Initial public offering and financial news
Jia-Lang Seng, Pi-Hua Yang and Hsiao-Fang Yang
Department of Accounting, College of Commerce, National Chengchi University, Taipei City, Taiwan, ROC

ABSTRACT
This paper explores the relation between the media sentiment and initial public offerings’ (IPOs) underpricing which means whether the media were responsible for the phenomenal rise and fall in the market value of IPOs’ shares from 2010 to 2014. This paper does not focus on what drives the media coverage of IPO firms but focuses on the effect of pre-IPO media coverage on the first day’s return. In this paper, we use sentiment analysis to calculate the media sentiment and use the sentimental words to classify the news articles. Then, we find that the media sentiment is significantly related to initial return (IR). We also find significant results that IR is related to market return compared to media sentiment in the non-electronics industry. Furthermore, as this paper expected, the number of positive (negative) news articles has a positive (negative) influence on the IR in the electronics industry.

1. Introduction
Corporations may raise capital in the primary market by way of initial public offerings (IPOs). The IPOs¹ are a type of public offering in which shares of stock in a company usually are sold to institutional investors that in turn sell to the general public, on a securities exchange, for the first time. The main IPO methods are book-building, auction method, and public offer. The book-building method has the advantage of reducing the degree of underpricing. Most companies in Taiwan securities market are using the combination of book-building and public offer. The book-building models began with Benveniste and Spindt (1989). Benveniste and Wilhelm (1990) argued that the underwriter’s control of both price and allocations may be used to induce investors to reveal their private information. The book-building bids and allocation data are not publicly accessible. It is difficult for us to observe how the final shares are allocated. Hence, it is difficult to research on the implications of the book-building model. Accordingly, this paper tries to use information flows as the key factor to observe this pricing process.

Consistent with recent studies, IPO firms tend to show themselves as those with good prospects. These issuers also intend to make use of every opportunity to maximize their own personal wealth. Arbel and Strebel (1983) suggested that individual investors are

ARTICLE HISTORY
Received 23 March 2017
Accepted 11 June 2017

KEYWORDS
IPOs’ underpricing; media sentiment; sentiment analysis

1. Introduction
Corporations may raise capital in the primary market by way of initial public offerings (IPOs). The IPOs¹ are a type of public offering in which shares of stock in a company usually are sold to institutional investors that in turn sell to the general public, on a securities exchange, for the first time. The main IPO methods are book-building, auction method, and public offer. The book-building method has the advantage of reducing the degree of underpricing. Most companies in Taiwan securities market are using the combination of book-building and public offer. The book-building models began with Benveniste and Spindt (1989). Benveniste and Wilhelm (1990) argued that the underwriter’s control of both price and allocations may be used to induce investors to reveal their private information. The book-building bids and allocation data are not publicly accessible. It is difficult for us to observe how the final shares are allocated. Hence, it is difficult to research on the implications of the book-building model. Accordingly, this paper tries to use information flows as the key factor to observe this pricing process.

Consistent with recent studies, IPO firms tend to show themselves as those with good prospects. These issuers also intend to make use of every opportunity to maximize their own personal wealth. Arbel and Strebel (1983) suggested that individual investors are

CONTACT Jia-Lang Seng seng@nccu.edu.tw Department of Accounting, College of Commerce, National Chengchi University, Taipei City 11605, Taiwan, ROC

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
usually uninformed, poorly trained, and inclined to rely on noise in the market. Especially, unlike the regulatory environment of an equity offering in the United States, there is no restriction on news articles’ announcements of issuers around IPOs in Taiwan. It also allows firms to release news articles that can generate favourable views at their discretion. For instance, a firm can announce its potential operation and business strategies via news articles. These news articles confirm that the firms attempt to present their positive image and attract investors. Therefore, examining the financial news articles of firms with IPOs in Taiwan allows us to investigate whether the investors are influenced by media sentiment. In addition, emerging markets are usually operated in an environment of information opacity as pointed out by Morck, Yeung, and Yu (2000); stock prices in these markets are quite sensitive to political events and rumours. This paper has been motivated by studies (Baker, Stein, & Wurgler, 2002; Neal & Wheatley, 1998) that investigated the role of investor sentiment. This paper focuses on whether the firms with IPOs can engage in other qualitative types of reporting to influence investors’ perception and their demand for the new issue shares. Then, this paper collects the news articles from the Internet news articles database and categorizes them. Azar argued that through the sentiment analysis process, information is converted from a textual form into a numeric form. Thus, this paper uses sentiment analysis to convert the qualitative news articles into quantitative measures. This paper is related to Liu, Sherman, and Zhang’s (2009); Seng, Yang, and Yang’s (2016) research studies that categorize the news articles as positive, negative, and neutral. In this paper, we investigate whether the media sentiment is related to IPOs’ underpricing. The empirical results show that the media sentiment is associated with IPOs’ underpricing. Furthermore, the number of positive (negative) news articles has a positive (negative) influence on the initial return (IR) in the electronics industry.

The remainder of this paper is organized as follows. Section 2 introduces relevant literature and theory. Section 3 introduces the research method, the data set, and all the variables used in the model. In Section 4, it presents the empirical results and the relation between the media coverage and underpricing. Finally, Section 5 concludes the research findings.

