

# Natural Language Processing Literature Survey

---

In client conversations, Natural Language Processing (NLP) and the analysis of unstructured data is a topic of regular conversation. S&P Global Market Intelligence offers several unstructured datasets garnering market attention. The first is transcripts of earnings, guidance, industry conferences, M&A, sales/trading and special calls. All will include unique speaker IDs to identify who is speaking on the call. The second data set is the text content delivered in the financial filings. In advance of a publication of Quantamental primer on NLP next month which will take readers through the process of handling unstructured data and generating sentiment scores, we offer this literature survey. What follows are nine papers that the team has identified as being of particular interest to investors on this topic<sup>1</sup>.

## Authors

Eric Johnstone  
Quantitative Specialist  
212-438-8620  
[ejohnstone@spglobal.com](mailto:ejohnstone@spglobal.com)

Vivian Ning, CFA  
Quantamental Research  
312-233-7148  
[vning@spglobal.com](mailto:vning@spglobal.com)

Temu Oyeniyi, CFA  
Quantamental Research  
312-233-7051  
[toenyiyi@spglobal.com](mailto:toenyiyi@spglobal.com)

Dave Pope, CFA  
Managing Director of  
Quantamental Research  
617 530-8112  
[dpope@spglobal.com](mailto:dpope@spglobal.com)

Dan Sandberg, PhD, CFA  
Quantitative Specialist  
212-438-1407  
[daniel.sandberg@spglobal.com](mailto:daniel.sandberg@spglobal.com)

Richard Tortoriello  
Quantamental Research  
212-438-9506  
[richard.tortoriello@spglobal.com](mailto:richard.tortoriello@spglobal.com)

Frank Zhao  
Quantamental Research  
617-530-8107  
[fzhao@spglobal.com](mailto:fzhao@spglobal.com)

## Earnings Call Transcripts

1. [Tips and Tells From Managers: How Analysts And The Market Read Between The Lines Of Conference Calls – Druz, Wagner, Zeckhauser 2015](#)
2. [Playing Favorites: How Firms Prevent the Revelation of Bad News – Cohen, Lou, and Malloy 2014](#)

## 10-K and Management Disclosure & Analysis

3. [Lazy Prices – Cohen, Malloy, and Nguyen 2016](#)
4. [Word Power: A New Approach for Content Analysis – Jegadeesh and Wu 2013](#)
5. [Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports? – Li 2006](#)
6. [Management's Tone Change, Post Earnings Announcement Drift and Accruals – Feldman, Govindaraj, Livnat, Segal 2009](#)

## News/Press Release

7. [News versus Sentiment: Predicting Stock Returns from News Stories – Heston and Sinha 2016](#)
8. [BlackRock Applied Finance Project: Using Natural Language Processing techniques for Stock Return Predictions – Chew, Puri, Sood, and Wearne 2017](#)
9. [News vs. Sentiment: Predicting Stock Returns from News Stories – Heston and Sinha 2017](#)

---

<sup>1</sup> By clicking the links below you hereby acknowledge and understand that you are leaving an S&P Global Market Intelligence web page and will be directed to a third-party website in order to access the desired content. S&P Global Market Intelligence, its affiliates, and/or their suppliers shall not be subject to any damages or liability arising from your access to or use of such third-party website, including, without limitation, any content, data or information contained therein.

## 1. Earnings Call Transcripts

### **Tips and Tells from Managers: How Analysts and the Market Read Between the Lines of Conference Calls – Druz, Wagner, Zeckhauser 2015**

**Abstract:** Stock prices react significantly to the tone (negativity of words) managers use on earnings conference calls. This reaction reflects reasonably rational use of information. “Tone surprise” – the residual when negativity in managerial tone is regressed on the firm’s recent economic performance and CEO fixed effects – predicts future earnings and analyst uncertainty. Prices move more, as hypothesized, in firms where tone surprise predicts more strongly. Experienced analysts respond appropriately in revising their forecasts; inexperienced analysts overreact (underreact) to tone surprises in presentations (answers). Post-call price drift, like post-earnings announcement drift, suggests less-than-full-use of information embedded in managerial tone.

