January 15, 2019

BEYOND FAKE NEWS

Extracting Signals from Noisy News, Social Media, and Events – Welcome to NICE

- **Revisiting RavenPack Analytics.** In this research, we continue our quest for alternative data based stock-selection signals by revisiting RavenPack Analytics – a leading data vendor on news sentiment and text analytics. The goal is to develop an orthogonal signal for longer term investors. This research comprises the third element of our text mining model suite. Combined with our two existing models – the SPEC (Systematic Profiling of EDGAR Composite) and SMEC (Systematic Mining of Earnings Calls), we have a complete library of NLP (Natural Language Processing) stock selection tools.
- **Event-Based Sentiment and Behavioral Signal.** Extracting true signal from noise is never easy, especially in the case of news and social media. Not only do we need to filter out so-called "fake news", but also we strive to understand the complex interactions among news coverage, sentiment, market behavioral biases, along with corporate events. Taking advantage of RavenPack's event identification algorithm, we focus our attention on less common but critical news events such as legal & regulatory issues, labor disputes, shareholder disclosures and executive appointments. We use the LASSO and xgBoost machine learning algorithms to dissect the complex and nonlinear relationship among news and corporate events.
- **Introducing the NICE.** Named after the French coastal city, the NICE model (News with Insightful Categorical Events) is constructed by combining multiple event-based sentiment and behavioral factors into one composite signal, using the elastic net penalized regression framework. The NICE model outperforms conventional quantitative factors on a risk adjusted basis, with long investment horizon, covering 90% of the Russell 3000 universe in the US, more than 1,500 stocks in other developed markets, and 1,500 firms in EM. The NICE model is relatively uncorrelated to our two existing NLP models (SPEC and SMEC) and traditional stock selection factors.

Source: Wolfe Research Luo's QES, RavenPack

Gaurav Rohal, CFA GRohal@wolferesearch.com
Yin Luo, CFA, CPA YLuo@wolferesearch.com
Javed Jussa JJussa@wolferesearch.com
Miguel Alvarez MAlvarez@wolferesearch.com
Sheng Wang SWang@wolferesearch.com
Jason Zhong, PhD JZhong@wolferesearch.com
Zhao Jin ZJin@wolferesearch.com
QES Desk: 1.646.582.9230 Luo.QES@wolferesearch.com

Table of Contents

A Letter to Our Readers

Having previously explored Thomson Reuters News Analytics, Recorded Future, News Quantified and of course, our favorite RavenPack data; news and web-based signals are not new to us. In this research, we continue our quest for alternative data by revisiting RavenPack Analytics – a leading data vendor on news sentiment and text analytics. In particular, we focus on the complex interaction between news coverage, sentiment, market behavioral biases and corporate events. The goal is to develop an orthogonal signal for longer term investors.

While most text mining data vendors and academic research focus exclusively on news sentiment analysis, we have developed our proprietary suite of NLP (Natural Language Processing) models to analyze more informative textual contents and machine learning algorithms well beyond sentiment.

- SPEC (Systematic Profiling of EDGAR Composite). As elaborated in *Text Mining Unstructured Corporate Filing Data* (see Rohal, et al [2017a]), our SPEC model examines corporate regulatory filings (e.g., 10K/10Q) to identify red flags, process tone/sentiment, and diagnose other behavior bias.
- SMEC (Systematic Mining of Earnings Calls) Model. As discussed in *Tone at the Top? Quantifying Management Presentation* (see Rohal, et al [2018a]), we inspect management presentations and conference calls to quantify language complexity, management sentiment, executive personality, and key topics.

In this paper, we further enhance our NLP offerings by adding the third element – the NICE[1] (News with Insightful Categorical Events) model. Our three text mining models – the NICE, SPEC, and SMEC each covers different information contents, using very different model building techniques. As a result, they are relatively uncorrelated. The combined NICE+SPEC+SMEC model offers significant performance enhancement.

One of the unique angles of this research is to take advantage of RavenPack's event identification algorithm. We find how the market reaction to news and sentiment varies tremendously on the type of corporate events. For example, investors typically overreact to bad news related to layoffs and litigations, which leads to price reversal post news releases. On the other hand, the overwhelmingly "boring" news on share buybacks and dividends are often overlooked, which produces a momentum effect. Using a combination of LASSO and xgBoost machine learning techniques, we can effectively take advantage of these conditional (and likely nonlinear) relationships, for each event type.

The final NICE model further blends all the event-based sentiment signals. The NICE model outperforms most conventional stock-selection factors on a risk-adjusted basis, covering 90% of the Russell 3000 universe, over 1,500 stocks each in other developed markets, and in EM.

Regards,

Yin, Gaurav, and the QES team

[1] We coin the model by the French coastal city Nice (pronounced as /niːs/ in IPA), rather than the English word nice (pronounced as /naɪs/ in IPA).

RavenPack Analytics

RavenPack tracks millions of news stories and websites on all major public companies, government entities, currencies, commodities, and people worldwide in real time. RavenPack's main product uses NLP algorithms to quantify the sentiment of each news item and tag it to relevant entities. We have used RavenPack data extensively in our previous research:

- In *The Big and the Small Sides of Big Data* (see Luo, et al [2017a]), we discuss how RavenPack sentiment signals can be incremental to traditional factor investing.
- In *Banking on the Banks – Welcome to BALI* (see Luo, et al [2018b]), we find that news sentiment data has additional insights for banking stocks.
- In *Global TMT Stock Selection Models – Introducing TALIA* (see Rohal, et al [2018c]), we find that TMT (Technology, Media, and Telecom) stocks are particularly sensitive to news coverage.

Academic Literature Review

News and sentiment related data has generated considerable attention from academia. In one of the seminal papers, Tetlock [2005] studies daily data from the *Wall Street Journal* and finds that high media pessimism predicts downward pressure on stock price on the next day. Sprenger, et al [2011] extract news data from an online stock forum and study its impact on asset pricing. The authors classify events in few broad categories and then apply sentiment analysis. They find that aggregate news events by topic contain little information about future asset returns. Rather, news sentiment does have predictive power of future stock returns.

Boudoukh [2012] finds stock price reversals occur on "no news" days, while days identified as "news days" show the opposite effect – namely a strong degree of continuation. More recently Huynh, et al [2015] suggests that the performance of the standard price momentum factor can be enhanced by adding news signals derived from Thomson Reuters News Analytics. Shu [2016] finds that non-local Twitter posts exhibit negative return predictability, while local posts are more relevant for international markets.

Lastly, Kearney, et al [2013] provide a comprehensive literature review of news and sentiment related research. In summary, most of these academic studies suffer from limited data coverage and short data history. Furthermore, most existing research seems to suggest that the predictive power of textual data is only relevant for a short horizon usually consisting of a few days.

Introducing RavenPack Data

RavenPack Analytics transforms unstructured news and social media data into systematic indicators. It can be used as a source for alternative alpha as well as to manage event risk. RavenPack Analytics allows investors to incorporate news, social media, and sentiment in their investment process in real time. In addition, news events on natural disasters, political upheaval, regulatory uncertainties and many others can be precursors to volatility spikes at both individual company or aggregate market level. Sentiment analysis and especially event classification give investors an edge by shortening response time to breaking news.

Figure 1 shows a snapshot of the RavenPack database, with a few selected data items:

- **Time stamp.** Every news is properly time stamped to millisecond[2].
- **Entity name.** Each news is tagged to potentially many entities. An entity can be a company, a place, an individual, a country, a currency, etc. Entity recognition is one of RavenPack's key strengths. Each entity has its own unique identifier. For the purpose of this research, we most concern publicly traded companies, which are properly mapped into our global equity database.
- **Relevance.** RavenPack also provides a relevance score for each of the related entities. The relevance score is always between 0 (completely unrelated) to 100 (very relevant to the entity). In our experience, it is critical to build our signals only on these highly relevant news articles[3].
- **Novelty (Similarity Days).** RavenPack provides a granular number which indicates the number of days since a similar event was detected. Values range between 0 and 365. A value of 365 means that the most recent similar story may have occurred 365 or more days in the past. For many news stories, there are typically multiple updates, even from the same news provider (e.g., Dow Jones Newswire). The market normally only reacts to the first (few) release of the news. The market impact on the subsequent updates tends to be far more modest.
- **Headline.**
- **Sentiment.** RavenPack provides a suite of pre-computed sentiment measures based on its proprietary NLP algorithms:
    - The headline RavenPack news sentiment signal is called **ESS** (Event Sentiment Score), a score between 0 and 100[4]. It is based on RavenPack's NLP algorithms, and extensive database of time sensitive information about each entity.
    - **NIP** (News Impact Projections) measures the impact of a particular news on stock volatility in the following two-hour window.
    - **MCQ** (Multi Classifier for Equities) is only applicable towards the most relevant companies mentioned in a story. MCQ is constructed similar to the ESS, with some noticeable differences.
    - **CSS** (Composite Sentiment Score) sentiment is trained on how the market reacted to similar words/phrases in the past, using intraday tick data.
- **Event Category.** RavenPack automatically detects key news events and identifies the role played by the entity. There are four layers of event categories:
    - **Topic** is the highest level of event taxonomy.
    - There are multiple **Groups** within a topic category.
    - The third event level is called "**Type**".
    - The most granular level is the **Sub-Type**.
- We further extend these signals by computing aggregate sentiment, abnormal volume and market reaction, using a variety of rolling windows. We measure abnormal volume as the

[2] We only show date in this example. In this research, we focus on the long-term implications of new stories.
[3] A score of 75 and above is normally considered as significant.
[4] Zero represents the most negative sentiment, while 100 means the most bullish tone.

deviation of news volume from its long-term trend and market reaction to news announcements as the excess returns during the event week.

Figure 1 RavenPack Data Sample Snapshot

timestamp_utc	entity_name	relevance	ess	similarity_days	topic	grp	type	sub_type	headline	css	nip	mcq
05/01/2018 15:44:09	Aetna Inc.	100	-0.24	365	business	revenues	revenue	down	Aetna Inc. Q1 Total Revenues USD 15.34B Vs USD 15.49B	0	0	0
05/01/2018 15:44:09	Aetna Inc.	100	0.44	365	business	earnings	pretax-earnings-expectations	above-expectations	Aetna Inc. Reports Q1 Pre-Tax Income Adjusted USD 1.34B Vs Consensus USD	0	0	0
05/01/2018 15:44:09	Aetna Inc.	100	-0.27	365	business	revenues	revenue	below-expectations	Aetna Inc. [illegible] 34B Vs Consensus USD 15.35B	[illegible]	0	0
05/01/2018 15:44:09	Aetna Inc.	100	-0.21	365	business	earnings	pretax-earnings	down	Aetna Inc. Q1 [illegible] 34B Vs USD 1.45B	[illegible]	0	0
05/01/2018 15:44:09	Aetna Inc.	100	0.86	365	business	earnings	pretax-earnings	up	Aetna Inc. Q1 [illegible] 3B Vs USD -628.00M	[illegible]	[illegible]	0
05/01/2018 15:44:09	Aetna Inc.	100	0.46	365	business	earnings	earnings-expectations	above-expectations	Aetna Inc. Reports Q1 Net Profit Adjusted USD 973.00M Vs Co	[illegible]	[illegible]	0
05/01/2018 15:44:09	Aetna Inc.	100	0.4	365	business	earnings	earnings	up	Aetna Inc. Q1 Net Profit Adjusted USD 973.00M Vs USD 939.00	[illegible]	[illegible]	0
05/01/2018 15:44:09	Aetna Inc.	100	-0.26	365	business	earnings	earnings-per-share-expectations	below-expectations	Aetna Inc. Reports Q1 EPS-Non-GAAP USD 2.82 Vs Consensus	[illegible]	[illegible]	0
05/01/2018 15:44:09	Aetna Inc.	100	0.4	365	business	earnings	earnings-per-share	up	Aetna Inc. Q1 EPS-Non-GAAP USD 2.82 Vs USD 2.54	[illegible]	[illegible]	0
05/01/2018 15:44:09	Aetna Inc.	100	-0.23	365	business	earnings	ebit-expectations	below-expectations	Aetna Inc. Reports Q1 EBIT Adjusted USD 1.30B Vs Consensus USD 1.33B	0	0	0
05/01/2018 15:44:09	Aetna Inc.	100	0.37	365	business	earnings	ebit	up	Aetna Inc. Q1 EBIT Adjusted USD 1.30B Vs USD 1.26B	0	0	0
05/01/2018 15:44:09	Aetna Inc.	100	-0.2	365	business	earnings	ebitda	down	Aetna Inc. Q1 EBITDA Adjusted USD 1.50B Vs USD 1.63B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	0.27	365	business	earnings	ebitda-expectations	meet-expectations	Aetna Inc. Reports Q1 EBITDA USD 1.52B Vs Consensus USD 1.52B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	-0.22	365	business	earnings	ebitda	down	Aetna Inc. Q1 EBITDA USD 1.52B Vs USD 1.70B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	-0.2	365	business	earnings	ebita	down	Aetna Inc. Q1 EBITA USD 1.42B Vs USD 1.54B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	0.43	365	business	earnings	ebit-expectations	above-expectations	Aetna Inc. Reports Q1 EBIT USD 1.38B Vs Consensus USD 1.37B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	-0.2	365	business	earnings	ebit	down	Aetna Inc. Q1 EBIT USD 1.38B Vs USD 1.45B	0	0	0
05/01/2018 15:44:10	Aetna Inc.	100	0.38	365	business	dividends	dividend	up	Aetna Inc. Q1 Dividend Per Share USD 0.51 Vs USD 0.50	0	0	0
05/01/2018 15:44:10	Aetna Inc.	[illegible]	[illegible]	365	business	earnings	earnings-expectations	above-expectations	Aetna Inc. Reports Q1 Net Profit Reported USD 930.75M Vs Consensus USD 92	0	0	0
05/01/2018 15:44:10	Aetna Inc.	[illegible]	[illegible]	365	business	earnings	earnings	up	Aetna Inc. Q1 Net Profit Reported USD 930.75M Vs USD -381.00M	0	0	0
05/01/2018 20:05:43	Deutsche Bank AG	[illegible]	[illegible]	365	business	stock-prices	stock-price	loss	10:33940266816	NA	0	0
05/01/2018 20:05:43	Deutsche Bank AG	99	-0.58	147	business	analyst-ratings	analyst-ratings-change	negative	Deutsche Bank Aktiengesellschaft (DB) Hit 52-Week Low	0	-0.18	0
05/01/2018 20:07:04	Virgin Australia Holdings Ltd.	100	0.49	365	business	partnerships	partnership		Virgin Atlantic and Virgin Australia Expand Codeshare Relationship	0.04	-0.22	1
05/01/2018 20:07:04	Virgin Atlantic Airways Ltd.	100	0.49	365	business	partnerships	partnership		Virgin Atlantic and Virgin Australia Expand Codeshare Relationship	0.04	-0.22	1
05/01/2018 20:09:34	AllianceBernstein Holding L.P.	100	0	365	business	assets	facility	relocation	AllianceBernstein Move to Nashville to Begin Later This Year -- Sources	0	-0.22	0
05/01/2018 20:09:34	AllianceBernstein Holding L.P.	100	0.06	365	business	assets	headquarters-change	[illegible]	llianceBernstein Holding LP to Move Headquarters and Most Staff to Nashvil	0	-0.34	0
05/01/2018 20:07:12	Western Union Co.	100	0.59	365	business	equity-actions	buybacks	[illegible]	estern Union Reports First Quarter Results	0	0.12	0
05/01/2018 20:09:45	Sunesis Pharmaceuticals Inc.	100	0	365	business	investor-relations	conference-call	[illegible]	:33940293349	NA	0	0
05/01/2018 20:10:00	Five9 Inc.	100	0.53	365	business	labor-issues	executive-appointment		ive9 Appoints Industry Veteran Rowan Trollope As New CEO >FIVN	0.1	0.04	1
05/01/2018 12:22:40	Lightbridge Corp.	100	0	153	business	marketing	conference		Lightbridge Corporation to Present at the 3rd Annual Disruptive Growth and H	0.02	0.18	1
05/01/2018 12:26:16	FedEx Corp.	100	0.39	365	business	products-services	business-contract		BRIEF-Plug Power and Workhorse Provide FedEx With Fuel Cell-Powered Elec	0.02	-0.48	0