2. Literature review

2.1. IPOs’ underpricing and asymmetric information

Most models of IPO underpricing are based on asymmetric information. Baron (1982) used a model of underpricing where issuers delegated the pricing decision to underwriters. Baron (1982) presented that the investment banker is better informed about the capital market than is the issuer, and the issuer cannot observe the distribution effort expended by the banker. Rock (1986) presented the privileged investors whose information is superior to the issuer as well as other investors. In the model of Rock (1986) was that the issuer himself is also uninformed. This means that the offer price is not informative to any investor in the IPOs, so that uninformed investors face severe adverse selection in the model of Rock (1986). Rock (1986) theorized ‘winner’s curse’; ‘informed investors’ like investment institutions have superior information about the true value of the stock after the listing; on the contrary, ‘uninformed investors’ have more probability of buying the inferior offering. Beatty and Ritter (1986) extended Rock’s (1986) theory and found
evidence that risk does affect after-market returns. Sherman and Titman (2002) showed that the effect of media on IRs should be asymmetric and should be magnified by uncertainty.

2.2. IPOs’ underpricing and signalling theory

Past studies (Allen & Faulhaber, 1989; Grinblatt & Hwang, 1989; Welch, 1989) formalized the notion that good firms underprice to ‘leave a good taste in investors’ mouths’. Welch’s (1989) model implied that the high-quality firms underprice at the IPOs in order to obtain a higher price at a second offer. Grinblatt and Hwang’s (1989) two-parameter signalling model indicated that the firm’s value and the degree of underpricing are positively related. Allen and Faulhaber (1989) confirmed that underpricing can signal favourable prospects for the firm and that it is temporary, industry-specific, and associated with improvements in the profitability of entry. Trauten and Langer (2012) stated that in fixed-price offerings, the issuer can maintain investors’ propensity to produce information by appropriately adjusting the offering price even if information costs are high. Thus, firms with IPOs attempt to signal good information for inducing investors.

2.3. IPOs’ underpricing and optimistic investors

Degeorge, Derrien, and Womack (2004) found in the French IPO market that book-building issues attract more press than auctions, but only after the book-building route is selected. Past studies (Derrien, 2005; Ljungqvist, Nanda, & Singh, 2006) suggested that the issuer and regular customers of an investment banker benefit from the presence of sentiment investors; neither model considers that an investment banker might promote an issue in such a way as to induce sentiment investors into the market for an IPO. Cook, Kieschnick, and Van Ness (2006) implied that initial IPOs’ returns are positively correlated with pre-issue publicity. Barber and Odean (2008) found that retail investors are more likely to purchase attention-grabbing stocks. Similarly, Tetlock (2007) found that either the media report investor sentiment before the sentiment is fully incorporated into market prices or the media directly influence investors’ attitudes towards securities. Frieder and Subrahmanyam (2005) determined that individual investors are more likely to hold stock in highly visible companies. The above discussions suggest that an investment banker’s efforts to promote an IPO through increased media increase the retail interest in that stock.

2.4. Sentiment analysis

To easily define sentiment analysis, this paper quotes from the studies (Nasukawa & Yi, 2003; Yi, Nasukawa, Bunescu, & Niblack, 2003) entitled their paper title. Their studies may explain the sentiment analysis among communities self-identified as focused on natural language processing. A large number of papers mentioning sentiment analysis focus on the specific application of classifying reviews as to their polarity, a fact that appears to have caused some authors to suggest that the phrase refers specifically to this narrowly defined task. However, nowadays, many construe the term more broadly to mean the computational treatment of opinion, sentiment, and subjectivity in the text. Thus, when broad interpretations are applied, sentiment analysis and opinion
mining denote the same field of study. Henry (2008) quantified the tone by using computer-based analytical tools to measure the frequency of positive and negative words found in earnings press releases and concluded that the tone of earnings press releases influences investors’ reaction to earnings. Davis, Piger, and Sedor (2006) found a significant positive association between levels of optimistic tone in earnings press releases and both future return-on-assets and the initial market response. Tetlock (2007) concluded that high values of media pessimism induce downward pressure on market prices; unusually high or low values of pessimism lead to temporarily high market trading volume. This is consistent with sentiment theories under the assumption that media content is linked to the behaviour of individual investors, who own a disproportionate fraction of small stocks.

3. Research method

3.1. Hypothesis, data collection, and description

Cook et al. (2006) argued that initial IPO returns are positively correlated with pre-issue publicity. Liu et al. (2009) found that the role of the media in IPOs is positively related to underpricing. Therefore, examining the financial news articles of IPO firms in Taiwan allows us to investigate whether the investors are influenced by the positive sentiment of the media. In addition, the number of media is a signal that high media attention is more positive than low media attention. This study also examines the number of news articles. Finally, this paper proposes two hypotheses. The first hypothesis H1: the media sentiment is related to IPO underpricing (IR). The second hypothesis H2: the number of positive (negative) news articles positively (negatively) relates to the IR of IPO.

In order to test H1 and H2, this paper collects the IPO firms which are listed on the Taiwan Stock Exchange (TWSE) or the over-the-counter GreTai Securities Market (GTSM) for the first time from 2010 to 2014. We excluded IPOs of financial firms and Taiwan depository receipts. This paper identifies and extracts the sample companies from the list of the Taiwan Economic Journal database (TEJ), and the numbers of final sample firms with IPOs listed on TWSE and GTSM are 75 and 166, respectively. This study removed one firm that lacked news articles and two firms whose name changed. All the public offering data, financial, and return data were obtained from the TEJ database.

