International Review of
Accounting, Banking and Finance

IRABF
© 2012

Vol 4, No.3/4, Fall/Winter 2012 Pages 42~82

The Application of Public and Private Information to the Prediction of Abnormal Returns and Portfolio Management

Yang-Cheng Lu,[a] Yu-Chen Wei[b], Yi-Wei Lin[c]

a. Professor, Depatment of Finance, Ming Chuan University, Taipei, Taiwan
b. Assistant Professor, Department of Money and Banking, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan
c. Research Assistant in Local Value-Adding Academic Technology Development Programs of Ministry of Economic Affairs, Ming Chuan University, Taipei, Taiwan

ABSTRACT

By using the technology of text mining to capture the information content of Chinese news, this paper empirically investigates the correlation between the information content of Chinese news and announcement drift. News announced before the earnings announcement influences investor sentiment and the trend of the stock price. This study applies financial news-corps mining by referring to Vega (2006) and Demers and Vega (2011) to proxy the information content of news, for which the proxy variables are measured by the media coverage (*MEDIA*), public news surprises (*SUR*) and sentiment ratio (*SR*). The probability of informed trading (*PIN*), as proposed by Easley, Hvidkjaer and O'Hara (2002), is adopted as the proxy variable for private information. The abnormal return around the earnings announcement date is calculated by the Fama and French three factors model (1992). The standardized unexpected earnings (*SUE*) and turnover (*TURN*) are regarded as the control variables. The empirical results show that there is a positive (negative) relationship between the sentiment ratio (*SR*) and the cumulative abnormal returns before (after) the earnings announcement. Our results confirm that the market's response to news covers the relevant information regarding the company during the earnings announcement period. The application of public news sentiment to portfolio management suggests that "long stocks with low *SR* and high *SUR*" or "short stocks with high *SR* and high *MEDIA*"

before the earnings announcement generate a positive excess return. This study confirms that information sentiment extracted by public news contains relevant information content and plays an important role in the prediction of abnormal returns and portfolio management.

KEY WORDS: **Public News, Sentiment, Linguistic Text Mining, Abnormal Return, Earnings Announcement**

JEL: G12, G14, D82

---

1. Introduction

Which factors drive the process of price movements? According to research on market microstructure, the process of price movements is driven by informed traders having private information or an information advantage. Therefore, the impression is given that public information is mere noise or valueless. However, the availability of public news, which is regarded as a source of public information, means that it is easy and common for investors to obtain information related to companies, and some investors propose that relevant information is contained in the public news. Market trends or security prices will change because investors may have different explanations of the 'information' which will then result in under- or over-reaction.

Both academics and regulators are interested in how investors respond to public information and what the effects of their behaviour are on security price movements. Technical analysis and fundamental analysis are two common ways of analyzing the movement in securities prices, but both of them ignore the impact of public information on short-term (i.e., pre- and post-earnings announcement) variations in securities prices. Rather than adjusting immediately to news surprises, stock prices tend to drift over time in the same direction as the initial surprise. Post-earnings announcement drift (PEAD) or earnings

momentum is one of the anomalies. By extracting sentiment content from information, we may provide some ideas to explain the existence of anomalies.

Existing studies on post-earnings announcements suggest that earnings announcements include information content if the announcements can cause stock prices to change by affecting the investors' expectations regarding the future returns of the stocks, or else the stock prices respond quickly to earnings information after earnings announcements (Ball and Brown, 1968; Beaver, 1986). The trend of cumulative abnormal returns depends on two types of information following an earnings announcement: good or bad information. Companies owning good (bad) information make cumulative abnormal returns increase (decrease) continuously after earnings announcements (Ball and Brown, 1968).

Moreover, companies owning good information are likely to provide earnings predictions voluntarily, and to obtain better cumulative abnormal returns by conveying good information to the stock market (Lev and Penman, 1990). Empirical research regarding quantitative earnings announcement effects on the stock price (Ball and Brown, 1968; Skinner and Sloan, 2002), turnover (Beaver, 1968; Bailey, Li, Mao, and Zhong, 2003) and volatility (Landsman and Maydew, 2002; DeFond, Hung and Trezevant, 2007) also support the idea that earnings information is influential. Schrand and Walther (2000) focus on examining the manipulation of stock prices and find that managers announce news strategically to disclose information related to earnings, and achieve their goal of affecting the stock market. As a result, public information may disturb stock prices due to its nature, and is correlated with the process of cumulative abnormal returns.

With regard to post-earnings announcements, a vast body of empirical research indicates that quantitative information regarding fundamentals has a low degree of correlation with asset price movements (Shiller, 1981; Roll, 1988; Mitchell and Mulherin, 1994). Given that news is likely to disturb stock prices, and that news is correlated with the process of cumulative abnormal returns, does news provide further explanation for the trend of

cumulative abnormal returns? To our knowledge, this question had not yet been answered especially in the Chinese news. One possible reason is that the technology of text mining had not been applied to extract the Chinese language. However, this situation is changing. In recent years, the technology of text mining has been developed expeditiously with the evolution of the computer, and a large number of studies have attempted to probe into the correlation between stock returns and news. There are many sources of news including e-mails, news, annual reports and so on that affect the stock market. Regardless of the source, the information content of news plays an important role in investment strategy (Ahmad, Oliveira, Manomaisupat, Casey and Taskaya, 2002). Because words included in news convey news sentiment, the meanings of words constitute a key factor that influences market movements or stock returns. Through text mining and expert sifting, some key words are used to predict price trends (Wuthrich et al., 1998; Lavrenko et al., 2000). Furthermore, some studies extract the relevant information from public news by using linguistic text mining to construct a quantitative proxy for soft information (Vega, 2006; Demers and Vega, 2011).

Today, linguistic text mining is applied to extract sentiment content from a large number of news items, and to sort news statements into optimistic or pessimistic based on sentiment content (Tan and Zhang, 2008). A stock usually has a higher cumulative abnormal return if relevant news regarding the stock gives rise to higher net optimism. News with sentiment content has an impact on investors' sentiment and behavior, and net optimism in regard to news can be adopted as a factor that leads to expectations of cumulative abnormal returns (Tetlock, 2007; Demers and Vega, 2011). Recently, Engelberg and Parsons (2011) compare the behaviors of investors with access to different media coverage of the same information event in 19 mutually exclusive trading regions. For all earnings announcement of S&P 500 Index firms, the authors find that local media coverage strongly predicts local trading and local trading is strongly related to the timing of local reporting. The impacts of different media sources may be considered in the further study. Except for the analysis between news announcement and stock returns, Jiang, Konstantinidi and Skiadopoulos (2012) examines the effect of US and European news announcements on the volatility spillovers across the markets. Scheduled (unscheduled) news releases resolve (create) information uncertainty, leading to a

decrease (increase) in implied volatility. This study is the pioneer that investigates the spillover effect of volatility across the markets with the incorporation of news effect.

There are different algorithms used to measure the net optimism in managers' earnings announcements in English documents, such as General Inquirer (GI), version 6.0 of the Diction text-analysis program, and the Loughran-McDonald dictionaries (L&M).[3] Up to now, there has still been very little research conducted on information sentiment in the case of Chinese documents.

This study aims to empirically investigate the role of private and public information during the period of an earnings announcement. Specifically, the signals of information sentiment that result from digging out a large number of financial news items are probed by extending the traditional event study of the earnings announcement. We empirically measure the effect of private and public information on the pre- and post- earnings announcements, respectively. In this study, the proxy variable for private information is estimated using the probability of informed trading (*PIN*) proposed by Easley, Hvidkjaer and O'Hara (2002), simplified as EHO (2002) hereafter. The media coverage (*MEDIA*), public news surprises (*SUR*) and sentiment ratio (*SR*) suggested by Vega (2006) and Demers and Vega (2011) are adopted as the proxy variables for public information. Furthermore, the application of private and public information to portfolio management is also investigated by designing a two-dimensional portfolio.

The empirical results show that the market turnover is significantly negatively related to the pre-earnings announcement drift. Besides, there is a positive relationship between the sentiment ratio and the cumulative abnormal returns prior to the earnings announcement. The higher the degree of the sentiment ratio, the higher the cumulative abnormal return. We

[3] The readers could refer to Tetlock (2007) and Demers and Vega (2011) for further details.

also find that the leading effect of the information sentiment (sentiment ratio, *SR*) constructed based on public news prior to the earnings announcement continues for four weeks after the earnings announcement.

Moreover, there exists significantly positive correlation between *PIN* adopted as the proxy for private information and cumulative abnormal returns two to four weeks before with the pre-earnings announcement except for the non-significant correlation one week before with the pre-earnings announcement. In addition, there is a positive relationship between *PIN* and the post-earnings announcement.

Although some investors believe that public information is merely noise, other investors propose that there is relevant information contained in the public news. Therefore, we construct proxy variables for private and public information to measure their information content, and then combine these proxy variables to examine the correlation with the stock price dynamic. To the best of our knowledge, private information content has an impact on the stock price dynamics, which is consistent with the research on the market microstructure. However, this does not mean that public information is noise. Public information also comprises relevant information related to stock price movements.