#### **Analyst Note:**

- The authors study the relationship between tone surprise and variables of interest to investors, including future stock return, next quarter earnings and analyst revisions. The study is based on earnings transcripts for all stocks in the S&P 500 index between 2004 and 2012.
- The authors first compute a “negativity” score, defined as the number of negative words minus the number of positive words divided by the sum of negative and positive words plus 1. Tone surprise is the standardized residuals of a regression of “negativity” score on several control variables including stock return over the last quarter, earnings per share growth, return volatility and total assets.
- The authors found that tone surprise strongly predicts earnings per share in the following quarter, even after controlling for mean consensus analyst forecast. The results were asymmetric as the results were stronger for excess negativity (signals lower earnings in the future) than excess positivity (signals positive earnings in the future).
- Tone surprise was also a strong predictor of subsequent earnings per share adjustments by sell side analysts.
- Using both a regression framework and portfolio formation approach, the authors found tone surprise to be related to future excess stock return over both the short term (1-3 days after conference call) and long term (60 days after conference call). The long-short characteristics-adjusted return of going long (short) stocks with positive (negative) tone surprise was about 1%, 60 days after the earnings call.

### **Playing Favorites: How Firms Prevent the Revelation of Bad News – Cohen, Lou, and Malloy 2014**

**Abstract:** We explore a subtle but important mechanism through which firms can control information flow to the markets. We find that firms that “cast” their conference calls by disproportionately calling on bullish analysts tend to underperform in the future. Firms that call on more favorable analysts experience more negative future earnings surprises and more future earnings restatements. A long-short portfolio that exploits this differential firm behavior earns abnormal returns of up to 149 basis points per month, or almost 18 percent per year. We find similar evidence in an international sample of earnings call transcripts from the UK, Canada, France, and Japan. Firms with higher discretionary accruals, firms that barely meet/exceed earnings expectations, and firms (and their executives) that are about to issue equity, sell shares, and exercise options, are all significantly more likely to cast their earnings calls.

**Analyst Note:** Cohen examines the impact of ‘casting’, or calling on analysts who have favorable recommendations during conference calls. He shows 149bps/month long-short excess return to casting. He also shows that 17% of firms ‘cast’ each quarter. Firms that cast

have higher accruals, issue shares, issue options, or have senior management who will be exercising options. Cohen shows that there is virtually no reversal effect to casting.

### 2. 10-K and Management Disclosure and Analysis (MD&A)

#### **Lazy Prices – Cohen, Malloy, and Nguyen 2016**

**Abstract:** We explore the implications of a subtle “default” choice that firms make in their regular reporting practices, namely that firms typically repeat what they most recently reported. Using the complete history of regular quarterly and annual filings by U.S. corporations from 1995-2014, we show that when firms make an active change in their reporting practices, this conveys an important signal about the firm. Changes to the language and construction of financial reports have strong implications for firms’ future returns: a portfolio that shorts “changers” and buys “non-changers” earns up to 188 basis points per month (over 22% per year) in abnormal returns in the future. These reporting changes are concentrated in the management discussion (MD&A) section. Changes in language referring to the executive (CEO and CFO) team, or regarding litigation, are especially informative for future returns.

**Analyst Note:** Cohen and his co-authors analyzed the length of 10Ks in terms of word count and find that firms with sudden increase (decrease) in word count especially in the risk factor and MD&A sections underperformed (outperformed) by about 800bps per annum where the performance is quite robust and orthogonal to existing common alpha and risk factors. They attributed the excess returns to behavioral tendencies.

#### **Word Power: A New Approach for Content Analysis – Jegadeesh and Wu 2013**

**Abstract:** We present a new approach for content analysis to quantify document tone. We find a significant relation between our measure of the tone of 10-Ks and market reaction for both negative and positive words. We also find that the appropriate choice of term weighting in content analysis is at least as important as, and perhaps more important than, a complete and accurate compilation of the word list. Furthermore, we show that our approach circumvents the need to subjectively partition words into positive and negative word lists. Our approach reliably quantifies the tone of IPO prospectuses as well, and we find that the document score is negatively related to IPO underpricing.

**Analyst Note:** Jegadeesh and Wu develop a new approach to quantify the tone of 10Ks. They found that a weighted scheme on positive or negative words is more effective in predicting stock returns two weeks out than an equal weight scheme where the weights are empirically determined using market data. Using their new approach, they found that positive words now have forecasting power, which is contrary to the findings of most of the existing literature. Lastly, they claim that the term weighting of words is as important as, having a complete and accurate compilation of a lexicon.