The news articles resource was collected from the Knowledge Management Winner (KMW) database that is marketed and maintained by the largest financial news articles group in Taiwan. The KMW is the largest news articles database composed of China Times, China Times Express and Commercial Times. This study searched the news articles by using the companies’ abbreviated name, and 5876 news articles were collected. We did not limit our news articles only to those news articles in which companies are mentioned in the headline or in the lead paragraph, but kept news articles containing multiple companies. Because of the characteristic of the financial news articles and based on our analysis of the media coverage, this paper did not exclude a large volume of news articles that actually cover the firm. In general, the time window is from the day after listing on emerging stock market to the day prior to the issue date of the IPOs (the pre-IPO period), and this paper acknowledges that the pre-IPO companies tend to have more attraction and positive image to the investors after listing on the emerging stock market. However, there are two other reasons that this study uses another time window.
window, one month prior to the issue day. First, since the length of the period varies from the firm, this paper standardizes the media coverage measure into per-month measure and form another hypothesis to research the relation between the media coverage and underpricing under the time window. Second, this paper follows Liu et al.'s (2009) study by using the one-month window before the issue day.

Table 1 shows that the number of firms is almost evenly distributed in these five years, but 2011 (66 firms) has a large increase in IPO firms. The number of firms is concentrated from 2011 to 2013; the proportion of firms in these three years is 68.05%. The breakdown of the 241 IPOs exhibits an unevenly distributed pattern across 24 industries classified by the TSE industrial classification code. In particular, 41 IPOs come from biotechnology and medical care industry, 31 from optoelectronic industry, 29 from the electronic parts and components industry, 24 from semiconductor, 16 from the communications and internet industry, 11 from electronics and peripheral equipment industry, 10 from the information service industry, 10 from the tourism industry and the rest scatter over the remaining 14 industries with a maximum of 9 IPOs and a minimum of 1. As expected, electronics industry-related firms dominate the Taiwanese IPO market with 60.17% of the total sample (based on the two-digit Standard Industrial Classification Code). The distribution of these firms reflects the importance of the electronics products sector, which has become the main driver of the Taiwanese economy for a long time. Obviously, biotechnology and medical care (41 firms, 17.01%) is blooming up and promoted in Taiwan, becoming the star industry in these years.

Since the length of the period varies from the firm, this paper standardizes the media coverage measure into per-month measure to compare to the quantities of news articles

| Industry                          | 2010 | 2011 | 2012 | 2013 | 2014 | Total | Percent |
|-----------------------------------|------|------|------|------|------|-------|---------|
| Biotechnology & medical care      | 5    | 9    | 11   | 11   | 5    | 41    | 17.01%  |
| Optoelectronic                    | 9    | 9    | 7    | 3    | 3    | 31    | 12.86%  |
| Electronic parts & components     | 6    | 10   | 3    | 5    | 5    | 29    | 12.03%  |
| Semiconductor                     | 4    | 7    | 4    | 3    | 6    | 24    | 9.96%   |
| Other                             | 2    | 6    | 4    | 2    | 2    | 16    | 6.64%   |
| Communications and Internet       | 2    | 6    | 1    | 5    | 2    | 16    | 6.64%   |
| Electric machinery                | .    | 2    | 4    | 2    | 4    | 14    | 5.81%   |
| Electronics and peripheral equipment | 1 | 2 | 2 | 4 | 2 | 11 | 4.56% |
| Information service               | .    | .    | 2    | 1    | 3    | 10    | 4.15%   |
| Tourism                           | .    | 2    | 5    | 3    | .    | 10    | 4.15%   |
| Other electronic                  | 3    | 3    | 2    | .    | 1    | 9     | 3.73%   |
| Trading and consumers’ goods      | .    | 1    | 2    | .    | 2    | 5     | 2.07%   |
| Shipping and transportation       | 1    | 1    | .    | 1    | 1    | 4     | 1.66%   |
| Chemical                          | .    | 2    | .    | 1    | .    | 3     | 1.24%   |
| Cultural and creative industry    | .    | .    | 2    | 1    | .    | 3     | 1.24%   |
| Textiles                          | 1    | .    | 1    | .    | 1    | 3     | 1.24%   |
| Building material and construction | 1 | . | 1 | . | 1 | 2 | 0.83% |
| Plastics                          | .    | .    | 1    | .    | 1    | 2     | 0.83%   |
| Electrical and cable              | 1    | .    | 1    | .    | 1    | 2     | 0.83%   |
| Iron and steel                    | 1    | .    | .    | 1    | .    | 2     | 0.83%   |
| Automobile                        | .    | .    | 1    | .    | .    | 1     | 0.41%   |
| Glass and ceramics                | .    | .    | .    | 1    | .    | 1     | 0.41%   |
| Foods                             | .    | .    | 1    | .    | .    | 1     | 0.41%   |
| Electronic products distribution  | .    | 1    | .    | .    | .    | 1     | 0.41%   |
| Total                             | 37   | 66   | 51   | 47   | 40   | 241   | 100%    |
across from the IPO firms. In addition, this paper also counts the number of news articles prior to the IPO issue date. Table 2 shows the characteristics of 5876 news articles collected in the pre-IPO period. This study divides the news articles into three types: positive, neutral, and negative. News articles are considered to have a positive (negative) impact on IPO firms. The news articles that are considered to have a trivial impact is classified as neutral news articles. Table 2 presents that a majority of news articles have a positive impact on the IPO firms, in that positive news articles show up 91.08%. Moreover, the proportion of positive news articles accounts for 93.28% one month before the IPO issue date. It is not surprising that all of the three industries have a higher percentage of positive news articles. As expected, electronics industry-related firms cover with 56.11% of the total news articles.