Our contribution to the existing literature is two-fold. Firstly, this study extracts the information content related to the sentiment from the Chinese news by using linguistic text mining. The quantitative indicators describing the sentiment ratio (*SR*) and the media coverage (*MEDIA*) that are available to investors prior to the earnings announcements are constructed from Chinese financial news. Then *SR* and *MEDIA* are combined with the financial quantitative dataset to examine the relationship with abnormal returns and to predict the post-earnings announcement drift around the announcement period. Secondly, this study also contributes to the literature by comparing private information with public information in the Taiwan Stock Market. Consequently, this study enhances the understanding of the behavior of stock price reactions to private and public information. In addition, public information

containing relevant content is regarded as one of the factors driving the process of price movements.

The remainder of this paper proceeds as follows. In Section 2, we describe the data and the different measures of public and private information. In Section 3, we present the ideas underlying the regression analysis. Section 4 summarizes the empirical results including the summary statistics, cross-sectional regression analysis and portfolio management. Finally we conclude in Section 5.

2. Sample and Data Description

Our analysis is conducted based on the corporate earnings announcements over the period 2001-2009. The financial data were obtained mainly from the quarterly financial statements of listed companies in the Taiwan Economic Journal (TEJ) database and the news-corpus information was collected from the InfoTimes database which collects the daily news reports.[4]

The financial regulators in Taiwan stock market is the Securities and Futures Bureau (SFB) in Financial Supervisory Commission Executive Yuan. The SFB has stipulated the policy directives for the purpose of administering and supervising securities issuance, securities trading and futures trading. The companies listed in the Taiwan Stock Exchange (TWSE) and GreTai Securities Market (GTSM) are required to announce the quarterly financial statements regularly and these financial data and earnings announcement date would be inquired publicly in the Market Observation Post System (MOPS). Interested readers could refer to the web site of Market Observation Post System (MOPS) for comprehensive details.[5] Except for collecting the information of each samples individually from the MOPS, there are

[4] Details of the introduction of the TEJ and InfoTimes database can be found at http://www.tej.com.tw/ and http://www.infotimes.com.tw/, respectively.
[5] Details of the introduction can be found at http://emops.twse.com.tw/emops_all.htm.

representative databases gather the reports systematically. The database of Taiwan Economic Journal (TEJ) is one of the representative database providers in Taiwan.

News reports regarding the announcement samples are collected from the InfoTimes database. The news collected from InfoTimes database comprises the Commercial Times and Chinese Times published by China Times Group which is one of the representative media in Taiwan. China Times Group truly a typical media organization in Taiwan, since most all of the securities firm, institutions and investors subscribe the news report or review it in the website from China Times Group. United Daily News (udn) Group is the other group published the related financial news regarding the listed companies in the Taiwan Stock Market. We do not use the udn news in this paper. However, the further research could try to combine both the sources of financial news for the robustness check.

The proxy variables of public information were measured by the media coverage (*MEDIA*), public news surprises (*SUR*) and sentiment ratio (*SR*). The probability of informed trading (*PIN*) proposed by Easley, Hvidkjaer and O'Hara (2002) was adopted as the proxy variable of private information. We also incorporated other control variables, namely, turnover (*TURN*) and the unexpected earnings surprises (*SUE*).

2.1 Abnormal returns and cumulative abnormal returns

The abnormal return used in this study is calculated by considering the three factors model of portfolio beta (*PBeta*), firm size (*SIZE*) and book to market value (*BM*). The event date is the quarterly earnings announcement date. The cumulative abnormal returns T days before and after the event date are defined as $CAR_{[-T,-1]}$ and $CAR_{[1,\ \eta]}$, respectively.

We calculated the betas by following the approach of Fama and French (1992) and Lu and Wong (2009) who revise the appropriate calculation in the Taiwan Stock Market. The

market value of equity is the price multiplied by the number of outstanding shares as at the end of month t. The SIZE variable at month t is simply defined as the log of the market value of each firm. TEJ uses the most recently available value of the book value of common equity (reported quarterly) to construct the book-to-market (*BM*) variable. Similar to the Fama-French approach, firms with negative book values are excluded.

2.2 Proxy of Public Information

Public Information Measures

By referring to Vega (2006), we define media coverage (*MEDIA*) as the number of days a particular firm is mentioned in the news prior to its earnings announcement.

$$MEDIA_{i,t} = \sum_{k=1}^{30} NEWS_{i,t-k} \qquad (1)$$

where $NEWS_{i,t-k}$ is a dummy variable equal to one if firm *i* is mentioned in the headline or lead paragraph of an article on day *t*. To measure $MEDIA_{i,t}$ we use the InfoTimes database which includes the representative Chinese news.

In considering the stock market's reaction to headline news, we construct the measure of public news surprises (*SUR*) prior to a firm's earnings announcement by referring to Vega (2006). By incorporating the properties of the Taiwan stock market, *SUR* is adjusted to capture the abnormal return and abnormal market turnover as,

$$SUR_{i,t} = \sum_{k=1}^{K} \{NEWS_{i,t-k} \left[I(AR_{i,t-k} \geq AR_{i,XT}) + I(AR_{i,t-k} \leq AR_{i,XB}) \right]\} \times \left\{ I(DV_{i,t-k} \geq DV_{i,YT}) + I(DV_{i,t-k} \leq DV_{i,YB}) \right\}, \qquad (2)$$

Where *K* is the calendar days prior to the earnings announcement and *K* equals to 7, 14, 21 and 28. $I(AR_{i,t-k} \geq AR_{i,XT})$ and $I(AR_{i,t-k} \leq AR_{i,XB})$ are indicator functions equal to one if the abnormal stock return for firm *i* on day *t* or day *t*+1 is above the top *XT*% and below the

bottom *XB*%, respectively, of daily abnormal stock returns for that firm. $I(DV_{i,t-k} \geq DV_{i,YT})$ **and** $I(DV_{i,t-k} \leq DV_{i,YB})$ **are indicator functions equal to one if the turnover for firm *i* on day *t-k* or day *t-k*+1 is above the top *YT*% and below the bottom *YB*%, respectively. The percentages of *XT*% and *YT*% are in the top 20% and the percentages of *XB*% and *YB*% are in the bottom 20%.**

2.3 Public Information Sentiment (SR)

There is no representative system revealing or constructing the information sentiment for Chinese documents. Therefore, we collect the Chinese financial news and conduct the process of information quantification presented in Figure 1.

Referring to Diction, optimism is defined as 'language endorsing some person, group, concept or event or highlighting their positive entailments'. The Diction formula for net optimism is [praise + satisfaction + inspiration] - [blame + hardship + denial] which is the difference between 'optimism' and 'pessimism'. In considering that there is no certain 'optimism' and 'pessimism' in Chinese documents, we translate the classification of sentiment and collect the related sentiment words.[6] The net optimism used in this study is measured as:

$$SR_{i,t} = \frac{\sum_{p=1}^{P} tf_{ip} - \sum_{n=1}^{N} tf_{in}}{TF_{i,t}} \times 100\% \quad \textbf{(3)}$$

where $SR_{i,t}$ is the net optimism of the *i*th firm at time *t*, tf_{ip} is the term frequency of the *p*th optimism characteristic terms of the *i*th firm, tf_{in} is the term frequency of the *n*th pessimism characteristic terms of the ith firm, and $TF_{i,t}$ is the total term frequency of the *i*th firm at time *t*. The selected optimism and pessimism characteristic terms are listed on Table 1.

[6] We translate the key words of 'optimism' and 'pessimism' by using the Academia Sinica Bilingual Ontological Wordnet. (http://bow.sinica.edu.tw/)

[Insert Figure 1 Here]

[Insert Table 1 Here]

2.4 Proxy for Private Information: The probability of informed trading (PIN)

The distinguishing feature for the Taiwan equity market lies in that there is no market maker at the Taiwan Stock Exchange and the market is purely order driven. So the probability of informed trading (*PIN*) model by EHO (2002) can be viewed as a microstructure model in which market participants observe trading activities and draw inferences about the underlying true value of an asset. Each trade conveys private information that updates the beliefs of market participants, who in turn set the trading prices based on the updated beliefs. Over time, the process of trading, learning and price setting results in prices converging to full information levels.

Figure 2 depicts the simple sequential trade tree diagram considered by EHO (2002). First, a news event occurs with probability α. There is a $(1-\delta)$ chance that the news is classified as good, and so a good news event occurs with probability $\alpha(1-\delta)$. Similarly, a bad news event occurs with probability $\alpha\delta$. Traders are assumed to arrive according to Poisson processes throughout the day. Orders from informed traders arrive at rate μ, whereas orders from uninformed buyers and sellers arrive at rates ε_b and ε_s, respectively. Basically, informed traders buy if they perceive good news and sell if the news is perceived to be otherwise. Market participants update their beliefs by extracting information from the buy and sell trades. New prices are formed, trades evolve, and the price dynamics reflect the market participants' changing beliefs.

