

Fei, Qingqing

**Article**

## Noise from online discussion forums: Evidence from IPOs in China

China Journal of Accounting Research

**Provided in Cooperation with:**

Sun Yat-sen University

*Suggested Citation:* Fei, Qingqing (2021) : Noise from online discussion forums: Evidence from IPOs in China, China Journal of Accounting Research, ISSN 1755-3091, Elsevier, Amsterdam, Vol. 14, Iss. 3, pp. 231-255,  
<https://doi.org/10.1016/j.cjar.2021.05.001>

This Version is available at:

<https://hdl.handle.net/10419/241838>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

HOSTED BY



ELSEVIER

Contents lists available at ScienceDirect

## China Journal of Accounting Research

journal homepage: [www.elsevier.com/locate/cjar](http://www.elsevier.com/locate/cjar)

# Noise from online discussion forums: Evidence from IPOs in China

Qingqing Fei<sup>\*</sup>*School of Accounting, Southwestern University of Finance and Economics, China*

## ARTICLE INFO

*Article history:*

Received 14 September 2020

Accepted 18 May 2021

Available online 6 June 2021

*Keywords:*

Online forum

IPO valuation

Underpricing

Overpricing

Quiet period

Online posting

## ABSTRACT

In this study, we use initial public offerings (IPOs) in China to investigate how online stock forums influence information asymmetry and IPO valuation. The empirical analysis isolates the underpricing and overvaluation components of initial returns. The number of forum comments, postings, and readings are positively associated with initial returns and the degree of underpricing, implying that forums create noise that exacerbates information asymmetry during IPOs. This effect is amplified by the quiet period regulation, which drives investors to rely on online discussion forums to obtain information. Through sentiment analyses of forum posts and media coverage, we find that the negative effect of online forums is more prominent when bad news prevails. We clarify the role of online stock forums in IPO pricing and information asymmetry by separating underpricing from overvaluation in initial returns.

© 2021 Sun Yat-sen University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

With the development of Internet technology in China, online stock forums are increasingly popular among investors. However, whether investors benefit from online postings and comments is unclear, and testing this issue is challenging under normal circumstances. Therefore, we use initial public offering (IPO) data to investigate how online stock forums influence information asymmetry and IPO valuations. The theoretical support for this study is partially from IPO rational theory. In the rational theory framework, underpricing is caused by information asymmetry and uncertainty (Rock, 1986; Benveniste and Spindt, 1989; Loughran et al., 1994; Chen et al., 2004). Lower levels of IPO underpricing indicate greater information transparency. The relation-

<sup>\*</sup> Corresponding author at: School of Accounting, Southwestern University of Finance and Economics, 555 Liutai Avenue, Chengdu, Sichuan 611130, China.

E-mail address: [feiqqing@outlook.com](mailto:feiqqing@outlook.com)

<https://doi.org/10.1016/j.cjar.2021.05.001>

1755-3091/© 2021 Sun Yat-sen University. Production and hosting by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

ship between information in online forums and IPO underpricing is tested to investigate the research topic. Unlike in other mature capital markets, an IPO's first-day return in China is not a good proxy for underpricing. Therefore, we must isolate the underpricing component of the first-day return before conducting the empirical analyses.

The research design considers the feasibility of using the level of IPO underpricing as a proxy for information asymmetry. Chinese IPO initial returns are both attention driven and information based (Gao, 2010; Song et al., 2014). An IPO offer price is often below its fair market price, whereas the short-run aftermarket equilibrium price is often above market value because investor sentiment is generally high in China. Therefore, rational theory alone cannot explain first-day returns that contain both under- and overpricing components. Behavioral theory that focuses on investor sentiment is thus appropriate (Ritter and Welch, 2002). Considering both rational and behavioral theories, we isolate the underpricing and overvaluation components to conduct the empirical analyses. Given the once fairly high level of IPO first-day returns in China, it is practical to test the influence of online discussions in an IPO setting.

Online discussion forums provide investors with places to converse by posting articles and comments. Forum information is one of the determinants of stock trading (Wysocki, 1999; Spiegel et al., 2010; Delort et al., 2011), but its effects on IPO under- and overpricing are unclear. Both rational and behavioral explanations are given regarding online postings (Wysocki, 1999). We investigate whether online discussions alleviate or exacerbate information asymmetry and investor sentiment. With underpricing and overpricing isolated from initial returns, the effect of online forums on the two components of initial returns are tested. We find that the number of forum comments, article postings, and article readings are positively related to initial returns and the level of underpricing. This result confirms our hypothesis that online forums create noise that exacerbates information asymmetry during IPOs.