Multiple updates of the same earnings release

Multiple sentiment measures

Including non-US

Including non-earnings/revenue news

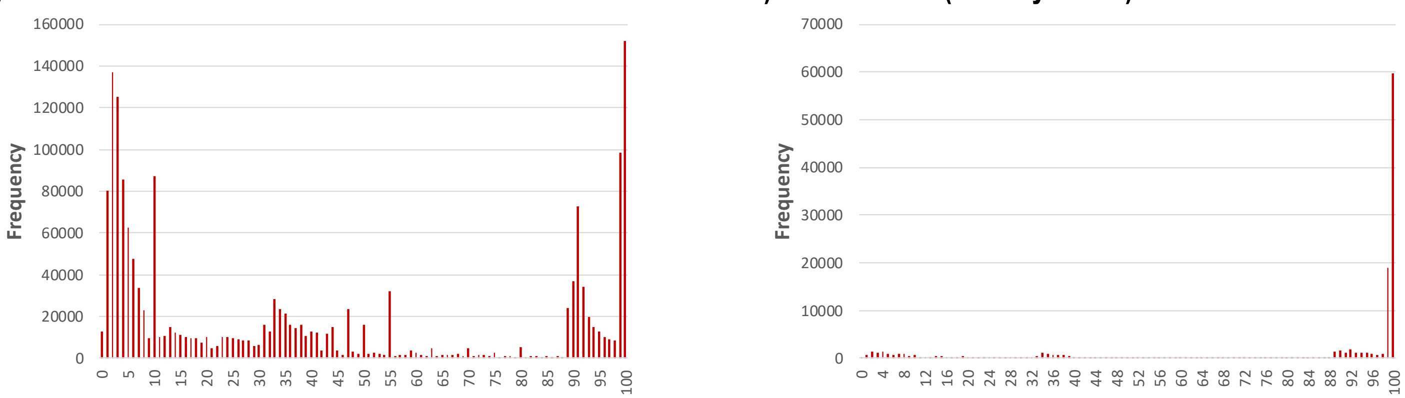
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Relevance

To construct predictive signals, we typically limit our analysis to only those news stories that are directly relevant and fresh to a given company. As shown in Figure 2, once we remove those "old" news (i.e., news with similar coverage in the previous 30 days), the vast majority stories tend to be highly relevant. For this research, we exclude irrelevant and duplicate news items by restricting our analysis to stories with a Relevance Score of 90 and above. Furthermore, we use only those unique/novel stories within a 30-day period at each point-in-time.

Figure 2 The Distribution of RavenPack Relevance Score

A) All News | **B) "New" News (Novelty > 30D)**



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

SIGNAL PROCESSING: ENTITY, EVENT AND TOPIC DETECTION

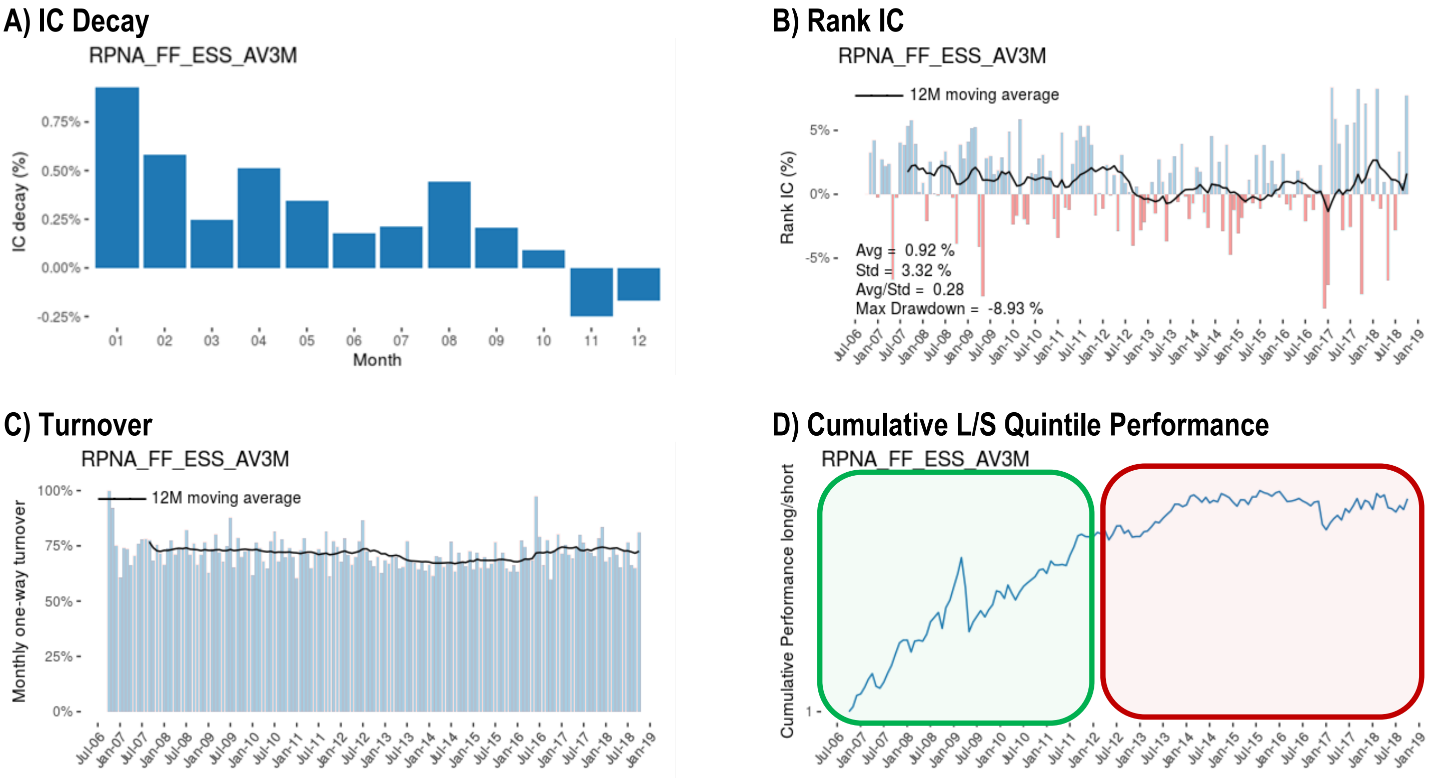
In RavenPack Analytics, news is received and processed through an automated real-time data stream made up of several parallel software components. Events in RavenPack's Taxonomy are defined using thousands of proprietary template programs and Part-of-Speech tagging. RavenPack systematically categorizes stories into a simple set of themes (i.e., Topics, Groups, Types and Sub-Types), which is essential for investment research. Next, the context in which the entities are mentioned is analyzed to determine what role they played in the story. Events can have more than one participant with only one being the principal.

The average overall latency of a processed news story at any time is at 300 milliseconds. The latency is defined as the average time it takes a story to get through the system from the time RavenPack receives it, to the time at which it leaves the Event Distribution Server at the RavenPack Data Center. RavenPack Analytics delivers sentiment analysis on more than 192,000 entities in over 130 countries and covers over 98% of the investable global market. It provides exhaustive coverage by entity type with more than 100,000 places, 43,000 companies, 1,000 products, more than 100 nationalities, currencies and commodities.

THE HEADLINE SENTIMENT SIGNAL – ESS – ONLY HAS MODEST PREDICTIVE POWER

Figure 3 shows the historical performance of the headline RavenPack sentiment signal (ESS) in the US (i.e., the Russell 3000 universe). In this case, the sentiment signal is computed using a three-month moving average. Similar to our previous experience, the performance of the naïve sentiment factor is not particularly strong. At the monthly frequency, the sentiment factor loses significant predictive power, especially since 2012.

Figure 3 Event Sentiment Score (ESS), Three-Month Aggregate Factor Performance (US)

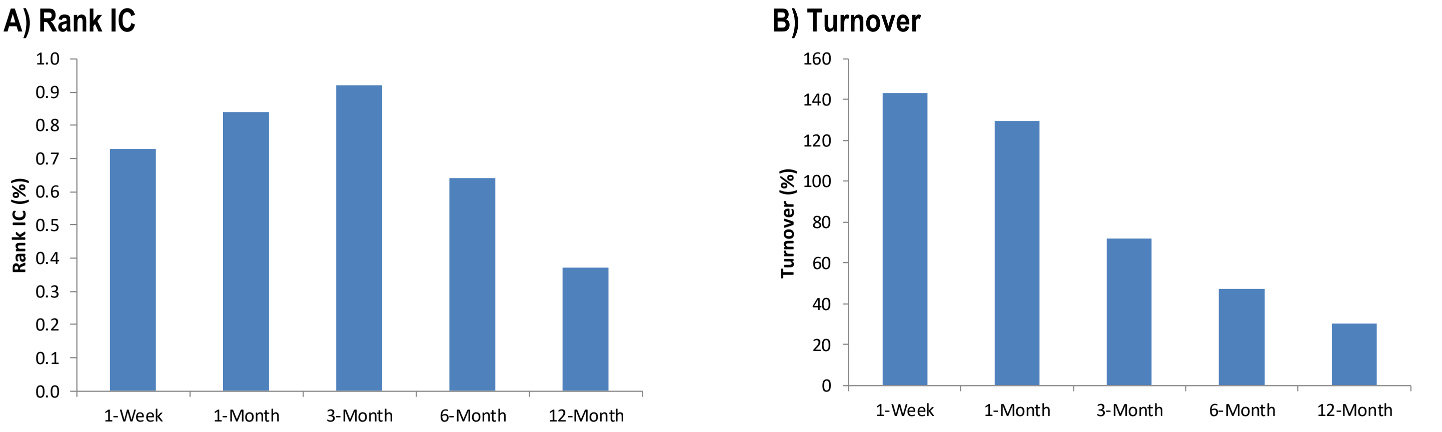


Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Signal Aggregation Window

Heston, et al [2015] finds that daily news is a short-term signal and can mostly predict stock returns for one to two days. They further propose that a longer aggregation window can increase the longevity. As shown in Figure 4, we test various rolling windows to compute our sentiment signal, from one week to 12-month. The three-month window is chosen, due to its highest Sharpe ratio and modest turnover.

Figure 4 Event Sentiment Score (ESS), Performance with Different Aggregation Window (US)



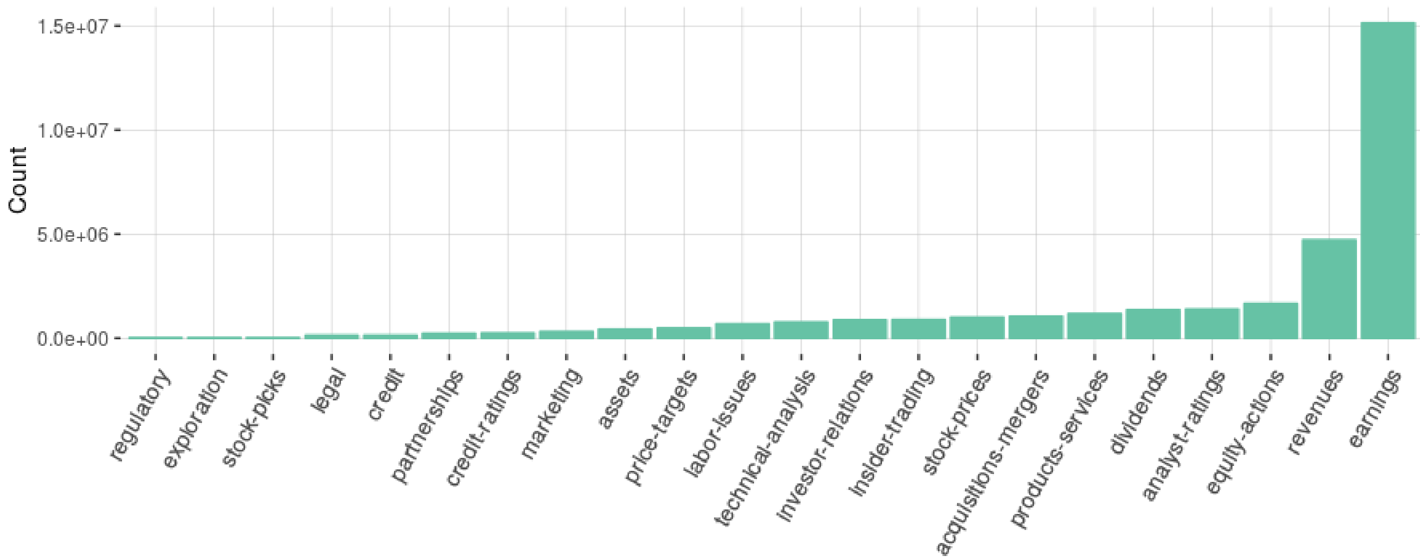
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

RAVENPACK EVENT CATEGORY

As elaborated in the previous section, RavenPack tags each news story to an entity and event category. There are four levels of events – Topic, Group, Type, and Sub-Type. As shown in Figure 5, at the Group level, news stories are overwhelmingly dominated by earnings and revenue related events. In fact, Earnings, revenue and stock-price related stories account for almost 90% of events. As a result, our previously discussed sentiment factor is highly skewed towards these few categories. Most of such news are already captured by various traditional analyst estimate and price momentum/reversal factors and therefore, may not provide much additional information.