[Insert Figure 2 Here]

Crucial to the *PIN* model is its estimate of the probability of trade based on private information regarding the stock. The calculation of *PIN* is referred to in EHO (2002) and Lu and Wong (2009). Let *B* and *S* denote the number of buy and sell trades in a single trading day,

respectively. We use the mid-quote and transaction price rules of Lee and Ready (1991) and Lu and Wei (2009) to calculate the required *B* and *S* for each firm on each day.[7] According to EHO (2002), the likelihood function induced by this simple model of the trade process for a single trade day is given below:

$$L(\theta \mid B,S) = (1-\alpha)\cdot e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}\cdot e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + \alpha\delta\cdot e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}\cdot e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^S}{S!} + \alpha(1-\delta)\cdot e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^B}{B!}\cdot e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!}.$$

(4)

In the above equation, the first line refers to the likelihood weighted by the probability of a day with no event (1 – *α*), the second line refers to that of a "bad news day" (*αδ*) and the last line refers to that of a "good news day". Let $\theta = (\alpha, \mu, \varepsilon_b, \varepsilon_s, \delta)$ be the parameter vector and use (B_i, S_i) to denote the buy-sell trade data for period $i \in (1,\ldots,I)$. Assuming sufficient independence conditions are held across the *I* trading days and the *I* equals to 60 days prior to the event date. The likelihood function for the period is

$$V = L(\theta \mid M) = \prod_{i=1}^{I} L(\theta \mid B_i, S_i),$$

(5)

where $M = \{(B_i, S_i)\}_{i=1}^{I}$ refers to the data set. Maximizing the likelihood above will give us the ML estimator for θ, from which we can estimate the probability that the trade is information-based as follows:

$$PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_b + \varepsilon_s}.$$

(6)

[7] See also Ellis et al. (2000) for further results on the calculation of buys and sells.

2.5 Unexpected Earnings Surprise

Following Demers and Vega (2011), we use a seasonal random walk model to capture the unexpected earnings surprise. We define the unexpected earnings as $UE_{iq}=A_{iq}-E_{iq}$, where A_{iq} is the earnings per share of firm i for fiscal quarter q, and E_{iq} is the proxy for the market's expectation of earnings which can be measured by the previous year's same quarter earnings per share for the TEJ sample. We standardize the unexpected earnings by dividing the surprise by the firm-specific standard deviation of the forecast error, and then the standardized unexpected earnings associated with firm i for quarter q is defined as SUE_{iq}.

3. Empirical Analysis

3.1 Summary Statistics

Before analyzing the relationship between the soft/hard information and the pre-/post-earnings announcement drift, we first show the summary statistics of the variables. To further classify the relevance of soft information before the earnings announcement, we construct different frequencies of the proxies of soft information including 7-, 14-, 21-, and 28- calendar-days prior to the earnings announcement. Table 2 presents the minimum, maximum, mean, median and standard deviation of the soft information, hard information and the control variables. Table 3 shows the correlation coefficient between the soft/hard information and the cumulative abnormal return before and after the earnings announcement. The soft information presented in Table 3 is constructed by using the 7-, 14-, 21-, and 28- days public information prior to the earnings announcement.

[Insert Table 2 Here]

[Insert Table 3 Here]

3.2 Cross-Sectional Regression Analysis between CAR and Relevant Information

We first investigate the cumulative abnormal returns' response to the hard and soft information surprises contained before the earnings announcement by using equation (7). The dependent variable is defined as the *t*-trading-day cumulative abnormal returns which correspond to the period in which we collect the public financial news. Then we examine the leading effect of the hard and soft information surprises by analyzing the relationship between hard and soft information collected before the earnings announcement day and the cumulative abnormal return after the earnings announcement by using equation (8). This leads to the following pooled regression model:

$$CAR_{i,[-T,-1]} = \beta_0 + \beta_1 SR_{i,t-1} + \beta_2 MEDIA_{i,t-1} + \beta_3 SUR_{i,t-1} + \beta_4 PIN_{i,t-1} + \beta_5 TURN_{i,t-1} + \beta_6 SUE_{i,t-1} + \varepsilon_{i,t}$$

(7)

$$CAR_{i,[1,T]} = \beta_0 + \beta_1 SR_{i,t-1} + \beta_2 MEDIA_{i,t-1} + \beta_3 SUR_{i,t-1} + \beta_4 PIN_{i,t-1} + \beta_5 TURN_{i,t-1} + \beta_6 SUE_{i,t-1} + \varepsilon_{i,t}$$

(8)

where the subscript *t*-1 indicates that the variable is estimated during the pre-announcement period, $CAR_{i,[-T,-1]}$ is the cumulative abnormal return estimated during the pre-announcement period, $CAR_{i,[1,\eta]}$ is the cumulative abnormal return estimated during the post-announcement period, and 0 is the earnings announcement day.

Panel A in Table 4 presents the cross-sectional regression analysis 7-calendar days prior to the earnings announcement. This means that soft information, including *MEDIA*, *SUR* and *SR*, are constructed by collecting public news one week or 7-calendar days prior to the earnings announcement. Panel B, Panel C, and Panel D show the cross-sectional regression analysis two weeks, three weeks, and one month prior to the earnings announcement day.

The empirical results show that the market turnover is negatively significant in relation to the pre-earnings announcement drift. In other words, there is a significant negative relationship between the market turnover and the cumulative abnormal return regardless of whether the study

period prior to the earnings announcement is one week, two weeks, three weeks, or one month. Model (2) in Table 4 indicates that there is a significant relationship between the proxy for private information, *PIN*, and the pre-earnings announcement drift except for one week prior to the earnings announcement.

For the proxies of the soft information, *MEDIA* and *SUR* refer to the degree of the disclosure of the specific firm in the public news. Models (4) and (5) in Table 4 reveal that *MEDIA* and *SUR* exhibit significant explanatory power in relation to the pre-earnings announcement drift for two to four weeks, while *MEDIA* and *SUR* are included in the model individually. Model (3) in Table 4 depicts the empirical results for the proxy for information sentiment, namely, the net optimism of the specific firm revealed in the public news. The analysis shows that there is a positive relationship between the net optimism and the cumulative abnormal returns prior to the earnings announcement. The higher the degree of the net optimism, the higher the cumulative abnormal returns.

Model (6) in Table 4 compares the explanatory power of the soft and hard information prior to the earnings announcement day. *SR* is positively significant 7- and 14-days prior to the earnings announcement, *SUR* is negatively significant 21- and 28-days prior to the earnings announcement, *MEDAI* is positively significant for all periods studied prior to the earnings announcement, and *PIN* is positively significant 14-, 21-, 28-days prior to the earnings announcement.

To further investigate the signal effect of the soft and hard information in relation to the post-earnings announcement drift, we use the 1-day to 22-day post-announcement period to calculate the cumulative abnormal return. The significance of the soft and hard information prior to the earnings announcement to the 1-day to 22-day post-announcement market response is presented in the **Figure 3**. Panel A to Panel D in **Figure 3** represents the relationship between the one week to one month private/public information prior to the earnings announcement and the one day to one month post-earnings announcement drift. We can find that the leading effect of the information sentiment (sentiment ratio, *SR*) constructed based on the one week public news prior to the earnings announcement continues until one month after the earnings announcement. The

information sentiment collected two weeks, three weeks or one month prior to the earnings announcement also has the same leading effect.

MEDIA and *SUR* constructed one week, two weeks, three weeks or one month prior to the earnings announcement exhibit a significant leading effect. However, the findings for *MEDIA* and *SUR* are quite different. There is a positive (negative) relationship between *MEDIA* (*SUR*) and the post-earnings announcement drift and the relationship is significant after one week of the earnings announcement.

The proxy for private information of *PIN* shows that although there is no significant relationship between *PIN* and the pre-earnings announcement one week before the earnings announcement, there is a significant positive relationship between *PIN* and the post-earnings announcement. The leading effect of *PIN* does not exist significantly until two weeks later after the earnings announcement. In other words, informed traders would like to hold the stocks before the earnings announcement and then close their positions after the earnings have been announced.

[Insert Figure 3 Here]

[Insert Table 4 Here]

3.3 Portfolio Management

If we can confirm the information content of the soft and hard information, we can then provide the average excess returns for portfolios of stocks sorted independently, for example by net optimism (*SR*) and public news surprises (*SUR*).

We combine the proxy variables of private and public information to investigate whether the relevant information could be an indicator for portfolio management. Different portfolios are designed and simulated to analyze the leading effect of private and public information. For each quarter of earnings announcement, stocks are sorted individually into five groups by

referring to the private and public proxy variable, ranging from low to high, and then 25 portfolios are constructed. Portfolios constructed from private and public information content reveal an important link between relevant information and average excess returns. The variations in the average excess returns during the 25 portfolios support the application of the relevant information before the earnings announcement to the trading strategies.