On May 1, 2009, the China Securities Regulatory Commission (CSRC) started requiring IPOs in the ChiNext sector to follow the quiet period regulation.<sup>1</sup> On May 18, 2012, the CSRC introduced the quiet period regulation<sup>2</sup> to the main stock exchange. During the quiet period, issuers are not allowed to release information or opinions about the firm. We propose that the quiet period makes investors rely on stock forums for information, which amplifies the impact of online discussions. A split-sample design is used to test the influence of online forums on firms subject to the quiet period regulation versus those not subject to it. Consistent with our hypotheses, forum information is more influential on the IPO observations subject to the regulation.

In additional analyses, we test the influence of the sentiment conveyed by online postings using a split-sample design. A tone variable for online forum postings is constructed as the basis for splitting the sample. To further support the idea of tone information, we construct a tone variable for media coverage, which we also use to split the sample. The results imply that the negative effect of online forums is more prominent when bad news prevails, probably because negative information ferments and spreads to a greater extent.

Our findings indicate that online discussion forums create noise during IPOs that leads to greater underpricing. Firms with high information asymmetry may drive investors to online forums to gather information and discuss the firm. Therefore, reverse causality is a potential endogeneity concern. We test the correlation between online discussions and media coverage in the same period and find a significantly positive relationship, which means that less media coverage does not drive investors to online forums, thereby alleviating this endogeneity concern. In addition, we view the quiet period regulation as a shock to investors' dependence on external unofficial information sources and apply the difference-in-differences method to further alleviate concerns about reverse causality.

Our key contribution is our use of an IPO setting to investigate whether online forum discussions in China improve information transparency. We contribute to the literature by shedding new light on the interaction between Internet information and IPO pricing, and we provide a clearer picture of this interaction by isolating the under- and overpricing components. The evidence of noise in forums suggests that investors should not put too much faith in stock forum discussions. The implication for regulatory authorities is that firms need channels through which they can provide accurate information and respond to rumors and fake news during

<sup>1</sup> [http://www.csrc.gov.cn/pub/shanghai/ztzl/ggpx/zcfg/200906/t20090614\\_107440.htm](http://www.csrc.gov.cn/pub/shanghai/ztzl/ggpx/zcfg/200906/t20090614_107440.htm)

<sup>2</sup> [http://www.csrc.gov.cn/pub/zijhpublic/G00306201/201205/t20120521\\_210397.htm](http://www.csrc.gov.cn/pub/zijhpublic/G00306201/201205/t20120521_210397.htm)

IPOs to alleviate information asymmetry. Also, better supervision and self-correction mechanisms for online forums are necessary to protect investors.

The remainder of this paper is organized as follows. [Section 2](#) provides the institutional background of IPOs in China. [Section 3](#) discusses the related literature and development of the hypotheses. [Section 4](#) introduces the data and research methods. [Section 5](#) presents and discusses the empirical results. [Section 6](#) concludes the paper.

## 2. Institutional background of IPOs in China

To better understand IPO pricing in China, we must consider its institutional background. China's IPO pricing mechanism has experienced several stages of government regulation. In 1990, Chinese investors began trading shares in domestic stock markets. At the time, IPOs followed a fixed-price system under strict government control. Until 1996, IPO pricing was decided by the government based on book value. From January 1996 to June 1999, a controlled P/E ratio pricing model was used. Thus, China's IPO pricing was not market driven at that stage.

The Securities Law was the first law in China to specifically regulate the securities market. It plays an important role in promoting the development of the capital market. From 2000 to 2001, after the law's formal implementation, the CSRC issued new rules allowing issuing firms to negotiate pricing with underwriters. China thus began exploring market-oriented pricing.

In practice, market-oriented pricing reform did not achieve its goal. The original intention of implementing market-based pricing was to issue new shares at a price that reduced the profit from subscription and reduced secondary market speculation. However, the secondary market experienced even greater price speculation.