#### **Do Stock Market Investors Understand the Risk Sentiment of Corporate Annual Reports? – Li 2006**

**Abstract:** I test the stock market efficiency with respect to the information in the texts of annual reports. More specifically, I examine the implications of corporate annual reports’ risk sentiment for future earnings and stock returns. I measure the risk sentiment of annual reports by counting the frequency of words related to risk or uncertainty in the 10-K filings. I find that an increase in risk sentiment is associated with lower future earnings: Firms with a larger increase in risk sentiment have more negative earnings changes in the next year. Risk sentiment of annual reports can predict future returns in a cross-sectional setting: Firms with a large increase in risk sentiment experience significantly negative returns relative to those firms with little increase in risk sentiment in the twelve months after the annual report filing date. A hedge portfolio based on buying firms with a minor increase in risk sentiment of

annual reports and shorting firms with a large increase in risk sentiment generates an annual Alpha of more than 10% measured using the four-factor model including the Fama-French three factors and the momentum factor.

**Analyst Note:** The author shows that by concentrating on only risk related words in the 10-K, an annual long-short return of 10% can be earned. The author also demonstrates that higher-risk firms have subsequent lower future earnings. The alpha is shown to be significant even when adjusted for Fama French factors, as well as other common risk factors. The author goes to great length to demonstrate robustness.

### **Management's Tone Change, Post Earnings Announcement Drift and Accruals – Feldman, Govindaraj, Livnat, Segal 2009**

**Abstract:** This study explores whether the management discussion and analysis (MD&A) section of Forms 10-Q and 10-K has incremental information content beyond financial measures such as earnings surprises and accruals. It uses a classification scheme of words into positive and negative categories to measure the tone change in the MD&A section relative to prior periodic SEC filings. Our results indicate that short window market reactions around the SEC filing are significantly associated with the tone change of the MD&A section, even after controlling for accruals and earnings surprises. We show that management's tone change adds significantly to portfolio drift returns in the window of two days after the SEC filing date through one day after the subsequent quarter's preliminary earnings announcement, beyond financial information conveyed by accruals and earnings surprises. The drift returns are affected by the ability of the tone change signals to help predict the subsequent quarter's earnings surprise but cannot be completely attributed to this ability. We also find that the incremental information of management's tone change depends on the strength of the firm's information environment.

### 3. News

#### **News versus Sentiment: Predicting Stock Returns from News Stories – Heston and Sinha 2016**

**Abstract:** This paper uses a dataset of more than 900,000 news stories to test whether news predicts stock returns. We measure sentiment with the Harvard psychosocial dictionary used by Tetlock, Saar-Tsechansky, and Macskassy (2008), the financial dictionary of Loughran and McDonald (2011), and a proprietary Thomson-Reuters neural network. Simpler processing techniques predict short-term returns that are quickly reversed, while more sophisticated techniques predict larger and more persistent returns. Daily news predicts stock returns for only 1 to 2 days, confirming previous research. Weekly news, however, predicts stock returns for one quarter. Positive news stories increase returns quickly, but negative stories have a long-delayed reaction.

#### **Analyst Note:**

- The study investigated the effectiveness of textual processing for forecasting stock returns. The authors applied a sophisticated neural network to a broad dataset of news stories in Thomson Reuters new system over the calendar years 2003-2010; The authors document that the neural networks appear to extract permanent information that is not fully incorporated into current stock prices.
- The authors explored the temporal pattern predictability in individual stock return (defined by the net sentiment – the return difference of highest and lowest sentiment portfolios) by aggregating the news over the following time horizons:
  - o Aggregating the news on one day
  - o Aggregating the news over one week

## Natural Language Processing Literature Survey

- They found that stocks with positive (negative) news on one day have predictably high (low) returns for the subsequent one to two days, both significant at the 95% level.
- The decile portfolios formed on the basis of weekly news produces a dramatic increase in predictability of returns over the subsequent 13 weeks.
- The authors further confirm that the positive returns of the news strategy are not subsumed by size, momentum and earnings surprise as the news strategy still yields significant positive return when controlling for size, momentum and earnings surprise effects.
- It appears that the market quickly incorporates positive information into returns. However negative news predicts low stock returns for up to one quarter, which is consistent with short-sale constraints that delay the incorporation of bad news.