### 3.2. Research design

This paper classifies each news article into one of three categories: positive, negative, or neutral. The classification is based on a dictionary-based approach to convert the qualitative news articles into a quantitative measure. Using a dictionary to compile sentiment words is an obvious approach because most dictionaries’ studies such as Miller, Beckwith, Fellbum, Gross, and Miller (1990) listed synonyms and antonyms for each word. Thus, a simple technique in this approach is to use a few seed sentiment words to bootstrap based on the synonym and antonym structure of a dictionary. This paper is also motivated by these studies. However, this paper did not extend our sentiment dictionary by using the dictionaries’ list synonyms and antonyms for each word. In this paper, three graduate students read the financial news articles from KMW news articles database and extracted the sentiment words. There are three steps about how to extract sentiment words. First of all, they pick sentiment words separately. Second, one of them merged and removed the duplicated sentiment words. Third, they discussed and filtered effective sentiment words. The way of building sentimental lexicons was based on Lin (2013). Then, this paper classified news articles by four rules. Rule one and rule two: when the number of positive (negative) words is more than the number of negative (positive) words, the

| Industry                         | News articles classification | The pre-IPO period | One month prior to the issue day |
|----------------------------------|------------------------------|--------------------|---------------------------------|
|                                  |                              | Frequency | Percent | Frequency | Percent |
| All sample                       | All                          | 5876      | 100     | 1606      | 100     |
|                                  | Positive                     | 5352      | 91.08   | 1498      | 93.28   |
|                                  | Neutral                      | 319       | 5.43    | 48        | 2.99    |
|                                  | Negative                     | 205       | 3.49    | 60        | 3.74    |
| Electronics industry             | All                          | 3297      | 100     | 844       | 100     |
|                                  | Positive                     | 2974      | 90.20   | 778       | 92.18   |
|                                  | Neutral                      | 189       | 5.73    | 28        | 3.32    |
|                                  | Negative                     | 134       | 4.06    | 38        | 4.50    |
| Biotechnology & medical care     | All                          | 1347      | 100     | 335       | 100     |
|                                  | Positive                     | 1265      | 93.91   | 320       | 95.52   |
|                                  | Neutral                      | 49        | 3.64    | 4         | 1.19    |
|                                  | Negative                     | 33        | 2.45    | 11        | 3.28    |
| Other industry                   | All                          | 1232      | 100     | 427       | 100     |
|                                  | Positive                     | 1113      | 90.34   | 400       | 93.68   |
|                                  | Neutral                      | 81        | 6.57    | 16        | 3.75    |
|                                  | Negative                     | 38        | 3.08    | 11        | 2.58    |

Table 2. News articles statistics.
news article is classified into positive (negative). Rule three: when the number of negative words is more than the number of positive words and the number of negative words is more than six, the news article is classified into negative. Rule four: when the number of positive words is equal to the number of negative words and the number of negative words is less than six, the news article is classified into neutral. In order to test our hypothesis, this paper uses a number of variables (shown in Table 3). The dependent variable is IR (shown in Equation (1)) and the independent variables TONE_RATIO (shown in Equation (2)) and TONE (shown in Equation (3)) are calculated according to studies (Bautin, Vijayarenu, & Skiena, 2008, March–April; Godbole, Srinivasaiah, & Skiena, 2007, March).

\[
IR_i = \frac{(CP_i - OP_i)}{OP_i}, \quad (1)
\]

where \(CP_i\) is the closing price of the issue day of firm \(i\), the \(OP_i\) is the offering price of firm \(i\).

\[
\text{TONE}_RATIO_{ij} = \left(\frac{PW_{ij}}{POS_{DIC}}\right) - \left(\frac{NW_{ij}}{NEG_{DIC}}\right), \quad (2)
\]

where \(PW_{ij}\) is the count of positive words that appear in the news article \(j\) of firm \(i\), the \(NW_{ij}\) is the count of negative words that appear in the news article \(j\) of firm \(i\), the \(POS_{DIC}\) is the total number of positive words that appear in dictionary, and the \(NEG_{DIC}\) is the total number of negative words that appear in the dictionary.

\[
\text{TONE}_{ij} = \left(\frac{PW_{ij} - NW_{ij}}{PW_{ij} + NW_{ij}}\right), \quad (3)
\]

where \(PW_{ij}\) is the count of positive words that appear in the news article \(j\) of firm \(i\), and the \(NW_{ij}\) is the count of negative words that appear in the news article \(j\) of firm \(i\).

We used different models shown in Table 4 to test our hypothesis between two different periods. We also tested our hypothesis under different industries. At the industry level, the importance of the electronic products sector, which has become the main driver of the Taiwanese economy for a long time, is also a key characteristic of IPO firms. In Taiwan, the electronics industry has the overwhelming advantage of the cluster effect of media. If one of the electronics supply chains blooms up, the rest of the supply chain would benefit from