To alleviate this problem, from 2002 to 2004, China's IPOs reverted to a controlled P/E ratio pricing scheme. The P/E ratio of IPOs was required to be no more than 20. Obviously, returning to controlled P/E ratio pricing interrupted the market-oriented reform and induced disadvantages. For issuing firms with good growth prospects, this pricing mechanism often leads to underpricing.

In 2005, the book-building system was introduced to the primary market. In this system, the issuing firm and underwriter decide the initial offer price range. Then the underwriters determine institutional investors' demand for the stock and revise the final offer price according to that demand. The price had to be approved by the CSRC and the P/E ratio was limited to less than 30. Although pricing was still not totally market driven at this stage, more market forces were involved.

From June 2009, the CSRC launched IPO reforms to relax the constraints in the inquiry process. The controlled P/E ratio regulation was lifted. The CSRC reduced its administrative guidance for IPO pricing. However, the new reform resulted in pricing below the offer price and high P/E ratios in new markets. Thus, the CSRC began to reimpose price controls on new issues. In 2014, issue price was controlled within a P/E ratio of 23. Given this institutional background, to ensure our empirical analysis is comparable to those in the literature regarding other markets, we limit our IPO sample to before 2014.

## 3. Related literature and hypothesis development

### 3.1. Rational and behavioral theories

The most commonly used theoretical models for research on underpricing are based on rational theories. Information asymmetry theory is among the most popular of such theories. [Rock \(1986\)](#) and [Benveniste and Spindt \(1989\)](#) believe that IPOs are deliberately underpriced to compensate investors for information asymmetry. Uncertainty about the IPO firm is considered the most important factor in underpricing. It is well documented that ex ante uncertainty and underpricing are positively related ([Loughran et al., 1994](#); [Chen et al., 2004](#)). [Hanley \(1993\)](#) provides empirical support to [Benveniste and Spindt \(1989\)](#) by showing that IPOs with upward offer price adjustments tend to have higher levels of underpricing, which implies that investors are compensated by greater initial returns. Issuers and underwriters offer IPO shares at a discount to encourage participation and price adjustment. The findings of [Hanley and Hoberg \(2010\)](#) and [Loughran and McDonald](#)

(2013) in the U.S. market provide empirical evidence for the theoretical models of uncertainty, book-building, and prospect theory.

However, another branch of the literature uses irrational theory based on investor sentiment to explain abnormal IPO initial returns (Ritter and Welch, 2002; Song et al., 2014; Mumtaz et al., 2016). The anticipation of a positive trend increases the demand for IPO stocks, which in turn leads to a high closing price. Ritter and Welch (2002) believe that future progress on this issue will be generated from the irrational and agency conflict explanations. Abundant evidence indicates that an IPO's closing price may not reflect its fair value and that overreaction may result in a high closing price. For instance, Miller (1977) develops a model under short-selling constraints and finds that divergence in investor opinions drives an IPO's price higher than its intrinsic value because of optimism. Ritter (1991) and Loughran and Ritter (1995) provide empirical evidence that long-term investors who buy shares of a firm immediately after its IPO may realize abnormal negative returns. Purnanandam and Swaminathan (2004) find that some U.S. IPO firms are overpriced; however, their research method is contradicted by Zheng (2007). As mentioned by Zheng (2007), to demonstrate overpricing, they should not focus on the difference of long-run risk-adjusted returns between high and low overpricing firms. Instead, one should compare the long-run risk-adjusted returns of overpricing IPOs with their peer firms. Cornelli et al. (2006) document that individual investor demand results in high first-day abnormal returns and low long-run returns in Europe. Using the actual when-issued trades of a sample of clients at a large German retail brokerage during 1999 and 2000, Dorn (2009) finds that IPOs that are aggressively bought by individuals driven by sentiment exhibit high first-day returns and poor after-market returns compared with similar stocks. Mezhoud and Boubaker (2011) find that the initial returns during the listing period can explain not only IPO underpricing but also overpricing. To date, consensus on whether an IPO's initial return represents rational underpricing, irrational sentiment, or both has not been reached.

### 3.2. China's online stock forum environment and first-day returns

With the development of Internet technology in China, more and more financial websites, such as Eastmoney, Snowball, and Hexun, have set up stock forums for investors to discuss securities investment. Browsing and posting on stock forums has become a daily routine for many investors, which demonstrates investors' demand for in-time information. When information disclosure is not standard or complete, investors use stock forums to reduce the cost of information collection. Online forums allow individual investors to exchange opinions. We build on the literature that examines the effect of online forums in the context of market trading (Wysocki, 1999; Antweiler and Frank, 2004; Spiegel et al., 2010; Delort et al., 2011; Dong and Wu, 2019; Yang et al., 2020).