### **BlackRock Applied Finance Project: Using Natural Language Processing techniques for Stock Return Predictions – Chew, Puri, Sood, and Wearne 2017**

**Abstract:** Our Applied Finance Project aims to develop a framework to determine if financial news headlines have meaningful impact on stock prices. This framework is a novel structure that primarily leverages on existing Natural Language Processing, including Name Entity Recognition, and Global Vector for Word Representation (GloVe) model, before combining them with techniques such as k-means clustering and portfolio optimization. The subsequent study on events with predictive abilities could be of interest to institutional investors.

Starting with 1.8 million financial news headlines obtained from the Internet Archive: Wayback Machine, we successfully identified several events with meaningful post-event drifts. These events include situations where the equities of a firm are oversold, approval is given to a firm, a deal or agreement is signed as well as when an advisor is hired. The out-of-sample information ratios for these events are between a range of -1.02 and 0.76. The events we identified are by no means exhaustive, signifying the potential of our model.

**Analyst Note:** The authors of this paper use a variety of natural language processing techniques to predict stock returns based on events stated in press release headlines for S&P 500 companies. We view this paper as a “proof of concept” study for using computers to parse text and infer denotation (explicit meaning), versus the plethora of studies that use sentiment-based lexicons to infer connotation (implied meaning or emotional content). The authors parse over 1.8 million news headlines obtained from the Internet Archive: Wayback Machine, focusing on PR Newswire press releases. They then use a variety of techniques to a) identify the subject company from a wide variety of potential forms/spellings and map it to a ticker; and b) determine the verb/object combination and its generalized meaning (e.g., declares dividend, announces investment). Techniques used/discussed include name entity recognition, sentence part of speech mapping, subject-verb-object extraction, verb/phrase vectors, verb vector clustering, and mapping vector centroids (the last three techniques are used to identify verb/object combinations that have similar meanings). The authors identify five verb/object “events” that meaningfully predict stock returns over a 20 – 60 day time horizon: declaring a dividend, oversold equity condition, receiving an approval, signing an agreement, hiring an advisor. They conduct event studies that show statistically significant information ratios and hit rates for each of these five event types. The authors believe that the event studies are by no means exhaustive, signaling good potential for further research.

**News vs. Sentiment: Predicting Stock Returns from News Stories– Heston and Sinha 2017**

**Abstract:** The authors used a dataset of more than 900,000 news stories to test whether news can predict stock returns. They measured sentiment with a proprietary ... neural network and found that daily news predicts stock returns for only one to two days, confirming previous research. Weekly news, however, predicts stock returns for one quarter. Positive news stories increase stock returns quickly, but negative stories receive a long-delayed reaction. Much of the delayed response to news occurs around the subsequent earnings announcement.

**Analyst Note:** The authors compare the relative power of 4 wordlist dictionaries on measuring tone in corporate earnings press releases. The wordlists studied include:

- Henry (2008): developed for use in the domain of financial disclosure;
- Diction: drawn from the field of politics and mass media.
- General Inquirer: drawn from the field of social psychology.
- Loughran and McDonald's (2011) generated from an analysis of 10-K filings.

The authors find that dictionaries from a domain specific wordlist outperform general dictionaries by almost 2 to 1. The research also found an asymmetrical reaction to positive and negative tone although the degree of asymmetry was not as pronounced with domain specific wordlists. The authors also compare equal vs weighted tone measurement methodologies and find equally weighted scheme generally performs better.

## Our Recent Research

### **June 2017: Research Brief: Four Important Things to Know About Banks in a Rising Rate Environment**

With the Fed signaling further rate hikes ahead, bank investors may want to know which investment strategies have worked best in a rising rate environment historically. This paper leverages our empirical work on the SNL Bank fundamental data to aid investors in selecting bank stocks as rates rise.

### **April 2017: Banking on Alpha: Uncovering Investing Signals Using SNL Bank Data**

This study leverages S&P Global Market Intelligence's SNL Financial data to answer three questions of importance to bank investors: 1. Which widely-used investment strategies have historically been profitable? 2. Which lesser-known strategies deserve wider attention? 3. How do these strategies perform across varying macro environments: rising vs. falling interest rates and above- vs. below-average financial stress?

### **March 2017: Capital Market Implications of Spinoffs**

Spinoff activities have picked up in recent years. In 2015, more than \$250 billion worth of spinoff transactions were closed globally - the highest level in the last 20 years. This report analyzes the short- and long-term performance of spun-off entities and their parent companies in the U.S. and international markets. We also examine a related but distinct corporate restructuring activity – equity carve-outs, which separate a subsidiary through a public offering.

