)	23.33	9.71	13.50	(8.40)	(0.22)	9.27	7.46	26.96	(0.72)
43 Mean recommendation revision, 3M	(4.61)	(20.92)	(1.80)	(2.42)	(5.35)	(2.74)	1.19	(6.83)	(1.34)	(3.65)	(4.45)
44 Target price implied return	(11.13)	(27.23)	(16.12)	(9.71)	(9.93)	(1.85)	(13.12)	(4.39)	(12.96)	(28.69)	(16.52)
45 Mean target price revision, 3M	(16.68)	6.80	(15.00)	(2.91)	(17.55)	21.85	4.61	(21.77)	(6.17)	(12.78)	13.49
5. Quality											
46 ROE, trailing 12M	12.26	(3.23)	28.34	11.67	23.83	24.31	8.05	26.93	6.69	20.38	6.18
47 Return on invested capital (ROIC)	12.45	(6.13)	25.98	12.09	26.90	31.11	12.88	24.38	9.46	29.68	12.64
48 Sales to total assets (asset turnover)	(13.92)	11.64	(3.27)	(7.63)	(6.95)	(18.45)	(2.66)	(2.43)	(13.42)	47.34	(0.85)
49 Operating profit margin	15.46	(3.13)	25.61	6.09	9.59	11.40	18.84	21.76	17.55	47.67	(0.89)
50 Current ratio	(1.72)	4.85	28.19	(2.47)	(5.22)	(7.64)	(1.80)	10.41	(5.86)	7.41	41.25
51 Long-term debt/equity	(7.17)	9.47	(0.91)	12.17	15.87	0.14	(6.06)	7.50	20.51	8.43	(3.70)
52 Altman's z-score	7.41	15.41	2.63	16.48	10.66	11.53	7.98	19.34	15.32	10.49	6.03
53 Merton's distance to default	23.34	10.52	22.84	16.84	26.53	24.08	23.94	19.61	22.69	53.20	45.53
54 Ohlson default model	6.70	(0.08)	(3.60)	8.12	11.43	(6.59)	15.96	(0.22)	7.16	2.49	17.51
55 Campbell, Hilscher, and Szilagyi model	17.83	7.04	26.35	(21.67)	28.77	(29.52)	20.81	(23.11)	20.81	20.77	(11.24)
56 Accruals (Sloan 1996 def)	(2.44)	(7.53)	(13.44)	3.60	4.40	(10.38)	1.27	(13.03)	(5.27)	(20.73)	15.89
57 Firm-specific discretionary accruals	3.79	(9.36)	(9.77)	8.39	(0.63)	(17.90)	(2.17)	NA	(8.11)	9.52	(23.23)
58 Hist 5Y operating EPS stability, coef of detern	(1.61)	(5.83)	2.28	(2.46)	(1.26)	(5.09)	6.36	(11.58)	6.59	(2.94)	6.36
60 IBES FY1 EPS dispersion	14.77	3.81	31.18	17.61	24.90	4.54	13.36	9.20	11.52	17.70	(29.65)
61 Payout on trailing operating EPS	8.10	11.01	0.11	(7.23)	(0.57)	22.16	(5.33)	(6.72)	0.69	12.79	(16.76)
62 YoY change in # of shares outstanding	2.74	6.39	(1.99)	(1.69)	8.52	(8.78)	3.37	21.16	(9.33)	5.70	(8.32)
63 YoY change in debt outstanding	(8.00)	(12.07)	7.58	3.00	(6.11)	14.82	(5.51)	(22.53)	10.48	14.24	(3.16)
64 Net external financing/net operating assets	(5.13)	16.28	2.07	(2.30)	(4.60)	14.68	2.62	7.58	(13.20)	(1.01)	(17.98)
65 Piotroski's F-score	(1.88)	0.06	(8.48)	1.28	(3.83)	(4.43)	0.79	(11.55)	(4.64)	23.43	(21.63)
66 Mohanram's G-score	11.33	NA	NA	14.22	34.35	31.12	10.15	NA	8.62	NA	NA
6. Technicals											
68 CAPM beta, 5Y monthly	35.07	16.88	(40.63)	24.93	36.48	(21.73)	13.77	(38.21)	26.27	57.51	(43.00)
69 CAPM idosyncratic vol, 1Y daily	20.21	2.49	12.07	10.68	26.70	16.39	10.61	24.51	4.70	40.35	33.93
70 Realized vol, 1Y daily	23.67	8.33	23.09	12.70	29.29	16.98	9.24	29.88	7.15	43.78	37.94
71 Skewness, 1Y daily	6.08	(0.35)	(12.49)	(6.30)	(1.49)	(13.97)	5.37	12.60	3.02	(12.73)	2.57
72 Kurtosis, 1Y daily	(1.27)	(4.42)	(4.79)	(6.62)	(5.31)	4.26	(7.60)	(5.47)	(1.98)	21.47	(4.61)
73 Idiosyncratic vol surprise	(4.14)	(14.23)	2.57	(7.74)	(6.21)	7.88	(5.65)	(16.78)	3.47	29.90	(12.21)
74 Normalized abnormal volume	(3.64)	(1.02)	6.40	4.93	(3.22)	7.71	(6.27)	(2.95)	14.44	12.28	(8.49)
75 Float turnover, 12M	18.19	7.44	20.31	4.06	18.55	0.46	7.14	(23.20)	6.99	24.92	(25.74)
76 Moving average crossover, 15W-36W	(14.91)	8.95	(15.46)	1.25	(22.94)	5.82	(6.75)	(23.27)	(5.11)	(22.26)	(1.12)
77 Log float-adj capitalization	(6.45)	(12.20)	(9.10)	(7.41)	2.51	6.47	(0.21)	(16.77)	(16.32)	8.31	(2.63)
78 # of month in the database	(15.06)	(9.65)	(14.57)	(14.45)	(20.76)	(14.19)	(1.89)	21.86	(22.77)	7.74	2.60