The opinions of other investors can affect an individual's decision-making and lead to converging viewpoints and behaviors (Das and Sisk, 2005). Online discussion forums can facilitate information dissemination and reduce information asymmetry. Investors with rational incentives read and post online to gather information (Wysocki, 1999). Using online forums for discussion allows investors with common interests to exchange ideas quickly. Therefore, online stock forums have the potential to alleviate information asymmetry and, according to rational theory, reduce first-day returns.

*H1a:* Online forum discussions tend to alleviate information asymmetry, reducing IPO first-day returns.

However, stock forums are imperfect as informal information release and dissemination platforms. For instance, an endless stream of irrelevant advertising and fraud have emerged in stock forums, making the overall information quality relatively low. Stock price manipulation can be achieved through online forums by disseminating false information. Furthermore, online forums lack effective self-correction mechanisms. As long as a posting does not involve plagiarism, personal attacks, or other negative content, it is allowed. Hence, forums have perhaps played a role in increasing information asymmetry by amplifying sentiment and accelerating rumor dissemination.

Sentiment information can spread quickly through online forums. Wysocki (1999) examines whether variation in message-posting volume is noise or is related to firm characteristics and stock market activity. Online discussions can be sentiment driven. The high posting volume for certain “glamour stocks” is probably driven by irrational fixation. With public information announcements controlled, Yang et al. (2020) find that sentimental information from investors can trigger abnormal trading and significantly affect stock price crashes. Among recent studies, Dong and Wu (2019) examine whether investor attention to online forums is a risk pricing factor.

Even worse, rumors can spread quickly through online forums and affect abnormal stock returns (Spiegel et al., 2010). The market reacts to rumors, and the impact is stronger for single, initial, or realized rumors. Delort et al. (2011) test manipulation in online forums and the associated market reactions. Consistently, they find that even with manual supervision of stock discussions, Internet users are not effectively protected from message manipulation. Even if the listing firms want to dispel rumors with explanations, disagreement among the posted messages is associated with increased trading volume and market volatility (Antweiler and Frank, 2004). Thus, forum postings and discussions can create noise that increases first-day returns.

*H1b:* Online forum discussions exacerbate information asymmetry, increasing IPO first-day returns.

### 3.3. IPO under- and overpricing in China

China’s market experienced fairly high IPO first-day returns before the first-day return restriction was issued. Investment behavior in China differs from that in other major markets (Tang and Li, 2013; Jiang and Akbar, 2018; Jiang et al., 2018). According to the classic rational theories (Rock, 1986; Benveniste and Spindt, 1989; Benveniste and Wilhelm, 1990), there must be a large degree of information asymmetry to cause such a substantial discount to the initial price. An assumption underlying this viewpoint is that the first-day closing price represents the fair value of the IPO firm. However, China’s market differs because it is full of irrational investors. The initial price following an IPO is always inflated by overreaction in the secondary market and corrects to its fundamental level in the long run. Therefore, information asymmetry theory is inadequate to explain a first-day return that contains both overpricing and underpricing.

Research on the Chinese IPO market reveals strong evidence of overpricing in the secondary market. For example, Gao (2010) provides strong evidence supporting the behavioral argument regarding overpricing in the Chinese IPO market. Using a sample of 506 Chinese IPOs issued during the 1998–2003 period, Shen et al. (2014) find that the offer price can reflect underpricing, whereas the short-run equilibrium price in the aftermarket can reflect overvaluation due to investor sentiment. Song et al. (2014) document that value uncertainty in IPOs is positively related to both underpricing and overvaluation and that investor sentiment has a positive effect on overvaluation but has no effect or a negative effect on underpricing. Huang et al. (2018) build regression models to explore the determinants of IPO overpricing and find that Internet data, such as online stock forums and search engines, contribute to an increase in the adjusted  $R^2$  value of the model. Therefore, according to the literature, both under- and overpricing compose China’s IPO first-day returns.