### **January 2017: U.S. Stock Selection Model Performance Review 2016**

2016 proved to be a challenging year for active investing. Against a backdrop of a sharp selloff in equities at the beginning of the year and political uncertainty over the course of the year, valuation was the only fundamental investing style that delivered positive excess returns. In this report, we review the performance of S&P Global Market Intelligence's four U.S. stock selection models in 2016.

### **November 2016: Electrify Stock Returns in U.S. Utilities**

The U.S. utilities sector has performed especially well in the past several years as the Federal Reserve and central banks around the world enacted accommodative monetary policies to spur growth. As global active investors flock to the U.S. utilities sector in search of yields and high risk-adjusted returns, we explore a number of utility-specific metrics from a unique database that is dedicated to the utilities sector – S&P Global Market Intelligence's Energy (Source: SNL Energy) – to ascertain whether investors could have historically made stock selection decisions within the sector to achieve excess returns.

### **October 2016: A League of their Own: Batting for Returns in the REIT Industry - Part 2**

SNL Financial's ("SNL") 1 global real estate database contains property level and geographical market-based demographic information that can be difficult for investors to obtain. These unique data points are valuable to investors seeking an understanding of the relationship between property level information and future stock price movement. In this report, we demonstrate how investors can use these data points as alpha strategies. Our back-tests suggest that metrics constructed from property level information may provide insights about future price direction not captured by fundamental or estimates data. Investors may want to consider incorporating information on a REIT's property portfolio when building a robust REIT strategy

### **September 2016: A League of their Own: Batting for Returns in the REIT Industry - Part 1**

This month REITs (Real Estate Investment Trusts) have been separated from the GICS (Global Industry Classification Standard) Financial sector into a sector of their own. Even prior to the sector reclassification, investors have been attracted to REITs' strong performance and attractive yield. REITs differ from traditional companies in several

important ways. Metrics that investors typically use to value or evaluate the attractiveness of stocks such as earnings yield or book-to-price are less meaningful for REITs. For active investors interested in understanding their REITs portfolio, an understanding of the relationship between REIT financial ratios and price appreciation is instructive. Is dividend yield relevant? What about funds from operations (“FFO”), one of the most widely used metrics?

**August 2016: Mergers & Acquisitions: The Good, the Bad and the Ugly (and how to tell them apart)**

In this study we show that, among Russell 3000® firms with acquisitions greater than 5% of acquirer enterprise value, post-M&A acquirer returns have underperformed peers in general. Specifically, we find that:

- Acquirers lag industry peers on a variety of fundamental metrics for an extended period following an acquisition.
- Stock deals significantly underperform cash deals. Acquirers using the highest percentage of stock underperform industry peers by 3.3% one year post-close and by 8.1% after three years.
- Acquirers that grow quickly pre-acquisition often underperform post-acquisition.
- Excess cash on the balance sheet is detrimental for M&A, possibly due to a lack of discipline in deploying that cash.

**July 2016: Preparing for a Slide in Oil Prices -- History May Be Your Guide**

**June 2016: Social Media and Stock Returns: Is There Value in Cyberspace?**

**April 2016: An IQ Test for the “Smart Money” – Is the Reputation of Institutional Investors Warranted?**

**March 2016: Stock-Level Liquidity – Alpha or Risk? - Stocks with Rising Liquidity Outperform Globally**

**February 2016: U.S. Stock Selection Model Performance Review - The most effective investment strategies in 2015**

**January 2016: What Does Earnings Guidance Tell Us? – Listen When Management Announces Good News**

**December 2015: Equity Market Pulse – Quarterly Equity Market Insights Issue 6**

**November 2015: Late to File - The Costs of Delayed 10-Q and 10-K Company Filings**

**October 2015: Global Country Allocation Strategies**

**September 2015: Equity Market Pulse – Quarterly Equity Market Insights Issue 5**

**September 2015: Research Brief: Building Smart Beta Portfolios**

**September 2015: Research Brief – Airline Industry Factors**

**August 2015: Point-In-Time vs. Lagged Fundamentals – This time i(t)'s different?**

**August 2015: Introducing S&P Capital IQ Stock Selection Model for the Japanese Market**

**July 2015: Research Brief – Liquidity Fragility**

**June 2015: Equity Market Pulse – Quarterly Equity Market Insights Issue 4**

**May 2015: Investing in a World with Increasing Investor Activism**

**April 2015: Drilling for Alpha in the Oil and Gas Industry – Insights from Industry Specific Data & Company Financials**

