Market Reaction to iPhone Rumors

Chong, Terence Tai Leung and Wu, Zhang and Liu, Yuchen

The Chinese University of Hong Kong, Columbia Univeristy

25 January 2019

Market Reaction to iPhone Rumors *

Zhang Wu† Terence Tai-Leung Chong‡ Yuchen Liu§

Abstract

The paper studies the effects of new product rumors about the iPhone on the stock price of the Apple company. We scrape iPhone rumors from Macrumors.com, and obtain a dataset covering 1,264 articles containing 180 words on average between January 2002 and December 2015. Moreover, we construct a market-decided lexicon to transform qualitative information into quantitative data, and analyze what type of words and what information embedded in the rumors are apt to impact on Apple's stock price. Unlike previous studies, we do not rely on the widely-adopted Harvard-IV-4 dictionary, as the coefficients of the words from the dictionary are neither significant nor consistent with their polarities, compared with our results. The paper obtains three main findings. First, the spread of rumors has a significant impact on the stock price. Second, positive words, rather than negative words, play an important role in affecting the stock price. Third, the stock price is highly sensitive to the words related to the appearance of the iPhone.

Keywords: stock price, iPhone, rumors, market-decided lexicon

*We would like to thank Joon Park, Aman Ulah, Jun Yu, Victoria Zinde-Walsh as participants of the 3rd Dongbei Econometrics Workshop held in Dongbei University of Economics and Finance for helpful comments. We would also like to thank Liugang Sheng, Chih Sheng Hsieh and Julan Du as seminar participants at the Chinese University of Hong Kong for constructive discussions and comments, and thank Sophia Lok and Galvin Chia for their research assistant. Any remaining errors are ours.

†Department of Economics, The Chinese University of Hong Kong, Hong Kong. Email address: zhangwu1224@gmail.com

‡Department of Economics and Lau Chor Tak Institute of Global Economics and Finance, The Chinese University of Hong Kong, Hong Kong; Department of International Economics and Trade, Nanjing University, Nanjing, China. Email address: chong2064@cuhk.edu.hk

§Department of Industrial Engineering and Operations Research, Columbia University.

1 Introduction

As global outsourcing and production develop further, rumors about new products being planned or developed are more easily spread throughout industries and markets. Thanks to progress in communication technology, rumors spread to the public quickly. Consumers and investors are now less likely to be surprised when official announcements of new products take place owing to information leakage.

The literature contains relatively few studies on the topic of rumors about potential new products. The majority of these studies analyzed rumors concerning takeovers (Pound and Zeckhauser [1990] and Bettman, Hallett and Sault [2011]) and market efficiency (Tumarkin and Whitelaw [2001]). Among them, the most relevant research on the topic at hand concerned new product announcements and their effects on stock market returns (Chaney, Devinney and Winer [1991]). Other related studies paid attention to the impact of news on the stock market returns of the companies concerned. Particular emphasis was placed on topics such as stock splits, earnings announcements, mergers or takeovers, macroeconomic fundamentals, regulatory changes and bond ratings (Fama et al. [1969], Pinches and Singleton [1978], Eckbo [1983], Malatesta and Thompson [1985], Mikkelson and Partch [1986], Horsky and Swyngedouw [1987], Goh and Ederington [1993], McQueen and Roley [1993], Lane and Jacobson [1995], MacKinlay [1997], and Benbunan-Fich and Fich [2004], for example).

The information underneath rumors could be obtained and efficiently analyzed when the knowledge of big data develops. Big data refers to a large volume of data that can be used for more accurate estimations and predictions. It is noteworthy that this type of data not only presents itself on a large scale but also permeates minute aspects of daily life, showing a variety of novel types. In addition to the supply of quantitative information, the underlying qualitative information plays a significant but typically neglected role in stock price prediction (Tetlock [2007]). By employing big data, many studies tended to focus on social media sources, such as Twitter (Bollen, Mao and Zeng [2011] and Sprenger et al. [2014]) and Sina Weibo (Bai, Chen and Zhan [2014] and Dong et al. [2015]), forums (Antweiler and Frank [2004] and Das and Chen [2007]) and search engines (Vosen and

Schmidt [2011], Carrière-Swallow and Labbé [2013], Varian [2014], and Smith [2016]). However, the content on these platforms tends to have character limits. Beyond the sources that previous research have tapped, big data can be obtained through other channels, such as articles and reports on websites. The rumor data used in the current chapter are crawled from a pool of such online articles.

A suitable case for applying big data in terms of qualitative information sourced from rumors is that of Apple's iPhone. As a technology giant, Apple's new product announcements have consistently attracted worldwide attention since the iPhone was introduced in 2007. The long queues outside Apple Stores demonstrate the significant degree of consumer interest. Such interest continues to build even when purchase restrictions are applied several months after products are launched. Prior to official announcements, rumors, such as what the next-generation iPhone might look like and what new features it might bring, circulate online in response to popular discussion.

Among Apple's products, the iPhone has been a huge success and has become the main source of Apple's revenue. It interests not only consumers but also investors. As Figure 1 shows, the revenue attributed to the iPhone, illustrated by the solid line, has increased considerably and has accounted for more than 50% of Apple's total revenue since 2012 and over 70% in recent years.

[Figure 1 about here.]

At the same time, the market capitalization of Apple has grown noticeably since the release of the first iPhone. Considering the importance of the iPhone in Apple's total revenue, rumors related to the iPhone are assumed to be relevant to Apple's stock returns. Moreover, Jung and Shiller [2005] proved that the stock market in the US was micro-efficient, indicating that the efficient market hypothesis was more applicable to individual stocks than to the overall stock market. This study thus intends to examine the relevance of iPhone rumors to the stock price of the Apple company through an in-depth study. We attempt to answer the following questions: (1) Are the rumors important enough to have an impact on Apple's stock price? (2) If so, what types of words are more important to the market? (3) What information can be uncovered from the words that are apt to exert the influence?

In the study, the term "rumor" is defined as information that appears before Apple's official announcement of each new generation of iPhones, regardless of whether it is true or not ex post. Such a definition differs from that of Blanchard, L'Huillier and Lorenzoni [2013], in which information or rumors were taken as "news" if they were proven to be true ex post or "noise" if they turned out to be false. Specifically, rumors from the production supply chain or commercial partners are chosen. Although Apple maintains strict confidentiality regarding new products, the public can still access new product information beforehand through rumors. These leakages are mainly from Apple's supply chain, commercial partners and trademark or patent application documents. Some rumors, such as those concerning the specification of new products, can be exposed voluntarily by commercial partners since these partners are able to obtain detailed information through business and technological cooperation. These factors guarantee the rumors' credibility and authenticity. Given this, this study uses the terms "rumor", "leakage", and "news" interchangeably.

The methodology of the current study is as follows. Firstly, an event study is employed to investigate whether the spread of rumors has effects on Apple's returns in the stock market. In particular, we utilize the abnormal returns to measure and evaluate such impacts. The difference between rumor and non-rumor days is also investigated using the bootstrap method. After the influence of rumors on the stock price is confirmed, the rumors leading to the abnormal returns on these days are used to build a lexicon for further study. Secondly, to answer the question of which types of words are important to the market, sentiment analysis is adopted. Following Mao et al. [2014], a market-decided lexicon with two polarities (positive and negative) is constructed in four steps: processing text, selecting seed words, targeting expanded words, and identifying final words. Based on this lexicon, the different effects between positive and negative words on the next-day abnormal and stock returns are compared. At the same time, the performance of a widely accepted but fixed dictionary is examined. Finally, along with the relationship between this constructed lexicon and the fixed dictionary, the informational content and features of the positive and negative words in our lexicon are analyzed.

One of the major methodological challenges concerns how qualitative messages can be transformed into quantitative information. Owing to the improvement in sentiment analysis by Pang, Lee and Vaithyanathan [2002], Turney [2002], and Pang and Lee [2008], studies using embedded qualitative information expanded in the area of finance. Antweiler and Frank [2004] utilized the naive Bayes method to analyze a discussion on an online message board and found that messages were abundant in information and useful in predicting stock volatility. The authors split the data into "buy", "sell", and "hold" types and aggregated them into an indicator. Das and Chen [2007] analyzed messages from Yahoo's bulletin board and linked them to the stock market. By extracting and transforming the posted words into sentiment scores, the authors measured the opinions of small investors on the market status and its future, or specifically, the predictions of the investors on whether the market would be bullish or bearish. Combining firm earnings and stock returns, Tetlock, Saar-tsechansky and Macskassy [2008] examined articles from the Dow Jones News Service and the Wall Street Journal, and found that underlying hard-to-quantify information existed in linguistic media content and that the proportion of negative words had a high degree of predictability. Luss and d' Aspremont [2015] employed support vector machines and multiple kernel learning methods to test their forecasting abilities of price movements in financial assets. They found that the text performed better in predicting returns than the past return data did.

This paper studies the effects of rumors about the next-generation iPhone on Apple's stock price and makes contributions to the literature in two ways. First, it examines the effects of the rumors of new products on the stock market. Previous studies on the impacts of rumors focused on important events of companies, such as earning announcements, stock splits, and takeovers; while studies on new products were confined to the announcement effects. The paper bridges the gap by collecting rumors on the next-generation products and examining the effects. Moreover, this paper emphasizes on exploring the qualitative information in rumors instead of simply counting the occurrence of the rumors. Due to different features of disciplines, we also construct a market-decided lexicon rather than employing any existing and general dictionary to transform the qualitative messages embedded in rumors into quantitative data.

Three conclusions are reached through the analysis of rumors. First, by analyzing the abnormal returns, we identify significant positive effects of the spread of rumors on the stock market. Second, on the basis of the abnormal returns, this study constructs a lexicon consisting of positive and negative words from rumors. Positive words, rather than negative words, are found to play an important role in the market in terms of prediction performance. This market-decided lexicon is confirmed to outperform the widely used Harvard-IV-4 dictionary. Third, the market is found to be most sensitive to the words that contain information related to the outward appearance of the iPhone rather than to the hardware configurations.

The rest of the study is structured as follows. The methodology and data description are introduced in Sections 2 and 3. In Section 4, the impacts of the spread of rumors on Apple's abnormal returns are analyzed. The effects of words and the analysis of words to which the market is sensitive are also presented in Section 4. The last section concludes the results.

2 Methodology

The abnormal return is employed to investigate and evaluate the effects of the spread of rumors. This is a widely used method in event studies to measure the degree of impact. The method for building a market-dependent lexicon proposed by Mao et al. [2014] is also briefly discussed in this part.

2.1 Abnormal Return

The method has been widely adopted (Pound and Zeckhauser [1990], Chaney, Devinney and Winer [1991], Kiymaz [2001], and Lucca and Moench [2015]). As shown in Equation 1, the abnormal return AR_t at time t depends on the difference between the actual and the expected returns (represented by R_t and $E(R_t|X_t)$, respectively), based on the conditioning information X_t.

$$AR_t = R_t - E(R_t|X_t) \tag{1}$$

Since assumptions of each model are different, R_t takes varying forms. The market model in Equation 2 assumes that the return of one firm is a fixed and linear ratio of a market index return, where R_{mt} is the return of a market index R_m at time t. ϵ_t is the abnormal return and satisfies the first- and second-moment conditions listed in the last line of Equation 2.

$$\begin{gathered} R_t = \alpha + \beta R_{mt} + \epsilon_t \\ E(\epsilon_t) = 0, \quad var(\epsilon_t) = \sigma_\epsilon^2 \end{gathered} \tag{2}$$

Considering the risk-free interest rate, the risk-adjusted version of Equation 2 can be rewritten as $R_t - R_{ft} = \alpha + \beta(R_{mt} - R_{ft}) + \epsilon_t$, where R_{ft} is the risk-free rate of return R_f at time t. However, this CAPM model cannot explain multidimensional stock risk. To avoid this limitation, a three-factor model was developed by Fama and French [1993] based on the CAPM model. The authors added two additional factors to control for the size and the value of firms, and suggested that this three-factor model, as shown in Equation 3, is able to overcome the limitation.

$$\begin{gathered} R_t - R_{ft} = \alpha + \beta(R_{mt} - R_{ft}) + \gamma SMB_t + \delta HML_t + \epsilon_t \\ E(\epsilon_t) = 0, \quad var(\epsilon_t) = \sigma_\epsilon^2 \end{gathered} \tag{3}$$

The term $(R_{mt} - R_{ft})$ represents the excess return on the market, while the term SMB_t indicates the performance difference between small and large stocks and is used to measure the size premium. HML_t measures the spread of returns between high-value and growth stocks and captures the value effect. As defined by the authors, size of stocks is measured in terms of market equity; a high book-to-market ratio signifies high-value stocks and a low ratio suggests growth stocks. Similar to the market model, ϵ_t is our focus and satisfies the two moment conditions above.

In this paper, we adopt the three-factor model in Equation 3 for our analysis. We also estimate a five-factor model (Fama and French [2015]). The model adds extra measures of profitability (the return spread between the most profitable firms and the least profitable ones) and investment (the return difference between firms investing conservatively and aggressively) to accommodate cross-sectional data better. However, we consider that this improvement is of trivial relevance to the current study, as our main concern is the excess return.