| Variable     | Description                                                                 |
|--------------|-----------------------------------------------------------------------------|
| IR           | The initial return from IPOs offers price to the closing price of the first day |
| TONE_RATIO   | The sentiment score that the ratio of the frequency count of positive words divided by the positive sentimental dictionary minus the ratio of the frequency count of negative words divided by the negative sentimental dictionary |
| TONE         | The sentiment score as (positive word – negative word) divided by (positive word + negative word) |
| TLHIT        | The total words count excluding symbol words in each news article using a computer program |
| NUM_POS      | The number of positive news articles about the firm I published during the period from the one month prior to the issue day |
| NUM_NEG      | The number of negative news articles about the firm I published during the period from the one month prior to the issue day |
| LN_FIRMSIZE  | The natural logarithm of the market value of the IPO firm at the offering and is noted in millions of dollars in Taiwan |
| LN_OFFERSIZE | The natural logarithm of the market value of the IPOs firm at the offering and is noted in millions of New Taiwan dollars |
| IPO_PERIOD   | The natural logarithm of difference in the number of days between the date of listing on emerging stock market and the issue date |
| SCALED_OFFER| Dummy variable where an SCALED_OFFER is given a value of 1 if revision is more than 1 or a 0 if a revision is less than or equal to 0 |
| MKTRET       | Market return for 30 trading days prior to issue day |
the trend. Hence, this paper also researches whether the positive sentiment of media coverage about the companies in the electronics industry positively relates to underpricing compared with those in the non-electronics industry.

Model 1 and Model 2 were used to test hypothesis H1 which investigated the association between underpricing and the sentiment of media before IPO issue date, and we used control variable LN_FIRMSIZE in Model 1 and Model 3, but substituted the LN_FIRMSIZE for LN_OFFERSIZE in Model 2 and Model 4. There is a reason to explain why this paper used a different model to test hypothesis H1. Because the LN_FIRMSIZE and LN_OFFERSIZE have a highly positive association, and it is more reasonable to use LN_OFFERSIZE control variable when testing the samples of one month prior to the issue day based on the timing of the pre-IPO period. Then, we used two approaches to confirm whether the IPO underpricing reflects the public media sentiment: one approach is TONE_RATIO, and the other is TONE. The Model 5 is meant to test the hypothesis H2 and to find the association between underpricing and the number of positive (negative) news articles before the IPO issue date, but the number of positive or negative news articles means the media coverage and media visibility of the target firm.

4. Research results and findings

Table 5 provides descriptive statistics regarding all variables for 241 samples during 2010–2014. On average, our sample has the IR of 29.29% and the first quartile shows 7.93% representing that at least 75% of IPO IRs are larger than 7.93%. It shows that the underpricing phenomenon is quite prevailing in Taiwan.

The table provides descriptive statistics regarding all variables. IR is the IR from IPOs’ offer price to the closing price of the first day. TONE_RATIO is the sentimental score of the ratio of the frequency count of positive words divided by the positive sentimental dictionary minus the ratio of the frequency count of negative words divided by the negative sentimental dictionary. TONE is calculated as (Positive word – Negative word) divided by

| Variables | N  | Mean | Std dev | Min   | Max  | Q1   | Q2   | Q3   |
|-----------|----|------|---------|-------|------|------|------|------|
| IR        | 241| 29.29| 31.12   | -18.49| 204.76| 7.933| 23.478| 43.00|
| TONE_RATIO| 241| 0.024| 0.009   | -0.003| 0.049 | 0.018| 0.025 | 0.030|
| TONE      | 241| 0.720| 0.125   | 0.212 | 0.958 | 0.652| 0.739 | 0.806|
| TLHIT     | 241| 659.589| 87.391  | 314.00| 1023.30| 611.333| 658.100| 713.857|
| IPO_PERIOD| 241| 6.461| 0.804   | 3.332 | 8.283 | 5.938| 6.349 | 6.914|
| LN_FIRMSIZE| 241| 8.049| 0.937   | 6.236 | 10.980| 7.328| 7.993 | 8.609|
| SCALED_OFFER| 241| 0.340| 0.475   | 0     | 1     | 0    | 0    | 1    |
| MKTRET    | 241| 0.320| 4.457   | -13.581| 10.066| -2.747| 1.353 | 3.367|
(Positive word + Negative word). TLHIT is the total words count excluding symbol words in each news article using a computer program. IPO_PERIOD is the natural logarithm of difference in the number of days between the date of listing on Emerging Stock Market and the issue date. LN_FIRMSIZE is the natural logarithm of the market value of the IPO firm at the offering and is noted in millions of New Taiwan dollars. SCALED_OFFER equals one if the offer price is set above the average of initial filing price range, and zero otherwise. MKTRET is the market return for 30 trading days prior to issue day, in per cent.

There are two firms without any news article in one month prior to the issue day; the descriptive statistics are shown in Table 6.

This paper uses two-time windows (TW1: pre-IPO period, TW2: one month prior to the issue day) and four models to test hypothesis H1. Our main results are given in Tables 7–10. Table 7 shows that IR has a significantly positive association with media sentiment. The TONE_RATIO and TONE have a positive coefficient, indicating that the positive media sentiment is more likely to influence on IPOs’ underpricing. The SCALED_OFFER is a significantly positive association with IR in four models. The SCALED_OFFER is the proxy for investors’ demand; the empirical results show that when the investors’ demand is higher, it will increase the likelihood of the IR going up. The MKTRET is a significantly positive association with IR in four models; it means that the market return can predict future