**March 2015: Equity Market Pulse – Quarterly Equity Market Insights Issue 3**

**February 2015: U.S. Stock Selection Model Performance Review - The most effective investment strategies in 2014**



January 2015: [Research Brief: Global Pension Plans - Are Fully Funded Plans a Relic of the Past?](#)

January 2015: [Profitability: Growth-Like Strategy, Value-Like Returns Profiting from Companies with Large Economic Moats](#)

November 2014: [Equity Market Pulse – Quarterly Equity Market Insights Issue 2](#)

October 2014: [Lenders Lead, Owners Follow - The Relationship between Credit Indicators and Equity Returns](#)

August 2014: [Equity Market Pulse – Quarterly Equity Market Insights Issue 1](#)

July 2014: [Factor Insight: Reducing the Downside of a Trend Following Strategy](#)

May 2014: [Introducing S&P Capital IQ's Fundamental China A-Share Equity Risk Model](#)

April 2014: [Riding the Coattails of Activist Investors Yields Short and Long Term Outperformance](#)

March 2014: [Insights from Academic Literature: Corporate Character, Trading Insights, & New Data Sources](#)

February 2014: [Obtaining an Edge in Emerging Markets](#)

February 2014: [U.S Stock Selection Model Performance Review](#)

January 2014: [Buying Outperformance: Do share repurchase announcements lead to higher returns?](#)

October 2013: [Informative Insider Trading - The Hidden Profits in Corporate Insider Filings](#)

September 2013: [Beggars Thy Neighbor – Research Brief: Exploring Pension Plans](#)

August 2013: [Introducing S&P Capital IQ Global Stock Selection Models for Developed Markets: The Foundations of Outperformance](#)

July 2013: [Inspirational Papers on Innovative Topics: Asset Allocation, Insider Trading & Event Studies](#)

June 2013: [Supply Chain Interactions Part 2: Companies – Connected Company Returns Examined as Event Signals](#)

June 2013: [Behind the Asset Growth Anomaly – Over-promising but Under-delivering](#)

April 2013: [Complicated Firms Made Easy - Using Industry Pure-Plays to Forecast Conglomerate Returns.](#)

March 2013: [Risk Models That Work When You Need Them - Short Term Risk Model Enhancements](#)

March 2013: [Follow the Smart Money - Riding the Coattails of Activist Investors](#)

February 2013: [Stock Selection Model Performance Review: Assessing the Drivers of Performance in 2012](#)

January 2013: [Research Brief: Exploiting the January Effect Examining Variations in Trend Following Strategies](#)

December 2012: [Do CEO and CFO Departures Matter? - The Signal Content of CEO and CFO Turnover](#)

November 2012: [11 Industries, 70 Alpha Signals -The Value of Industry-Specific Metrics](#)

October 2012: [Introducing S&P Capital IQ's Fundamental Canada Equity Risk Models](#)

September 2012: [Factor Insight: Earnings Announcement Return – Is A Return Based Surprise Superior to an Earnings Based Surprise?](#)

August 2012: [Supply Chain Interactions Part 1: Industries Profiting from Lead-Lag Industry Relationships](#)

July 2012: [Releasing S&P Capital IQ's Regional and Updated Global & US Equity Risk Models](#)

June 2012: [Riding Industry Momentum – Enhancing the Residual Reversal Factor](#)

May 2012: [The Oil & Gas Industry - Drilling for Alpha Using Global Point-in-Time Industry Data](#)

May 2012: [Case Study: S&P Capital IQ – The Platform for Investment Decisions](#)

March 2012: [Exploring Alpha from the Securities Lending Market – New Alpha Stemming from Improved Data](#)

January 2012: [S&P Capital IQ Stock Selection Model Review – Understanding the Drivers of Performance in 2011](#)

January 2012: [Intelligent Estimates – A Superior Model of Earnings Surprise](#)

December 2011: [Factor Insight – Residual Reversal](#)

November 2011: [Research Brief: Return Correlation and Dispersion – All or Nothing](#)

October 2011: [The Banking Industry](#)

September 2011: [Methods in Dynamic Weighting](#)

September 2011: [Research Brief: Return Correlation and Dispersion](#)

July 2011: [Research Brief - A Topical Digest of Investment Strategy Insights](#)