We provide Table 5 to Table 8 below which present the average excess returns for the portfolios of stocks sorted independently by *SR* and *MEDIA*, *SR* and *SUR*, *SR* and *PIN*, *MEDIA* and *PIN*, and *SUR* and *PIN* to confirm the application of public and private information.

Each Table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by public and private information proxies. The soft information is collected 7 days prior to the earnings announcement. *PIN* is calculated by using the 60 trading days data prior to the earnings announcement. The holding periods of the excess returns are four weeks after the earnings announcement since the empirical results in the Figure 3 supports the four weeks leading effect of soft information. The Panel A in Table 5 and Table 6 show that the application of public and private information to portfolio management suggests that "long stocks with low *SR* and high *SUR*" or "short stocks with high *SR* and high *MEDIA*" before the earnings announcement generate a positive excess return.

The portfolios of stocks sorted independently by soft information and *PIN* in Table 7 to Table 9 does not present the superior performance than the portfolios that are calculated by the combinations of soft information in Table 5 and Table 6. The reasons might lie in that the soft information used to simulation the portfolios with *PIN* is 7 days prior to the earnings announcement, however Panel A in Figure 3 shows that the leading effect of *PIN* is weak while it is combined with the short term 7 days soft information of *SR*, *MEDIA* and *SUR*.

[Insert Table 5 Here]

[Insert Table 错误！未找到引用源。 Here]

[Insert Table 6 Here]

[Insert Table 7 Here]

[Insert Table 8 Here]

4. Conclusions

By using the technology of text mining to capture the information content of Chinese news, we are able to proxy the information content of public news and analyze the behavior of stock responses to public news from the pre- to the post-announcement period. Furthermore, this research also examines the effect of private information on the announcement drift, and compares this with public information to analyze the movement of private and public information. The purpose of this study is to confirm whether the public information is news or noise.

The proxy variables for public information are constructed by using the linguistic text mining referred to in Vega (2006) and Demers and Vega (2011), and are measured by the media coverage (*MEDIA)*, public news surprises (*SUR*) and sentiment ratio (*SR*). The probability of informed trading (*PIN*), proposed by Easley, Hvidkjaer and O'Hara (2002), is adopted as the proxy variable for private information.

The results, which extend the traditional event study of the earnings announcement, indicate that not only the signals of information sentiment based on digging out a large number of financial news items but also private information based on adopting the probability of informed trading are correlated with cumulative abnormal returns for pre-announcements and exhibit a leading effect on post-announcements. Our results suggest that there is a positive relationship between the sentiment ratio (*SR*) and the cumulative abnormal return before the earnings announcement which confirms that the market response covers the relevant information regarding the company. A higher sentiment ratio (*SR*) reflects a higher certainty about earnings information, which is consistent with the findings of Demers and Vega (2011).

Our study also provides empirical evidence that the probability of informed trading (*PIN)* has a notably positive connection with the accumulation of an abnormal return after the earnings announcement, which indicates that informed traders would like to conceal the trading activities prior to the disclosure of earnings information and close their positions after the earnings are announced.

The application of information sentiments to portfolio management suggests that "long stocks with low *SR* and high *SUR*" or "short stocks with high *SR* and high *MEDIA*" before the earnings announcement generate positive excess returns.

In conclusion, our findings suggest that the public news owning information content can be regarded as an important factor in the price discovery process. Information sentiments regarding individual stocks could be applied to the warning model of abnormal returns and portfolio management.

Acknowledgement

This study was supported by a grant from the Department of Industrial Technology at the Ministry of Economic Affairs in Taiwan (Grant number 99-EC-17-A-34-S1-149). We would also like to express our appreciation for the support provided by the Chinese Knowledge and Information Processing Group of the Institute of Information Science and the Institute of Linguistics at Academia Sinica

References

Ahmad, K., P.C.F.D. Oliveira, P. Manomaisupat, M. Casey and T. Taskaya. 2002. "Description of Events: An Analysis of Keywords and Indexical Names," *the 3rd International Conference on Language Resources and Evaluation*, pp. 29-35.

Bailey, W., H. Li, C.X. Mao and R. Zhong. 2003. "Regulation Fair Disclosure and Earnings Information: Market, Analyst, and Corporate Responses," *Journal of Finance*, Issue 58, Vol.6, pp. 2487-2514.

Ball, R. and P. Brown. 1986. "An Empirical Evaluation of Accounting Income Numbers," *Journal of Accounting Research*, Issue 6, Vol.2, pp. 159-178.

Beaver, W. 1986. "The Information Content of Annual Earnings Announcements," *Journal of Accounting Research*, Vol.6, pp. 67-92.

DeFond, M., M. Hung and R. Trezevant. 2007. "Investor Protection and the Information Content of Annual Earnings Announcements: International Evidence," *Journal of Accounting and Economics*, Vol.43, pp. 37-67.

Demers, E. and V. Clara. (2011), "Linguistic Tone in Earnings Announcements: News or Noise?", FRB International Finance Discussion Paper No. 951.

Easley, D., S. Hvidkjaer and M. O'Hara. 2002. "Is Information Risk a Determinant of Asset Returns," *Journal of Finance*, Issue 57, Vol.5, pp. 2185-2221.

Engelberg, J. E. and C. A. Parsons. (2012), "The Causal Impact of Media in Financial Markets," *Journal of Finance*, Vol. 66, pp. 67-97.

Jiang, G. J., E. Konstantinidi and G. Skiadopoulos. (2012), "Volatility Spillovers and the Effect of News Announcements," *Journal of Banking & Finance*, Vol. 36, pp. 2260-2273.

Landsman, W.R. and D.L Maydew. 2002. "As the Information Content of Quarterly Earnings Announcements Declined in the Past Three Decades," *Journal of Accounting Research*, Vol.40, pp. 797-808.

Lavrenko, V., M. Schmill, D. Lawrie, P. Ogilvei, D. Jensen and J. Allan. 2000. "Mining of Concurrent Text and Time Series," the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining Workshop on Text Mining, pp.37-44.

Lee, C.M.C. and M.J. Ready. 1991. "Inferring Trade Direction from Intraday Data," *Journal of Finance*, Vol.46, pp.733-746.

Lev, B. and S. Penman. 1990. "Voluntary Forecast Disclosure, Nondisclosure, and Stock Prices," ***Journal of Accounting Research*****, Vol.28, pp.49-76.**

Lu, Y.C. and Y.C. Wei. 2009. "Classification of Trade Direction for an Equity Market with Price Limit and Order Match: Evidence from the Taiwan Stock Market," ***Investment Management and Financial Innovations*****, Issue 6, Vol.2, pp. 23-35.**

Lu, Y.C. and W.K. Wong. 2009. "Probability of Information-Based Trading as a Pricing Factor in the Taiwan Stock Market," ***International Research Journal of Finance and Economics*****, Issue 33, pp. 31-49.**

Schrand, C.M. and B. Walther. 2000. "Strategic Benchmarks in Earnings Announcements: The Selective Disclosure of Prior-Period Earnings Components," ***The Accounting Review*****, Vol.75, pp. 151-177.**

Skinner, D. and R. Sloan. 2002. "Earnings Surprises, Growth Expectations, and Stock Returns or Don't Let an Earnings Torpedo Sink Your Portfolio," ***Review of Accounting Studies*****, Vol.7, pp. 289-312.**

Tan, S. and J. Zhang. 2008. "An Empirical Study of Sentiment Analysis for Chinese Documents," ***Expert System with Applications*****, Issue 34, Vol.4, pp. 2622-2629.**

Tetlock, P.C. 2007. "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," ***Journal of Finance*****, Vol.62, pp.1139-1168.**

Vega, C. 2006. "Stock Price Reaction to Public and Private Information," ***Journal of Financial Economics*****, Vol.82, pp.103-133.**