Based on this rationale, this paper uses the three-factor model to estimate abnormal returns (Tetlock, Saar-tsechansky and Macskassy [2008]). The five-factor model and the market model are used for robustness check of the empirical results generated from the three-factor model.

Next, the parameters of the model are estimated. The timeline presented in Figure 2 shows three time ranges: before, during and after an event.

[Figure 2 about here.]

The first grey area in the time between T_1 and T_2 is the estimation window, in which the coefficients in the models based on known returns can be estimated. The time interval between T_3 and T_4 is the event window, in which the coefficients obtained in the estimation window are employed to calculate the abnormal returns using the actual stock returns in this period. $T = 0$ indicates the exact day when the event happens. This paper uses a nine-day event window (including the event day and four days before and after the event day), as Kiymaz [2001] found that this window was more appropriate for studies related to rumors. Lastly, the interval shown in the last shaded area from T_5 onwards indicates the post-event window. Specifically, since the three-factor model is used, the parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\delta}$ are estimated in the estimation window. The abnormal returns in the event window are calculated as shown in Equation 4.

$$A\hat{R}_t = R_t - R_{ft} - \hat{\alpha} - \hat{\beta}(R_{mt} - R_{ft}) - \hat{\gamma}SMB_t - \hat{\delta}HML_t \tag{4}$$

Additionally, according to the notation of event studies introduced previously, AR_0 represents the abnormal return on the event day. AR_{-t} indicates the abnormal return t day(s) prior to the event day, and AR_t indicates the abnormal return t day(s) after the event day.

Another important term is the cumulative abnormal return (CAR), which is the aggregation of abnormal returns. If we assume that the time interval concerning the event window is between τ_1 and τ_2 and satisfies $T_3 \leq \tau_1 \leq \tau_2 \leq T_4$, then the cumulative abnormal return from τ_1 to τ_2 is the sum of the abnormal returns within the periods, i.e., $CAR(\tau_1, \tau_2) = \sum_{t=\tau_1}^{\tau_2} \hat{AR}_t$.

2.2 Construction of Lexicon

From previous studies on sentiment analysis, two approaches can be identified: the lexicon-based and the classifier-based approaches. The lexicon based sentiment analysis involves finding sentiment bearing words from a predefined sentiment word lexicon in a document to classify the polarity of the document. The sentiment analysis with a machine learning approach trains a classifier which is able to classify documents automatically in terms of sentiment. However, for the machine learning based sentiment analysis which needs manually labelled training data, selecting training sets among a number of news articles is time-consuming, and the results of the selection are often inconsistent among screeners.

For the lexicon-based approach, dictionaries such as the well-accepted General Inquirer Harvard-IV-4 (H-IV-4) and Loughran and McDonald Sentiment Word Lists are usually employed. Such word lists are widely used across various disciplines, including psychology and sociology. However, they were criticized for their loose application to fields in social science, in which the sentiment for the same word could be different or even antithetical. As Loughran and McDonald [2011] demonstrated, the domain was crucial for conducting sentiment analysis. The authors found that more than 70% of "negative words" in the commonly used H-IV-4 dictionary were misclassified as negative in the financial context. Therefore, this paper does not rely on any currently used dictionaries. Instead, following the method suggested by Mao et al. [2014], we construct a lexicon in this study. We let the stock market decide whether the words from rumors are positive or negative and compile them in a lexicon based on abnormal returns.

With reference to the H-IV-4 dictionary and Tetlock, Saar-tsechansky and Macskassy [2008], it can be presumed that only two word types exist: positive and negative. Building our lexicon involves four steps: processing text, selecting seed words, targeting expanded words, and identifying final words. In the first step, articles are collapsed into a matrix for further processing, and seed words, which are most sensitive to abnormal returns, are selected. Based on the similarity to the seed words, other words from the matrix are chosen and added to the lexicon. In the last stage, words that are in conflict with certain rules (as discussed below) are eliminated to ensure that the lexicon is consistent with the market's reactions.

In terms of text processing, according to Das and Chen [2007], Tetlock [2007], and Tetlock, Saar-tsechansky and Macskassy [2008], the construction of the lexicon in this paper is based on the content of news, instead of headlines as used by Chan [2003] and Mao et al. [2014]. In this regard, the qualitative information contained in the articles needs to be collapsed into a document-term matrix (DTM) first before it is fully transformed into quantitative data. In this matrix, each row includes a rumor, while each column contains a word that is extracted from the data set. [1]

Furthermore, all the words in the rumors are transformed into their corresponding lower case before being collapsed into a DTM.[2] The bag-of-words scheme is adopted in the further processing of the matrix.

Contrary to the approach that focuses on whether words appear in the matrix or not, the scheme that we adopt counts the frequency of words' occurrence instead. Methods applied in the area of natural language processing (NLP) are used here as well. To illustrate this point, a word is included in the matrix if it contains more than three

[1]For instance, if there are two documents: (a) "I like rumors" and (b) "I hate rumors", then the matrix would be:

	I	like	hate	rumors
(a)	1	1	0	1
(b)	1	0	1	1

[2]The words shown in the paper are adjusted to their proper formats for the convenience of reading.

letters (punctuation is excluded, except for intra-word hyphens) and if it appears in at least two documents. However, stop words (including articles and pronouns), which have a high frequency but are meaningless, are removed.

Beyond that, it is worth explaining that the method of stemming is not employed in this paper when rumors are collapsed into a DTM. Stemming indicates that words derived from the same root are considered to be the same and are transformed into the same form by affix removal. According to Liu [2007], current stemming algorithms could change the meaning of original words and link irrelevant terms. Because of this potential limitation, this method is not used. Additionally, since stop words have been eliminated, the weighting scheme applied in this paper to the matrix is based on the term frequency (TF) approach, which measures how often a term occurs in each rumor. Another popular weighting scheme is the term frequency-inverse document frequency (TF-IDF) method, which punishes the TF if the rumors contain many of the same terms. Although the TF-IDF method could fulfill the task as well, it would generate similar results while involving more complex work. Hence, the TF scheme is used.

Next, seed words are selected. Unlike Mao et al. [2014], who employed the raw stock returns, this paper uses the abnormal returns to represent the market reaction to the news. This is because abnormal returns are free from the influence related to market and industry volatility. In terms of specific procedures, the abnormal returns are normalized first. More precisely, they are centralized and divided by the standard deviation. Based on the assumption that positive (or negative) words are most likely to be found in the rumors that are associated with significant positive (or negative) abnormal returns, seed words are chosen from the extreme quantiles of the returns. Specifically, positive seed words are selected from rumors with the top 5% of abnormal returns, while negative seed words are chosen from the bottom 5%. (A change in these numbers, for instance, to 10%, would not have any obvious influence on the conclusion.) According to these criteria, the top 100 seed words with the highest TF are selected from the articles with extreme market responses. Following Mao et al. [2014], 30 words for each type of polarity are further picked out. The number of positive seed words and that of negative seed words should be the same. This assures that the judgments made in the next stages are free of bias. Moreover, the selected positive and negative seed words are assigned 1 and -1 as initial scores, respectively. This suggests that the words chosen in the next stage will have sentiment scores in the range of $[-1, 1]$. Additionally, the economic value (EV) proposed by Mao et al. [2014] is used to measure and verify the significance and the impact of the words in the lexicon on the stock market. It is shown in the following equation:

$$EV(w) = \frac{\sum_{i=1}^{n} r_i}{n}$$

where $EV(w)$ denotes the economic value for word w, and n is the number of rumors including word w. r_i, originally the raw stock return, is now replaced by the abnormal return and refers to the excess return on the day when the rumor containing word w appears. The extreme values of EV are adjusted to eliminate their influence. Values above the ninety-fifth percentile are fixed to the value of the ninety-fifth percentile, and those below the fifth are fixed to the fifth.

In the third stage, expanded words are targeted. More words are incorporated into the lexicon based on their similarity to the seed words. With reference to Mao et al. [2014], the word similarity is defined in Equation 5, where w_s and w_o refer to seed words and all the other words in the columns of the DTM, respectively.

$$sim(w_s, w_o) = \frac{\mathbb{F}(w_s, w_o)}{\mathbb{F}(w_s) + \mathbb{F}(w_o) - \mathbb{F}(w_s, w_o)} \tag{5}$$

Given the set of seed words and other words, $sim(w_s, w_o)$ measures the Jaccard similarity coefficient between every word from each set. $\mathbb{F}(w_s, w_o)$ denotes the co-occurrence frequency between words from two sets, while $\mathbb{F}(w_s)$ and $\mathbb{F}(w_o)$ show the frequency of seed words and other words existing in the data set, respectively.

Based on the results of word similarity, the semantic orientation (SO) for each of the other words can be calculated from Equation 6:

$$SO(w_o) = \sum_{w_s \in pos} sim(w_s, w_o) - \sum_{w_s \in neg} sim(w_s, w_o) \tag{6}$$

Specifically, the value of the SO for a word is the sum of the similarity coefficients between the word and all the seed words (including positive and negative ones). If the value is less than zero, the word is negative; if the value is greater than zero, the word is positive.

Finally, words that fail to satisfy certain conditions are removed. The most important condition is whether the SO is consistent with the effect of the word on the market. To achieve consistency, the EV is applied. If the direction of the SO of a word is opposite to that of the EV, the word is excluded. The significance of a word for the stock price is another condition. Therefore, only words with an EV that deviates from zero at the 10% level are included in the final set. In addition, words with an absolute EV value of less than 0.01 are excluded, since the number implies that the word has a minor and insignificant impact on the stock price. Lastly, as the TF weighting scheme is adopted, the upper limit of the word frequency applies to the final words as well. To eliminate noise, words with a frequency above the ninety-fifth percentile are excluded.

3 Data

3.1 Qualitative Data

The iPhone's production is highly globalized: it is designed in the US, but its components are ordered internationally, assembled by millions of workers overseas, and tested by worldwide partners. Although non-disclosure agreements are signed as required, information leakage from links along the production chains still occur and are hard to avoid. Useful information on the next-generation iPhone can usually be extracted from such leakages. For instance, information about the production of the low-cost plastic iPhone 5C was available eight months before its release date, and information about the new rose gold color option for the iPhone 6S was leaked half a year before its official announcement. The following subsection describes the source and scope of the qualitative data used in the present study.

The qualitative data are generated using a web crawler written in Python from articles on the MacRumors.com website. By collecting and tracking the latest news, leakages and reports, the website has concentrated on rumors related to Apple's products for over 16 years. Thanks to its success and impact in this field, it now has more than 865,000 registered users and 22 million posts in its forums. The website publishes or reposts new and detailed rumors promptly: it was one of the earliest websites to predict the change in the naming scheme for the iPhone 4S, and it was also the first website to publish articles disclosing the external appearance of an assembled iPhone 6.

In total, 25,414 articles are obtained from the Macrumors.com website published between January 2002 and December 2015. The sample of articles are presented in Appendix A. In addition to the iPhone, information on Apple's major hardware and software products, such as the iMac desktop computer, MacBook laptop computer series and OS X operating system, is included. The original qualitative data set also contains important news on other IT companies, influential industry events and introductions to third-party accessories.

The data are further processed to ensure that the selected rumors focus on Apple and the iPhone. In addition, because rumors and unestablished facts are the concerns, articles that cover Apple's launch events, annual reports, reportage of its patent fights with other enterprises, reports on user experiences and coverage of the iPhone's market share are excluded. Moreover, round-up articles are omitted because they do not provide new information. Eventually, 1,264 rumors related to the iPhone covering the period from August 2002 to December 2015 are selected. The information in these rumors is provided by a range of sources, such as supply chain and commercial partners, including manufacturers, carriers, retailers and bureaus responsible for patent and trademark applications.

The final set of chosen rumors revolves around three main questions: (1) What would the new iPhone be like? (2) When would it be available? (3) Where can it be bought? The first question involves not only the officially designated name of the new iPhone but also its appearance, hardware configurations and features. These leakages may come directly from, for instance, assembly lines and source codes, or indirectly from new technologies announced by supply partners and patent application files prepared by Apple itself. The second question concerns when the next-generation iPhone would be released and when there would be sufficient supply. Among the leaked information in this category, the debut date of the new iPhone attracts the most attention from investors and consumers. The date can be inferred from its current production stage (such as whether it is undergoing testing or is in mass production) or from the actions taken by commercial partners. Pre-orders of the new iPhone and restrictions on vacation requests from Apple's employees are other indicators. Additionally, because of limits on consumer orders, the question of when the new iPhone could be purchased receives wide attention. Thus, rumors about potential supply constraints are also of interest and are included. Lastly, the final question is specifically related to market expansion and is particularly pertinent in markets like the US and China. This includes issues such as the markets in which the new iPhone would be available and the new local carriers with which Apple would cooperate.

On average, each selected rumor contains 14.49 sentences and 180.34 words. Rumors are depicted in Figure 3 according to different time spans.

[Figure 3 about here.]