| Table 6. Descriptive statistics (one month prior to the issue day). |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| **Variable**      | **N**             | **Mean**          | **Std Dev**       | **Min**           | **Max**           | **Q1**         | **Q2**         | **Q3**         |
| IR                | 239               | 29.531            | 31.136            | −18.488           | 204.762           | 8.000          | 23.800         | 43.333         |
| TONE_RATIO        | 239               | 0.019             | 0.011             | −0.011            | 0.056             | 0.012          | 0.019          | 0.027          |
| TONE              | 239               | 0.749             | 0.164             | −0.023            | 1.000             | 0.657          | 0.759          | 0.870          |
| TLHIT             | 239               | 660.381           | 85.684            | 314.000           | 1023.300          | 611.333        | 658.100        | 713.857        |
| NUM_POS           | 239               | 6.268             | 3.834             | 1                 | 30                | 4             | 5             | 8              |
| NUM_NEG           | 239               | 0.251             | 0.554             | 0                 | 4                 | 0             | 0             | 0              |
| IPO_PERIOD        | 239               | 6.458             | 0.803             | 3.332             | 8.283             | 5.938          | 6.349          | 6.914          |
| LN_OFFERSIZE      | 239               | 12.379            | 0.883             | 10.115            | 14.989            | 11.735         | 12.326         | 12.934         |
| SCALED_OFFER      | 239               | 0.343             | 0.476             | 0                 | 4                 | 0             | 0             | 0              |
| MKTRET            | 239               | 0.322             | 4.463             | −13.581           | 10.066            | −2.747         | 1.353          | 3.367          |

| Table 7. Empirical result for testing H1 (Model 1 to Model 4). |
|-------------------|-------------------|-------------------|-------------------|-------------------|
| **Variable**      | **The pre-IPO period** | **Model 1 Coef.** | **Model 2 Coef.** | **The issue day** | **Model 3 Coef.** | **Model 4 Coef.** |
| TONE_RATIO        | 1329.4*** (5.81)   | 39.10*** (2.49)   | 647.6*** (3.91)   | 32.16*** (2.76)   |
| TONE              | −0.0649*** (−2.77) | −0.0135 (−0.60)   | −0.0312 (−1.36)   | −0.0204 (−0.89)   |
| TLHIT             | 8.567*** (3.80)    | 9.101*** (3.75)   | 5.866*** (2.52)   | 6.201*** (2.62)   |
| IPO_PERIOD        | 5.144*** (2.61)    | 2.488 (1.23)      | 3.764* (−1.77)    | −4.016* (−1.86)   |
| LN_FIRMSIZE       | 10.33*** (2.84)    | 11.53*** (3.01)   | 10.79*** (2.89)   | 11.31*** (2.98)   |
| SCALED_OFFER      | 1.162*** (2.96)    | 1.068*** (2.57)   | 1.199*** (2.96)   | 1.164*** (2.82)   |
| MKTRET            | −60.99*** (−2.17)  | −73.40*** (−2.32) | 41.96 (1.13)      | 24.02 (0.62)      |
| **N**             | 241               | 241               | 239               | 239               |
| **Adj. R²**       | 0.2098            | 0.1190            | 0.1660            | 0.1394            |

*p < .10.
**p < .05.
***p < .01.
underpricing in the hot market. The IPO_PERIOD is a significantly positive association with IR in four models; it means that the longer the pre-IPO period, the higher the visibility in the media is. Positive IR may occur as a result of investor optimism driving up.

Table 8 shows the empirical results for testing H1 under different industry levels. There is a significantly positive association between IR and media sentiment in the electronics industry. However, Table 6 indicates that the positive sentiment in the electronics industry performs significantly related to IR than those in the non-electronics industry. The adjusted $R^2$ of the Model 1 is higher than the Model 2 in two industries (in TW1: 0.24 > 0.18, in TW2: 0.16 > 0.14).

About hypothesis H2, this paper uses Model 5 to test it. The empirical results are given in Table 9. Table 9 indicates that IR has a significantly positive association with the number of positive news articles (NUM_POS).

Table 10 shows the empirical results for testing hypothesis H5 under different industry levels. As shown in Table 10, the association of underpricing and the media sentiment is a significant difference between the electronics and non-electronics industry. Furthermore, Table 10 indicates that IR has a significantly positive (negative) association with the number of positive (negative) news articles in the electronics industry. In terms of

Table 8. Empirical result for testing H1 at the industry level.

| Variable         | Electronics industry | Non-electronics industry |
|------------------|----------------------|--------------------------|
|                  | The pre-IPO period   | One month prior to the   | The pre-IPO period   | One month prior to the   |
|                  | Model (1) Coef.      | issue day Coef.          | Model (2) Coef.      | issue day Coef.          |
| TONE_RATIO       | 1540.0*** (5.15)     | 819.2*** (3.79)          | 1069.747*** (2.87)   | 460.779* (1.65)          |
| TLHIT            | −0.0390 (−1.17)      | −0.0452 (−1.25)          | −0.065* (−1.78)      | −0.007 (−0.23)           |
| IPO_PERIOD       | 6.566*** (2.07)      | 4.040 (1.21)             | 9.675*** (2.89)      | 6.768** (1.98)           |
| LN_FIRM_SIZE     | 6.317** (2.20)       |                          | 5.055* (1.78)        |                          |
| LN_OFFERSIZE     | −3.841 (−1.19)       |                          | −3.111 (−1.06)       |                          |
| SCALED_OFFER     | 14.57*** (2.89)      | 16.16*** (3.10)          | 4.807 (0.91)         | 4.014 (0.75)             |
| MKTRET           | 0.946* (1.72)        | 0.852 (1.49)             | 1.466*** (2.59)      | 1.603*** (2.76)          |
| INTERCEPT        | −78.03* (−1.92)      | 61.32 (1.10)             | −61.522 (−1.54)      | 15.280 (0.29)            |
| $N$              | 134                  | 133                      | 107                  | 106                      |
| Adj. $R^2$       | 0.2379               | 0.1795                   | 0.1642               | 0.1428                   |

*p < .10.
**p < .05.
***p < .01.