This dilemma encourages us to follow Song et al. (2014) and find a way to isolate these two components and empirically analyze their effects. After decomposing IPO first-day returns into the under- and overpricing components, we extend our hypotheses to the component level. If the information posted in stock forums is real and valid instead of sentiment driven, according to behavioral theory, forum information can reduce overpricing. Otherwise, the sentiment information and rumors tend to reduce information transparency and lead to greater under- and overpricing.

*H2a:* Online forum discussions alleviate information asymmetry and investor sentiment, thereby reducing underpricing and overpricing.

*H2b:* Online forum discussions exacerbate information asymmetry and investor sentiment, thereby increasing underpricing and overpricing.



## 4. Methods

### 4.1. Model specification

To test the first parts of our hypotheses, we first examine the effect of online forums on initial returns. The model is as follows:

$$IR = \beta_0 + \beta_1 ComList + \alpha.FirmIPO + \gamma.Rational + \theta.Behavioral + \varepsilon \quad (1)$$

In model 1, the dependent variable is initial return (*IR*) and the key independent variable is the number of online forum discussions before the listing date (*ComList*), which is expressed as *ComList1*, *ComList2*, and *ComList3*. Following the literature, we include three groups of control variables, which are defined in Section 4.2 and Table 1. The regression includes *Industry* and *Year* dummies.

Rational theory indicates that high initial returns imply information asymmetry and uncertainty (Rock, 1986; Benveniste and Spindt, 1989; Loughran et al., 1994; Chen et al., 2004). As discussed in Section 3, if an IPO's initial return is an appropriate proxy for underpricing and if online forum discussions exacerbate information asymmetry, the coefficient of *ComList* should be significantly positive. However, the initial returns of Chinese IPOs may not be a good proxy for underpricing because overpricing also constitutes part of the initial return. Hence, the coefficient of *ComList* contains both under- and overpricing. If online forum

Table 1  
Variable definitions.

Variable	Definition
<b>Dependent variables</b>	
<i>IR</i>	Initial returns = (first-day closing price – offer price)/offer price
<i>Underpricing</i>	Underpricing = (intrinsic price – offer price)/offer price
<i>Overpricing</i>	Overpricing = (closing price – intrinsic price)/offer price
<b>Independent variables</b>	
<i>ComList1</i> , <i>ComList2</i> , <i>ComList3</i>	Number of comments within 7, 14, and 60 days before the listing date
<i>ComIssue1</i> , <i>ComIssue2</i> , <i>ComIssue3</i>	Number of comments within 7, 14, and 60 days before the offer price declaration date
<i>Posting1</i> , <i>Posting2</i> , <i>Posting3</i>	Number of forum articles posted within 7, 14, and 60 days before the listing date
<i>Reading1</i> , <i>Reading2</i> , <i>Reading3</i>	Number of times articles read within 7, 14, and 60 days before the listing date
<i>News1</i> , <i>News2</i> , <i>News3</i>	Number of news articles within 7, 14, and 60 days before the listing date
<b>Control variables</b>	
<i>Gap</i>	Natural logarithm of the number of days between the issue date and listing date
<i>IssueSize</i>	Ln (funds raised through IPO)
<i>Underwriter</i>	Whether the IPO is underwritten by the top 10 underwriters
<i>EPS</i>	Earnings per share reported in the annual report before an IPO
<i>Board</i>	Whether the firm lists in the small and medium-sized enterprise sector or the ChiNext sector
<i>PriceRange</i>	(upper price limit – lower price limit)/(mid-range price)
	Mid-range price = mean value of upper and lower price limits
<i>Revision</i>	(offer price – mid-range price)/(mid-range price)
<i>Prospectus</i>	Ln (number of sentences in the IPO prospectus)
<i>NewsListing</i>	(positive news – negative news)/(positive news + negative news) before the listing date
<i>NewsIssue</i>	(positive news – negative news)/(positive news + negative news) before the offer price declaration date
<i>Oversubscription</i>	Ln [(number of shares effectively subscribed)/(number of shares issued)]
<i>Volume</i>	Trading volume in tens of millions on the listing date
<i>MomentumList</i>	Market returns over the 30 days before the listing date
<i>MomentumIssue</i>	Market returns over the 30 days before the offer price declaration date
<i>Quiet</i>	Whether the IPO firm is subject to the quiet period regulation
<i>SentiPosts1</i> , <i>SentiPosts2</i> , <i>SentiPosts3</i>	(positive posts – negative posts)/(positive posts + negative posts) within 7, 14, and 60 days before the offer price declaration date
<i>SentiNews1</i> , <i>SentiNews2</i> , <i>SentiNews3</i>	(positive news – negative news)/(positive news + negative news) within 7, 14, and 60 days before the offer price declaration date
<i>Industry</i>	Dummy variables for industry effects
<i>Year</i>	Dummy variables for year effects

discussions exacerbate information asymmetry, investor sentiment, or both, the coefficient should be positive. Models 4 and 5 should provide a clearer picture.