The underrepresented, low frequency events such as regulatory, legal and labor-issues can be quite relevant to asset returns and are generally not well captured by traditional stock selection factors.

Figure 5 Frequency Count by RavenPack Event Groups (Level II)

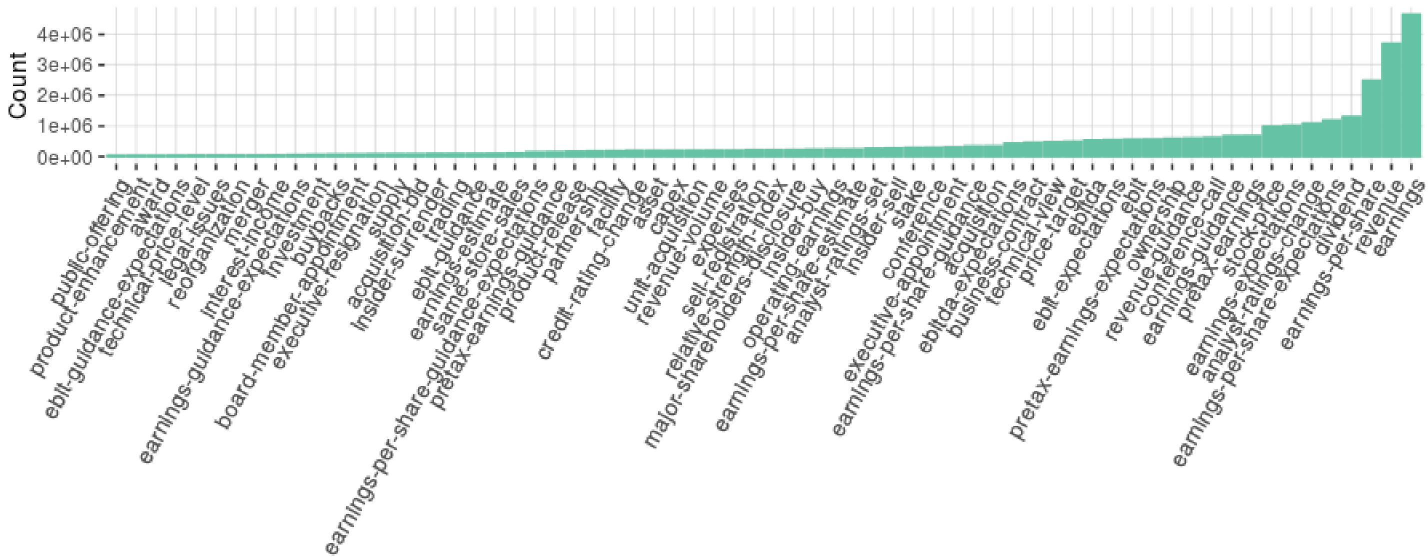


Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

The next event category level, i.e., Event Type, includes more interesting critical news such as legal-issues, executive resignation, partnership, product-release, reorganization, credit-rating change, and buyback (see Figure 6).

Zooming into event-based sentiment could be promising. Some event types may draw attention but are not informative in asset pricing. Equity analyst or creditor opinions are certainly market moving events, but may not reveal any new information, especially if the views are not materially different from analysts' previous commentaries. As highlighted by Barber and Odean [2008], individual investors are more likely to buy attention-grabbing stocks irrespective of the arrival of any new information. Such price action tends to reverse quickly. On the other hand, price moves driven by new information are likely to persist in the long term. For each category of news, it is also important to track the volume of news flow and its changes.

Figure 6 Frequency Count by RavenPack Event Types (Level III)



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

RavenPack Aggregate Sentiment Score by Event Groups

RavenPack has a few pre-defined sentiment measures. In this research, we study four of them in detail:

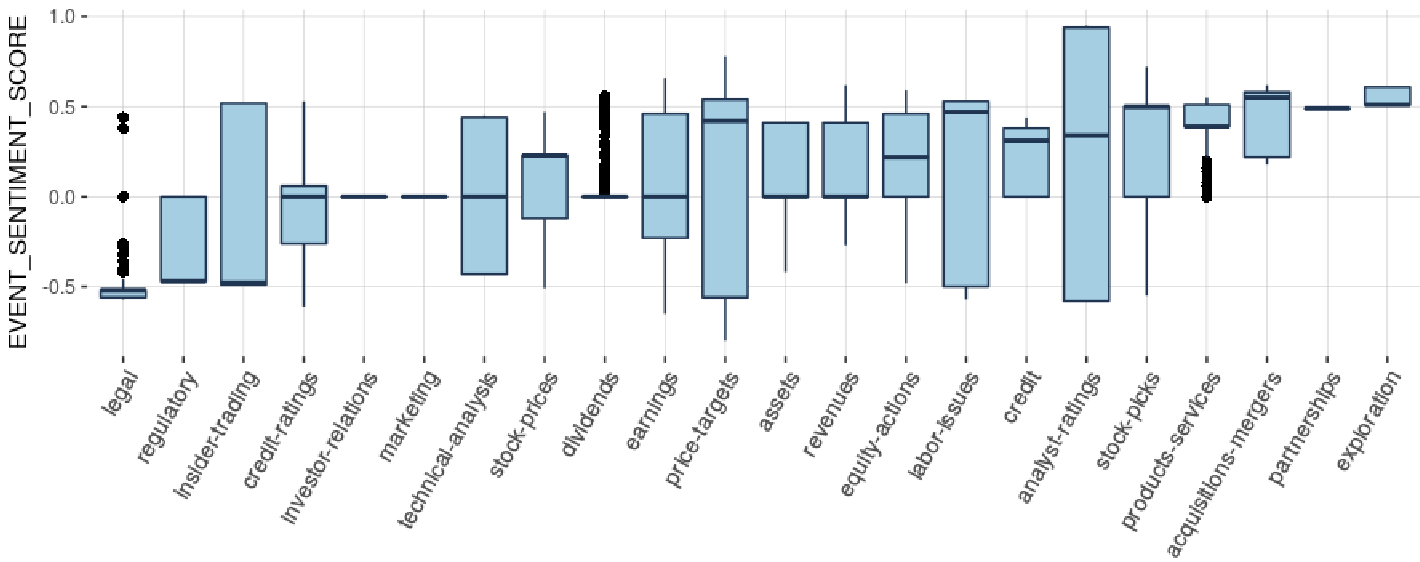
- ESS – Event Sentiment Score
- CSS – Composite Sentiment Score
- MCQ – Multi Classifier for Equities
- NIP – News Impact Projections

Event Sentiment Score (ESS)

The Event Sentiment Score is determined by systematically matching each new story with a predefined library of classified entity specific events. The score ranges from -1 to $+1$. The average sentiment across event categories shows strong variations (see Figure 7). Some events have predominantly negative sentiment, e.g., regulatory, legal, credit rating. On the other hand, partnership, M&A, product and services related events are mostly positive. Surprisingly, labor-issues also have a positive average sentiment but with some large negative outliers. It is possible that firms are more likely to make labor-issues related disclosures, when there is a resolution than otherwise.

Figure 7 RavenPack Average Event Sentiment Score by Event Groups

A) ESS by Group



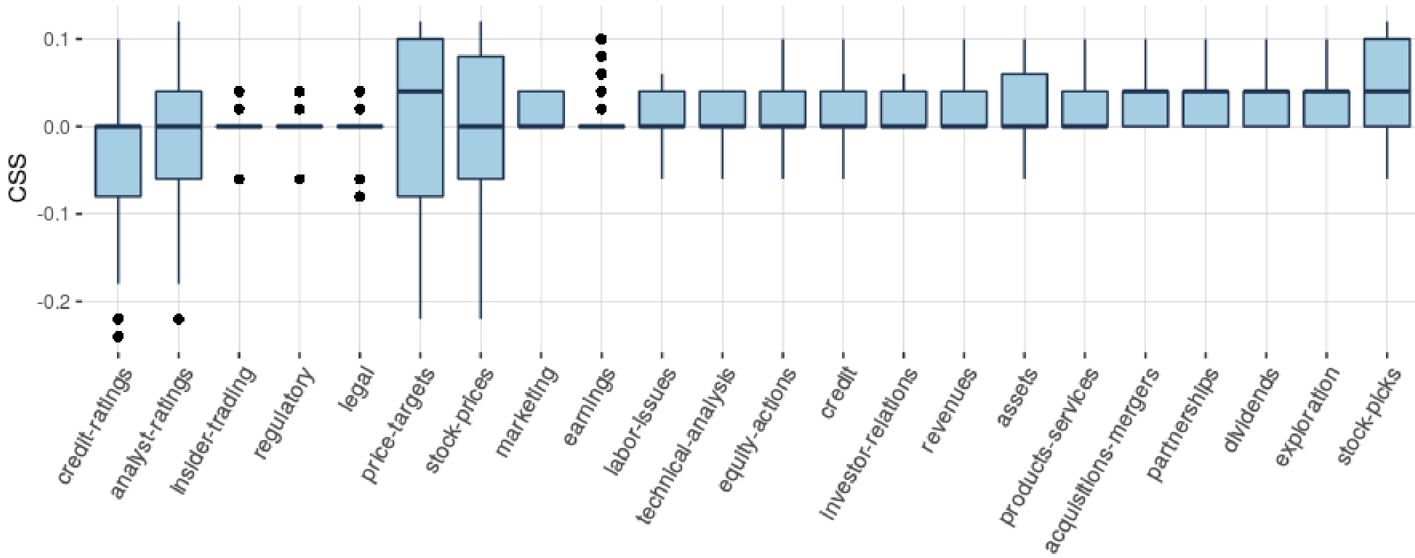
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Composite Sentiment Score (CSS)

The second sentiment measure we study is the CSS (Composite Sentiment Score). It combines five RavenPack sentiment analytics (PEQ, BEE, BMQ, BAM, and BCA) using a pre-defined set of rules. The direction of the CSS score is determined by looking at emotionally charged words and phrases and by matching stories typically rated by experts as having short-term positive or negative share price impact. The model was trained using tick data from approximately 100 large cap stocks. Similar to ESS, the CSS score is also granular and varies between −1 and +1. The average CSS scores for credit- and analyst-rating issues are both negative (see Figure 8). Because the CSS was developed and calibrated using share prices, it is possibly due to that analyst downgrades have larger (negative) impact than analyst upward revisions.

Figure 8 RavenPack Average Composite Sentiment Score by Event Groups

A) CSS by Group



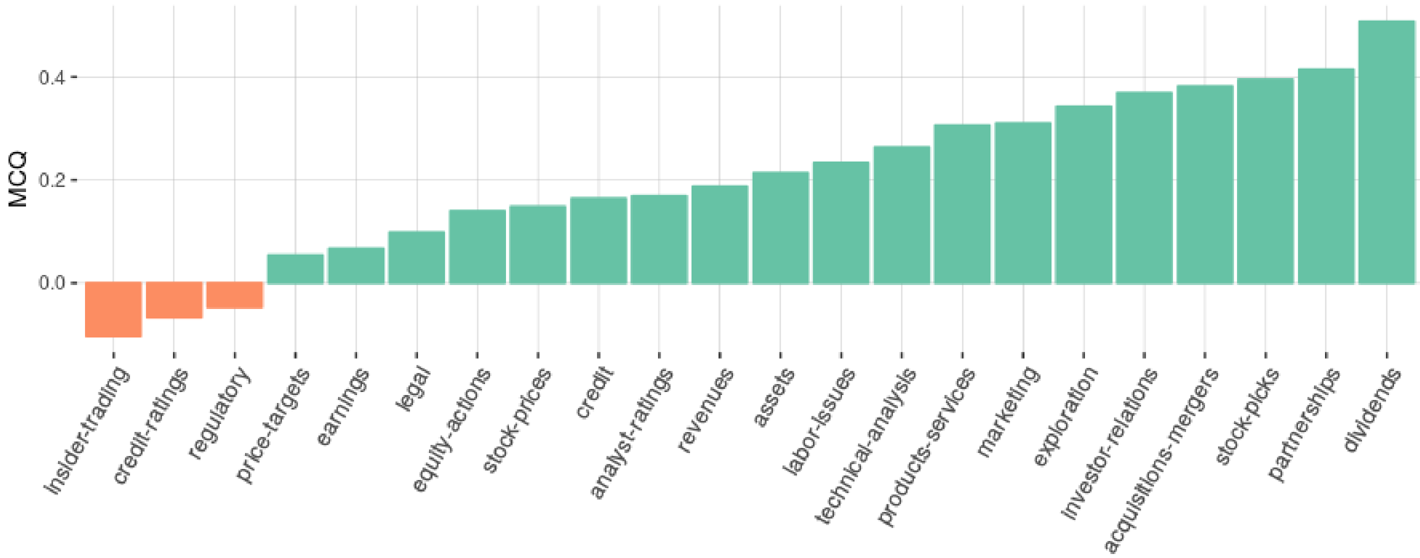
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Multi Classifier for Equities (MCQ)

The last sentiment variation we explore is the MCQ (Multi Classifier for Equities). It classifies news stories into three categories – negative, neutral, or positive, respectively. It represents the news sentiment based on the tone for the most relevant entities mentioned in a story. Similar to CSS, it is also derived from a combination of other RavenPack matrices namely BMQ, BEE, BCA, and ANL-CHG. Not directly related news (i.e., relevance score below 90) are assigned a neutral MCQ score. As shown in Figure 9, the average MCQ score is largely in-line with the ESS sentiment measure, in that analyst ratings and labor issues are more likely to be positive.