The top left picture indicates that rumors are published more frequently on weekdays than at weekends, and that Thursday is the day with the most pieces of news. The top right figure presents a growing trend of iPhone rumors on a year-over-year basis. The bottom left chart shows that leakages occur intensively in the second and third quarters of the year. Denoting the release months of the new iPhone as zero and labeling the months leading to the product launch as negative numbers, the bottom right figure implies that the rumors leading up to a new product launch cycle can be divided into three groups chronologically. Half a year before each announcement, rumors start appearing, and they gradually increase in number in the first four months after the release of the first rumor. Finally, sudden growth is typically seen during the last two months and in the month when the new iPhone is released.

3.2 Quantitative Data

With reference to Tetlock, Saar-tsechansky and Macskassy [2008], the estimation window is set at [-252, -31], i.e., starting from one calendar year to one and a half months before a rumor appears. Correspondingly, the time range of quantitative data is thus set from August 2001 to December 2015, and 3,627 observations for each variable are collected. However, the data on the abnormal returns will only be aligned with the qualitative data of the same date for further study. More specific procedures for the initial data processing are shown below.

Firstly, to ensure that there is sufficient time for the articles to be noticed fully, the rumors published before the last 30 minutes of each trading day are considered as occurring on the current day; otherwise, they are classified as articles released on the next day. Secondly, rumors are combined into one piece if multiple articles are posted on the same day. Lastly, leakages released at weekends are linked to the data of the next trading day. Given these adjustments, the occurrence dates of the rumors are decided. Applying the Fama-French three-factor model, we show the details of the abnormal returns on the days when rumors occur in Figure 4.

[Figure 4 about here.]

The dashed horizontal line indicates the mean of AR. It is clear that the AR on rumor days is remarkable in both directions, although its average value is close to zero.

Apart from Apple's adjusted stock returns (R_a) and data for the three-factor model, the financial data set also contains the market value (ME), book-to-market value (BM), share turnover (ST), and standardized unexpected earnings (SUE), with 963 observations for each variable. The data for the three-factor model include the excess returns on the market ($R_m - R_f$), the performance spread between small stocks and big ones (SMB), and the performance spread between the growth and the value stocks(HML). The BM is calculated from the ME and the book value, and the book value is the difference between the total assets and the liabilities. The ST denotes the ratio of the number of shares traded to the number of outstanding shares. In addition, because of data limitations, the book value, ST and SUE are all updated quarterly. R_a and data for the Fama-French three-factor model are collected from the Bloomberg database and Kenneth French's website, respectively, whereas the remaining variables are obtained from Wharton Research Data Services (WRDS).

Additionally, with R_a and the data from the three-factor model, the AR is estimated via Equation 4. $FFCAR$ represents the fact that the CAR is derived from the Fama-French three-factor model. Based on Tetlock, Saar-tsechansky and Macskassy [2008], $FFCAR_{-30,-3}$, $FFCAR_{-2,-2}$, $FFCAR_{-1,-1}$ and $FFCAR_{0,0}$ are employed to ensure that the full impact is estimated completely. The first variable indicates the abnormal returns accumulated from 30 days to 3 days prior to the event day, and the remaining ones show the abnormal returns on days -2, -1 and 0, respectively. $FFAlpha_{-252,-31}$ represents the intercept in the event study regression with the estimation window [-252,-31].

To provide further information, the data description for each financial series is provided in Table 1. It can be seen that the mean of AR and R_a on non-rumor days is smaller than that on rumor days.

[Table 1 about here.]

4 Empirical Result

In this section, the significance of the occurrence of rumors on Apple's abnormal returns is verified. We also examine the specific impacts of the rumors from two perspectives: how the SO of the words in the rumors affects the stock returns and what information can be revealed by words to which the market is sensitive.

4.1 The Impact of Rumors on Abnormal Return

Based on the methods outlined in Subsection 2.1, the results of the event study show an obvious jump in the AR from one day ahead ($t = -1$) to the occurrence day ($t = 0$). Mathematically, it is a 0.158% increase, which is the largest rise in the event window, compared with the growth of 0.033% one day before. After the occurrence day, the growth rate of the AR for rumors drops. It only rises by 0.120% in total during the next three days.

Apart from the rumors, each announcement of the next-generation iPhone is included and serves the purpose of comparison. The pattern above is not observed on the day of Apple's official iPhone announcement. The AR increases by an average of 0.068% on the announcement day, which is clearly smaller than that on the occurrence day of the rumor. This implies that the information of the announcement has already been revealed through the historical stock price, and thus no new information is released on the day of the product launch.

[Figure 5 about here.]

Before verifying the significance of the effects, distribution of AR is studied. Figure 5 shows the features of the AR for rumor and non-rumor days. The upper chart shows the relationship between the actual quantiles and the theoretical ones for the AR. The dark dashed line and light dotted line represent the theoretically expected normal distribution quantiles for rumor and non-rumor days, respectively.

As Figure 5 indicates, the AR for rumor and non-rumor days are not normally distributed. The tails of AR under these two conditions are more dispersed than the corresponding expected normal distribution quantiles, while the middle parts conform well. This implies a fat-tailed feature, which means that there are more extreme values at both ends than theoretically expected — specifically, in low quantiles the values of the AR are more negative than the normal distribution suggests, and in high quantiles they are more positive. However, the AR on non-rumor days shows more extreme values than that on rumor days in both upper quantiles and lower quantiles. In addition, the values of the Shapiro ratio for rumor and non-rumor days are 0.939 and 0.93, respectively, also rejecting the hypothesis that the AR for rumor and non-rumor days follow a standard normal distribution at the 1% significance level.

The graph on the bottom depicts almost identical density of the AR on rumor and non-rumor days, and the AR is more clustered around the mean for occurrence days. Along with the kurtosis of the AR in Panels A and B of Table 1, the high peak feature of rumor and non-rumor days is clearly presented. Similarly, the AR for announcement days does not follow a standard normal distribution neither. Its Shapiro ratio is 0.792, rejecting the null hypothesis of normal distribution as well.

Given the non-normal distribution and based on Lucca and Moench [2015], we employ the bootstrap approach in a statistical significance test for the AR on the event day (day 0). Based on 10,000 replications, the lower bound of the 95% confidence interval is 0.04%, while the upper bound is 0.26%. The results support the conclusion that the AR is significant on the same day that rumors occur.

Furthermore, whether the AR on rumor days differs from the AR on non-rumor days is of importance. Considering that the AR on both rumor and non-rumor days do not follow the normal distribution, and given a 10,000 iteration bootstrap, the difference in means is 0.15%, which is significant at the 5% level. These results clearly show that the AR on the occurrence days of the rumor is significant and that the AR between rumor and non-rumor days is also significantly different. In addition, the tests are applied to the AR on iPhone announcement days, but the results indicate neither that the AR is significantly different from zero nor that the AR on release days differs from that on non-release days at the 5% significance level. According to the previous results, the impact of rumors on Apple's abnormal returns can be verified. The lexicon is built on this basis, and further analysis is conducted based on the AR for the occurrence days of the rumor.

4.2 Polarities of Words

Based on the methods described in Subsection 2.2, 1,264 pieces of rumors are collapsed into a matrix with 6,433 columns, and over 6,000 words are initially extracted. Given the rumors' association with the top and the bottom 5% AR, the corpus is further processed, and, correspondingly, positive and negative seed words are found. Specifically, the positive rumors include 102.73 words on average, whereas the negative ones contain 128.63 words. From the top 100 words, 30 seed words for each SO are identified according to TF.

[Table 2 about here.]

As Table 2 shows, words that are commonly defined as "positive" are recognized by the market as positive seed words as well, including "advantage", "agreement", "faster", "improve", and "holiday". In addition, Apple's intelligent assistant "Siri" appears in the list. The word "leaks" is also included, which implies that the market seems to welcome rumor leakages and sees them as a positive signal. This is also consistent with the result of the event study, which concludes that rumors can have statistically significant effects on the AR.

Surprisingly, except for terms such as "costs" and "delays", which are commonly labeled as negative, most negative seed words turn out to adhere closely to the technologies that the iPhone adopts. These include "TFT-LCD", "backlight", and "FinFET"(technologies in the screen), "LTE" (a communication standard) and "ARM Cortex" (a type of processor). The iPhone's competitors and manufacturers are also included in the negative category, such as "Android" (Google's rival operating system), "Qualcomm" and "Pegatron" (hardware firms involved in the iPhone's design and manufacture). Based on a comparison between the two sides of the table, it is worth noticing that the term "difference" is positive, while "familiar" is negative. This suggests that the market desires distinct, novel products rather than similar ones with which the public is familiar.

Given these seed words, the average EV for positive seed words is 236.72 basis points, and it is -198.91 basis points for negative seed words. According to Mao et al. [2014], the 435.62 basis points difference between these values is significant and reflects the AR trends efficiently.

The degree of similarity between the seed words and the rest of the words in the DTM is then calculated. On this basis, similarity coefficients are assigned to each of the remaining 6,373 words before they are incorporated into the expanded word set. As Table 3 shows, the difference in the EV between the two polarities is notably narrowed, owing to the increasing number of words.

[Table 3 about here.]

After excluding the words with SO values that are opposite to their corresponding EV and the words with values of EV that are insignificant and unnoticeable, 314 words are obtained to form the final set. As a result of our control of the word number for both polarities in seed words selection, the positive words account for approximately 50.32% in the final set, which slightly exceeds the number of negative words. This pattern is consistent with that of Mao et al. [2014]. Moreover, in terms of the impact on the AR, these two types of words differ from each other. According to Table 3, the average EV of all the positive words in the final set is 35.02, while that of negative terms is -48.81. Although the gap between positive and negative words in the final set is smaller than that in the seed word set, this difference is still shown to be significant.

To verify the validity and the effectiveness of the final set of words, the evaluation method proposed by Mao et al. [2014] is used here:

$$Senti_j = \frac{\sum_{i=1}^{N} SO(w_i) \times f_i}{\sum_{i=1}^{N} f_i} \tag{7}$$

The sentiment score ($Senti_j$) on a given day j is calculated through Equation 7, where $SO(w_i)$ refers to the SO for word w_i derived from Equation 6. f_i represents the frequency of word w_i, and N is the total number of the final words in the combined rumors on day j.

To assess the performance of the final words, all the rumors with different occurrence dates are randomly split into two sets, with a ratio of three to two. Hence, the first set includes 577 observations and the second set contains 386. Through this measurement, the sentiment scores for each rumor are calculated and divided into five quantiles. Based on the scores, the average AR for each group is obtained. Given the AR and the sentiment scores, a positive relationship is shown in Figure 6.

[Figure 6 about here.]

It is clear that the monotonically increasing patterns in both sets demonstrate that the final words are valid in measuring the AR, and that the qualitative information embedded in rumors reflects the preferences of the market. Considering this, the final words are used as the lexicon that we construct for the current study.

To transform these qualitative words into quantitative information, the positive and the negative words on a given day are counted and divided by the total number of words of the combined rumor. Following Tetlock, Saar-tsechansky and Macskassy [2008], the resulting proportions are then transformed into standardized ratios by subtracting the mean and dividing by the standard deviation, and the standardized ratios for the positive and negative words are hereafter referred to as POS and NEG, respectively. The impacts of the qualitative information on the stock price are investigated using the H-IV-4 dictionary and our constructed lexicon. The results, with the next-day abnormal returns and stock returns being the dependent variables, are reported in Table 4 and Table 5.

[Table 4 about here.]

[Table 5 about here.]

In Table 4, the performance of factors on the AR is shown. Column one lists several control variables and serves as a benchmark model, which captures the effects of past returns and earnings, firm size, and trading volume. These variables are commonly adopted as Tetlock, Saar-tsechansky and Macskassy [2008] suggested.

First of all, the standardized ratio is used, and the results are presented in columns two and three. It is obvious that the proportion of the positive words in our lexicon, or POS, is highly significant. It also implies that a one standard deviation increase results in a growth of 10.3 basis points in the next-day abnormal returns. However, the proportion of the negative words is insignificant and small.

In addition to the standardized ratio, the widely used sentiment index $\frac{P-N}{P+N}$ (where P refers to the positive words and N represents the negative words) is adopted as well. The notation IDX denotes the sentiment index derived from our lexicon, whereas IDX_H is the corresponding index from H-IV-4. The results of adopting the indices are shown in columns four and seven. The impact of IDX is significant but weaker than that of POS in terms of statistical significance.

Other variables, such as the previous month's returns, book-to-market ratio, and share turnover, are significant across these three models, whereas the past returns of two days ahead of the occurrence of rumors are significant in some models.

Meanwhile, the standardized ratios of the positive and negative words against the total number of words for the H-IV-4 dictionary are denoted by POS_H and NEG_H, respectively, and are listed in columns five to seven. Clearly, POS_H is insignificant and displays values that are opposite to the expected direction. The influence of the words on returns is supposed to be in the same direction as their coefficients suggest. Although the coefficient of NEG_H is slightly significant in column six, its value also displays a false leaning towards the negative polarity. Moreover, the coefficient of IDX_H is tiny and insignificant.

Compared with the benchmark, there are improvements from 0.032 to 0.040 in the adjusted R^2 in the models with positive words derived from our lexicon. The models with negative words from the H-IV-4 dictionary only improve the adjusted R^2 marginally, from 0.032 to 0.034. The two models that involve negative words from the final set and positive words from the H-IV-4 dictionary perform slightly worse than the benchmark. This means that the negative words in the final set and the positive words in the H-IV-4 dictionary are redundant in the models.