Source: Compustat, IBES, Russell, S&P, Thomson Reuters, and Deutsche Bank

Source: Compustat, IBES, S&P, Russell, Deutsche Bank

Appendix 1

Important Disclosures

Additional information available upon request

For disclosures pertaining to recommendations or estimates made on a security mentioned in this report, 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**.**

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. Rochester Cahan/Yin Luo/Javed Jussa

Hypothetical Disclaimer

Backtested, hypothetical or simulated performance results discussed on page 10 herein and after 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.

Regulatory Disclosures

1. Important Additional Conflict Disclosures

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.

2. Short-Term Trade Ideas

Deutsche Bank equity research analysts sometimes have shorter-term trade ideas (known as SOLAR ideas) that are consistent or inconsistent with Deutsche Bank's existing longer term ratings. These trade ideas can be found at the SOLAR link at http://gm.db.com.

3. Country-Specific Disclosures

Australia: This research, and any access to it, is intended only for "wholesale clients" within the meaning of the Australian Corporations Act.

EU countries: Disclosures relating to our obligations under MiFiD can be found at http://globalmarkets.db.com/riskdisclosures.

Japan: Disclosures under the Financial Instruments and Exchange Law: Company name - Deutsche Securities Inc. Registration number - Registered as a financial instruments dealer by the Head of the Kanto Local Finance Bureau (Kinsho) No. 117. Member of associations: JSDA, 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.

New Zealand: This research is not intended for, and should not be given to, "members of the public" within the meaning of the New Zealand Securities Market Act 1988.

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.

Deutsche Bank Securities Inc.

