THE IMPACT of non-scheduled news on S&P/ASX 50 stocks

With the large increase in the availability of news, growing numbers of financial market participants are relying on analytics software to evaluate news events. This paper explores whether indicators processed by analytics software are effective in assigning sentiment, and whether it is possible to trade profitably using such measures. Based on a sample of 33 highly liquid S&P/ASX 50 stocks, the results indicate that while the sentiment indicators correctly classify news, it is unlikely that trading strategies with time horizons of 30 seconds or more (and even moderate transaction costs) will produce statistically significant abnormal returns.

The advent of modern technology has facilitated a large increase in the flow of news available to financial market participants and made it costly for traders to process all asset-specific news. As a result, an increasing number of participants are starting to rely on textual algorithms and pre-processed news analytics provided by news vendors to evaluate news events. With such innovative data playing an increasingly important role in the trading of financial assets it seems appropriate to ask whether indicators of sentiment are both useful and reliable.

This paper primarily addresses a single key question: do indicators of sentiment matter, and if so, does preprocessed data do a good job of assigning such indicators in a way that allows investors to trade profitably? Using a sample of 33 highly liquid S&P/ASX 50 stocks over an 11-year time period, I examine market activity around 345,977 non-scheduled news announcements. Empirical evidence is supportive of analytics software correctly classifying news releases; there is some evidence that returns react to news prior to its arrival on the newswire but the level of returns is not sufficient to be profitable after considering transaction costs. An asymmetric response is found whereby negative news has a greater impact than positive news. Further, the empirical evidence suggests that while the relationship between news sentiment and returns is in the expected direction, this relationship is statistically insignificant (owing to large standard errors) and so it would not be possible for traders without access to extremely low transaction costs, or the ability to trade at higher frequencies, to consistently generate abnormal returns using a news sentiment based strategy.

The quantification of sentiment in news has enabled the identification of firm responses and market reactions across a wide range of news events. Negative stories are found to predict both low market returns and low firm earnings (Tetlock 2007; Tetlock et al. 2008). Several recent studies have focused on the predictability of returns based on news sentiment with Sinha (2011) and Dzielinski (2011) reporting that positive (negative) news predicts above (below) average returns, although Garcia (2013) suggests this predictability is confined to recessions. Hafez (2011) shows that news sentiment outperforms price momentum when predicting future stock market returns, while the applicability of this research in industry is highlighted by O’Hara (2014), who claims that news sentiment tools are frequently adopted in the trading strategies of high-frequency traders.

The impact of news events on other market measures such as volatility and volume has been investigated in the literature. Kalev et al. (2004) note a positive and significant impact on trading activity following the arrival of non-scheduled news, which Von Beschwitz et al. (2013) report occurs in the first few seconds after the event. Groß-Klußmann and Hautsch (2011) observe that this effect is applicable only to news identified as ‘relevant’. The importance of firm-specific news sentiment is demonstrated by Ho et al. (2013) who show that the effect on volatility persistence is greater than that for macro-economic news.

The extant literature has also recognised an asymmetric effect relating to the reaction of financial markets to news events. Hafez (2013) identifies low (high) volatility stocks as having the strongest reaction to positive (negative) sentiment. More generally, Tetlock (2007), Leinweber and Sisk (2011), LEE SMALES, Senior Lecturer, Department of Finance and Banking, School of Economics and Finance, Curtin Business School, Curtin University

JASSA The Finnia Journal of Applied Finance ISSUE 2 2014 17
and Riordan et al. (2013) emphasise the importance of negative news, reporting that the reaction is both more vigorous and more informative than that for positive news.

Data
RavenPack provides real-time news analytics tools to market participants in order to capture important characteristics of the large amounts of news that flow over the Dow Jones Newswire throughout the trading day. For every new instance in which a company is reported in the news, RavenPack produces a company level record that contains 28 fields including a timestamp (to the nearest millisecond), company identifier, and a unique identifier for each news story. The software also provides scores for relevance to the given company, novelty, and sentiment; the scores are determined using pattern recognition algorithms and are available in real-time to subscribers.

The empirical literature (e.g. Groß-Klußmann and Hautsch 2011; Hafez 2011; Riordan et al. 2013) highlights the importance of novel information with high relevance and this provides the focus of this article; stories with a relevance score lower than 90 and a novelty score greater than 1 are therefore removed from the sample, leaving a sample of 345,977 news headlines for the 33 stocks in my sample.