Furthermore, the effects of other factors are evaluated. First, dummy variables are set for Apple's products excluding the iPhone. Considering that the abnormal returns may be attributed to rumors about other products as well, the dummies serve as control variables. However, since Apple's product line is large and covers a range of items from computers to watches, it is inevitable that news about other products will be reported on the same day as the release of iPhone rumors. Therefore, the dummy variables are adjusted to take a value of one if there is a rumor about a new product, excluding general updates to an existing product, and a value of zero if the rumor does not involve information on any new products.

In total, three dummy variables are used to represent Apple's primary products. The dummy variable for the iPad includes not only the rumors of the first-generation iPad, which made its debut in 2010, but also those of the other two important models, the iPad mini and iPad Pro. The dummy for the MacBook coincides with both the MacBook Air, launched in 2008, and the ultra-thin MacBook, introduced in 2015. The higher-end MacBook Pro is not included because its first release in 2006 was earlier than the iPhone's. Lastly, the Apple Watch, the latest of Apple's products, is considered when the co-occurrence dummy is calculated. The number of pictures depicting a potential new iPhone model is also counted and added to the model. This additional variable's function is to evaluate the credibility of the rumors. Generally, rumors published with photos, especially those concerning the parts of a product, provide visualized and catchy information and are thus more likely to be deemed authentic. The results are shown from columns eight to 12, and column eight is the benchmark for comparison. In terms of word sentiment, the results are similar to those listed in columns two to seven and indicate that POS plays an important role in affecting abnormal returns. Column nine states that POS is statistically significant at the 1% significance level.

The coefficient is economically significant as well. The estimated growth of abnormal returns is $4 \times 0.093 = 0.372$ standard deviations when POS increases from two standard deviations below the mean to two standard deviations above the mean. However, NEG in column ten is insignificant both statistically and economically, similar to what is shown in column three. By combining POS and NEG, the complete results can be obtained in column 11. The signs of the coefficients of POS and NEG are consistent with their polarities, respectively. POS is highly significant, while NEG is insignificant. However, column 12 indicates that POS_H and NEG_H are problematic because the signs of the coefficients are opposite to the polarities of words.

Apart from the sentiment, the coefficients of the control variables are worthy of note. Other Apple products, including the MacBook and the Apple Watch, contribute little to Apple's abnormal returns. Even though the iPad seems to create some impact, it is only marginally significant. Furthermore, unlike Apple's other products, the negative sign on the MacBook potentially demonstrates that the post-PC era is coming, given that smaller and more portable products, instead of laptops, are increasingly favored by the market. Moreover, including photos in rumors has a significant and positive impact on Apple's abnormal returns across the models. This reflects the fact that leaked photos efficiently enhance the credibilities of rumors and thus affect expectations and abnormal returns.

Moreover, including POS improves the model performance. The adjusted R^2 of 0.038 in the original benchmark increases to 0.045 and 0.044 in the models with POS only and in the one with both POS and NEG, respectively. Compared with the benchmark, the model with POS_H and NEG_H merely causes a marginal increase in the adjusted R^2, from 0.038 to 0.041.

In addition to the influence on the abnormal returns, the impact of words from the rumors on Apple's stock returns is investigated and presented in Table 5. From columns one to six, in which there are no control variables for other products and images, the main results are the same. The POS is significant at the 5% significance level. Every increase in its standard deviation causes a 7.8 basis points rise in stock returns. In contrast, the NEG is insignificant. Similar to the coefficients of POS_H and NEG_H in Table 4, the results indicate that the coefficients are insignificant, and that their signs are opposite to the polarities of words. This implies that the sentiment calculation, based on a fixed dictionary that neglects the characteristics of disciplines, can result in mistakes.

In contrast to the corresponding columns in Table 4, the cumulative abnormal returns do not affect the stock returns here. This reflects the difference in the dependent variables between the stock returns and the abnormal returns. In terms of the adjusted R^2, adding POS to regression improves the model performance from 0.003 to 0.007, whereas the best performance of models with words from the H-IV-4 dictionary is only the same as that of the benchmark. After adding extra control variables, POS still plays a significant role, as the case of the abnormal returns suggests in Table 4. A one standard deviation increase in POS results in a growth of roughly seven basis points in the stock returns. However, NEG, POS_H and NEG_H have no impact on stock returns. Apart from the ratios of the words, the coefficients of the control variables are different from those in Table 4. Coefficients of all the other products, including the iPad, which is marginally significant in affecting the next-day abnormal returns, are insignificant in influencing the stock returns. In addition, the sign of the MacBook coefficient remains negative. This implies that the market has changed its focus from computers to more portable electronic devices. Furthermore, as a symbol of the credibility of rumors, the coefficients corresponding to pictures in rumors turn out to be insignificant. In terms of performance, adding these control variables and POS improves the adjusted R^2 from 0.004 in the benchmark to 0.008, whereas the models with the words from the H-IV-4 dictionary gain nothing relative to the benchmark. These results demonstrate that the words from our constructed lexicon outperform those from the H-IV-4 dictionary. Furthermore, the positive words in the lexicon perform better than the negative ones.

Given this conclusion, it can be seen that the comparison between our constructed lexicon and the H-IV-4 dictionary is of importance. The latter contains 4,206 word entries, including 1,915 positive words and 2,291 negative words. Table 6 shows the words that have the same polarity in both the H-IV-4 dictionary and our lexicon, as well as those that have opposite orientations.

[Table 6 about here.]

In particular, the words in the top left and the bottom right corners have the same polarity in both dictionaries. In comparison, the top right and bottom left areas contain words of which the sentiments in the H-IV-4 dictionary are opposite to those in our lexicon. According to the previous regression results, the positive words, instead of the negative ones, from our lexicon play a significant role in affecting the next-day abnormal returns and stock returns. Therefore, the words in the right-hand column are less important than those in the left-hand one.

Moreover, the words in the constructed lexicon are investigated further to identify the specific information to which the market is sensitive. According to the SO values and frequency, words are further selected from the lexicon and displayed in Table 7.

[Table 7 about here.]

From the table, it is observed that words which describe the appearance of iPhone figure more prominently in the positive category of our lexicon. The words "colors", "black", and "gold" imply that the market focuses more on the colors of the new iPhone. Along with the colors, "appearance", "thinner", and "large screen" indicate that the iPhone's outward appearance is more important to the market than its configurations and hardware components. This result implies that the iPhone is a product which represents the latest fashion; thus, only the features that differ remarkably from those of either other products or its previous models and the features that can be identified easily are attractive to the market. In the positive category, terms such as "early", "image", and "show" reflect expectations that the rumors about the new iPhone contain visual information. Additionally, the positive category of "developers" implies the importance of the iOS ecosystem to the iPhone, and the word "Jobs" suggests the notable influence of Steve Jobs.

In the negative category, the most negative word is "Samsung", which is both Apple's main supplier and strong competitor. Samsung supplies components to Apple, such as the iPhone's screen and flash memory, but Samsung and Apple also compete in the global smartphone market. Moreover, some months are perceived as negative, for example "January" and "July"; however, "November" and "fall" are identified as positive words. This may be due to the release dates of the new iPhone, which are usually announced in the fall and become available to most of the world in November. These words should be considered positive if the next-generation iPhone is still released according to the schedule above. However, January and July tend to be the times of the year when the iPhone enters testing and mass production stages, and when problems and delays often occur. Other negative words are in line with those patterns found in seed words, namely, manufacturing partners and hardware components. They are classified as negative, perhaps because of the common association with potential problems that may delay the launch of the iPhone.

4.3 Robustness Check

Based on the constructed lexicon, the five-factor model and the market model are employed to check the robustness of the results derived from the three-factor model. Tables 8 and 9 present the results for the five-factor model, while Tables 10 and 11 display the results from the market model. Since the Nasdaq index concentrates on the technology sector, to which Apple belongs and of which the industry features may present patterns that are different from the total stock market index, the Nasdaq index is adopted in the market model.[3]

[Table 8 about here.]

[Table 9 about here.]

[Table 10 about here.]

[Table 11 about here.]

It is clear that POS is still highly significant at the 1% or 5% levels across these models. Considering all the control variables, if POS increases by one standard deviation, the next-day abnormal returns rise by 10.2 and 8.3 basis points for the five-factor model and the market model, respectively, whereas the next-day stock returns increase by 7.1 and 7.3 basis points. However, NEG is insignificant across the models. Furthermore, POS_H and NEG_H perform poorly, as suggested by the three-factor model, in which they are statistically insignificant and have coefficient signs opposite to the corresponding polarities.

In contrast to the results from the three-factor model and the market model, in which news concerning the iPad marginally affects Apple's abnormal returns, the results of the five-factor model show that the effect is not robust. Although not reported in this paper, the coefficients of the market model using the S&P 500 index similarly suggest that the impact of the leakages concerning the iPad is insignificant.

However, regardless of which models are employed, the control variable for the number of leaked images in rumors is still remarkably significant in influencing Apple's next-day abnormal returns. This demonstrates that the occurrence and the number of images in rumors play a key role in affecting the abnormal returns. An increase in the number of pictures shown in rumors would lead to growth of at least ten basis points in the abnormal returns. However, this pattern is not observed in the stock returns.

Moreover, the performance of the models with POS in terms of the adjusted R^2 exceeds that of the models with POS and NEG in Tables 8 to 11. In addition, the

[3] Although not reported here, the S&P 500 is employed and tested as the total market index as well; the coefficients change little and the conclusion remains the same.

performance on the next-day stock returns of the market model outweighs that of other models in terms of the adjusted R^2. This improvement may be attributed to the characteristics of the information technology industry included in the market model.

Robustness testing shows that, regardless of the different models employed, POS rather than NEG plays a significant role in affecting the next-day abnormal returns and stock returns. However, the ratio derived from the words in the H-IV-4 dictionary does not follow the same pattern. The analysis reveals that the H-IV-4 dictionary performs poorly because of its cross-disciplinary feature. Moreover, the number of images in rumors has a notable impact on the abnormal returns but not on the stock returns. Although the spread of the iPad rumors marginally affects Apple's abnormal returns based on the three-factor model, the result is not supported by the five-factor model, and thus it is not robust.

5 Conclusion

Thanks to big data, information revealing itself in various formats and permeating daily life can be mined. The current paper, which uses big data by extracting rumors from the leading MacRumors.com website, aims to explore the effect of rumors on the market by focusing on the case of Apple's iPhone. Data crawling via Python enables us to collect a large amount of qualitative information and transform such information into quantitative data. The final results show that qualitative information is valuable and cannot be neglected, which is in line with the findings of Tetlock [2007] and Tetlock, Saar-tsechansky and Macskassy [2008]. The paper contributes to the field by studying the effects of rumors of new products on the stock market, a topic that few previous studies discussed. Also, instead of relying on an existing dictionary as what was involved in the majority of previous studies, a market-decided lexicon is constructed to examine such effects of rumors.

Specifically, three main results are obtained. First, based on an event study, this paper verifies that the spread of rumors about the next-generation iPhone significantly influences Apple's abnormal returns. The abnormal returns, which are also found to be significant on the occurrence days of rumors, are considerably different from those on non-rumor days, based on 10,000 bootstrap iterations. However, this pattern is not observed in official announcements: the bootstrap method indicates that the abnormal returns on the day when new products are released are neither significant nor substantially distinct from those on non-release days. This result demonstrates that the new information has already been received and absorbed by the market through previous leakages.

Second, the paper constructs a market-decided lexicon based on the rumors extracted from the website. In particular, the words extracted from the rumors are categorized into two polarities relative to the abnormal returns: positive words and negative words. It is found that the impacts of positive words on the next-day abnormal and stock returns, rather than those of negative words, are statistically and economically significant. This conclusion challenges the results in the literature, which has found the impacts of negative words to be more significant. We think the conclusion is justified in the sense that technology news belongs to a special domain, and that 70% of negative words are misclassified in the H-IV-4 dictionary (Loughran and McDonald [2011]). In addition, it is found that our constructed lexicon outperforms the H-IV-4 dictionary.

Third, by examining the content of the positive and negative words from the constructed lexicon, we discover that positive words are primarily related to the outward appearance of iPhone. This is consistent with the iPhone's image as a lifestyle product. In addition, the words related to competitors, suppliers, and hardware components are recognized as negative by the market. This result implies that the market responds more positively to outward appearances than to hardware configurations.