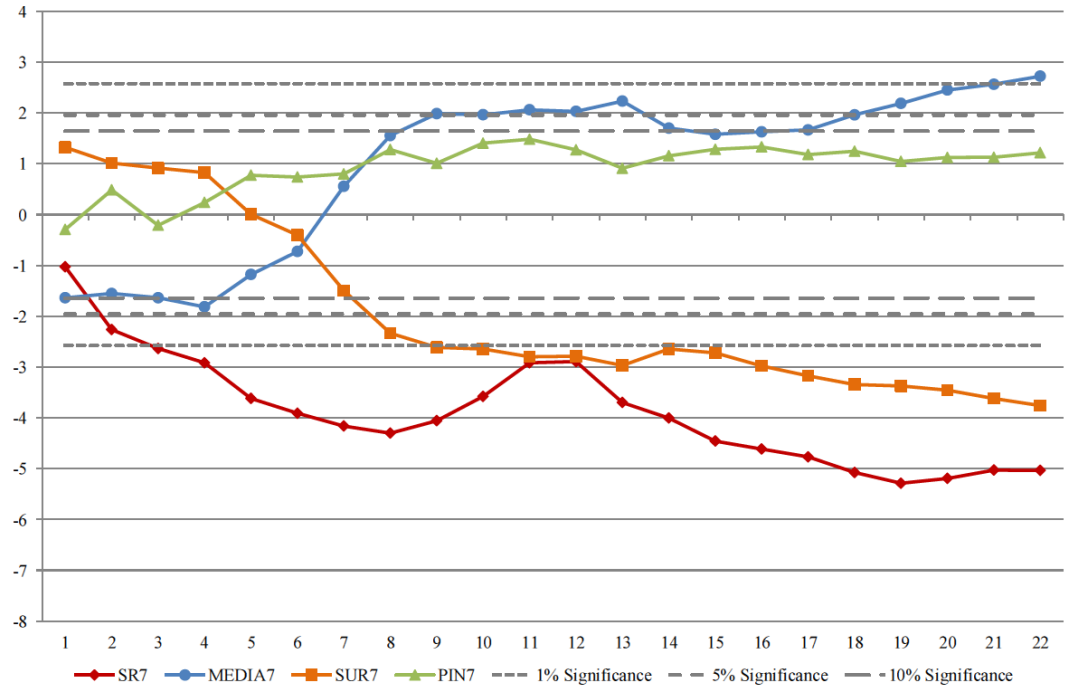
Wuthrich, B., V. Cho, S. Leung, D. Permunetilleke, K. Sankaran, J. Zhang and W. Lam. 1998. "Daily Stock Market Forecast from Textual Web Data." ***Paper presented at the IEEE International Conference on Systems, Man and Cybernetics*****, pp. 2720-2725.**

Figure 1 Diagram of the process of information quantification

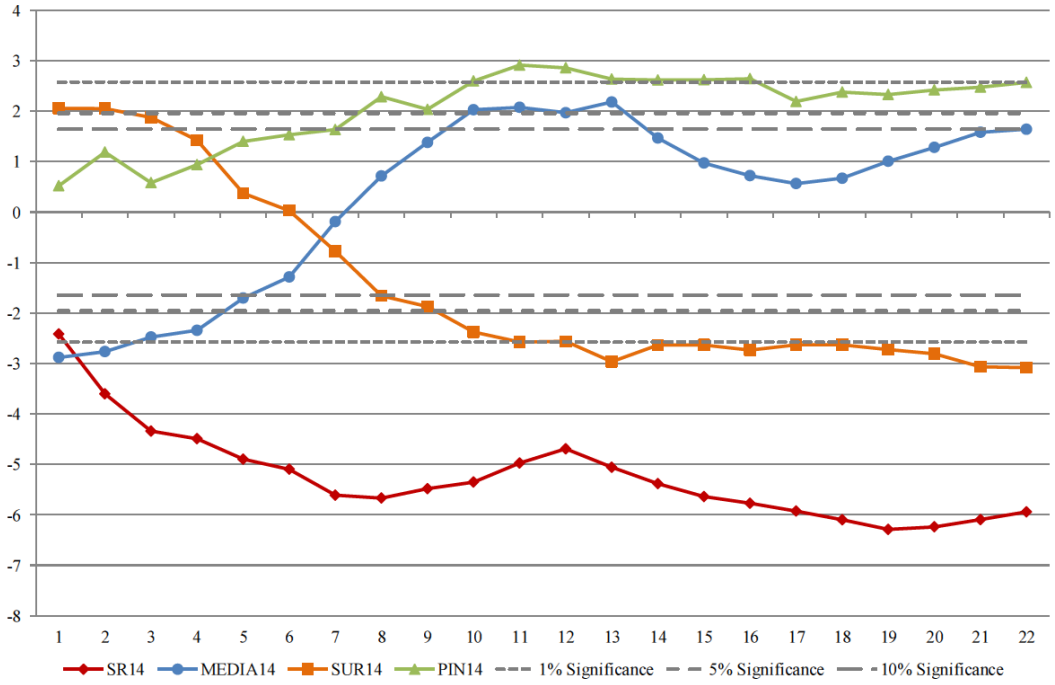
Figure 2 Microstructure tree diagram of the trading process

Notes: In the diagram above, α refers to the probability of an information event whereas δ refers to the probability of a bad news event. μ, ε_b and ε_s are the arrival rates of informed trade, uninformed buy and uninformed sell orders, respectively. Events to the left of the dotted line occur once per day.

Panel A The relationship between the private and public information 7-days prior to the earnings announcement and the post-earnings announcement drift



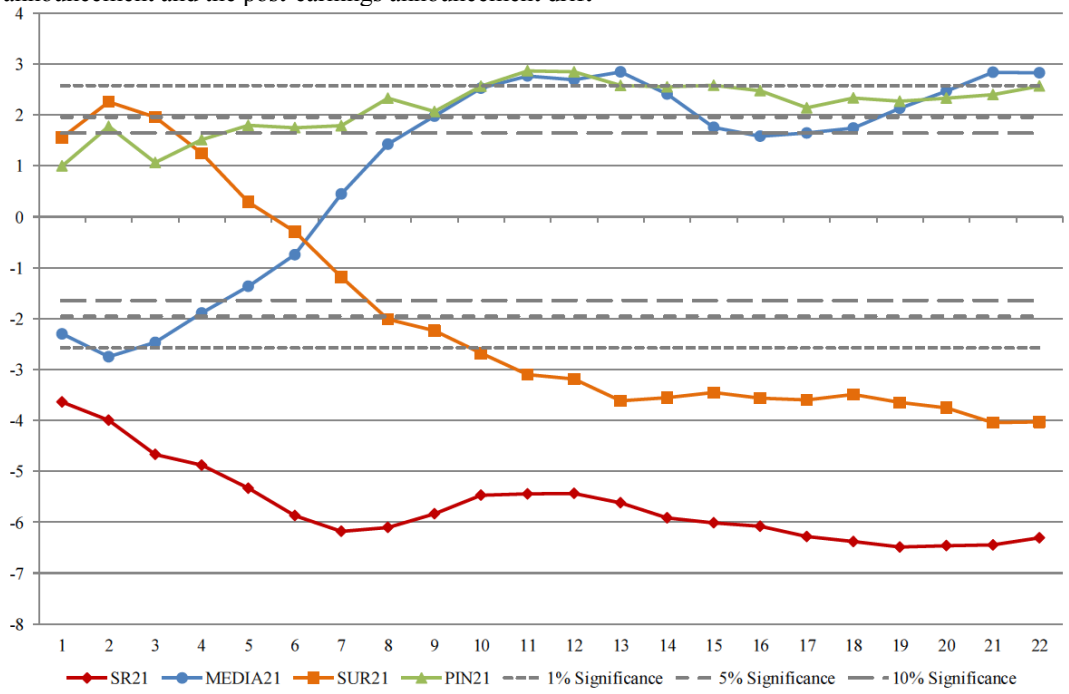
Panel B The relationship between the private and public information 14-days prior to the earnings announcement and the post-earnings announcement drift



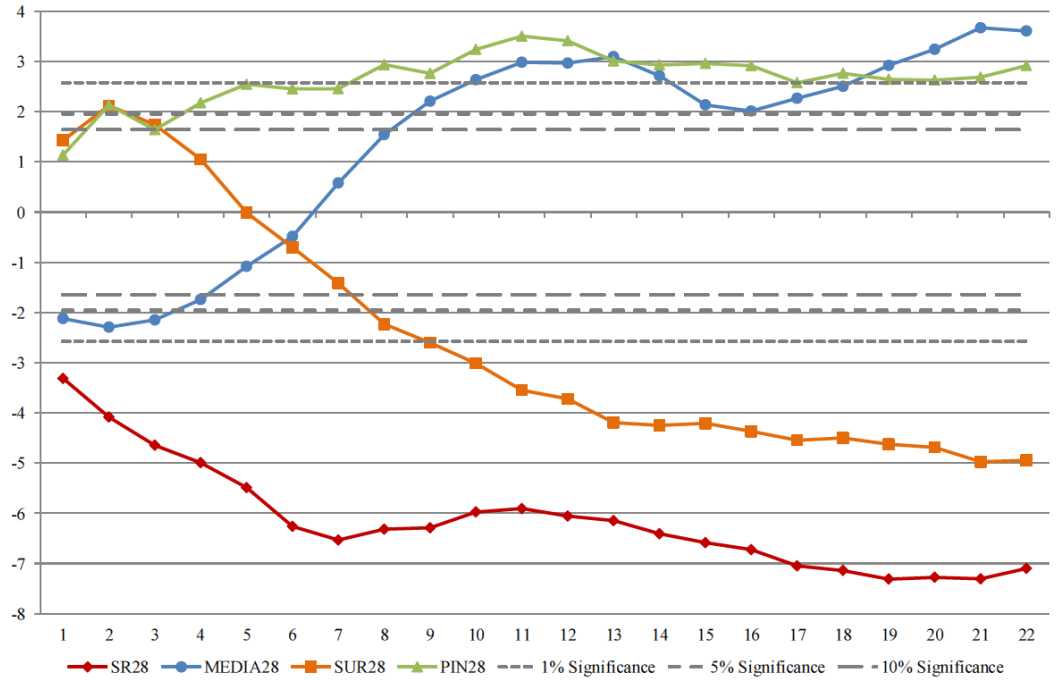
Notes: The vertical axis represents the *t*-statistics and the horizontal axis represents the trading days after the earnings announcement date. The dots are the *t*-values examined by the regression analysis between the $CAR_{[-1,-T]}$ and the public and private information, T is 1, 2, …22. The variable is positive (negative) significant to the $CAR_{[-1,-T]}$ if the *t*-value is above (below) the significance level.