June 2011: [A Retail Industry Strategy: Does Industry Specific Data tell a different story?](#)

May 2011: [Introducing S&P Capital IQ's Global Fundamental Equity Risk Models](#)

May 2011: [Topical Papers That Caught Our Interest](#)

April 2011: [Can Dividend Policy Changes Yield Alpha?](#)

April 2011: [CQA Spring 2011 Conference Notes](#)

March 2011: [How Much Alpha is in Preliminary Data?](#)

February 2011: [Industry Insights – Biotechnology: FDA Approval Catalyst Strategy](#)

January 2011: [US Stock Selection Models Introduction](#)

January 2011: [Variations on Minimum Variance](#)

January 2011: [Interesting and Influential Papers We Read in 2010](#)

November 2010: [Is your Bank Under Stress? Introducing our Dynamic Bank Model](#)

October 2010: [Getting the Most from Point-in-Time Data](#)

October 2010: [Another Brick in the Wall: The Historic Failure of Price Momentum](#)

July 2010: [Introducing S&P Capital IQ's Fundamental US Equity Risk Model](#)

## Natural Language Processing Literature Survey

Copyright © 2017 by S&P Global Market Intelligence, a division of S&P Global Inc. All rights reserved.

These materials have been prepared solely for information purposes based upon information generally available to the public and from sources believed to be reliable. No content (including index data, ratings, credit-related analyses and data, research, model, software or other application or output therefrom) or any part thereof (Content) may be modified, reverse engineered, reproduced or distributed in any form by any means, or stored in a database or retrieval system, without the prior written permission of S&P Global Market Intelligence or its affiliates (collectively, S&P Global). The Content shall not be used for any unlawful or unauthorized purposes. S&P Global and any third-party providers, (collectively S&P Global Parties) do not guarantee the accuracy, completeness, timeliness or availability of the Content. S&P Global Parties are not responsible for any errors or omissions, regardless of the cause, for the results obtained from the use of the Content. THE CONTENT IS PROVIDED ON "AS IS" BASIS. S&P GLOBAL PARTIES DISCLAIM ANY AND ALL EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE OR USE, FREEDOM FROM BUGS, SOFTWARE ERRORS OR DEFECTS, THAT THE CONTENT'S FUNCTIONING WILL BE UNINTERRUPTED OR THAT THE CONTENT WILL OPERATE WITH ANY SOFTWARE OR HARDWARE CONFIGURATION. In no event shall S&P Global Parties be liable to any party for any direct, indirect, incidental, exemplary, compensatory, punitive, special or consequential damages, costs, expenses, legal fees, or losses (including, without limitation, lost income or lost profits and opportunity costs or losses caused by negligence) in connection with any use of the Content even if advised of the possibility of such damages.

S&P Global Market Intelligence's opinions, quotes and credit-related and other analyses are statements of opinion as of the date they are expressed and not statements of fact or recommendations to purchase, hold, or sell any securities or to make any investment decisions, and do not address the suitability of any security. S&P Global Market Intelligence may provide index data. Direct investment in an index is not possible. Exposure to an asset class represented by an index is available through investable instruments based on that index. S&P Global Market Intelligence assumes no obligation to update the Content following publication in any form or format. The Content should not be relied on and is not a substitute for the skill, judgment and experience of the user, its management, employees, advisors and/or clients when making investment and other business decisions. S&P Global Market Intelligence does not act as a fiduciary or an investment advisor except where registered as such. S&P Global keeps certain activities of its divisions separate from each other in order to preserve the independence and objectivity of their respective activities. As a result, certain divisions of S&P Global may have information that is not available to other S&P Global divisions. S&P Global has established policies and procedures to maintain the confidentiality of certain non-public information received in connection with each analytical process.

S&P Global may receive compensation for its ratings and certain analyses, normally from issuers or underwriters of securities or from obligors. S&P Global reserves the right to disseminate its opinions and analyses. S&P Global's public ratings and analyses are made available on its Web sites, [www.standardandpoors.com](http://www.standardandpoors.com) (free of charge), and [www.ratingsdirect.com](http://www.ratingsdirect.com) and [www.globalcreditportal.com](http://www.globalcreditportal.com) (subscription), and may be distributed through other means, including via S&P Global publications and third-party redistributors. Additional information about our ratings fees is available at [www.standardandpoors.com/usratingsfees](http://www.standardandpoors.com/usratingsfees).