Table 1: Descriptive Statistics on Quantitative Data

Variables	Mean	Max	Min	St. Dev.	Skewness	Kurtosis	Obs.
Panel A: Rumors							
AR	0.002	0.078	-0.126	0.017	-0.198	8.573	963
R_a	0.003	0.100	-0.197	0.020	-0.861	15.595	963
$FFAlpha_{-252,-31}$	-0.002	0.002	-0.019	0.005	-1.986	5.801	963
$FFCAR_{-30,-3}$	0.025	0.488	-0.452	0.121	0.390	4.758	963
ME	382.598	774.691	4.803	203.260	-0.166	1.874	963
BM	0.219	0.874	0.101	0.068	4.111	34.279	963
ST	1243.482	3408.332	493.679	615.154	1.208	4.177	963
SUE	2.568	11.931	-1.784	2.639	1.249	5.304	963
Panel B: non-Rumors							
AR	0.000	0.137	-0.131	0.020	0.619	9.392	2481
R_a	0.001	0.130	-0.132	0.023	0.147	6.214	2481
$FFAlpha_{-252,-31}$	-0.005	0.002	-0.019	0.006	-0.814	2.430	2481
$FFCAR_{-30,-3}$	0.007	0.428	-0.466	0.129	-0.094	3.824	2481

Table 2: Selected Seed Words

Positive			Negative		
ability	actually	advantage	adopt	android	arm
agreement	apparently	appearance	backlight	capacitive	cases
black	cited	confirmation	cause	cortex	costs
curved	developers	difference	delays	designed	developing
discovered	drawings	edges	familiar	finfet	graphics
fact	faster	gold	lte	major	manufacturers
holiday	idea	improve	noted	panels	pegatron
leaks	longer	originally	possibility	preparing	processors
revealing	shown	siri	qualcomm	require	strength
top	updates	weekend	technologies	testing	tftlcd

Table 3: Average EV of Lexicons

Type	Positive (bps)	Negative (bps)	Number of Words
Seed Words	236.72	-198.91	60
Expanded Words	3.01	-12.29	6373
Final Words	35.02	-48.81	314

Table 4: Effects of Words on Abnormal Return in the Three-Factor Model

	Dependent variable:											
	AR_{+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.103^{***}							0.093^{***}		0.093^{***}	
		(0.033)							(0.033)		(0.034)	
NEG			−0.027							−0.015	−0.003	
			(0.033)							(0.033)	(0.034)	
IDX				0.126^{**}								
				(0.056)								
POS_H					−0.030							−0.040
					(0.032)							(0.033)
NEG_H						0.057^{*}						0.062^{*}
						(0.032)						(0.032)
IDX_H							0.006					
							(0.004)					
$FFCAR_{0,0}$	0.025	0.001	0.018	0.007	0.024	0.026	0.027	0.023	0.001	0.019	0.001	0.022
	(0.033)	(0.033)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.032)	(0.033)	(0.034)	(0.034)	(0.032)
$FFCAR_{-1,-1}$	0.009	0.008	0.009	0.004	0.010	0.010	0.011	0.013	0.012	0.013	0.012	0.016
	(0.032)	(0.032)	(0.032)	(0.033)	(0.033)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)	(0.032)
$FFCAR_{-2,-2}$	-0.056^{*}	−0.052	-0.054^{*}	-0.067^{**}	-0.055^{*}	-0.057^{*}	-0.058^{*}	-0.055^{*}	−0.052	-0.054^{*}	−0.052	-0.056^{*}
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	−0.051	−0.044	−0.054	−0.020	−0.047	−0.049	−0.054	−0.058	−0.050	−0.059	−0.051	−0.050
	(0.051)	(0.051)	(0.051)	(0.053)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
$FFCAR_{-30,-3}$	0.173^{***}	0.161^{***}	0.171^{***}	0.173^{***}	0.175^{***}	0.171^{***}	0.169^{***}	0.179^{***}	0.168^{***}	0.178^{***}	0.168^{***}	0.180^{***}
	(0.035)	(0.035)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.036)	(0.035)
SUE	0.007	0.014	0.008	0.013	0.006	0.009	0.008	0.010	0.016	0.010	0.016	0.011
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	0.024	0.022	0.026	−0.023	0.021	0.022	0.025	−0.012	−0.011	−0.010	−0.011	−0.015
	(0.045)	(0.045)	(0.045)	(0.048)	(0.045)	(0.045)	(0.045)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)
Log(Book/Market)	0.374^{**}	0.363^{**}	0.371^{**}	0.308^{**}	0.382^{**}	0.379^{**}	0.369^{**}	0.300^{*}	0.297^{*}	0.300^{*}	0.297^{*}	0.319^{**}
	(0.152)	(0.151)	(0.152)	(0.154)	(0.152)	(0.152)	(0.152)	(0.153)	(0.153)	(0.153)	(0.153)	(0.153)
Log(Share Turnover)	0.064^{*}	0.063^{*}	0.062^{*}	0.083^{**}	0.068^{*}	0.066^{*}	0.061^{*}	0.065^{*}	0.064^{*}	0.063^{*}	0.064^{*}	0.072^{**}
	(0.035)	(0.035)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
iPad								0.320^{**}	0.314^{*}	0.317^{*}	0.313^{*}	0.319^{**}
								(0.162)	(0.161)	(0.162)	(0.162)	(0.162)
MacBook								−0.414	−0.331	−0.402	−0.329	−0.426
								(0.403)	(0.403)	(0.404)	(0.404)	(0.403)
iWatch								0.118	0.122	0.121	0.123	0.110
								(0.300)	(0.299)	(0.300)	(0.299)	(0.300)
Picture								0.120^{**}	0.107^{**}	0.118^{**}	0.106^{**}	0.112^{**}
								(0.050)	(0.050)	(0.051)	(0.051)	(0.050)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.032	0.040	0.031	0.039	0.031	0.034	0.033	0.038	0.045	0.038	0.044	0.041

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 5: Effects of Words on Stock Return in the Three-Factor Model

	Dependent variable:											
	R_{a+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.078**							0.071**		0.069**	
		(0.034)							(0.034)		(0.034)	
NEG			−0.036							−0.026	−0.017	
			(0.034)							(0.034)	(0.034)	
IDX				0.141**								
				(0.057)								
POS_H					−0.022							−0.027
					(0.033)							(0.034)
NEG_H						0.029						0.032
						(0.032)						(0.033)
IDX_H							0.002					
							(0.004)					
$FFCAR_{0,0}$	−0.028	−0.046	−0.037	−0.054	−0.029	−0.027	−0.028	−0.030	−0.046	−0.036	−0.050	−0.030
	(0.033)	(0.034)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.035)	(0.033)
$FFCAR_{-1,-1}$	−0.025	−0.026	−0.025	−0.025	−0.024	−0.025	−0.025	−0.022	−0.023	−0.022	−0.023	−0.021
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFCAR_{-2,-2}$	−0.041	−0.038	−0.039	−0.049	−0.041	−0.042	−0.041	−0.040	−0.038	−0.039	−0.037	−0.041
	(0.033)	(0.033)	(0.033)	(0.034)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	0.004	0.009	0.001	0.028	0.007	0.005	0.003	−0.001	0.004	−0.003	0.003	0.004
	(0.052)	(0.052)	(0.052)	(0.054)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)	(0.052)
$FFCAR_{-30,-3}$	0.024	0.015	0.021	0.029	0.026	0.023	0.023	0.028	0.019	0.025	0.018	0.029
	(0.036)	(0.036)	(0.036)	(0.037)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
SUE	0.041	0.047	0.043	0.042	0.041	0.042	0.042	0.043	0.048	0.045	0.049	0.044
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	−0.001	−0.002	0.002	−0.037	−0.003	−0.002	−0.0004	−0.023	−0.022	−0.020	−0.020	−0.025
	(0.046)	(0.046)	(0.046)	(0.048)	(0.046)	(0.046)	(0.046)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
Log(Book/Market)	0.360**	0.351**	0.356**	0.307*	0.365**	0.362**	0.358**	0.308**	0.305*	0.307**	0.305*	0.319**
	(0.154)	(0.154)	(0.154)	(0.157)	(0.154)	(0.154)	(0.154)	(0.156)	(0.156)	(0.156)	(0.156)	(0.156)
Log(Share Turnover)	0.080**	0.079**	0.077**	0.094**	0.083**	0.081**	0.079**	0.079**	0.078**	0.077**	0.077**	0.084**
	(0.035)	(0.035)	(0.036)	(0.037)	(0.036)	(0.035)	(0.035)	(0.036)	(0.035)	(0.036)	(0.036)	(0.036)
iPad								0.194	0.190	0.190	0.187	0.195
								(0.165)	(0.165)	(0.165)	(0.165)	(0.165)
MacBook								−0.498	−0.435	−0.476	−0.422	−0.507
								(0.410)	(0.411)	(0.411)	(0.412)	(0.411)
iWatch								0.022	0.025	0.026	0.027	0.016
								(0.305)	(0.305)	(0.305)	(0.305)	(0.305)
Picture								0.082	0.072	0.077	0.069	0.077
								(0.051)	(0.051)	(0.051)	(0.052)	(0.051)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.003	0.007	0.003	0.011	0.002	0.003	0.002	0.004	0.008	0.004	0.007	0.004

Note: *p<0.1; **p<0.05; ***p<0.01

Table 6: Comparison between Words in the Final Set and in the H-IV-4 Dictionary

H-IV-4 Dictionary	Final Set					
	Positive			Negative		
Positive	actual	agreement	fitness	capability	contribute	cooperate
	gold	hit	natural	enable	familiar	fellow
	reconcile	steady	subscription	pay	proprietary	respectable
				super		
Negative	box	chronic	default	corrosion	false	force
	edge	fall	get	load	serve	
	hit	show	wait			

Table 7: Selected Final Words

Positive			Negative		
early	image	show	samsung	supplier	testing
thinner	indeed	original	force	people	january
gold	says	black	july	nano	manufacturer
jobs	fall	cable	costs	familiar	noting
developers	included	final	ago	designed	adopt
colors	large screen	firm	relatively	manufactured	think
edge	limited	get	pay	sometime	processing
test	provided	left	developed	enable	korea
actual	agreement	payments	taiwanbased	probably	measures
potentially	appearance	november	drawbacks	singlechip	canceled

Table 8: Effects of Words on Abnormal Return in the Five-Factor Model

	Dependent variable:											
	AR_{+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.112***							0.102***		0.103***	
		(0.033)							(0.033)		(0.034)	
NEG			−0.024							−0.011	0.004	
			(0.033)							(0.033)	(0.034)	
IDX				0.126**								
				(0.056)								
POS_H					−0.045							−0.054
					(0.032)							(0.033)
NEG_H						0.046						0.054*
						(0.032)						(0.033)
IDX_H							0.004					
							(0.004)					
$FFCAR_{0,0}$	0.013	−0.012	0.007	−0.003	0.011	0.013	0.013	0.011	−0.011	0.008	−0.011	0.009
	(0.033)	(0.033)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.033)
$FFCAR_{-1,-1}$	0.021	0.018	0.021	0.015	0.023	0.022	0.022	0.026	0.023	0.026	0.023	0.029
	(0.032)	(0.032)	(0.032)	(0.033)	(0.032)	(0.032)	(0.033)	(0.033)	(0.032)	(0.033)	(0.032)	(0.032)
$FFCAR_{-2,-2}$	−0.025	−0.023	−0.025	−0.036	−0.024	−0.025	−0.027	−0.022	−0.020	−0.022	−0.020	−0.021
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	−0.054	−0.046	−0.056	−0.020	−0.047	−0.054	−0.056	−0.062	−0.053	−0.063	−0.053	−0.054
	(0.052)	(0.052)	(0.053)	(0.055)	(0.053)	(0.052)	(0.053)	(0.052)	(0.052)	(0.053)	(0.052)	(0.053)
$FFCAR_{-30,-3}$	0.135***	0.126***	0.134***	0.133***	0.139***	0.133***	0.133***	0.141***	0.132***	0.140***	0.132***	0.142***
	(0.035)	(0.035)	(0.035)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
SUE	−0.008	0.001	−0.006	0.00002	−0.009	−0.006	−0.007	−0.004	0.003	−0.003	0.003	−0.003
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	0.017	0.015	0.018	−0.028	0.012	0.016	0.018	−0.010	−0.009	−0.009	−0.009	−0.014
	(0.046)	(0.046)	(0.046)	(0.049)	(0.046)	(0.046)	(0.046)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
Log(Book/Market)	0.330**	0.311*	0.327**	0.249	0.340**	0.334**	0.327**	0.274*	0.263	0.273*	0.264	0.294*
	(0.160)	(0.159)	(0.160)	(0.164)	(0.160)	(0.160)	(0.160)	(0.161)	(0.161)	(0.161)	(0.161)	(0.161)
Log(Share Turnover)	0.059*	0.057	0.057*	0.074**	0.066*	0.061*	0.058*	0.059*	0.057	0.058*	0.057	0.067*
	(0.035)	(0.035)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
iPad								0.151	0.144	0.149	0.144	0.152
								(0.164)	(0.163)	(0.164)	(0.163)	(0.164)
MacBook								−0.647	−0.550	−0.637	−0.554	−0.669
								(0.407)	(0.407)	(0.409)	(0.408)	(0.407)
iWatch								−0.043	−0.028	−0.040	−0.029	−0.054
								(0.303)	(0.301)	(0.303)	(0.302)	(0.302)
Picture								0.115**	0.100*	0.113**	0.100**	0.105**
								(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.016	0.027	0.016	0.021	0.017	0.017	0.016	0.021	0.029	0.020	0.028	0.023