Figure 3 The relationship between the private and public information and the post-earnings announcement drift.

Panel C The relationship between the private and public information 21-days prior to the earnings announcement and the post-earnings announcement drift



Panel D The relationship between the private and public information 28-days prior to the earnings announcement and the post-earnings announcement drift



Notes: The vertical axis represents the *t*-statistics and the horizontal axis represents the trading days after the earnings announcement date. The dots are the *t*-value examined by the regression analysis between the $CAR_{[-1,-T]}$ and the public and private information, T is 1, 2, …22. The variable is positive (negative)

significant to the $CAR_{[-1,-T]}$ if the *t*-value is above (below) the significance level.

Figure 3 The relationship between the private and public information and the post-earnings announcement drift (Continued)

Table 1 The optimism and pessimism characteristic terms

Optimism characteristic terms			
Abundant (豐富)	**Active** (積極)	**Advantage** (優勢)	**Amazing** (驚人)
Benefit (受惠)	**Best** (最佳)	**Breakthrough** (突破)	**Conducive** (有助於)
Exceed (超越)	**Excellent** (出色)	**Explode** (激增)	**Favor** (看好)
Fluency (暢旺)	**Growth** (成長)	**Increase** (提升)	**Influx** (湧進)
Lead (領先)	**New Height** (新高)	**Optimistic** (樂觀)	**Overbought** (買超)
Profit (獲利)	**Prospect** (展望)	**Rebound** (反彈)	**Rise** (上漲)
Steady (穩定)	**Stimulate** (激勵)	**Success** (成功)	**Transcend** (優於)
Turn Loss Into Gain (轉虧為盈)	**Upgrade** (調升)	**Upturn** (好轉)	**Upward Price Limit** (漲停)
Pessimism characteristic terms			
Bankruptcy (破產)	**Capital-Reducing** (減資)	**Cheapen** (跌價)	**Crisis** (危機)
Decline (衰退)	**Decrease** (減少)	**Deficit** (赤字)	**Depreciate** (貶值)
Depression (不景氣)	**Deteriorate** (惡化)	**Dispirit** (不振)	**Down** (下滑)
Downgrade (調降)	**Downward Price Limit** (跌停)	**Encumber** (拖累)	**Fail** (告吹)
Fall (下跌)	**Impact** (衝擊)	**Involve** (波及)	**Loss** (虧損)
Mournful 悽慘	**Negative** (負面)	**New Low** (新低)	**Not Good Enough** (不佳)

One Disaster After Another **(雪上加霜)**	**Over-Fall** **(跌破)**	**Oversold** **(賣超)**	**Pessimistic** **(悲觀)**
Plunge **(重挫)**	**Reverse** **(失利)**	**Serious** **(嚴重)**	**Weak** **(疲弱)**

Note: The characteristic terms represent the "terms" which can describe the characteristics of the optimism and pessimism groups. The terms in the parentheses are the Chinese for each of the special terms. A complete list of special terms in each group is available from the authors upon request.

Table 2 Summary Statistics

	Min	Max	Mean	Median	Std
Panel A: Sample Data 7 days prior to the earnings announcement					
SUE	-3.1767	3.0425	-0.0818	-0.0271	0.9304
TURN	0.3467	223.2666	31.3362	23.2500	28.5055
PIN	0.0000	0.5452	0.0997	0.0958	0.1164
SR	-47.0588	48.2759	8.0115	6.9498	9.3703
MEDIA	0.0000	10.0000	0.8059	1.0000	0.6330
SUR	0.0000	53.0000	1.7632	1.0000	2.6331
$CAR_{[-7,-1]}$	-39.6043	32.9032	-1.0872	-1.2004	7.5656
Panel B: Sample Data 14 days prior to the earnings announcement					
SUE	-3.1767	3.0425	-0.0613	-0.0099	0.9280
TURN	0.3467	223.2666	30.9903	22.6627	28.2073
PIN	0.0000	0.5452	0.1003	0.0971	0.1165
SR	-65.7895	50.5263	8.7654	7.8443	8.9872
MEDIA	0.0000	10.2143	0.8041	1.0000	0.5494
SUR	0.0000	95.0000	2.4292	1.0000	3.7998
$CAR_{[-14,-1]}$	-54.8056	59.2386	-1.3263	-1.9581	11.9301
Panel C: Sample Data 21 days prior to the earnings announcement					
SUE	-3.1767	3.0425	-0.0790	-0.0271	0.9169
TURN	0.2136	223.2666	29.9630	21.8971	27.5167
PIN	0.0000	0.5452	0.0988	0.0953	0.1161
SR	-65.7895	48.2759	9.1050	8.3455	8.5792
MEDIA	0.0000	10.0000	0.7865	1.0000	0.5111
SUR	0.0000	102.0000	3.0396	2.0000	4.7962

$CAR_{[-21,-1]}$	-86.7296	94.5900	-1.2932	-1.8286	15.7743
Panel D: Sample Data 28 days prior to the earnings announcement					
SUE	-3.1767	3.0425	-0.0860	-0.0355	0.9153
TURN	0.2136	223.2666	29.3862	21.2524	27.1701
PIN	0.0000	0.5640	0.0989	0.0954	0.1161
SR	-65.7895	48.2759	9.1243	8.4850	8.3394
MEDIA	0.0000	10.0000	0.7764	0.8571	0.4860
SUR	0.0000	102.0000	3.6108	2.0000	5.8078
$CAR_{[-28,-1]}$	-103.8704	125.1082	-1.3943	-1.5096	18.3741

Note: *SUE* is the unexpected earnings surprises. *TURN* is the market turnover. *PIN* is the probability of informed trading. *SR* is the sentiment ratio. *MEDIA* is the media coverage and *SUR* is the measure of public news surprises considering the stock market's reaction to headline news. $CAR_{[-7,-1]}$, $CAR_{[-14,-1]}$, $CAR_{[-21,-1]}$, and $CAR_{[-28,-1]}$ are the cumulative abnormal returns 7-, 14-, 21-, and 28-days before the earnings announcement, respectively.

Table 3 Correlation Analysis

Panel A: Sample Data 7 days prior to the earnings announcement

	TURN	*PIN*	*SR*	*MEDIA*	*SUR*	$CAR_{[-7,-1]}$
SUE	0.0787***	0.0342*	0.1982***	0.0047	-0.0177	-0.0173
TURN		0.1526***	0.1106***	-0.0698***	-0.056***	-0.0514***
PIN			0.0456**	0.0747***	0.055***	-0.0067
SR				0.1229***	-0.0161	0.1227***
MEDIA					0.7333***	0.0517***
SUR						0.0117

Panel B: Sample Data 14 days prior to the earnings announcement

	TURN	*PIN*	*SR*	*MEDIA*	*SUR*	$CAR_{[-14,-1]}$
SUE	0.0786***	0.0295*	0.1914***	0.0037	-0.0074	-0.0625***
TURN		0.164***	0.1067***	-0.0559***	-0.0307*	-0.0537***
PIN			0.0356**	0.0824***	0.0676***	0.0302*
SR				0.1506***	-0.0464***	0.0795***
MEDIA					0.6129***	0.0903***
SUR						0.038**

Panel C: Sample Data 21 days prior to the earnings announcement

	TURN	*PIN*	*SR*	*MEDIA*	*SUR*	$CAR_{[-21,-1]}$
SUE	0.0793***	0.0265*	0.1802***	0.0057	0.0113	-0.0879***
TURN		0.1553***	0.1026***	-0.04***	-0.0053	-0.0571***
PIN			0.0414***	0.0939***	0.0725***	0.026*
SR				0.1782***	-0.0434***	0.0074
MEDIA					0.5675***	0.0913***
SUR						0.029**

Panel D: Sample Data 28 days prior to the earnings announcement

	TURN	*PIN*	*SR*	*MEDIA*	*SUR*	$CAR_{[-28,-1]}$

SUE	**0.0773*** **	**0.0247***	**0.1822*** **	**0**	**0.0131**	**-0.1034*** **
TURN		**0.145*** **	**0.0918*** **	**-0.0348** **	**0.0056**	**-0.0627*** **
PIN			**0.0362*** **	**0.0948*** **	**0.0693*** **	**0.0389*** **
SR				**0.1896*** **	**-0.0352** **	**-0.0117**
MEDIA					**0.5363*** **	**0.0884*** **
SUR						**0.0328** **