Figure 9 RavenPack Average MCQ Score by Event Groups

A) MCQ by Group



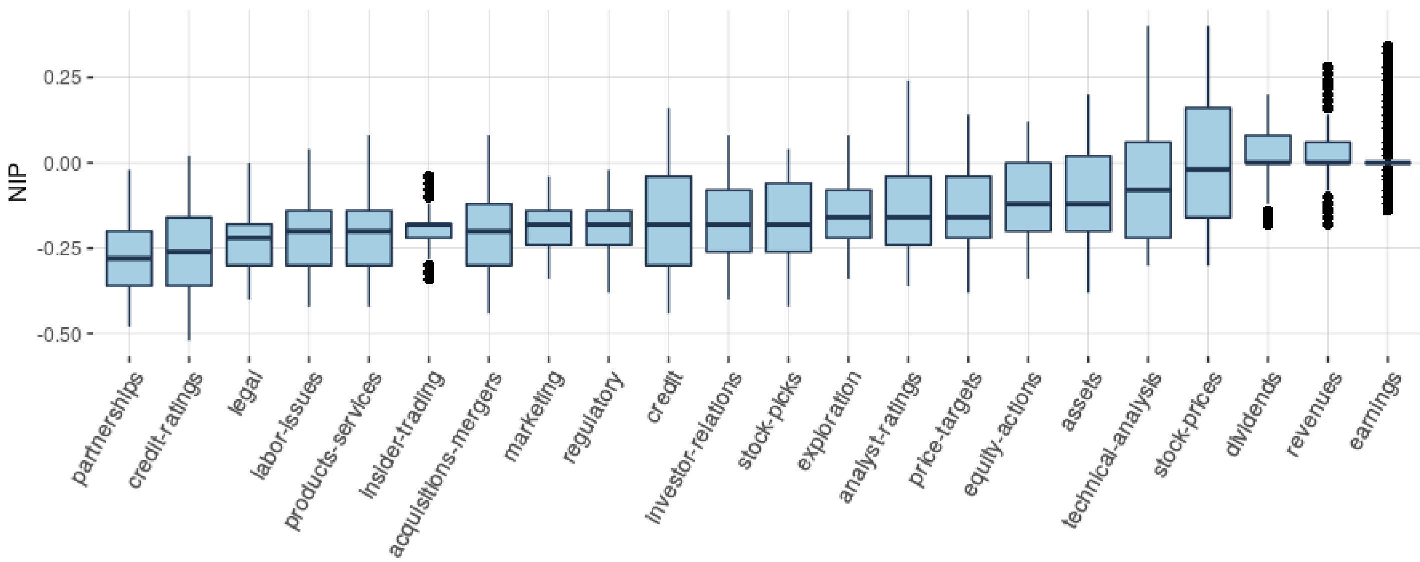
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

News Impact Projections (NIP)

Unlike the previous three sentiment measures, the NIP is designed to predict short-term (i.e., the next two hours) volatility impact. The NIP score tries to predict whether relative volatility of the entity is high or low in the next two hours following a news-flash given the language of the text. The NIP was also calibrated on a pre-classified set of large cap companies. It is a granular score between -1 and $+1$, where values above 0 indicate higher impact on volatility. The higher the NIP, the higher the confidence based on the NIP. As shown in Figure 10, earnings and revenue related events are likely to lead to higher volatility (with significant outliers), while other events typically reduce uncertainty.

Figure 10 RavenPack Average NIP Score by Event Groups

A) NIP by Group

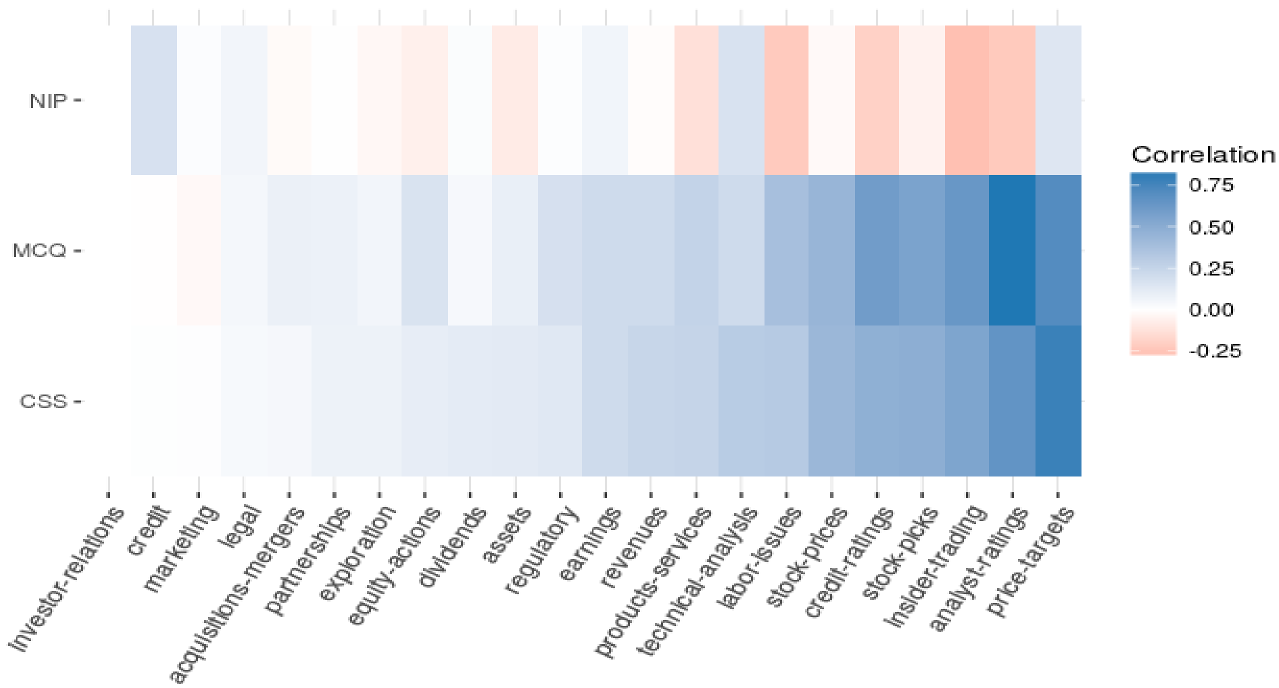


Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

CORRELATION AMONG THE FOUR RAVENPACK FACTORS

Figure 11 shows the correlation of the MCQ, CSS and NIP with the ESS. Since ESS, MCQ and CSS are all sentiment measures, it is not surprising to see that they are positively correlated. Similarly, since the NIP is designed to predict volatility – negative news is more likely to lead to higher volatility; therefore, the NIP is mostly negatively correlated to the ESS. Analyst-ratings and price target events are typically more correlated, while the four models produce very different predictions for investor-relation, credit, marketing and legal news.

Figure 11 Correlation of the CSS, MCQ and NIP with the ESS (Global)



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

A Comprehensive News-Based Event Study

To assess the accuracy and predictive power of RavenPack's news event classification and sentiments algorithms, in this section we conduct a series of event studies. Major corporate actions such as earnings announcements and partnerships are certainly market moving events. However, around the news announcement date, there are often other company-specific factors impacting a stock's return. By combining many companies that underwent similar type of events at different points in time, these company-specific idiosyncrasies are largely eliminated. Therefore, we can better extract the common patterns around event date. Event studies are helpful to understand investor behavior and therefore design more effective investment strategies. For example, if we see a significant and persistent negative (positive) price trend leading to negative (positive) news, it may indicate information leakage prior to public release of news. A sustained price drift in the same direction post the event day can be attributed to market underreaction and slow information diffusion, which presents profitable investment opportunities. Similarly, the market may overreact to news, which leads to price reversal.

In the next few sections, we present an event study on a few prominent event types. For each event, we also need to break down the sample into positive and negative stories. In general stocks move in the direction of expected positive/negative news well before announcement. This may either be due to information leakage or sequential release of similar news.

Merger and Acquisition

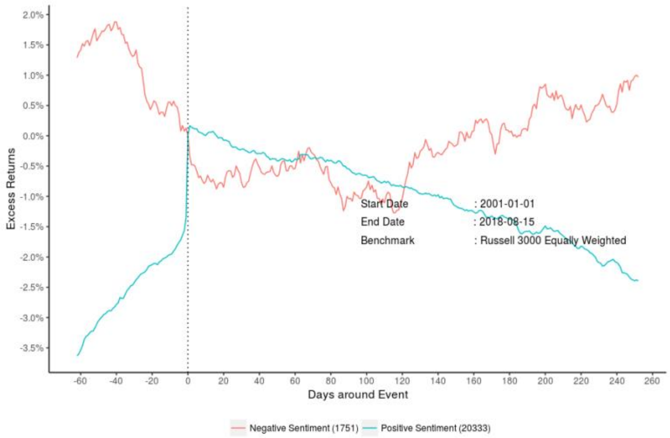
In a classic M&A transaction, one party – the acquirer (e.g., a company, a private equity firm, a government entity) attempts to take control of the other party – the target (e.g., a company, a division of a company), which often involves cash, stocks and other considerations. On the takeover announcement day, the target's share often jumps significantly higher on the news (see *Machine Learning Takeovers*, Sheng, et al [2017b], where we introduce our takeover prediction model SMAP). Obviously, not all takeovers eventually succeed. Failed deals are likely to lead to significant losses to risk arbitrageurs who have bought the target's stocks (see *Systematic Alpha from Risk Arbitrage*, Sheng et al [2018a]), where we develop our systematic risk arbitrage strategy SARA).

In the RavenPack data, for simplicity (and to maximize sample size), we focus on all M&A related news and the associated sentiment scores. As shown in Figure 12, we split all M&A news into "Mergers" and "Acquisitions". Furthermore, within each event category, we track the pre- and post-announcement performance by "positive" and "negative" news sentiment.

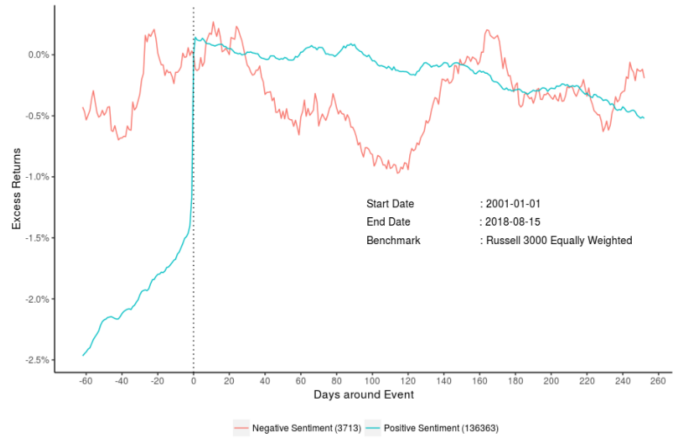
In the RavenPack database, most M&A news have positive sentiment. Stocks run up significantly before the announcement such positive news, which may indicate both information leakage and/or consecutive news in the same direction. Stocks rally on the announcement day (the average announcement-day return is almost 150bps). Post-announcement day, however, stocks consistently underperform, which suggest potential mean reversal trading opportunities. For negative M&A news, the underperformance during the pre-announcement period and on the announcement day is far more modest. The post-announcement drift (in the same direction, i.e., downward) is more persistent, hinting at price momentum effect.

Figure 12 Event Study for M&A Event (US)

A) Merger



B) Acquisition



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

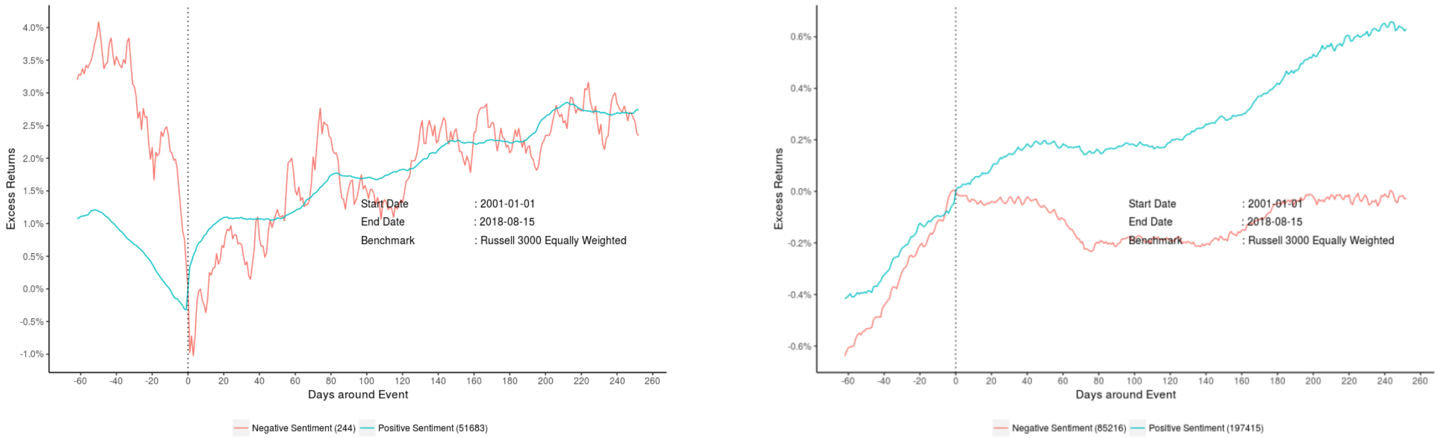
BUYBACK AND OWNERSHIP

As documented in our previous research, companies tend to announce share buybacks when the management believes their own stocks to be undervalued. Most buyback news has a positive sentiment. Buybacks are typically initiated on the back of some short-term underperformance; therefore, the pre-announcement period return is mostly negative, regardless whether the news is positive or negative. Similarly, the announcement-day return and post-announcement drift are positive on average – for both positive and negative news stories (see Figure 13A).

Sentiment for ownership related events is more evenly distributed between positives and negatives. Before ownership news, stocks have already rallied. The news sentiment is quite relevant in this case, as the positive announcement drift is significant (and positive) only for positive news (see Figure 13B). In *Port@ble Ownership* (see Alvarez, et al [2018]), we discuss ways of using detailed ownership data and our performance attribution tools to generate alpha.