Note: *p<0.1; **p<0.05; ***p<0.01

Table 9: Effects of Words on Stock Return in the Five-Factor Model

	Dependent variable:											
	R_{a+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.079**							0.071**		0.069**	
		(0.033)							(0.034)		(0.034)	
NEG			−0.038							−0.028	−0.018	
			(0.033)							(0.034)	(0.034)	
IDX				0.141**								
				(0.056)								
POS_H					−0.021							−0.027
					(0.033)							(0.034)
NEG_H						0.029						0.032
						(0.032)						(0.033)
IDX_H							0.002					
							(0.004)					
$FFCAR_{0,0}$	−0.027	−0.044	−0.036	−0.050	−0.028	−0.027	−0.027	−0.029	−0.045	−0.035	−0.048	−0.030
	(0.033)	(0.034)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.034)	(0.033)
$FFCAR_{-1,-1}$	−0.005	−0.007	−0.005	−0.008	−0.004	−0.004	−0.004	0.0002	−0.002	−0.001	−0.003	0.002
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFCAR_{-2,-2}$	−0.024	−0.022	−0.023	−0.033	−0.024	−0.024	−0.025	−0.022	−0.021	−0.022	−0.021	−0.022
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	−0.008	−0.002	−0.011	0.018	−0.005	−0.008	−0.009	−0.014	−0.008	−0.016	−0.009	−0.010
	(0.053)	(0.053)	(0.053)	(0.055)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)	(0.053)
$FFCAR_{-30,-3}$	−0.002	−0.008	−0.004	0.003	−0.0005	−0.004	−0.003	0.003	−0.003	0.002	−0.004	0.003
	(0.036)	(0.036)	(0.036)	(0.037)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)
SUE	0.038	0.044	0.041	0.040	0.038	0.040	0.039	0.041	0.046	0.042	0.047	0.041
	(0.034)	(0.034)	(0.034)	(0.035)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	0.005	0.004	0.008	−0.033	0.003	0.005	0.006	−0.016	−0.015	−0.013	−0.013	−0.018
	(0.047)	(0.046)	(0.047)	(0.049)	(0.047)	(0.047)	(0.047)	(0.048)	(0.048)	(0.049)	(0.048)	(0.048)
Log(Book/Market)	0.369**	0.356**	0.365**	0.304*	0.374**	0.372**	0.368**	0.324**	0.317*	0.322**	0.316*	0.335**
	(0.161)	(0.161)	(0.161)	(0.165)	(0.162)	(0.161)	(0.161)	(0.163)	(0.162)	(0.163)	(0.163)	(0.163)
Log(Share Turnover)	0.077**	0.076**	0.074**	0.090**	0.080**	0.078**	0.076**	0.077**	0.075**	0.074**	0.074**	0.081**
	(0.035)	(0.035)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
iPad								0.199	0.194	0.195	0.192	0.199
								(0.165)	(0.165)	(0.165)	(0.165)	(0.165)
MacBook								−0.503	−0.435	−0.478	−0.422	−0.513
								(0.411)	(0.412)	(0.412)	(0.413)	(0.412)
iWatch								0.015	0.025	0.022	0.029	0.009
								(0.305)	(0.305)	(0.306)	(0.305)	(0.306)
Picture								0.081	0.071	0.076	0.068	0.076
								(0.051)	(0.051)	(0.052)	(0.052)	(0.052)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.001	0.006	0.001	0.008	0.0004	0.001	0.0002	0.002	0.006	0.002	0.005	0.002

Note: *p<0.1; **p<0.05; ***p<0.01

Table 10: Effects of Words on Abnormal Return in the Market Model

	Dependent variable:											
	AR_{+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.093***							0.083**		0.082**	
		(0.033)							(0.033)		(0.034)	
NEG			−0.030							−0.017	−0.005	
			(0.034)							(0.034)	(0.034)	
IDX				0.130**								
				(0.057)								
POS_H					−0.016							−0.020
					(0.032)							(0.033)
NEG_H						0.039						0.040
						(0.032)						(0.033)
IDX_H							0.007					
							(0.004)					
$FFCAR_{0,0}$	0.012	−0.009	0.005	−0.006	0.011	0.013	0.014	0.010	−0.008	0.006	−0.010	0.011
	(0.033)	(0.033)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.032)	(0.033)	(0.033)	(0.034)	(0.032)
$FFCAR_{-1,-1}$	−0.019	−0.020	−0.020	−0.021	−0.019	−0.018	−0.017	−0.017	−0.017	−0.017	−0.017	−0.015
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.032)	(0.033)	(0.033)	(0.033)
$FFCAR_{-2,-2}$	−0.067**	−0.064*	−0.066**	−0.074**	−0.067**	−0.068**	−0.070**	−0.067**	−0.064**	−0.066**	−0.064**	−0.067**
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	−0.047	−0.043	−0.051	−0.023	−0.046	−0.045	−0.048	−0.056	−0.051	−0.058	−0.052	−0.051
	(0.048)	(0.048)	(0.048)	(0.050)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
$FFCAR_{-30,-3}$	0.054	0.047	0.050	0.069*	0.055	0.054	0.049	0.054	0.049	0.052	0.048	0.056
	(0.036)	(0.036)	(0.037)	(0.038)	(0.036)	(0.036)	(0.036)	(0.036)	(0.036)	(0.037)	(0.037)	(0.036)
SUE	0.016	0.024	0.018	0.026	0.016	0.018	0.018	0.019	0.026	0.020	0.026	0.021
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	−0.026	−0.019	−0.026	−0.032	−0.025	−0.024	−0.027	−0.072*	−0.063	−0.071*	−0.063	−0.069*
	(0.037)	(0.037)	(0.037)	(0.038)	(0.037)	(0.037)	(0.037)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Log(Book/Market)	0.167	0.181	0.152	0.190	0.176	0.179	0.155	0.058	0.079	0.052	0.077	0.086
	(0.183)	(0.183)	(0.184)	(0.191)	(0.184)	(0.184)	(0.183)	(0.186)	(0.186)	(0.186)	(0.186)	(0.188)
Log(Share Turnover)	0.058*	0.056	0.055	0.064*	0.060*	0.059*	0.055	0.058*	0.057*	0.057	0.056	0.063*
	(0.035)	(0.034)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.034)	(0.035)	(0.035)	(0.035)
iPad								0.277*	0.271*	0.275*	0.271*	0.276*
								(0.164)	(0.164)	(0.164)	(0.164)	(0.164)
MacBook								−0.448	−0.375	−0.433	−0.371	−0.454
								(0.409)	(0.409)	(0.410)	(0.410)	(0.409)
iWatch								0.031	0.032	0.033	0.033	0.026
								(0.304)	(0.303)	(0.304)	(0.303)	(0.304)
Picture								0.134***	0.123**	0.131**	0.122**	0.129**
								(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.005	0.012	0.005	0.013	0.004	0.006	0.007	0.012	0.018	0.012	0.017	0.012

Note: *p<0.1; **p<0.05; ***p<0.01

Table 11: Effects of Words on Stock Return in the Market Model

	Dependent variable:											
	R_{a+1}											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
POS		0.080**							0.073**		0.071**	
		(0.033)							(0.034)		(0.034)	
NEG			−0.038							−0.029	−0.018	
			(0.034)							(0.034)	(0.034)	
IDX				0.140**								
				(0.057)								
POS_H					−0.021							−0.027
					(0.032)							(0.033)
NEG_H						0.029						0.032
						(0.032)						(0.033)
IDX_H							0.002					
							(0.004)					
$FFCAR_{0,0}$	−0.037	−0.055*	−0.046	−0.061*	−0.038	−0.036	−0.037	−0.038	−0.055	−0.045	−0.059*	−0.038
	(0.033)	(0.033)	(0.034)	(0.035)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.034)	(0.034)	(0.033)
$FFCAR_{-1,-1}$	−0.045	−0.045	−0.045	−0.042	−0.044	−0.044	−0.044	−0.043	−0.043	−0.043	−0.043	−0.041
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFCAR_{-2,-2}$	−0.025	−0.023	−0.024	−0.033	−0.025	−0.026	−0.026	−0.025	−0.023	−0.024	−0.022	−0.025
	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)	(0.033)
$FFAlpha_{-252,-31}$	0.024	0.028	0.020	0.042	0.026	0.026	0.024	0.020	0.024	0.017	0.022	0.024
	(0.048)	(0.048)	(0.048)	(0.050)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)	(0.048)
$FFCAR_{-30,-3}$	0.031	0.025	0.027	0.047	0.033	0.031	0.030	0.032	0.027	0.029	0.025	0.034
	(0.036)	(0.036)	(0.037)	(0.038)	(0.036)	(0.036)	(0.037)	(0.036)	(0.036)	(0.037)	(0.037)	(0.037)
SUE	0.043	0.050	0.045	0.047	0.043	0.044	0.043	0.044	0.050	0.046	0.051	0.045
	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)	(0.034)
Log(Market Equity)	0.014	0.020	0.013	0.003	0.014	0.015	0.014	−0.014	−0.006	−0.013	−0.006	−0.011
	(0.037)	(0.037)	(0.037)	(0.038)	(0.037)	(0.037)	(0.037)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
Log(Book/Market)	0.424**	0.436**	0.406**	0.456**	0.437**	0.433**	0.422**	0.352*	0.371**	0.342*	0.364*	0.381**
	(0.183)	(0.183)	(0.184)	(0.191)	(0.184)	(0.184)	(0.184)	(0.187)	(0.186)	(0.187)	(0.187)	(0.189)
Log(Share Turnover)	0.082**	0.080**	0.079**	0.094***	0.085**	0.084**	0.082**	0.081**	0.080**	0.079**	0.078**	0.086**
	(0.035)	(0.034)	(0.035)	(0.037)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
iPad								0.189	0.183	0.184	0.180	0.189
								(0.165)	(0.165)	(0.165)	(0.165)	(0.165)
MacBook								−0.502	−0.438	−0.478	−0.424	−0.513
								(0.410)	(0.410)	(0.411)	(0.411)	(0.411)
iWatch								0.025	0.027	0.029	0.029	0.019
								(0.305)	(0.304)	(0.305)	(0.304)	(0.305)
Picture								0.081	0.072	0.077	0.069	0.076
								(0.051)	(0.051)	(0.051)	(0.051)	(0.051)
Observations	963	963	963	927	963	963	963	963	963	963	963	963
Adjusted R^2	0.005	0.010	0.005	0.013	0.004	0.004	0.004	0.006	0.010	0.006	0.009	0.005
Note:												*p<0.1; **p<0.05; ***p<0.01

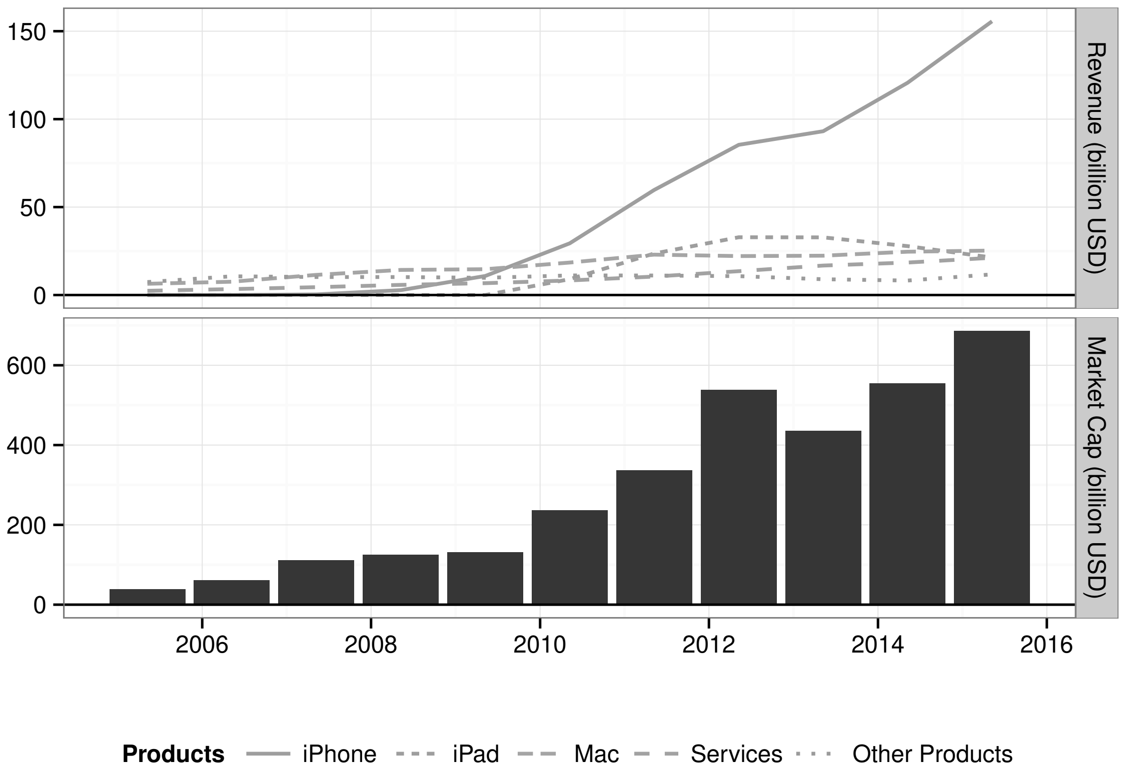

Figure 1: Apple's Revenue and Market Capitalization

Note: The solid line in the upper graph is Apple's revenue attributed to the iPhone, and the black bars in the lower graph show Apple's market capitalization.