Note: *SUE* is the unexpected earnings surprises. *TURN* is the market turnover. *PIN* is the probability of informed trading. *SR* is the sentiment ratio. *MEDIA* is the media coverage and *SUR* is the measure of public news surprises considering the stock market's reaction to headline news. $CAR_{[-7,-1]}$, $CAR_{[-14,-1]}$, $CAR_{[-21,-1]}$, and $CAR_{[-28,-1]}$ are the cumulative abnormal return 7-, 14-, 21-, and 28-days before the earnings announcement, respectively. The *, **, and * symbols denote significance at the 10%, 5% and 1% levels, respectively.**

Table 4 Cross-Sectional Regression Analysis between Cumulative Abnormal Returns and Relevant Information prior to the earnings announcement

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Cross-Sectional Regression Analysis between CAR and Relevant Information 7-days prior to the earnings announcement						
C	-0.677***	-0.6846***	-1.475***	-1.1726***	-0.725***	-1.7935***
	(-3.1963)	(-2.9257)	(-6.2201)	(-4.1074)	(-3.0758)	(-5.7922)
SUE	-0.1083	-0.1086	-0.3216**	-0.1123	-0.1073	-0.317**
	(-0.7096)	(-0.7110)	(-2.0867)	(-0.7368)	(-0.7032)	(-2.0569)
TURN	-0.0134***	-0.0134***	-0.0169***	-0.0125**	-0.0132***	-0.0159***
	(-2.6846)	(-2.6650)	(-3.3994)	(-2.4986)	(-2.6548)	(-3.1599)
PIN		0.0938				-0.2971
		(0.0761)				(-0.2425)
SR7			0.1111***			0.1052***
			(7.2364)			(6.7099)
MEDIA7				0.5792***		0.6373*
				(2.5860)		(1.9094)
SUR7					0.025	-0.0835
					(0.4645)	(-1.0527)
Panel B: Cross-Sectional Regression Analysis between CAR and Relevant Information 14-days prior to the earnings announcement						
C	-0.7291**	-1.0645***	-1.7916***	-2.3328***	-1.0188***	-3.253***
	(-2.5753)	(-3.4130)	(-5.4346)	(-5.7793)	(-3.2830)	(-7.4562)
SUE	-0.7539***	-0.7628***	-0.9918***	-0.7632***	-0.7516***	-0.9756***
	(-3.6788)	(-3.7241)	(-4.7790)	(-3.7378)	(-3.6694)	(-4.7119)

TURN	**-0.0208*** **	**-0.0236*** **	**-0.0247*** **	**-0.0187*** **	**-0.0203*** **	**-0.0247*** **
	(-3.0793)	**(-3.4552)**	**(-3.6617)**	**(-2.7729)**	**(-3.0108)**	**(-3.6152)**
PIN		**4.2118****				**3.4044****
		(2.5550)				**(2.0701)**
***SR*14**			**0.1334*** **			**0.1154*** **
			(6.2091)			**(5.2292)**
***MEDIA*14**				**1.9124*** **		**1.6883*** **
				(5.5537)		**(3.7757)**
***SUR*14**					**0.1134****	**-0.032**
					(2.2734)	**(-0.5018)**

Note: *SUE* is the unexpected earnings surprises. *TURN* is the market turnover. *PIN* is the probability of informed trading. *SR* is the sentiment ratio. *MEDIA* is the media coverage and *SUR* is the measure of public news surprises considering the stock market's reaction to headline news. Values in the table and the parentheses are regression beta and T-statistics, respectively. The *, **, and * symbols denote significance at the 10%, 5% and 1% levels, respectively.**

Table 4 Cross-Sectional Regression Analysis between Cumulative Abnormal Returns and Relevant Information prior to the earnings announcement (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Panel C: Cross-Sectional Regression Analysis between CAR and Relevant Information 21-days prior to the earnings announcement						
C	-0.541	-0.9399**	-0.9874**	-2.7881***	-0.8414**	-3.2294***
	-(1.5909)	-(2.5073)	-(2.4112)	-(5.6354)	-(2.2731)	-(6.0189)
SUE	-1.4442***	-1.4533***	-1.5302***	-1.458***	-1.4502***	-1.4827***
	-(5.7783)	-(5.8174)	-(6.0315)	-(5.8567)	-(5.8040)	-(5.8616)
TURN	-0.0289***	-0.0322***	-0.0304***	-0.0268***	-0.0288***	-0.0295***
	-(3.4720)	-(3.8202)	-(3.6350)	-(3.2301)	-(3.4604)	-(3.4925)
PIN		5.0195**				3.8712*
		(2.5218)				(1.9419)
*SR*21			0.0531*			0.0135
			(1.9552)			(0.4824)
*MEDIA*21				2.776***		3.183***
				(6.2307)		(5.6744)
*SUR*21					0.0976**	-0.1006*
					(2.0504)	-(1.7149)
Panel D: Cross-Sectional Regression Analysis between CAR and Relevant Information 28-days prior to the earnings announcement						
C	-0.4723	-1.1186***	-0.696	-3.0744***	-0.863**	-3.4691***
	(-1.2564)	(-2.6910)	(-1.5128)	(-5.5136)	(-2.1203)	(-5.7319)
SUE	-1.9911***	-2.0049***	-2.0331***	-1.9959***	-1.9999***	-1.9719***
	-(7.1711)	-(7.2284)	-(7.2072)	-(7.2149)	(-7.2059)	(-7.0104)

TURN	**-0.0372*** **	**-0.0421*** **	**-0.0378*** **	**-0.0352*** **	**-0.0373*** **	**-0.0386*** **
	(-3.9773)	**(-4.4610)**	**(-4.0315)**	**(-3.7700)**	**(-3.9907)**	**(-4.0751)**
PIN		**7.9833*** **				**6.7027*** **
		(3.6237)				**(3.0364)**
***SR*28**			**0.0261**			**-0.0193**
			(0.8433)			**(-0.6025)**
***MEDIA*28**				**3.2736*** **		**3.5739*** **
				(6.2989)		**(5.5958)**
***SUR*28**					**0.1089*** **	**-0.0618*** **
					(2.4963)	**(-1.1821)**

Note: *SUE* is the unexpected earnings surprises. *TURN* is the market turnover. *PIN* is the probability of informed trading. *SR* is the sentiment ratio. *MEDIA* is the media coverage and *SUR* is the measure of public news surprises considering the stock market's reaction to headline news. Values in the table and the parentheses are regression beta and *T*-statistics, respectively. The *, **, and * symbols denote significance at the 10%, 5% and 1% levels, respectively.**

Table 5 Portfolio Excess Returns Using Information Sentiment and *MEDIA*

SR*/*MEDIA	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
	Panel A: Excess Returns					**Panel B: Number of Stocks**				
Low	**-0.72**	**-1.81**	**-2.02**	**-5.67**	**-4.62**	**5.50**	**4.57**	**5.07**	**4.07**	**4.64**
2	**-4.59**	**-0.92**	**-3.97**	**-6.63**	**-6.92**	**5.71**	**4.36**	**4.43**	**3.71**	**5.64**
3	**-6.80**	**-6.56**	**-6.38**	**-10.69**	**-5.18**	**4.57**	**5.57**	**3.93**	**5.07**	**4.71**
4	**-6.10**	**-9.47**	**-5.44**	**-5.70**	**-8.34**	**4.14**	**4.79**	**4.64**	**4.79**	**5.50**
High	**-8.22**	**-6.11**	**-6.87**	**-7.84**	**-14.73**	**3.93**	**4.57**	**5.79**	**6.21**	**5.64**
	Panel C: *SR*					**Panel D: *MEDIA***				
Low	**-2.40**	**-3.18**	**-3.60**	**-3.21**	**-3.43**	**0.15**	**0.49**	**0.76**	**0.96**	**1.60**
2	**3.40**	**3.68**	**3.73**	**3.68**	**3.64**	**0.16**	**0.49**	**0.74**	**0.95**	**1.85**
3	**6.90**	**6.97**	**6.78**	**6.88**	**6.93**	**0.18**	**0.52**	**0.78**	**0.96**	**1.50**
4	**10.94**	**10.99**	**10.63**	**10.87**	**11.15**	**0.21**	**0.48**	**0.78**	**0.95**	**1.49**
High	**18.36**	**17.96**	**18.43**	**19.27**	**18.69**	**0.22**	**0.51**	**0.76**	**0.97**	**1.35**

Note: This table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by sentiment ratio (*SR*) and *MEDIA* constructed based on linguistic text mining. The holding periods of the excess returns are four weeks after the earnings announcement. The reported results are averages of the relevant variables over the sample period 2001 to 2009 for the Taiwan stock market. Panels A, B, C and D respectively report for each portfolio the average excess returns, the average number of stocks, and the average *SR* and *MEDIA* of stocks.