Figure 13 Event Study for Buyback and Ownership (US)

A) Buyback **B) Ownership**



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

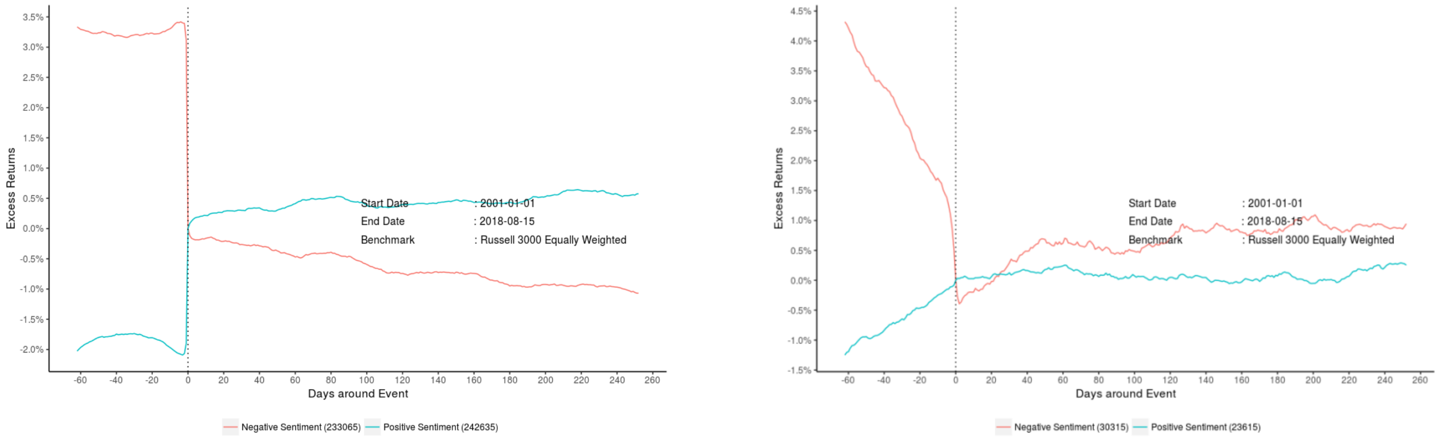
Rating Change

Analyst rating (these are predominantly sell-side equity analysts) and credit rating changes have an even split of positive and negative sentiment events. As expected, there is almost no pre-announcement effect for analyst rating changes. In the heightened regulatory environment, information leakage on the sell-side is extremely rare. The immediate market reaction to analyst revisions is as expected. Post-announcement drift is in the intuitive direction, albeit moderate and slightly larger for negative revisions (see Figure 14A).

Credit analysts, due to their nature of being conservative, tend to be more reactive than proactive. They typically downgrade (upgrade) after a period of share price underperformance (outperformance). The market reacts to negative news and downgrades far more strongly than upgrades. In the end, a going concern credit rating causes more damage to shareholders than a credit upgrade. Equity investors exhibit a classic "sell first and think about it later" behavioral pattern, in that the market overreacts to negative news, which leads to a reversal pattern post negative credit events (see Figure 14B).

Figure 14 Event Study for Analyst and Credit Rating Change (US)

A) Analyst Rating Change **B) Credit Rating Change**



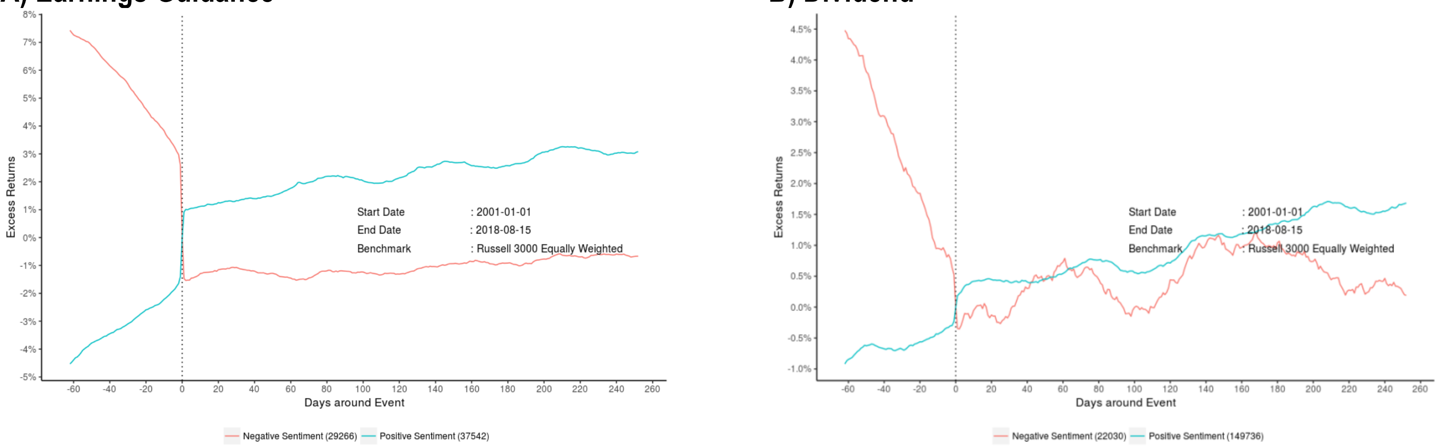
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Earnings Guidance and Dividend

Earnings guidance and dividend announcement closely resemble analyst rating changes. In the pre-announcement period, stocks have already shown strong outperformance (underperformance) for positive (negative news). In the post-announcement period, positive news tends to be more persistent (see Figure 15). In our previous study on crowdsourced earnings estimates from Estimize (see *Crowdsourcing Earnings and Revenue Estimates*, Sheng, et al [2017a]), we discuss how more accurate earnings estimates can better capture post-earnings announcement drift.

Figure 15 Event Study for Earnings Guidance and Dividend (US)

A) Earnings Guidance **B) Dividend**



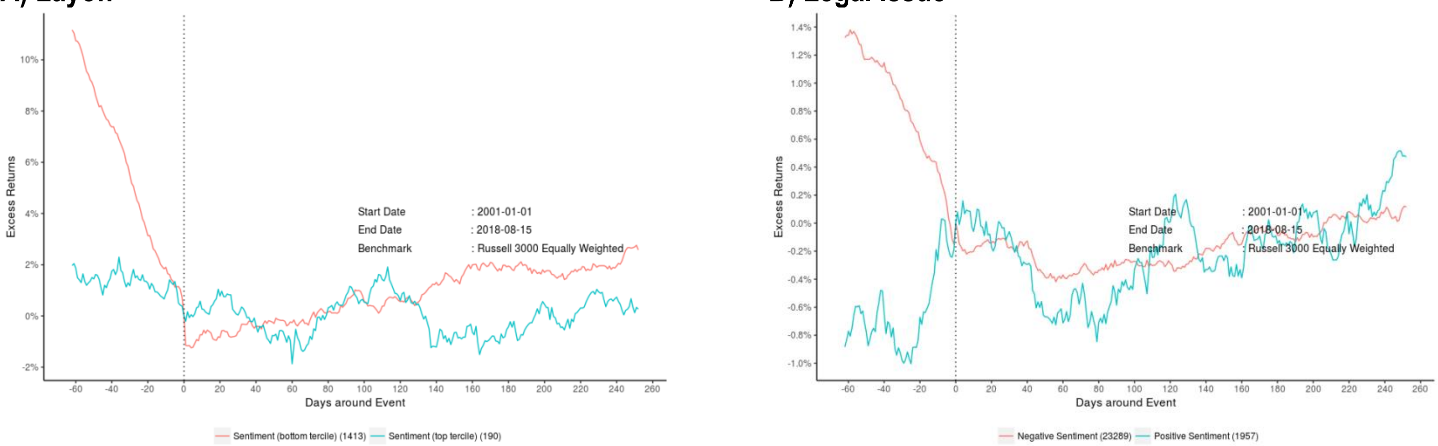
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

LAYOFFS AND LEGAL ISSUES

Some corporate events such as layoffs and legal issues are inherently negative. As a result, such news is also classified as mostly negative by RavenPack. Due to the inflexibility of the labor market, companies typically announce layoffs after their business has been suffering for a period of time. Similarly, because of the complexity in many legal issues, the market is likely to be aware of such problems before the official news break out. Therefore, the pre-announcement return is mostly negative for both event types (Figure 16A and B). Investors tend to overreact to bad news on the announcement day, which produces a positive post-announcement drift (i.e., a reversal pattern) for both layoffs and legal issues. For a more extensive study on corporate governance and accounting quality see *Quant CSI*, Jussa, et al [2017a].

Figure 16 Event Study for Layoff and Legal Issue (US)

A) Layoff **B) Legal Issue**



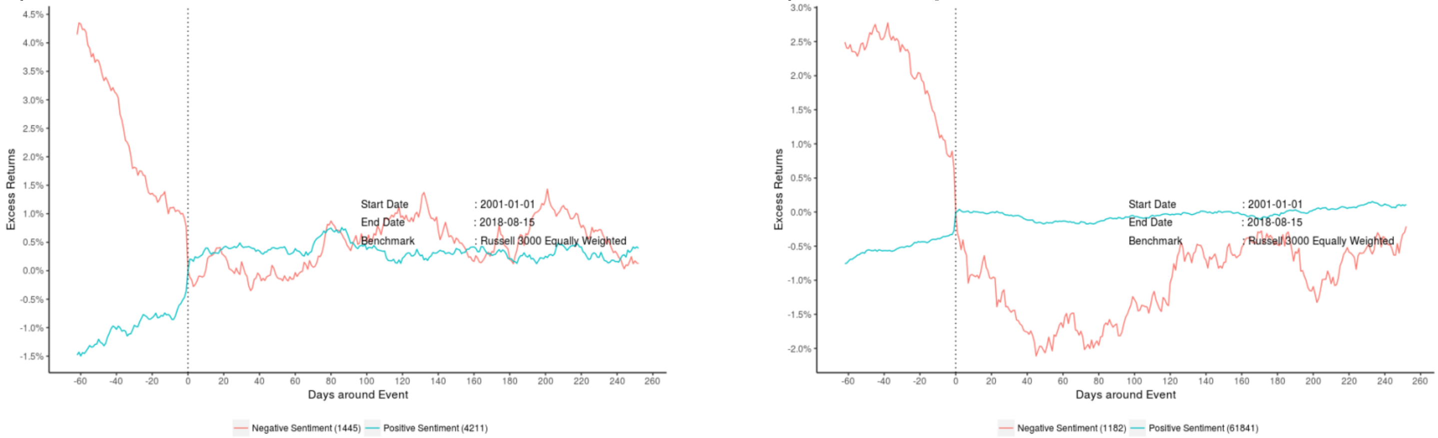
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

MARKET SHARE AND PARTNERSHIP

Market share and partnership news events offer significant event day abnormal returns but have limited scope for post-announcement alpha (see Figure 17).

Figure 17 Event study for Market share and Partnership (US)

A) Market Share **B) Partnership**



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Sector-Specific News Event (Clinical Trials and Same Store Sales)

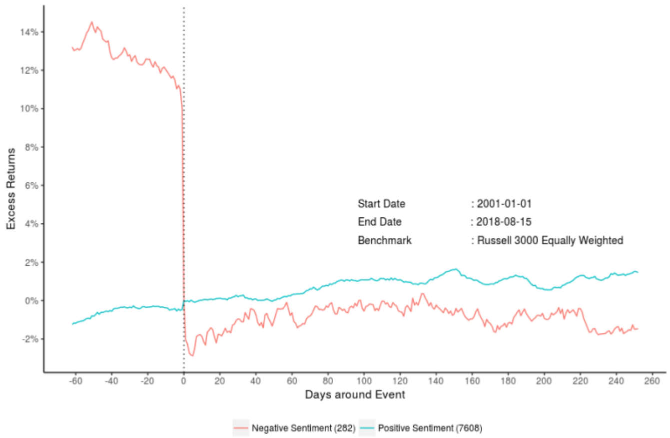
For two of the most dynamic industries – the BioPharma and retailing[5], RavenPack offers two interesting categories of news – clinical trials and same store sales. Clinical trials for pharmaceutical companies involve large capital investment and long multi-year development cycles; therefore, are typically market moving events. A negative outcome of a clinical trial study on average leads to large correction of -10% on the announcement day, albeit post-announcement drift is minimal (see Figure 18A). On the other hand, positive clinical trial outcomes are much less headline grabbing, with small announcement day return, but lead to more persistent post-announcement drift.

Same-store sales data is widely tracked for retail companies. Most retailers report their own sale store growth periodically. There are also third-party vendors that track retail activities using credit card, geolocation, and satellite data (see *Space –The Next Alpha Frontier*, Javed, et al [2017b], for how same store sales and satellite data can be used). Most of the information is leaked well in advanced for these events. We observe a strong momentum effect leading up to the same store sales announcement, as well as persistent post-announcement drift, especially for positive news (see Figure 18B)

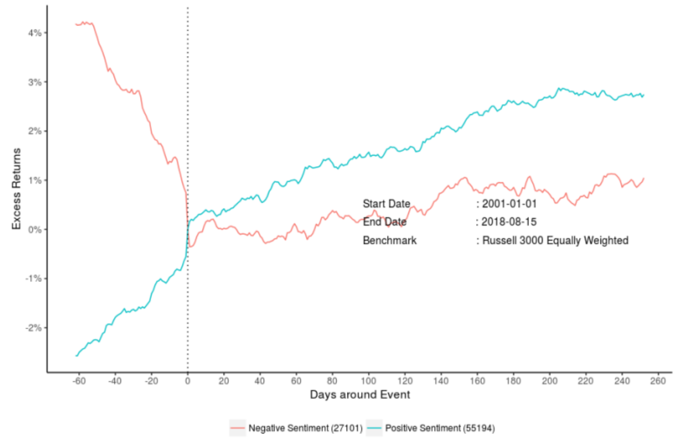
[5] For investors, stocks in these two sectors are also very liquid with wild swings, offering tremendous alpha opportunities.

Figure 18 Event Study for Sector Specific Event (US)

A) Clinical Trial



B) Same Store Sales



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

A Machine Learning Approach to Extract Event-Based Sentiment

The traditional way of using news sentiment signals ignores the fact that not all news stories are the same – different types of events may have very different implications to future asset returns. Similarly, conventional event studies may not necessarily consider the positive/negative nature of each event. Furthermore, most investors do not trade on corporate events on their own. In this section, we attempt to combine news, sentiment, event, news volume, and market behavioral biases together – all in a systematic way, using cutting edge machine learning techniques.

News Volume – Does No News Mean Good News?