RavenPack provides subscribers with several sentiment indicators including Composite Sentiment Score (CSS), News Impact Projections (NIP), and Editorials & Commentary (BMQ). However, the results reported in this article are for the Multi-Classifier for Equities (MCQ) indicator. MCQ is a multi-classifier score that represents the news sentiment according to the tone applicable only towards the most relevant companies mentioned in a story, with the score derived from a combination of values produced by classifiers using both a rule-based methodology that maps key words and phrases to pre-defined sentiment values, and an expert consensus methodology that trains classification algorithms on the results of financial experts manually tagging stories. The logic of focusing on this classifier is that it detects consistent sentiment classifications, and discards combinations where the classifiers have contradictory scores. An MCQ is provided only when the relevance score for a company is 90 or higher, and RavenPack assigns a score of 0 to negative sentiment, 50 to neutral and 100 to positive; my analysis scales the score to the more intuitive levels of -1, 0, +1. Figure 1 shows that news stories classified as neutral are predominant constituting around 55 per cent of all news items; the proportion of positive news items is almost double that of negative news items.

To enable a high-frequency study of market dynamics it is necessary to ensure that the stocks covered within the analysis are liquid and thus I focus on the stocks which make up the S&P/ASX 50; the leading 50 domestic stocks by market capitalisation trading on the Australian Securities Exchange (ASX). As data is required for the 2,525 trading days that make up the January 2001 to November 2011 sample, the sample is reduced to the 33 stocks listed for the entire sample period. Covering 75.7 per cent of the market capitalisation of the Australian All Ordinaries Index, the sample can be considered as being representative of the Australian stock market.

Underlying transaction data is obtained from Thomson Reuters Tick History (TRTH), via SIRCA, and aggregated into 30-second intervals. This aggregation level is a compromise between exploiting maximum information and making the analysis computationally tractable and this is especially important given that the analysis covers 11 years of data. While there may be a slight lag in TRTH data as compared to the RavenPack news release, the aggregation process and lengthy sample period should reduce any effect on the reported results. Market activity, volatility and liquidity are captured by the following variables computed over 30-second intervals. To account for intraday patterns, each of the variables is normalised by the 12-month rolling average of the corresponding underlying 30-second interval:

> Money value volume traded — calculated as price multiplied by volume traded in the given interval
> Volatility — calculated using the time-weighted standard-deviation of mid-point returns in each interval
> Trade order imbalance — defined as the value of the difference in cumulated buyer- and seller-initiated trades
> Bid-ask spread — defined as the average bid-ask spread over the given interval
> Returns — calculated using the mid-point of the bid-ask quote.
Analysis and results

Figure 2 shows standardised measures of market activity in the period around neutral, negative and positive news announcements for the whole sample period. Consistent with Dzielinski (2011), neutral news releases appear to result in no significant change in any of the variables around the time of the news release. In contrast, both positive and negative news events elicit a sharp spike in the normalised variables in the period around the event. The above-average activities start at least 15 minutes prior to the news release, move sharply higher in the two minutes immediately prior to the release, and quickly revert to lower levels following the release. Note that trading activity (volume and order imbalance) increases despite wider than average bid–ask spreads.

Consistent with Chen et al. (2003) the measures for all normalised variables are higher in the case of negative news; this is particularly so in the 30-second intervals either side of the news release.

**FIGURE 2: Stock market activity in period surrounding newswire releases**

<table>
<thead>
<tr>
<th>MV volume traded around news</th>
<th>Volatility around news</th>
<th>Order imbalance around news</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The profitability of trading on news items is first considered using an event study framework, with cumulative abnormal returns calculated using a standard market model and averaged over all stocks. Figure 3 shows the average cumulated abnormal returns around highly relevant news — disaggregated into positive and negative news events as defined by the MCQ. Starting 15 minutes before the disclosure I observe positive (negative) cumulated abnormal returns in reaction to positive (negative) news items. While price movements are observed prior to news releases there are only limited movements in the following periods and this suggests the possibility of the presence of information leakage, and potentially the clustering of news items. However, once moderate transaction costs are considered, the impact on returns is sufficiently small as to render trading on news items non-profitable.

**FIGURE 3: Cumulative abnormal returns around negative and positive newswire messages**

I investigate the relationship between non-scheduled news releases and changes in returns, and the market activity variables, at high-frequency 30-second intervals. In utilising a specification that is at least partially contemporaneous I do not wish to imply causality, but instead seek to examine the reliability of statistical relations. This relationship is estimated using an ordinary least square (OLS) specification:

\[ \Delta V_{kt} = \beta_0 + \beta_1 \text{News}_{kt} + \beta_2 \text{News}_{kt-1} + \epsilon_t \]
Where $\Delta IV_{kt}$ is the change in the independent variable (Returns, Volume, Volatility, Order Imbalance or BAS) and News$_{kt}$ is the sentiment of news occurring for company $k$ during the 30-second interval $t$. News releases are also disaggregated into negative and positive components in order to assess the possibility of an asymmetric response to news:

$$\Delta IV_{kt} = \beta_0 + \beta_1 \text{News}_t^- + \beta_2 \text{News}_{t-1}^- + \beta_3 \text{News}_t^+ + \beta_4 \text{News}_{t-1}^+ + \epsilon_t$$

Where News$_t^-$ is the sentiment of news items released during interval $t$, conditional on negative news (i.e. News$_t < 0$) and 0 otherwise, and News$_t^+$ is the sentiment of news items released during interval $t$, conditional on positive news (News$_t > 0$) and 0 otherwise.