Table 错误！未找到引用源。 Portfolio Excess Returns Using Information Sentiment and *SUR*

SR*/*SUR	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
	Panel A: Excess Returns					**Panel B: Number of Stocks**				
Low	**0.61**	**2.32**	**4.68**	**0.54**	**4.26**	**4.89**	**4.28**	**4.44**	**4.78**	**4.11**
2	**0.61**	**4.06**	**-0.28**	**-3.08**	**-1.40**	**5.33**	**3.50**	**4.00**	**4.17**	**5.50**
3	**-1.21**	**-3.83**	**-4.64**	**-5.90**	**-0.65**	**4.78**	**4.61**	**3.33**	**3.56**	**6.22**

4	**-0.75**	**-3.20**	**-3.41**	**-1.68**	**-2.29**	**3.61**	**4.67**	**4.33**	**4.61**	**5.28**
High	**-3.51**	**-2.63**	**-6.15**	**-9.91**	**-8.02**	**3.89**	**5.44**	**6.39**	**5.39**	**3.61**
	Panel C: *SR*					**Panel D: *SUR***				
Low	**-3.28**	**-4.78**	**-5.83**	**-4.40**	**-3.12**	**0.22**	**0.65**	**1.02**	**1.58**	**4.20**
2	**3.40**	**3.52**	**3.34**	**3.27**	**3.63**	**0.19**	**0.70**	**1.00**	**1.77**	**6.14**
3	**7.46**	**7.74**	**6.82**	**7.54**	**7.45**	**0.30**	**0.76**	**1.08**	**1.72**	**4.26**
4	**12.00**	**11.83**	**11.66**	**12.11**	**11.86**	**0.27**	**0.71**	**1.07**	**1.66**	**3.99**
High	**19.35**	**19.35**	**20.89**	**19.74**	**18.30**	**0.31**	**0.73**	**1.06**	**1.70**	**3.13**

Note: This table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by sentiment (*SR*) and *SUR* constructed based on linguistic text mining. The holding periods of the excess returns are four weeks after the earnings announcement. The reported results are averages of the relevant variables over the sample period 2001 to 2009 for the Taiwan stock market. Panels A, B, C and D respectively report for each portfolio the average excess returns, the average number of stocks, and the average *SR* and *SUR* of stocks.

Table 6 Portfolio Excess Returns Using Information Sentiment and PIN

SR*/*PIN	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
	Panel A: Excess Returns					**Panel B: Number of Stocks**				
Low	**-5.79**	**-4.23**	**-1.91**	**0.28**	**-2.94**	**4.79**	**5.07**	**4.93**	**4.29**	**5.50**
2	**-3.14**	**-3.21**	**-4.50**	**-5.90**	**-6.79**	**4.50**	**4.00**	**4.21**	**6.50**	**5.36**
3	**-8.91**	**-5.94**	**-5.18**	**-4.55**	**-10.07**	**5.00**	**4.50**	**5.79**	**4.07**	**5.21**
4	**-10.06**	**-5.68**	**-3.35**	**-6.22**	**-14.06**	**5.29**	**5.07**	**4.71**	**4.57**	**4.93**
High	**-10.40**	**-8.51**	**-7.40**	**-9.38**	**-11.63**	**5.00**	**5.93**	**4.93**	**5.14**	**6.07**
	Panel C: *SR*					**Panel D: PIN**				
Low	**-2.57**	**-3.12**	**-2.11**	**-2.62**	**-2.49**	**0.00**	**0.02**	**0.09**	**0.13**	**0.28**
2	**3.99**	**3.91**	**4.17**	**3.94**	**4.16**	**0.00**	**0.03**	**0.09**	**0.13**	**0.27**
3	**7.05**	**7.30**	**7.31**	**7.21**	**7.01**	**0.00**	**0.04**	**0.09**	**0.13**	**0.27**
4	**11.12**	**11.12**	**11.09**	**11.42**	**10.94**	**0.00**	**0.03**	**0.09**	**0.13**	**0.26**
High	**17.11**	**18.90**	**18.84**	**18.97**	**18.20**	**0.00**	**0.02**	**0.08**	**0.13**	**0.24**

Note: This table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by sentiment (*SR*) constructed based on linguistic text mining and *PIN*. The holding periods of the excess returns are four weeks after the earnings announcement. The reported results are averages of the relevant variables over the sample period 2001 to 2009 for the Taiwan stock market. Panels A, B, C and D respectively report for each portfolio the average excess returns, the average number of stocks, and the average *SR* and *PIN* of stocks.

Table 7 Portfolio Excess Returns Using *MEDIA* and PIN

MEDIA*/*PIN	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
	Panel A: Excess Returns					**Panel B: Number of Stocks**				
Low	**-4.21**	**-7.07**	**-5.48**	**-1.83**	**-10.10**	**4.44**	**3.44**	**3.78**	**5.00**	**6.89**
2	**-13.50**	**-6.87**	**-4.79**	**-0.11**	**-10.96**	**5.78**	**5.11**	**5.22**	**3.33**	**4.11**
3	**-10.20**	**-9.14**	**-3.87**	**-5.84**	**-8.82**	**3.89**	**4.00**	**6.22**	**4.00**	**5.44**

4	**-13.19**	**-8.20**	**-7.39**	**-2.27**	**-10.67**	**6.78**	**5.22**	**4.00**	**3.89**	**3.67**
High	**-8.10**	**-9.13**	**-4.93**	**-6.53**	**-6.31**	**2.67**	**5.78**	**4.33**	**7.33**	**6.00**
	Panel C: *MEDIA*					**Panel D: PIN**				
Low	**0.02**	**0.06**	**0.02**	**0.04**	**0.03**	**0.00**	**0.04**	**0.10**	**0.13**	**0.27**
2	**0.34**	**0.31**	**0.40**	**0.35**	**0.34**	**0.00**	**0.02**	**0.09**	**0.13**	**0.26**
3	**0.76**	**0.72**	**0.78**	**0.77**	**0.82**	**0.00**	**0.04**	**0.09**	**0.14**	**0.25**
4	**0.98**	**0.99**	**1.01**	**1.00**	**1.01**	**0.00**	**0.05**	**0.10**	**0.14**	**0.26**
High	**1.52**	**1.44**	**1.55**	**1.62**	**1.65**	**0.00**	**0.03**	**0.09**	**0.13**	**0.23**

Note: This table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by *MEDIA* constructed based on linguistic text mining and *PIN*. The holding periods of the excess returns are four weeks after the earnings announcement. The reported results are averages of the relevant variables over the sample period 2001 to 2009 for the Taiwan stock market. Panels A, B, C and D respectively report for each portfolio the average excess returns, the average number of stocks, and database of the average *MEDIA* and *PIN* of stocks.

Table 8 Portfolio Excess Returns Using *SUR* and PIN

SUR*/ *PIN	**Low**	**2**	**3**	**4**	**High**	**Low**	**2**	**3**	**4**	**High**
	Panel A: Excess Returns					**Panel B: Number of Stocks**				
Low	**0.59**	**8.89**	**-4.41**	**0.40**	**3.61**	**4.25**	**2.25**	**4.25**	**3.50**	**4.25**
2	**7.23**	**-1.68**	**-1.91**	**-5.42**	**-3.56**	**2.25**	**3.50**	**4.50**	**3.50**	**4.75**
3	**-9.96**	**-1.57**	**2.34**	**1.85**	**-7.45**	**7.00**	**2.75**	**2.50**	**2.25**	**4.00**
4	**-1.80**	**0.67**	**-4.27**	**0.75**	**-7.71**	**3.00**	**5.50**	**3.00**	**3.75**	**3.25**
High	**-0.76**	**0.55**	**-8.02**	**-5.00**	**-3.35**	**2.00**	**4.50**	**4.25**	**5.50**	**5.25**
	Panel C: *SUR*					**Panel D: PIN**				
Low	**0.25**	**0.42**	**0.17**	**0.06**	**0.20**	**0.00**	**0.02**	**0.08**	**0.10**	**0.22**
2	**1.25**	**0.94**	**1.04**	**1.25**	**1.28**	**0.00**	**0.04**	**0.08**	**0.12**	**0.26**
3	**1.50**	**1.50**	**1.67**	**1.63**	**1.50**	**0.00**	**0.03**	**0.08**	**0.10**	**0.19**
4	**2.25**	**2.17**	**2.33**	**2.47**	**2.38**	**0.00**	**0.01**	**0.08**	**0.12**	**0.26**
High	**6.17**	**6.00**	**6.90**	**6.33**	**5.99**	**0.00**	**0.02**	**0.08**	**0.12**	**0.20**

Note: This table contains results for portfolios of stocks sorted independently on each earnings announcement quarter by *SUR* constructed based on the linguistic text mining and PIN. The holding periods of the excess returns are four weeks after the earnings announcement. The reported results are averages of the relevant variables over the sample period 2001 to 2009 for the Taiwan stock market. Panels A, B, C and D respectively report for each portfolio the average excess returns, the average number of stocks, the average *SUR* and PIN of stocks.