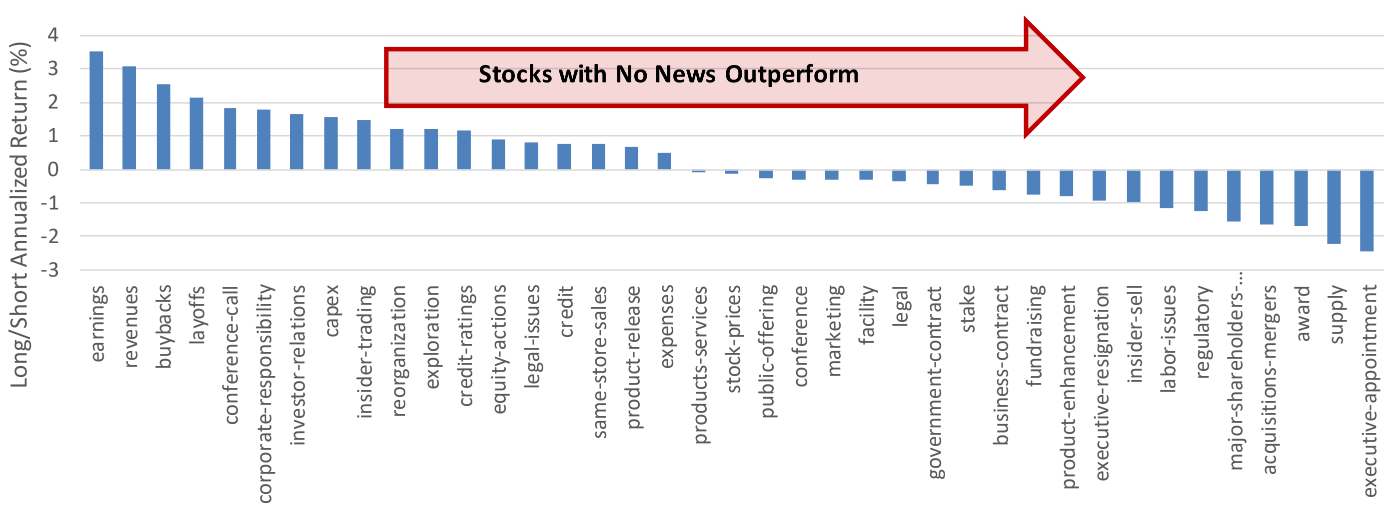
First of all, let's offer a preliminary analysis on news volume. The frequency and amount of information disseminated by companies and coverage by media can have significant impact on stock returns. In academia, there are conflicting views about whether news coverage itself is positive or negative for stock returns. Fang and Peress [2009] suggest that stocks with no media coverage earn higher returns, i.e., no news is good news. In contrast, Mitchell and Mulherin [1994] find a weak positive relationship between news volume and firm-specific returns.

Before we conduct our study, we want to highlight two issues:

- News coverage in absolute terms suffers from significant size bias. Larger firms are more likely to have high news coverage. Hence, our news volume factor needs to be adjusted for size.
- Furthermore, a simple count of all news stories fails to capture the huge variations in event types and the associated tone/sentiment.

As shown in Figure 19, the performance of our size-adjusted news volume factor varies considerably depending on the type of news event. For earnings, revenue and buybacks related news, volume itself is a positive signal. On the other hand, for executive appointment, M&A related stories, no news is good news.

Figure 19 News Volume Factor (Size Adjusted, US)



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

EVENT-BASED SENTIMENT FACTORS

As elaborated in the previous sections, there are complex interactions among news event categories, sentiment, volume, volatility and market impact. In this section, we use two advanced machine learning techniques to untangle the complex relationship among the following variables and build a predictive model of future stock returns:

- **News Event.** We rely on RavenPack's news event classification algorithm, by focusing on Groups (Level II) and Type (Level III).
- **News Sentiment.** We use the three pre-defined sentiment measures – E SS, MCQ, and CSS – all computed over different rolling windows.
- **News Volume.** We compute several news volume factors, i.e., frequency of news stories, over various look-back windows. New volume factors are all adjusted for size.
- **Volatility Impact.** RavenPack offers a pre-defined NIP factor, i.e., predicted impact on stock volatility.
- **Market Behavioral Bias.** To measure how the market reacts to each event, we compute a series of event-day return (and abnormal trading volume) factors. Event-day returns (and abnormal trading volumes) are typically calculated on a five-day window, i.e., two days before and two days after each event. In addition, since many news events are sporadic arriving at low frequency, we further aggregate our signals over a few rolling windows (one-, three-, and six-month).

Investment Universe

We divide the global equity market into three universes:

- US (Russell 3000)
- Developed ex US: Canada, Europe, UK, Developed Asia (HK and Singapore), ANZ (Australia and New Zealand)
- Emerging Markets

Please note that our universe definition is different from our traditional 10-region classification. RavenPack only covers English language news; therefore, the coverage is not particularly strong in certain countries. The three-region grouping above gives us a large enough sample to train our machine learning models.

A Machine Learning Approach

Similar to our TMT sector-specific model – TALIA (see *Global TMT Stock Selection Models – Introducing TALIA*, Rohal, et al [2018c]) – we train two machine learning models, for each Event Group (and Type), for each of the above three investment universes:

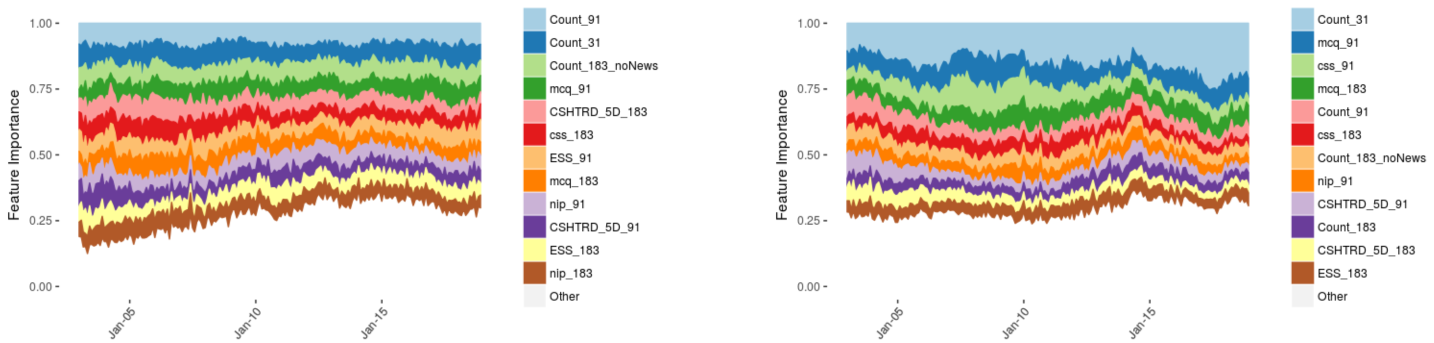
- A linear LASSO model
- A non-linear xgBoost model

An equally weighted LASSO and xgBoost model is further created for each event type, i.e., event-based sentiment factors.

As an example, Figure 20 shows the chosen factors by the xgBoost algorithm (i.e., feature importance) for Legal and Revenue events, respectively. The relative importance of factors varies greatly by event types. However, we do observe a robust mix of event count, sentiment, volatility, and market impact factors.

Figure 20 Feature Importance for Legal and Revenue Event Types (US)

A) Legal **B) Revenue**

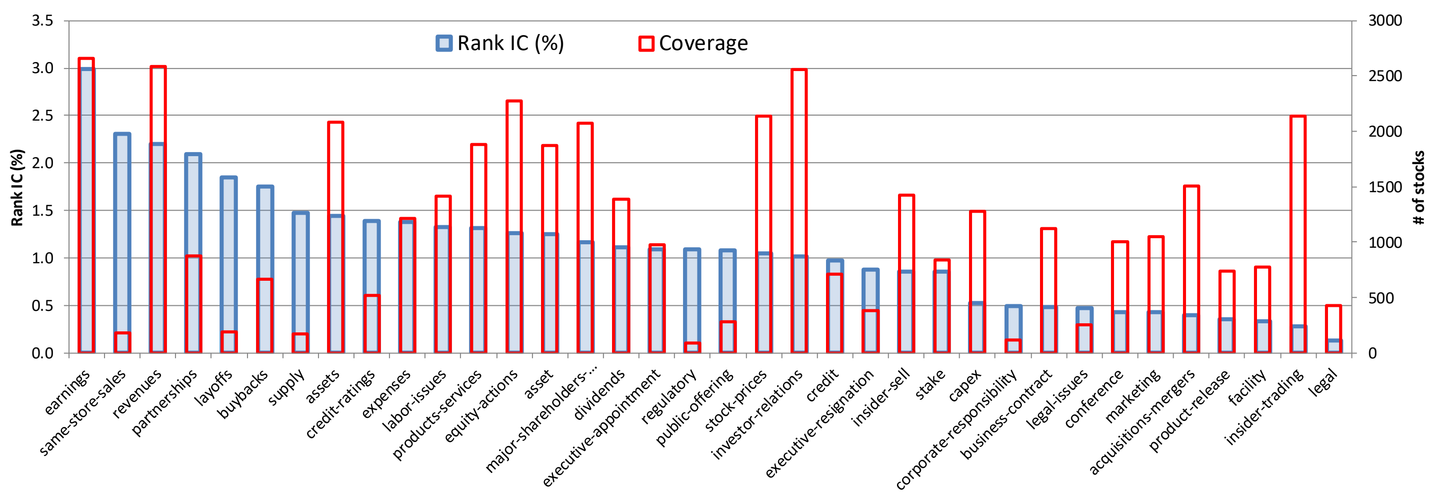


Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Model Performance

Figure 21 shows the performance (as measured by rank IC) and coverage of our event-based sentiment factor for each event type in the US. Both coverage and performance vary greatly across event types. For example, layoffs and same-store sales have low coverage, but strong performance. On the other hand, insider-trades and ownership have broader coverage, but poor cross-sectional predictive power. Earnings- and revenue-related events have both strong coverage and performance.

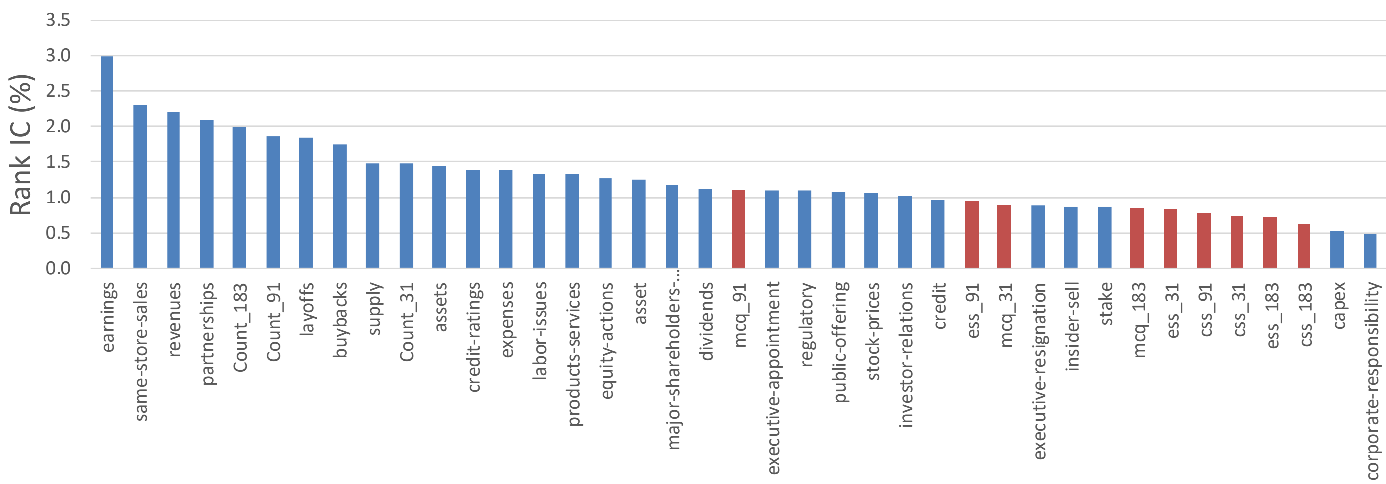
Figure 21 Performance of Event-Based Sentiment Factors in the US



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Furthermore, as shown in Figure 22, our event-based sentiment factors outperform the off-the-shelf sentiment signals (e.g., ESS, MCQ, CSS) considerably.

Figure 22 Rank IC for each Event Groups/Types Composite and Basic Aggregate Sentiment Factors



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Welcome to NICE (News with Insightful Categorical Events)

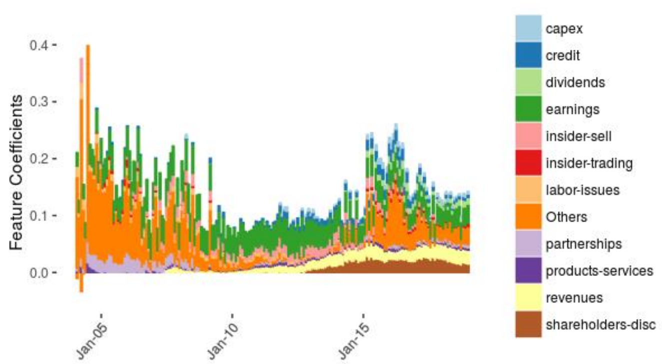
Finally, we use the elastic net machine learning algorithm to combine our suite of event-based sentiment factors together. The final model is called the NICE[6] (News with Insightful Categorical Events) model.

Elastic net is essentially a regularization technique, combining the Ridge (using sum of squared coefficients penalty) and the LASSO (using sum of absolute coefficients penalty) algorithms. The LASSO regularization shrinks many regression coefficients to zero, leaving us with just a couple of events linked to earnings and revenue. Although these event types have strong performance and great coverage, they are likely to be highly correlated to conventional accounting-based factors. On the other hand, the Ridge regression shrinks the regression coefficients towards zero, but not necessarily to zero.

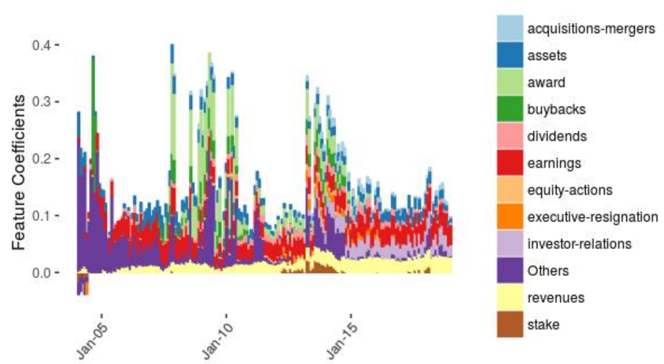
As shown in Figure 23, by blending the Ridge regression into the LASSO, the elastic net algorithm maintains a more balanced allocation among different event categories. Revenue and earning related event types still have the largest weights. Coefficients for other event types differ significantly from one region to another. Shareholder disclosure, partnership, products-services and labor-issues are prominent in the US (see Figure 23A). On the other hand, for Developed markets ex US, investor-relations, assets, award and buybacks events contribute to a meaningful extent to the NICE model (see Figure 23B).