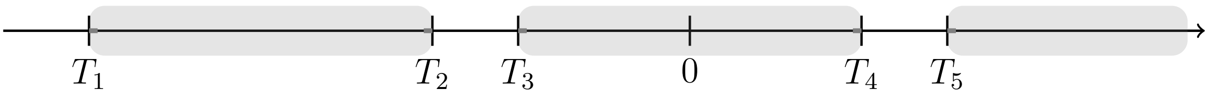

Figure 2: Timeline for Event Studies

Note: 0 in the timeline indicates the exact day when the event happens. The first grey area in the time between T_1 and T_2 is the estimation window, in which the parameters can be estimated. The time interval between T_3 and T_4 is the event window, in which the counterfactual effect is estimated based on the parameters obtained. The interval shown in the last shaded area from T_5 onwards indicates the post-event window.

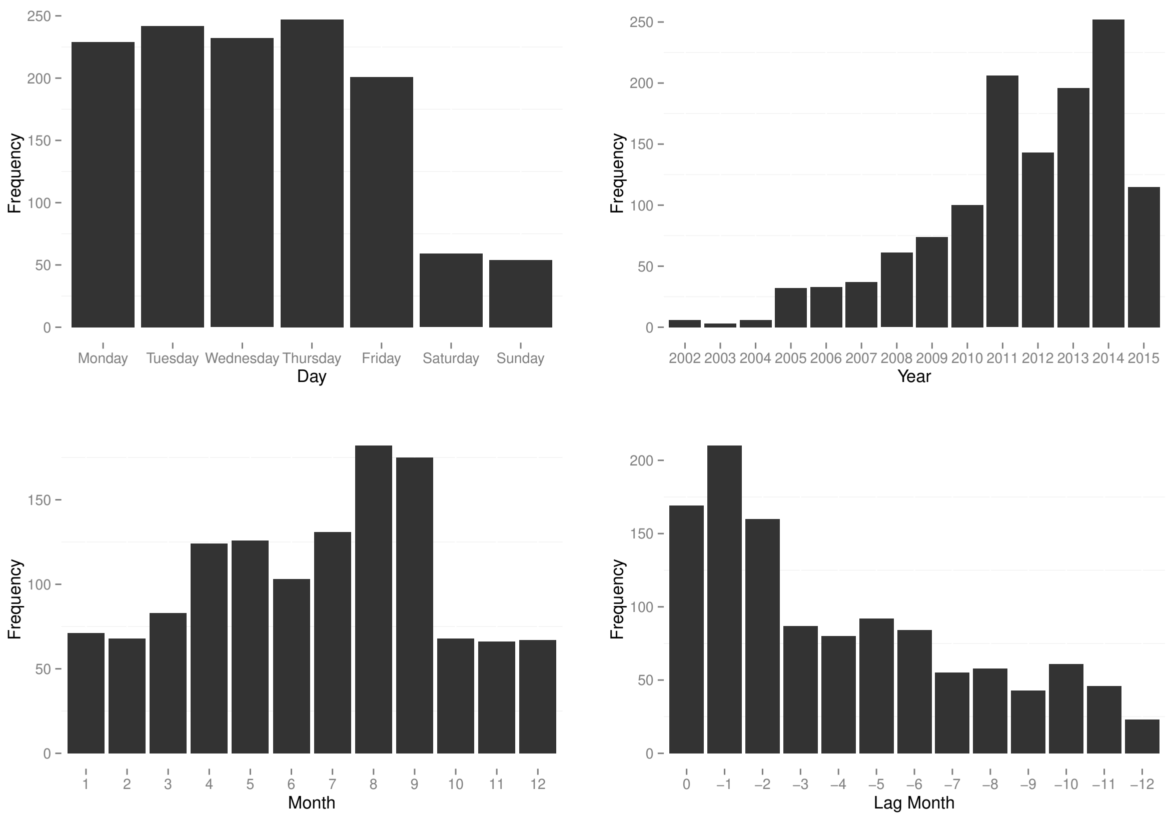

Figure 3: Rumor Data Description

Note: The black bars show the frequencies of rumors in terms of different time spans, including days, months, years, and lag months which refer to the months prior to each official announcement of iPhone.

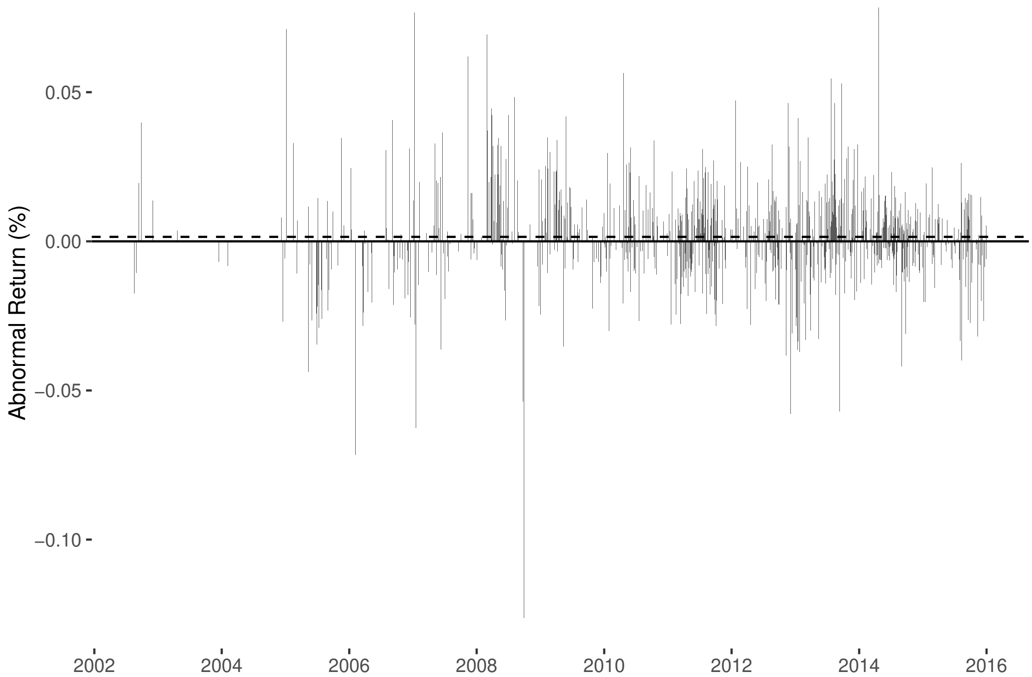

Figure 4: Abnormal Return on the Occurrence Days of Rumors

Note: The dashed horizontal line indicates the mean of AR. The black bars refer to the abnormal returns on the days when rumors occur.

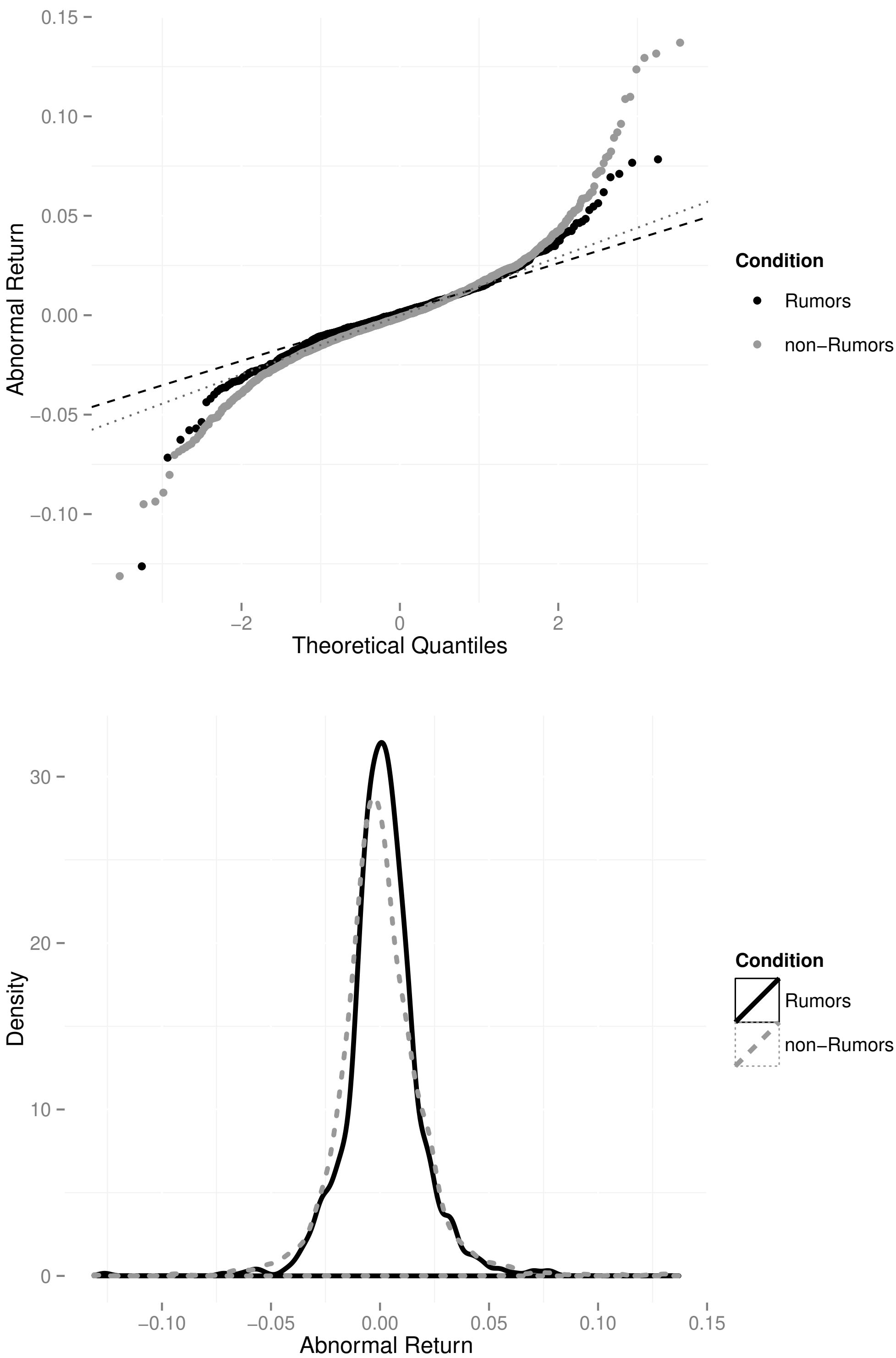

Figure 5: Features of Abnormal Return

Note: The upper graph and the lower graph illustrate the fat-tail feature and high-peak feature of Apple's abnormal returns. The upper chart shows the relationship between the actual quantiles and the theoretical ones for the AR. The dark dashed line and light dotted line represent the theoretically expected normal distribution quantiles for rumor and non-rumor days, respectively. The lower graph presents almost identical density of the AR on rumor and non-rumor days.

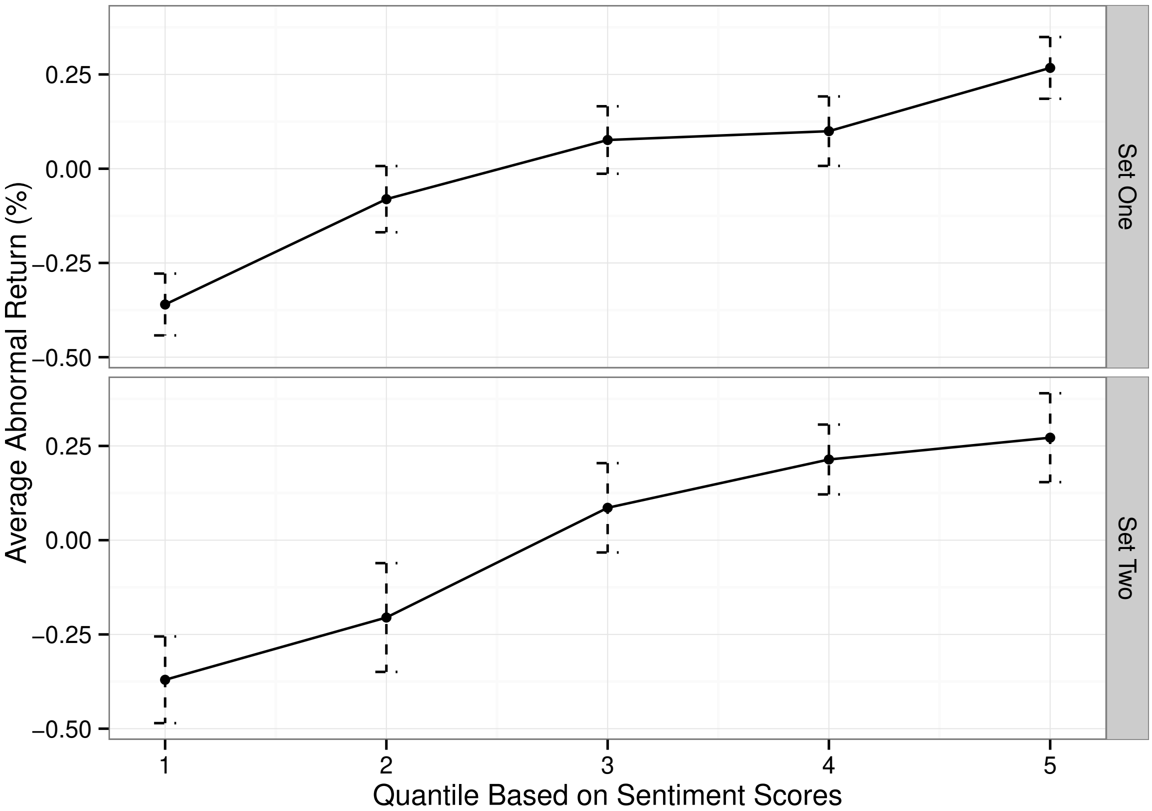

Figure 6: Evaluation of Final Words

Note: To assess the performance of the final words, all the rumors with different occurrence dates are randomly split into two sets, with a ratio of three to two. The sentiment scores for each rumor are calculated and divided into five quantiles. Based on the scores, a mean AR for each group is obtained. Given the AR and the sentiment scores, monotonically increasing patterns are observed in both sets.