Table 9. Regression result for testing H2 (Model 5).

| Variable     | One month prior to the issue day Coef. |
|--------------|----------------------------------------|
| NUM_POS      | 1.122** (2.00)                         |
| NUM_NEG      | −6.552 (−1.76)                         |
| IPO_PERIOD   | 7.288*** (3.03)                        |
| LN_OFFER_SIZE| −5.376* (−2.32)                        |
| SCALED_OFFER| 10.27*** (2.66)                        |
| MKTRET       | 1.187*** (2.87)                        |
| INTERCEPT    | 39.44 (1.16)                           |
| $N$          | 239                                    |
| Adj. $R^2$   | 0.1339                                 |

*p < .10.
**p < .05.
***p < .01.
economic statistics, the coefficient of NUM_POS (NUM_NEG) means that one piece of positive (negative) news articles increases the likelihood of IR driving up (going down) by about 2.045% (−12.43%); it indicates that the IR may react largely to negative news articles than positive news articles. Interestingly, it presents that here have noticeable results in a non-electronics industry that the IR is highly related to market return (MKTRET) in the non-electronics industry.

5. Conclusion and discussion

This paper focuses on the relation between the sentiment of the media and IPO underpricing based on the IPO firms’ listing on TWSE or GTSM. This paper finds a positive coefficient on the positive sentiment score of media coverage and statistically significant association, indicating that the positive sentiment of media coverage is more likely to influence on IPO underpricing. The interesting results that the association of underpricing and the sentiment of media significantly difference between the companies in electronics and those in the non-electronics industry. The results indicate that IR has a significantly positive (negative) association with the number of positive (negative) news articles in the electronics industry. On the contrary, the IR is highly related to market return in the non-electronics industry. The media is a market signal for investors, issuers, and underwriters. In Taiwan, a pre-IPO market has an organized trading platform – the emerging stock market in Taiwan. The pre-market plays an important part of price discovery, and the pre-market price should be useful; thus, we may analyse the media coverage into different categories, such as earnings and financials, strategy and policy. It may help us to further discuss the content of the media and to reach a detailed and consistent point.

In conclusion, this paper contributes to the literature in IPO underpricing and the sentiment analysis on media coverage. The media is a market signal for investors, issuers, and underwriters. Thus, the participants can understand the role of the media coverage in IPO markets from our research.

Notes
1. https://en.wikipedia.org/wiki/Initial_public_offering
2. The over-the-counter GreTai Securities Market (GTSM) officially changed the English title as Taipei Exchange (TPEx) on 24 February 2015.

Table 10. Regression result for testing H2 (Model 5) at the industry level.

| Variable         | Electronics industry coef. | Non-electronics industry coef. |
|------------------|----------------------------|--------------------------------|
| NUM_POS          | 2.045*** (2.60)            | 0.169 (0.21)                   |
| NUM_NEG          | −12.43** (−2.42)           | −1.584 (−0.29)                 |
| IPO_PERIOD       | 6.870** (2.03)             | 6.699* (1.90)                  |
| LN_OFFER_SIZE    | −6.211* (−1.84)            | −4.009 (−1.24)                 |
| SCALED_OFFER     | 15.36*** (2.89)            | 4.495 (0.81)                   |
| MKTRET           | 0.664 (1.14)               | 1.755*** (3.00)                |
| INTERCEPT        | 48.82 (1.05)               | 31.143 (0.44)                  |
| N                | 133                        | 106                            |
| Adj. $R^2$       | 0.1574                     | 0.1199                         |

*p < .10.
**p < .05.
***p < .01.
3. Lowry and Schwert (2004) implied a negative correlation between issue size and price update, and initial return.
4. Logue (1973) found a positive relationship between pre-IPO market return and IPO first-day return.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research is supported by NSC 102-2627-E-004-001, MOST 103-2627-E-004-001, MOST 104-2627-E-004-001, and MOST 105-2811-H-004-035.

**Notes on contributors**

**Jia-Lang Seng** is a distinguished professor in the department and graduate school of Accounting, College of Commerce, National Chengchi University, Taipei, Taiwan. Prof. Seng received her doctorate from the University of Maryland, College Park, MD, U.S.A. Her research interests are focused on data analytics, business intelligence, information integration, IFRS, and XBRL. Prof. Seng has published papers in the *Decision Support Systems, Information & Management, Information Processing and Management, Journal of Information System, Information Sciences, Data & Knowledge Engineering*, and other highly regarded journals.

**Pi-Hua Yang** is a master graduate in the department and graduate school of Accounting at National Chengchi University, Taipei, Taiwan. Ms. Yang is currently a senior auditor at the PwC public audit firm. Her research interest is in the areas of auditing, IPO, and capital market analysis.

**Hsiao-Fang Yang** is a postdoctoral researcher in the department of Accounting at National Chengchi University, Taipei, Taiwan. She received her Ph.D. degree in the department of Management Information Systems at National Chengchi University, Taiwan (R.O.C), in 2014. Her research interests are in the areas of human–computer interaction in product design, data/text mining in knowledge discovery and capital market analysis.

**References**

Allen, F., & Faulhaber, G. R. (1989). Signaling by underpricing in the IPO market. *Journal of Financial Economics, 23*(2), 303–323.

Arbel, A., & Strebel, P. (1983). Pay attention to neglected firms!. *The Journal of Portfolio Management, 9* (2), 37–42.

Baker, M., Stein, J. C., & Wurgler, J. (2002). When does the market matter? Stock prices and the investment of equity-dependent firms. *The Quarterly Journal of Economics, 118*(3), 969–1005.

Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *Review of Financial Studies, 21*(2), 785–818.

Baron, D. P. (1982). A model of the demand for investment banking advising and distribution services for new issues. *Journal of Finance, 37*(4), 955–976.

Bautin, M., Vijayarenu, L., & Skiena, S. (2008, March-April). *International sentiment analysis for news and blogs*. Paper presented at the proceedings of the International Conference on Weblogs and Social Media (ICWSM 2008), Seattle, WA, USA.

Beatty, R. P., & Ritter, J. R. (1986). Investment banking, reputation, and the underpricing of initial public offerings. *Journal of Financial Economics, 15*(1), 213–232.
Benveniste, L. M., & Spindt, P. A. (1989). How investment bankers determine the offer price and allocation of new issues. *Journal of Financial Economics*, 24(2), 343–361.

Benveniste, L. M., & Wilhelm, W. J. (1990). A comparative analysis of IPO proceeds under alternative regulatory environments. *Journal of Financial Economics*, 28(1), 173–207.

Cook, D. O., Kieschnick, R., & Van Ness, R. A. (2006). On the marketing of IPOs. *Journal of Financial Economics*, 82(1), 35–61.

Davis, A. K., Piger, J. M., & Sedor, L. M. (2006). *Beyond the numbers: An analysis of optimistic and pessimistic language in earnings press releases* (Working paper No. 2006-005A). St. Louis, MO: Federal Reserve Bank of St. Louis. EconBiz website. Retrieved from https://research.stlouisfed.org/wp/2006/2006-005.pdf

Degeorge, F., Derrien, F., & Womack, K. L. (2004). *Quid pro quo in IPOs: Why book-building is dominating auctions* (Working paper No. 011). LSE website. Retrieved from http://www.lse.ac.uk/fmg/research/RICAFE/pdf/RICAFE-WP11-Degeorge.pdf

Derrien, F. (2005). Ipo pricing in ‘hot’ market conditions: Who leaves money on the table? *The Journal of Finance*, 60(1), 487–521.

Frieder, L., & Subrahmanyam, A. (2005). Brand perceptions and the market for common stock. *Journal of Financial and Quantitative Analysis*, 40(1), 57–85.

Godbole, N., Srinivasahal, M., & Skiena, S. (2007, March). *Large-scale sentiment analysis for news and blogs*. Paper presented at the Proceedings of the International Conference on Weblogs and Social Media (ICWSM 2007), Denver, CO, USA.

Grinblatt, M., & Hwang, C. Y. (1989). Signalling and the pricing of new issues. *Journal of Finance*, 44(2), 393–420.

Henry, E. (2008). Are investors influenced by how earnings press releases are written? *Journal of Business Communication*, 45(4), 363–407.

Lin, I. H. (2013). *Creating and verifying sentiment dictionary of finance and economics via financial news* (Unpublished master’s thesis). Accounting, National Taiwan University. Taipei, Taiwan (R.O.C).

Liu, L. X., Sherman, A. E., & Zhang, Y. (2009). *Media coverage and IPO underpricing* (Working Paper Series). Retrieved from Research Gate website: https://ssrn.com/abstract=972776

Ljungqvist, A., Nanda, V., & Singh, R. (2006). Hot markets, investor sentiment, and IPO pricing. *The Journal of Business*, 79(4), 1667–1702.

Logue, D. E. (1973). On the pricing of unseasoned equity issues: 1965–1969. *Journal of Financial and Quantitative Analysis*, 8(1), 91–103.

Lowry, M., & Schwert, G. W. (2004). Is the IPO pricing process efficient? *Journal of Financial Economics*, 71(1), 3–26.

Miller, G. A., Beckwith, R., Fellburn, C., Gross, D., & Miller, K. J. (1990). Introduction to WordNet: An online lexical database. *International Journal of Lexicography*, 3(4), 235–312.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, 58(1), 215–260.

Nasukawa, T., & Yi, J. (2003, October). *Sentiment analysis: Capturing favorability using natural language processing*. Paper presented at the Proceedings of the Conference on Knowledge Capture (K-CAP), pp. 70–77.

Neal, R., & Wheatley, S. M. (1998). Do measures of investor sentiment predict returns? *Journal of Financial and Quantitative Analysis*, 33(4), 523–547.

Rock, K. (1986). Why new issues are underpriced. *Journal of Financial Economics*, 15(1–2), 187–212.

Seng, J.-L., Yang, P.-H., & Yang, H.-F. (2016). IPO and Financial News. In N. T. Nguyen, B. Trawiński, H. Fujita, & T.-P. Hong (Eds.), *Intelligent information and database systems: 8th Asian Conference, AICIDS 2016, Da Nang, Vietnam, March 14–16, Proceedings, Part II* (pp. 346–353). Berlin: Springer.

Sherman, A. E., & Titman, S. (2002). Building the IPO order book: Underpricing and participation limits with costly information. *Journal of Financial Economics*, 65(1), 3–29.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3), 1139–1168.

Trauten, A., & Langer, T. (2012). Information production and bidding in IPOs: An experimental analysis of auctions and fixed-price offerings. *Zeitschrift für Betriebswirtschaft*, 82(4), 361–388.
Welch, I. (1989). Seasoned offerings, imitation costs, and the underpricing of initial public offerings. *The Journal of Finance, 44*(2), 421–449.

Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003, November). *Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques*. Paper presented at the Third IEEE International Conference on Data Mining (ICDM), Melbourne, FL, USA.