Figure 23 Feature Importance for NICE Composite

A) US



B) Developed Markets ex US



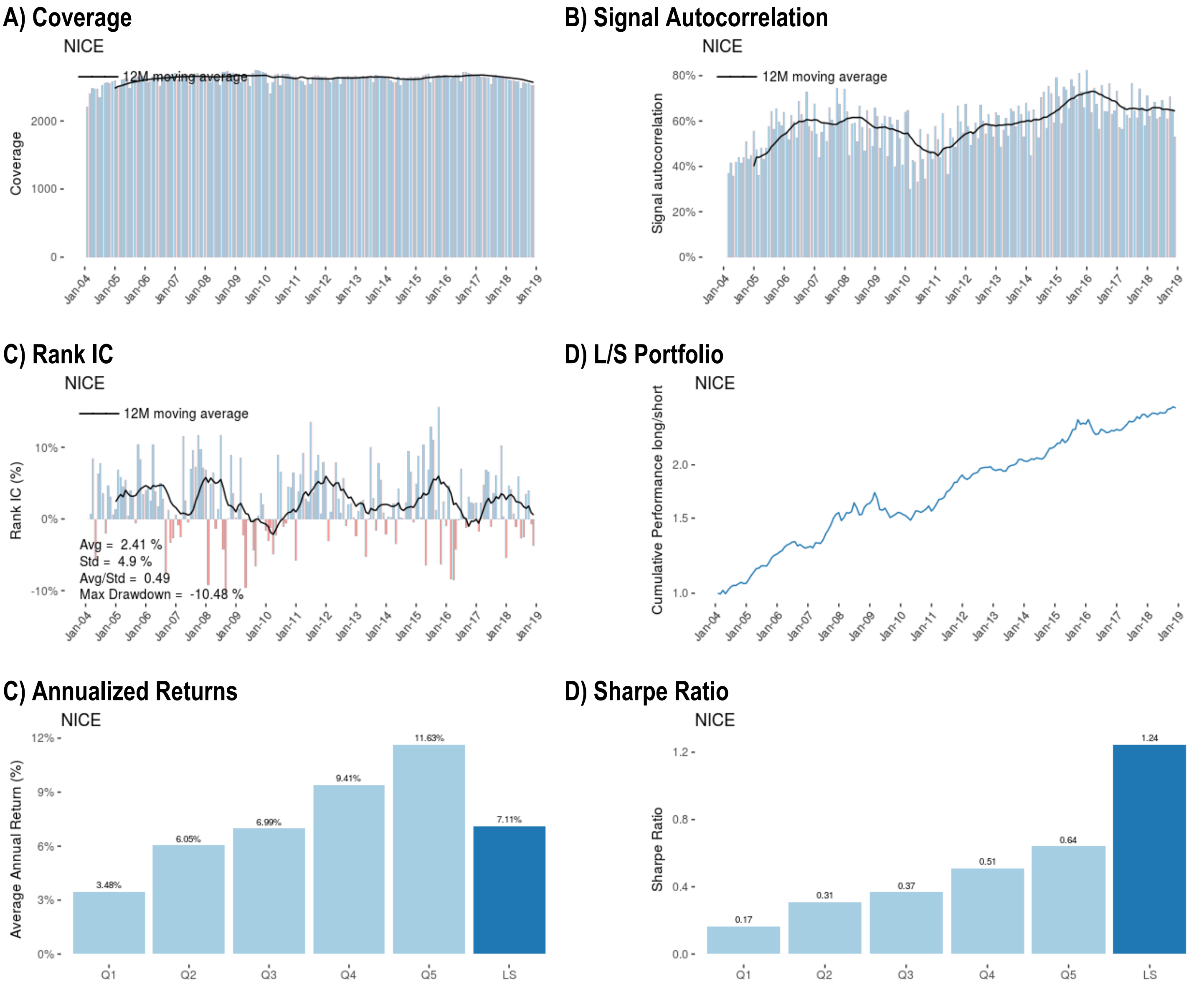
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Model Performance in the US

As shown in Figure 24(A), the NICE model has nearly complete coverage of the Russell 3000 universe in the US. Unlike conventional news sentiment factors, the NICE model's turnover is fairly modest (see Figure 24B), with a monthly autocorrelation of close to 70%, in line with price momentum factors. The NICE model performance (based on a long/short quintile portfolio) has been consistent over the past decade (see Figure 24C and D), with a Sharpe ratio of 1.2x (see Figure 24E).

[6] We coin the model the French city Nice (pronounced as /niːs/ in IPA), rather than the English word nice (pronounced as /naɪs/ in IPA).

Figure 24 NICE Model Performance (US)

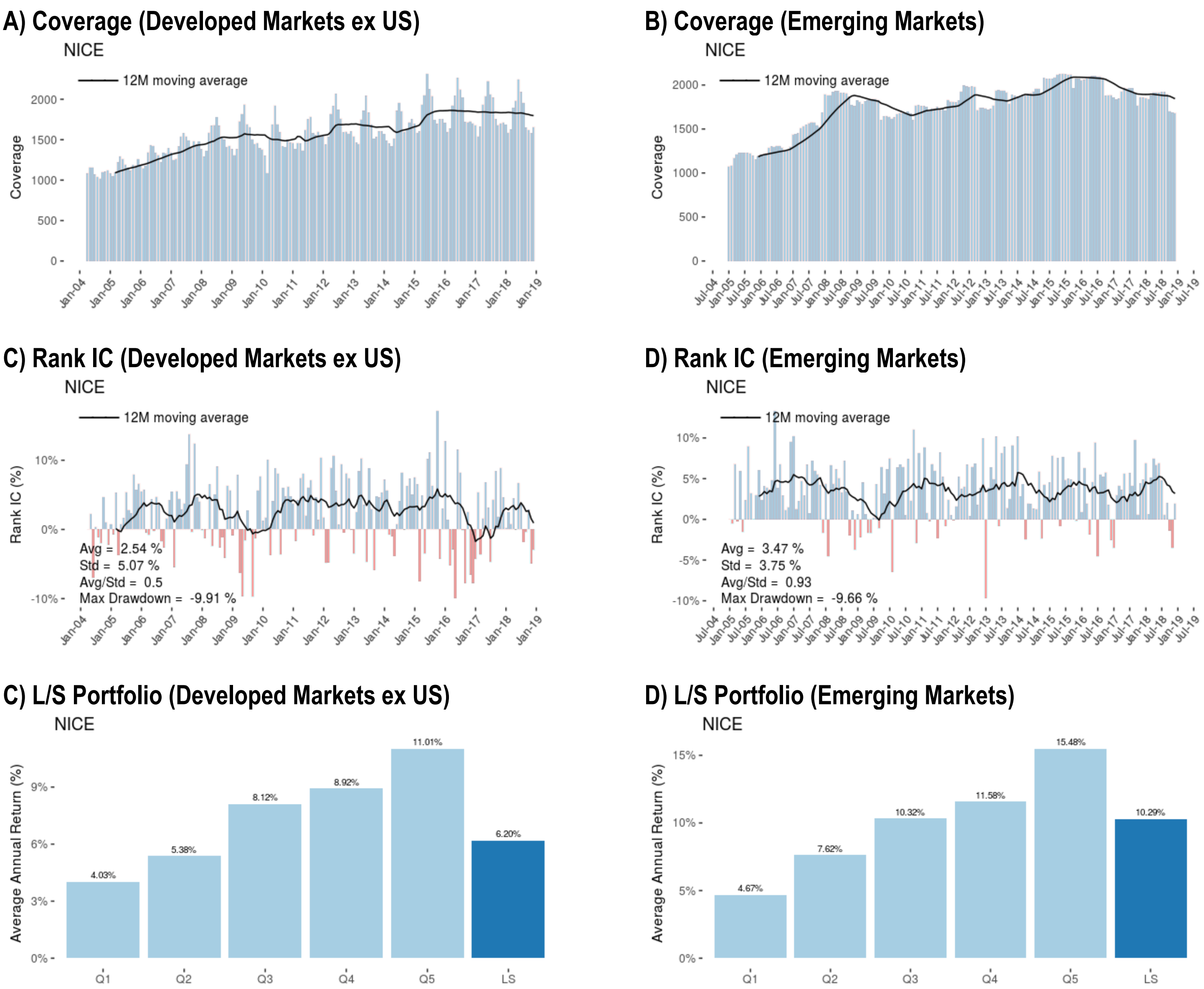


Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Model Performance – International Evidence

The NICE model's coverage outside of the US remains strong – with 1,500 stocks in developed markets (see Figure 25A) and over 1,500 companies in EM (see Figure 25B). Performance is exceptionally strong in EM (see Figure 25D), but also robust in developed markets (see Figure 25C).

Figure 25 NICE Model Performance (World ex US)



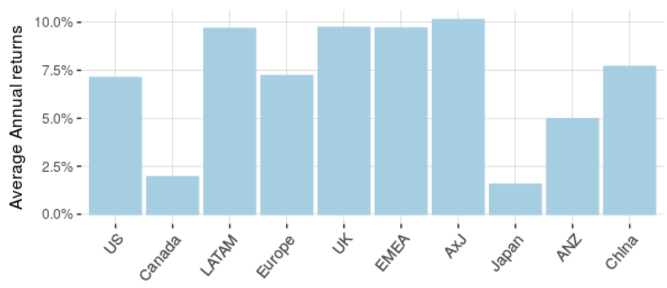
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

As shown in Figure 26, the NICE model is particularly effective in predicting future stock returns in Asia ex Japan, Europe, US, and emerging EMEA, while it struggles in Canada and Japan. While the excess returns are high in China and ANZ, the NICE model volatilities are also large in these two regions.

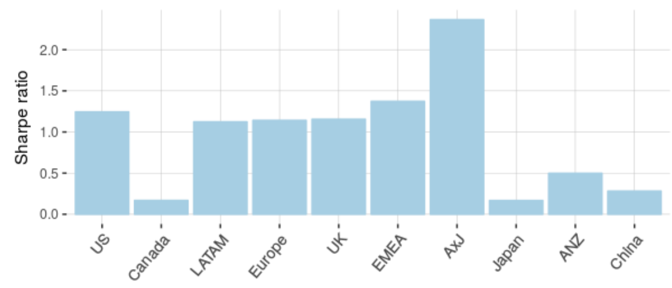
Although RavenPack data focuses exclusively on English language news, our NICE model has delivered exceptional performance in AxJ, with a Sharpe ratio of more than 2.0x.

Figure 26 NICE Model Global Performance

A) Average Annual Return



B) Sharpe Ratio



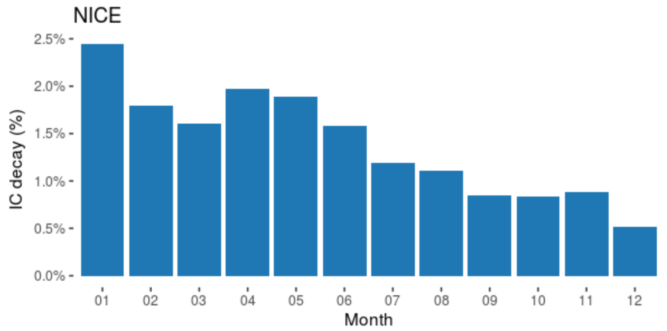
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Signal Decay

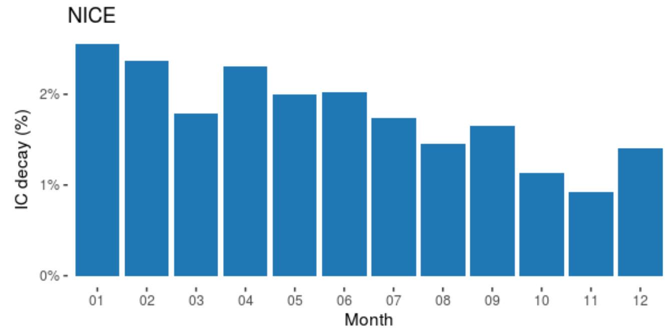
Unlike conventional news sentiment factors (which show no predictive power beyond a month), the NICE model demonstrates decent predictive power up to six months in the US and even longer outside of the US (see Figure 27).

Figure 27 NICE Model Signal Decay

A) US



B) Developed Markets ex US



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

NICE, SPEC, AND SMEC

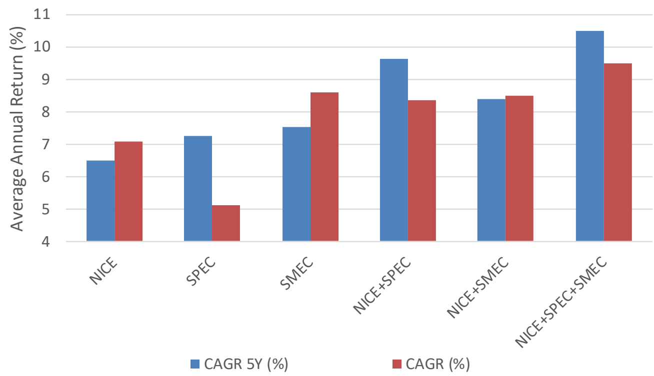
In addition to the NICE model, we have two existing models using NLP and text mining techniques.

- SPEC (Systematic Profiling of EDGAR Composite). As elaborated in *Text Mining Unstructured Corporate Filing Data* (see Rohal, et al [2017a]), our SPEC model examines corporate regulatory filings (e.g., 10K/10Q) to identify red flags, process tone/sentiment, and diagnose other behavior bias.
- SMEC (Systematic Mining of Earnings Calls) Model. As discussed in *Tone at the Top? Quantifying Management Presentation* (see Rohal, et al [2018a]), we inspect management presentations and conference calls to quantify language complexity, management sentiment, executive personality, and key topics.

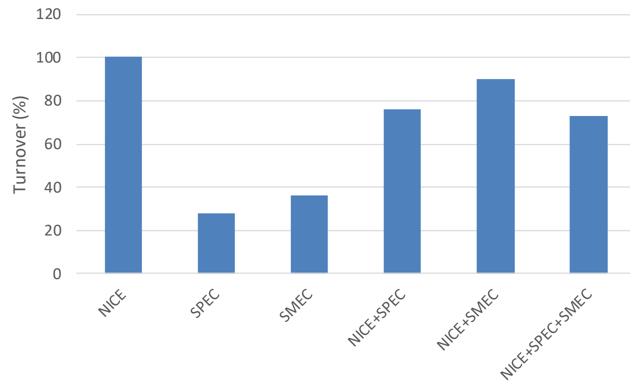
The three models – NICE, SPEC, and SMEC – use very different contents, with different NLP and machine learning algorithms. As a result, they are reasonably uncorrelated. Combining the three models can further improve performance. As shown in Figure 28(A), adding the NICE to the SPEC and SMEC can lift annual alpha by 3% and 1%, respectively. Furthermore, given the low turnover profile of the SPEC and SMEC, the combined model should be useful to long-term buy-and-hold managers as well (see Figure 28B).