Appendices

A Sample of Rumors

Format of the articles related to iPhone rumors in our data set employed includes the following part: date of publish, title of the article, content of the rumor.

Date	Title	Cotent
29/05/2010 1:00 pm	Next iPhone Screen Resolution Confirmed at 960x640	The next iPhone screen has been confirmed to be at running at a resolution of 960x640, according to SuperPhone.cz and magnified it under a microscope. From this procedure, they were able to count the RGB elements of the screen and found it to be exactly twice the density in both horizontal and vertical directions compared to the current iPhone. This would quadruple the total number of pixels and increase the resolution from the current 480x320 to 960x640. This would increase the pixel density to a whopping 320dpi – higher than any other device on the market. Rumors of this high resolution display were first reported back in March. The big advantage of exactly doubling the vertical and horizontal resolution of the iPhone's screen is that backward compatibility with existing apps would be trivial to implement. A similar pixel-doubling system is already in place for running iPhone applications on the iPad.
24/09/2011 10:32 am	Apple Blocking Out Vacations for October 14th iPhone 5 Retail Launch?	AppleInsider reports that Apple has started denying vacation requests for employees during the 2nd week of October. "Apple is quietly denying requests for employee vacations during the second week of October, hinting that the company currently anticipates an influx of customers to its stores around that time related to availability of its new iOS 5 and fifth-generation iPhone products." Specifically, dates from October 9th through 12th and October 14th through 15th are said to be restricted, suggesting Apple plans on launching new products during that time. The first set of dates from October 9th-12th could represent iOS 5 launch, as Apple has informed AppleCare to expect an influx of support calls starting on October 10th. Coincidentally, Twitter has announced their own developer session in London on the same date. The October 10th Twitter developer session is said to have a "heavy focus on the iOS 5 Twitter integration". Meanwhile, the 14th to 15th dates suggests that Apple may be preparing for a retail launch of their next generation iPhones on that Friday, the 14th. The timeframe is consistent with other circulating iPhone rumors as well as our knowledge of Apple's expected availability for the White iPod touch. Apple is expected to hold a media event on October 4th to announce the next generation iPhone as well as the official launch of iOS 5. Update: UK Apple Stores have been similarly told to block out the "first two weeks of October".

Continued on next page

Date	Title	Cotent
10/03/ 2015 6:21 pm	Apple to Add Force Touch and Pink Color Option to Next iPhones, Screen Size to Stay the Same	Apple will add pressure-sensing Force Touch technology to its next-generation iPhones, reports The Wall Street Journal in an article that covers a range of new details on the iPhone 6s and iPhone 6s Plus. Currently built into the upcoming Apple Watch and 12-inch Retina MacBook, Force Touch lets devices distinguish between a light tap and a hard press, enabling new gestures. According to the report, which is sourced from Apple suppliers, Apple's next-generation iPhones will continue to be available in 4.7 and 5.5-inch screen sizes, with plans to "keep the resolution similar." New colors are a possibility though, and Apple is said to be considering adding a pink option to its existing space gray, silver, and gold iPhone lineup. Production may begin on components for the next-generation iPhones as early as May, but The Wall Street Journal notes that Apple often tests technologies and designs with various suppliers that may not make it into finalized products. Today's report echoes several other reports that have also pointed towards Force Touch for the next-generation iPhone. Supply chain sources first hinted at Force Touch technology back in January, and those rumors seem more plausible now that the feature has been incorporated into both the Apple Watch and Apple's recently announced MacBook. Beyond the Force Touch rumors, little is known about the next-generation iPhones, which will likely be called the "iPhone 6s" and the "iPhone 6s Plus." The devices are expected to receive upgraded A9 processors and have been rumored to include features like camera upgrades, more RAM, and improved Touch ID. Apple will presumably release the new iPhones in September.
29/12/2015 10:03 am	'iPhone 7' Waterproof Rumors Persist Amid Claims of Hidden Antenna Bands	Catcher Technology will remain the largest chassis supplier for the upcoming line of next-generation iPhones, tentatively referred to as the "iPhone 7", according to the China-based Commercial Times. Sources noted that Catcher's non-Apple clients, representing about 40 percent of its overall sales, will keep it going until the majority of its output begins with the manufacturing of the iPhone 7 later in 2016. In total, Catcher Technology's manufacturing supply is estimated to account for 30 to 35 percent of the shipment numbers for the iPhone 7. The report also mentioned the continuing rumor that the iPhone 7 may be a completely waterproof device, building on the recent momentum that the iPhone 6s and iPhone 6s Plus had this year that showed improved water resistance. The Commercial Times also spoke of "new compound materials" that would be put in place to form a discreet housing for the iPhone 7's antenna, suggesting the possible removal of the bands from the back of the current iPhone generation. As a non S-generation year, the iPhone 7 is expected to be a big step-up from the iPhone 6s and 6s Plus, in terms of design and functionality, when it launches next year. Besides a waterproof design and now the possibility of a hidden antenna band, another rumor suggested Apple could be phasing out the 3.5mm headphone jack for an all-in-one Lightning connector port.

References

Antweiler, Werner, and Murray Z Frank. 2004. "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards." *The Journal of Finance*, 59(3): 1259–1294.

Bai, Xue, Fu Chen, and Shaobin Zhan. 2014. "A Study on Sentiment Computing and Classification of Sina Weibo with Word2vec." 358–363, Los Alamitos: Institute of Electrical and Electronics Engineers.

Benbunan-Fich, Raquel, and Eliezer M Fich. 2004. "Effects of Web Traffic Announcements on Firm Value." *International Journal of Electronic Commerce*, 8(4): 161–181.

Bettman, Jenni L, Aiden G Hallett, and Stephen Sault. 2011. "Rumortrage: Can Investors Profit on Takeover Rumors on Internet Stock Message Boards?"

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni. 2013. "News, Noise, and Fluctuations: An Empirical Exploration." *The American Economic Review*, 103(7): 3045–70.

Bollen, Johan, Huina Mao, and Xiaojun Zeng. 2011. "Twitter Mood Predicts the Stock Market." *Journal of Computational Science*, 2(1): 1 – 8.

Carrière-Swallow, Yan, and Felipe Labbé. 2013. "Nowcasting with Google Trends in an Emerging Market." *Journal of Forecasting*, 32(4): 289–298.

Chaney, Paul K, Timothy M Devinney, and Russell S Winer. 1991. "The Impact of New Product Introductions on the Market Value of Firms." *The Journal of Business*, 64: 573–610.

Chan, Wesley S. 2003. "Stock Price Reaction to News and No-news: Drift and Reversal After Headlines." *Journal of Financial Economics*, 70(2): 223 – 260.

Das, Sanjiv R, and Mike Y Chen. 2007. "Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web." *Management Science*, 53(9): 1375–1388.

Dong, Ying Hong, Hao Chen, Wei Ning Qian, and Ao Ying Zhou. 2015. "Micro-Blog Social Moods and Chinese Stock Market: The Influence of Emotional Valence and Arousal on Shanghai Composite Index Volume." *International Journal of Embedded Systems*, 7(2): 148–155.

Eckbo, B Espen. 1983. "Horizontal Mergers, Collusion, and Stockholder Wealth." *Journal of Financial Economics*, 11(1): 241–273.

Fama, Eugene F, and Kenneth R French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33(1): 3–56.

Fama, Eugene F., and Kenneth R. French. 2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics*, 116(1): 1 – 22.

Fama, Eugene F, Lawrence Fisher, Michael C Jensen, and Richard Roll. 1969. "The Adjustment of Stock Prices to New Information." *International Economic Review*, 10(1): 1–21.

Goh, Jeremy C, and Louis H Ederington. 1993. "Is a Bond Rating Downgrade Bad News, Good News, or No News for Stockholders?" *The Journal of Finance*, 48(5): 2001–2008.

Horsky, Dan, and Patrick Swyngedouw. 1987. "Does It Pay to Change Your Company's Name? A Stock Market Perspective." *Marketing Science*, 6(4): 320–335.

Jung, Jeeman, and Robert J Shiller. 2005. "Samuelson's Dictum and the Stock Market." *Economic Inquiry*, 43(2): 221–228.

Kiymaz, Halil. 2001. "The Effects of Stock Market Rumors on Stock Prices: Evidence from an Emerging Market." *Journal of Multinational Financial Management*, 11(1): 105–115.

Lane, Vicki, and Robert Jacobson. 1995. "Stock Market Reactions to Brand Extension Announcements: The Effects of Brand Attitude and Familiarity." *The Journal of Marketing*, 59: 63–77.

Liu, Bing. 2007. *Web data mining: exploring hyperlinks, contents, and usage data.* Springer Science & Business Media.

Loughran, Tim, and Bill McDonald. 2011. "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks." *The Journal of Finance*, 66(1): 35–65.

Lucca, David O, and Emanuel Moench. 2015. "The Pre-FOMC Announcement Drift." *The Journal of Finance*, 70(1): 329–371.

Luss, Ronny, and Alexandre d' Aspremont. 2015. "Predicting Abnormal Returns from News Using Text Classification." *Quantitative Finance*, 15(6): 999–1012.

MacKinlay, A Craig. 1997. "Event Studies in Economics and Finance." *Journal of Economic Literature*, 35: 13–39.

Malatesta, Paul H, and Rex Thompson. 1985. "Partially Anticipated Events: A Model of Stock Price Reactions with an Application to Corporate Acquisitions." *Journal of Financial Economics*, 14(2): 237–250.

Mao, Huina, Pengjie Gao, Yongxiang Wang, and Johan Bollen. 2014. "Automatic Construction of Financial Semantic Orientation Lexicon from Large-Scale Chinese News Corpus." *Institut Louis Bachelier*, 20(2): 1–18.

McQueen, Grant, and V Vance Roley. 1993. "Stock Prices, News, and Business Conditions." *The Review of Financial Studies*, 6(3): 683–707.

Mikkelson, Wayne H, and M Megan Partch. 1986. "Valuation Effects of Security Offerings and the Issuance Process." *Journal of Financial Economics*, 15(1): 31–60.

Pang, Bo, and Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis." *Foundations and Trends in Information Retrieval*, 2(1-2): 1–135.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. 2002. "Thumbs up?: Sentiment Classification Using Machine Learning Techniques." Vol. 10 of *EMNLP '02*, 79–86. Stroudsburg: Association for Computational Linguistics.

Pinches, George E, and J Clay Singleton. 1978. "The Adjustment of Stock Prices to Bond Rating Changes." *The Journal of Finance*, 33: 29–44.

Pound, John, and Richard Zeckhauser. 1990. "Clearly Heard on the Street: The Effect of Takeover Rumors on Stock Prices." *The Journal of Business*, 63: 291–308.

Smith, Paul. 2016. "Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data." *Journal of Forecasting*, 35: 263–284.

Sprenger, Timm O, Andranik Tumasjan, Philipp G Sandner, and Isabell M Welpe. 2014. "Tweets and Trades: The Information Content of Stock Microblogs." *European Financial Management*, 20(5): 926–957.

Tetlock, Paul C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market." *The Journal of Finance*, 62(3): 1139–1168.

Tetlock, Paul C., Maytal Saar-tsechansky, and Sofus Macskassy. 2008. "More Than Words: Quantifying Language to Measure Firms' Fundamentals." *The Journal of Finance*, 63(3): 1437–1467.

Tumarkin, Robert, and Robert F Whitelaw. 2001. "News or Noise? Internet Postings and Stock Prices." *Financial Analysts Journal*, 57(3): 41–51.

Turney, Peter D. 2002. "Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews." *ACL '02*, 417–424. Stroudsburg: Association for Computational Linguistics.

Varian, Hal R. 2014. "Big Data: New Tricks for Econometrics." *The Journal of Economic Perspectives*, 28(2): 3–27.

Vosen, Simeon, and Torsten Schmidt. 2011. "Forecasting Private Consumption: Survey-Based Indicators vs. Google Trends." *Journal of Forecasting*, 30(6): 565–578.