North American location

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

Deutsche Bank Securities Inc.
225 Franklin Street
25th Floor
Boston, MA 02110
Tel: (617) 988 6100

Deutsche Bank Securities Inc.
222 South Riverside Plaza
30th Floor
Chicago, IL 60606
Tel: (312) 537-3758

Deutsche Bank Securities Inc.
3033 East First Avenue
Suite 303, Third Floor
Denver, CO 80206
Tel: (303) 394 6800

Deutsche Bank Securities Inc.
1735 Market Street
24th Floor
Philadelphia, PA 19103
Tel: (215) 854 1546

Deutsche Bank Securities Inc.
101 California Street
46th Floor
San Francisco, CA 94111
Tel: (415) 617 2800

Deutsche Bank Securities Inc.
700 Louisiana Street
Houston, TX 77002
Tel: (832) 239-4600

International Locations

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

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

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

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
Level 55
Cheung Kong Center
2 Queen's Road Central
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 6701

Disclaimer

The information and opinions in this report were prepared by Deutsche Bank AG or one of its affiliates (collectively "Deutsche Bank"). 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 the accuracy or completeness of such information.

Deutsche Bank may engage in securities transactions, on a proprietary basis or otherwise, in a manner **inconsistent** with the view taken in this research report. In addition, others within Deutsche Bank, including strategists and sales staff, may take a view that is **inconsistent** with that taken in this research report.

Opinions, estimates and projections in this report constitute the current judgement 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 has no obligation to update, modify or amend this report or to otherwise notify a recipient thereof in the event that any opinion, forecast or estimate set forth herein, changes or subsequently becomes inaccurate. Prices and availability of financial instruments are subject to change without notice. This report is provided for informational purposes only. 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 judgement.

As a result of Deutsche Bank's recent acquisition of BHF-Bank AG, a security may be covered by more than one analyst within the Deutsche Bank group. Each of these analysts may use differing methodologies to value the security; as a result, the recommendations may differ and the price targets and estimates of each may vary widely.

Deutsche Bank has instituted a new policy whereby analysts may choose not to set or maintain a target price of certain issuers under coverage with a Hold rating. In particular, this will typically occur for "Hold" rated stocks having a market cap smaller than most other companies in its sector or region. We believe that such policy will allow us to make best use of our resources. Please visit our website at http://gm.db.com to determine the target price of any stock.

The financial instruments discussed in this report may not be suitable for all investors and investors must make their own informed investment decisions. Stock 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.

All prices are those current at the end of the previous trading session unless otherwise indicated. Prices are sourced from local exchanges via Reuters, Bloomberg and other vendors. Data is sourced from Deutsche Bank and subject companies.

Past performance is not necessarily indicative of future results. Deutsche Bank may with respect to securities covered by this report, sell to or buy from customers on a principal basis, and consider this report in deciding to trade on a proprietary basis.

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. 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.theocc.com/components/docs/riskstoc.pdf If you are unable to access the website please contact Deutsche Bank AG at +1 (212) 250-7994, for a copy of this important document.

The risk of loss in futures trading, foreign or domestic, can be substantial. As a result of the high degree of leverage obtainable in futures trading, losses may be incurred that are greater than the amount of funds initially deposited.

Unless governing law provides otherwise, all transactions should be executed through the Deutsche Bank entity in the investor's home jurisdiction. In the U.S. this report is approved and/or distributed by Deutsche Bank Securities Inc., a member of the NYSE, the NASD, NFA and SIPC. In Germany this report is approved and/or communicated by Deutsche Bank AG Frankfurt authorized by the BaFin. In the United Kingdom this report is approved and/or communicated by Deutsche Bank AG London, a member of the London Stock Exchange and regulated by the Financial Services Authority for the conduct of investment business in the UK and authorized by the BaFin. This report is distributed in Hong Kong by Deutsche Bank AG, Hong Kong Branch, in Korea by Deutsche Securities Korea Co. This report is distributed in Singapore by Deutsche Bank AG, Singapore Branch, and recipients in Singapore of this report are to contact Deutsche Bank AG, Singapore Branch 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), Deutsche Bank AG, Singapore Branch accepts legal responsibility to such person for the contents of this report. In Japan this report is approved and/or distributed by Deutsche Securities Inc. The information contained in this report does not constitute the provision of investment advice. In 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. Deutsche Bank AG Johannesburg is incorporated in the Federal Republic of Germany (Branch Register Number in South Africa: 1998/003298/10). 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 by any person for any purpose without Deutsche Bank's prior written consent. Please cite source when quoting.