Figure 28 Combining the NICE, SPEC, and SMEC (US)

A) Average Annual Return



B) Turnover



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

Correlation with Conventional Factors

Because earnings-, revenue-, and analyst-related news account for the majority of news stories in the RavenPack database, the final NICE model shows some modest correlation with Growth, Momentum and Analyst Sentiment factors, while appears to be relatively uncorrelated to Value and Quality (see Figure 29).

Figure 29 NICE Model Correlation with Traditional Factors (US)

	Earnings yield	YoY EPS growth	Momentum	3M EPS revision	ROE	Accruals	Volatility	1M Reversal	NICE	SPEC	SMEC
Earnings yield	100%	24%	20%	-7%	83%	-5%	-74%	-12%	26%	66%	49%
YoY EPS growth		100%	72%	74%	48%	-2%	-34%	-25%	67%	31%	58%
Momentum 12M_1M			100%	70%	49%	-26%	-42%	-38%	71%	38%	63%
3M EPS revision				100%	22%	-16%	-12%	-28%	65%	13%	41%
ROE					100%	-8%	-84%	-35%	49%	71%	66%
Accruals						100%	20%	19%	-20%	-10%	-16%
Volatility							100%	42%	-41%	-75%	-69%
1M Reversal								100%	-51%	-36%	-41%
NICE									100%	42%	65%
SPEC										100%	62%
SMEC											100%

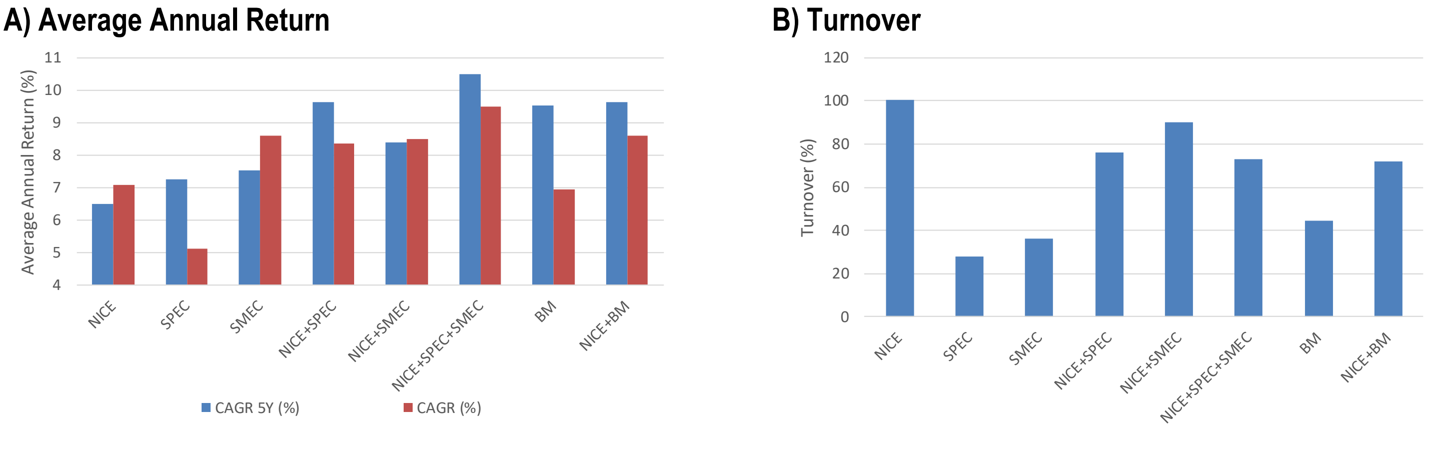
Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

ADDING THE NICE TO A CONVENTIONAL MULTIFACTOR MODEL

To show the potential incremental value of the NICE signal to a conventional multifactor quantitative model, we add it to our benchmark model BM. As defined in *Style Rotation, Machine Learning, and the Quantum LEAP* (see Luo, et al [2017c]), the benchmark model (BM) is a proxy for conventional multifactor quantitative models, by equally weighting eight common stock selection factors (Trailing Earnings Yield, Book-to-Market, Consensus FY1/FY0 EPS Growth, 12M Total Return, 3M EPS Revision, Return on Equity, Debt/Equity and Sloan's Accruals).

As shown in Figure 30A, an equally weighted BM+NICE boosts annualized alpha by 1.5%, albeit turnover also increases (see Figure 30B). More interestingly, our NLP-based composite NICE+SPEC+SMEC beats the benchmark model comfortably.

Figure 30 Combining the NICE and Benchmark Multifactor Model (US)



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

PERFORMANCE ATTRIBUTION

Lastly, to properly identify the drivers of risk and return, we take advantage of our suite of risk models and performance attribution tools – Port@ (see details in *Port@ble Ownership*, Alvarez, et al [2018] and *Risk, Portfolio Construction, and Performance Attribution*, Luo, et al [2017d]).

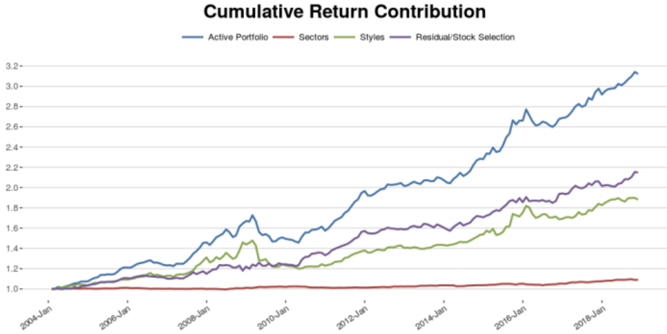
As a demonstration, we attribute the performance of a long/short quintile portfolio based on the combined NICE+SPEC model, rebalanced monthly.

As shown in Figure 31, the NICE/SPEC portfolio is relatively sector neutral[7]. As we would expect from this strategy, it does take positive risk exposure to the Revision (FY1 EPS analyst revision), Price Momentum (12-1 month), and Profitability (ROE) factors, from which it is compensated with positive return over time. More importantly, the "pure" alpha from the model, i.e., the residual/stock selection return (the ***purple*** line in Figure 31A) is more significant and consistent than the payoff from common factors.

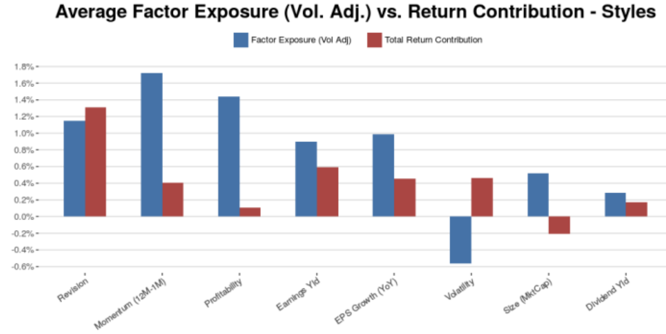
[7] Both the NICE and SPEC models are sector neutralized by design.

Figure 31 Performance Attribution for Quintile NICE+SPEC Model Portfolio (US)

A) Cumulative Return Contribution



B) Factor Exposure and Return Contribution



Sources: Bloomberg Finance LLP, FTSE Russell, Google, IDC, RavenPack, S&P Capital IQ, Thomson Reuters, SEC, Wolfe Research Luo's QES

BIBLIOGRAPHY

Alvarez, M., Luo, Y., Rohal, G., Wang, S., Jussa, J. and Zhong, J. [2018]. "Multi-Dimensional Alpha: Port@ble Ownership", Wolfe Research Luo's QES, June 19, 2018, available here

Boudoukh, J., Feldman, R., Kogan, S. and Richardson, M. [2012]. "Which News Moves Stock Prices? A TEXTUAL ANALYSIS", NBER Working Paper, available at NBER

Chen, H., De, P., Hu, Y. and Hwang, B. [2011]. "Wisdom of Crowds: The Value of Stock Opinions Transmitted Through Social Media", available at SSRN

Giannini, R., Irvine, P. and Shu, T. [2016]. "Nonlocal Disadvantage: A Direct Examination of Individual Investors' Information Set", The Review of Asset Pricing Studies, available here

Heston, S., and Sinha, N. [2015]. "News versus Sentiment: Predicting Stock Returns from News Stories", FEDS Working Paper, available at SSRN

Huynh, T. D. and Smith, D. R. [2015]. "Stock price reaction to news: the joint effect of tone and attention on momentum", FEDS working paper, available at SSRN

Jussa, J., Luo, Y., and Wang, S. [2017a]. "Quant CSI", Wolfe Research Luo's QES, March 13, 2017, available here

Jussa, J., Luo, Y., Wang, S., and Rohal, G. [2017b]. "Space – The Next Alpha Frontier", Wolfe Research Luo's QES, May 17, 2017, available here

Kearney, C. and Liu, S. [2013]. "Textual sentiment in finance: A survey of methods and models", SSRN Working Paper, available at SSRN

Luo, Y., Jussa, J., and Wang, S. [2017a]. "The Big and The Small Sides of Big Data", Wolfe Research Luo's QES, February 8, 2017, available here

Luo, Y., Jussa, J., and Wang, S. [2017b]. "Signal Research and Multifactor Models", Wolfe Research Luo's QES, February 16, 2017, available here

Luo, Y., Jussa, J., and Wang, S. [2017c]. "Style Rotation, Machine Learning, and The Quantum LEAP" Wolfe Research Luo's QES, February 24, 2017, available here

Luo, Y., Jussa, J., and Wang, S. [2017d]. "Risk, Portfolio Construction, and Performance Attribution", Wolfe Research Luo's QES, May 9, 2017, available at here

Luo, Y., Alvarez, M., Jussa, J., Wang, S. and Rohal, G. [2018a]. "Trade war, North Korea conflict, no war after all", Wolfe Research Luo's QES, March 15, 2018, available here

Luo, Y., Rohal, G., Wang, S., Alvarez, M., Jussa, J., and Zhong, J. [2018b]. "Banking on the Banks – Welcome to BALI", Wolfe Research Luo's QES, September 4, 2018, available here

Rohal. G., Luo, Y., Jussa, J., and Wang, S. [2017a]. "Text Mining Unstructured Corporate Filing Data" Wolfe Research Luo's QES, April 20, 2017, available here

Rohal, G., Luo, Y., Jussa, J., and Wang, S. [2017b]. "New Insights in Short Interest – A Global Perspective", Wolfe Research Luo's QES, August 15, 2017, available here

Rohal, G., Luo, Y., Jussa, J., and Wang, S. [2018a]. "Tone at the Top? Quantifying Management Presentation", Wolfe Research Luo's QES, January 23, 2018, available here

Rohal, G., Luo, Y., Jussa, J., and Wang, S. [2018b]. “Pay for Performance”, Wolfe Research Luo’s QES, May 22, 2018, available here

Rohal, G., Luo, Y., Jussa, J., and Wang, S. [2018c]. “Global TMT Stock Selection Models”, Wolfe Research Luo’s QES, Nov 6, 2018, available here

Sprenger, T. O. and Welpe, I. M. [2011]. “News or Noise? The Stock Market Reaction to Different Types of Company-Specific News Events”, available at SSRN

Tetlock, P. C. [2015]. “Giving Content to Investor Sentiment: The Role of Media in the Stock Market”, Journal of Finance, available at SSRN

Wang, S., Luo, Y., Jussa, J., and Rohal, G. [2017a]. “Crowdsourcing Earnings and Revenue Estimates”, Wolfe Research Luo’s QES, April 24, 2017, available here

Wang, S., Luo, Y., Jussa, J., Alvarez, M., and Rohal, G. [2017b]. “Machine Learning Takeovers”, Wolfe Research Luo’s QES, September 12, 2017, available here

Wang, S., Luo, Y., Alvarez, M., Jussa, J., Rohal, G. [2017c]. “The Silk Road to China”, Wolfe Research Luo’s QES, December 1, 2017, available here

Wang, S., Luo, Y., Alvarez, M., Jussa, J., Rohal, G. [2018a]. “Systematic Alpha from Risk Arbitrage (SARA)”, Wolfe Research Luo’s QES, May 15, 2018, available here

Wang, S., Luo, Y., Alvarez, M., Jussa, J., Rohal, G. [2018b]. “The Fundamental Impact of GICS Structural Change”, Wolfe Research Luo’s QES, July 10, 2018, available here

DISCLOSURE SECTION

Analyst Certification:

The analyst of Wolfe Research primarily responsible for this research report whose name appears first on the front page of this research report hereby certifies that (i) the recommendations and opinions expressed in this research report accurately reflect the research analysts' personal views about the subject securities or issuers and (ii) no part of the research analysts' compensation was, is or will be directly or indirectly related to the specific recommendations or views contained in this report.

Other Disclosures:

Wolfe Research, LLC does not assign ratings of Buy, Hold or Sell to the stocks it covers. Outperform, Peer Perform and Underperform are not the respective equivalents of Buy, Hold and Sell but represent relative weightings as defined above. To satisfy regulatory requirements, Outperform has been designated to correspond with Buy, Peer Perform has been designated to correspond with Hold and Underperform has been designated to correspond with Sell.

Wolfe Research Securities and Wolfe Research, LLC have adopted the use of Wolfe Research as brand names. Wolfe Research Securities, a member of FINRA (www.finra.org) is the broker-dealer affiliate of Wolfe Research, LLC and is responsible for the contents of this material. Any analysts publishing these reports are dually employed by Wolfe Research, LLC and Wolfe Research Securities.

The content of this report is to be used solely for informational purposes and should not be regarded as an offer, or a solicitation of an offer, to buy or sell a security, financial instrument or service discussed herein. Opinions in this communication constitute the current judgment of the author as of the date and time of this report and are subject to change without notice. Information herein is believed to be reliable but Wolfe Research and its affiliates, including but not limited to Wolfe Research Securities, makes no representation that it is complete or accurate. The information provided in this communication is not designed to replace a recipient's own decision-making processes for assessing a proposed transaction or investment involving a financial instrument discussed herein. Recipients are encouraged to seek financial advice from their financial advisor regarding the appropriateness of investing in a security or financial instrument referred to in this report and should understand that statements regarding the future performance of the financial instruments or the securities referenced herein may not be realized. Past performance is not indicative of future results. This report is not intended for distribution to, or use by, any person or entity in any location where such distribution or use would be contrary to applicable law, or which would subject Wolfe Research, LLC or any affiliate to any registration requirement within such location. For additional important disclosures, please see www.wolferesearch.com/disclosures.

The views expressed in Wolfe Research, LLC research reports with regards to sectors and/or specific companies may from time to time be inconsistent with the views implied by inclusion of those sectors and companies in other Wolfe Research, LLC analysts' research reports and modeling screens. Wolfe Research communicates with clients across a variety of mediums of the clients' choosing including emails, voice blasts and electronic publication to our proprietary website.