Table 1 presents the results for the regression with aggregated news effects. Newey-West standard errors are used to correct for possible autocorrelation in the variables. The relationship between news and returns is in the direction one would expect, i.e. more positive news results in higher returns and more negative news in lower returns. However, the size of the standard error for returns is so large that the estimated coefficients are insignificant. Volume, volatility, and order imbalance all seem to be significantly affected by the presence of news releases and the sign of the coefficients suggests that volume and volatility increases (decreases) in the periods of positive (negative) news. The positive coefficient for order imbalance implies that the number of buy (sell) orders outweighs the number of sell (buy) orders when there is positive (negative) sentiment. The bid–ask spread does not appear to deviate significantly from the standardised value of 1 during either period and is not significantly affected by the sentiment of newswire messages. The results suggest that although market activity is significantly related to news sentiment there is little scope for market participants to make significant trading profits, at least in the 30-second time intervals considered.

**Table 1: Regression results for news sentiment and standardised market measures**

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Constant</th>
<th>News</th>
<th>News (-1)</th>
<th>Adj R²</th>
<th>F-Stat</th>
<th>DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns</td>
<td>0.690</td>
<td>3.241</td>
<td>0.234</td>
<td>0.001</td>
<td>1.944</td>
<td>1.954</td>
</tr>
<tr>
<td></td>
<td>(3.383)</td>
<td>(38.028)</td>
<td>(36.043)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volume</td>
<td>0.916 ***</td>
<td>0.150 ***</td>
<td>0.068 **</td>
<td>0.019</td>
<td>18.90</td>
<td>1.693</td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.046)</td>
<td>(0.032)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volatility</td>
<td>0.635 ***</td>
<td>0.070 ***</td>
<td>0.039 **</td>
<td>0.010</td>
<td>125.20</td>
<td>1.423</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.021)</td>
<td>(0.018)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order imbalance</td>
<td>0.890 ***</td>
<td>0.051 ***</td>
<td>0.047 **</td>
<td>0.013</td>
<td>42.79</td>
<td>1.761</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.019)</td>
<td>(0.020)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bid–ask spread</td>
<td>1.000</td>
<td>0.019</td>
<td>0.020</td>
<td>0.002</td>
<td>2.514</td>
<td>1.142</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.011)</td>
<td>(0.014)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Dependent variables for this series of regression results are Returns (calculated as $100 \times \ln(P_t/P_{t-1})$), Money Value of Volume traded (calculated as Price x Volume), Volatility (calculated as standard deviation of mid-point changes), Order Imbalance (calculated as the MV Volume of buy initiated trades — MV Volume of sell initiated trades), and Bid-Ask Spread (calculated as the time-weighted average bid-ask spread) for each 30-second interval during the trading day. The variables are normalized to 1 using a 12-month rolling average. The independent variable, News, is news sentiment (and 1-lag) occurring in the given 30-second interval, $t$. Since the dependent variables are normalized to 1, using a 12-month rolling average, the t-test for the constant term is whether it differs from 1.

Sample period: 2 January 2001 – 1 November 2011
The reported evidence suggests that while the relationship between news sentiment and returns is in the expected direction, with negative news producing negative returns and vice-versa, this relationship is insignificant and it would not be possible for market participants to consistently generate excess returns using a news sentiment based strategy once transaction costs are incurred, i.e. market efficiency holds. Of course, this does not necessarily mean that investors able to trade at a greater speed than the 30-second intervals used in this study, or at very low transaction costs, will not be able to benefit from such strategies.

Therefore, it is likely that high frequency traders, many of whom receive transaction cost rebates from the exchange, would have the greatest potential to incorporate this type of news analytics into a successful trading strategy. Also, investors looking to execute large orders may be able to utilise news analytics to understand changes in market activity.
Endnotes

1 More extensive information on the RavenPack news analytics tool, complete with details on the methodology of assigning sentiment scores may be found at www.ravenpack.com

2 The processed data is available to market participants (at a cost) almost instantaneously; academic researchers are able to access this information only at a later stage.

3 RavenPack assigns each story a Relevance score in the range [0,100] with 0 indicating a passive mention and 100 indicating predominantly mentioned, while novelty is indicated by the Event_Novelty_Index. A new story, published for the first time has an index value equal to 1.

4 The analysis has also been repeated with the CSS measure and the results were qualitatively similar.

5 Details of the companies and related news releases are available from the author on request.

6 Negative news multiplied by the positive coefficient in Table 3 gives the expected negative returns.

7 Negative news multiplied by the positive coefficient in Table 3 give the positive (increase) in measures.

References


Hafez, PA 2013, Attention conditions stock market reaction to news sentiment, available at RavenPack.com