To test the second parts of the hypotheses, we must isolate overpricing to explore the link between online forum information and first-day valuation. Overreaction on the first day typically results in a high closing price far beyond the IPO's intrinsic value (Loughran and Ritter, 1995; Cornelli et al., 2006). We believe that the Chinese IPO market provides an ideal setting in which to investigate this topic. Separating underpricing from overpricing is econometrically unmanageable in developed markets in which IPO initial returns are relatively small. However, it is manageable in China's market. To isolate under- and overpricing, the closing price 6 months post-IPO is used as a proxy for intrinsic value. Under- and overpricing are calculated as follows:

$$\text{Underpricing} = (\text{Intrinsic price} - \text{Offer price}) / \text{Offer price} \quad (2)$$

$$\text{Overpricing} = (\text{Closing price} - \text{Intrinsic price}) / \text{Offer price} \quad (3)$$

To explore the effect of online forums on under- and overpricing, we use the following models:

$$\text{Underpricing} = \beta_0 + \beta_1 \text{ComIssue} + \alpha \cdot \text{FirmIPO} + \gamma \cdot \text{Rational} + \theta \cdot \text{Behavioral} + \varepsilon \quad (4)$$

$$\text{Overpricing} = \beta_0 + \beta_1 \text{ComList} + \alpha \cdot \text{FirmIPO} + \gamma \cdot \text{Rational} + \theta \cdot \text{Behavioral} + \varepsilon \quad (5)$$

where the dependent variables are the under- and overpricing components. The key independent variable is the number of relevant comments before the offer price declaration date (*ComIssue*), which is expressed as *ComIssue1*, *ComIssue2*, and *ComIssue3*. The regressions include *Industry* and *Year* dummies.

According to our theoretical analysis, both rational theory and behavioral theory are relevant. Under rational theory, a high level of underpricing implies information asymmetry and uncertainty (Rock, 1986; Benveniste and Spindt, 1989; Loughran et al., 1994; Chen et al., 2004), whereas under behavioral theory, investor sentiment and divergence in investor opinions can result in overpricing (Miller, 1977; Ritter and Welch, 2002; Cornelli et al., 2006; Mumtaz et al., 2016). Therefore, if online forum discussions exacerbate information asymmetry and investor sentiment, *Underpricing* and *Overpricing* should have positive coefficients.

In additional analyses, we use a split-sample design to test the effect of the quiet period regulation. *Quiet* is a dummy variable that is coded as 1 if the IPO observation is subject to the quiet period regulation. We compare the coefficients of *ComIssue* for the IPOs subject to the quiet period regulation with those not subject to the regulation. As the regulation restricts IPO firms from releasing information, investors may rely more on unofficial information sources, such as online forums, which increases the influence of online forum discussions. We thus conjecture a more significant effect of online discussions on IPOs that are subject to the quiet period regulation.

In addition to the influence of the amount of information available, we also consider the effect of sentiment information conveyed by online postings using a split-sample design. As shown in model 6, *SentiPosts* is constructed as a proxy for tone. It is measured as the difference between the number of positive and negative postings, scaled by the sum of positive and negative postings. We compare the coefficients of *ComIssue* for IPOs in the sample with *SentiPosts* values above and below its median value. According to the assumption that online forums tend to create noise, negative information ferments and spreads to a larger extent, increasing the effect of online comments. Therefore, we expect the coefficients of *ComIssue* to be more significant when the IPOs in the sample are exposed to negative tone.

$$\text{SentiPosts} = (\text{Positive posts} - \text{Negative posts}) / (\text{Positive posts} + \text{Negative posts}) \quad (6)$$

To further explore the idea regarding the tone of postings, we also construct a tone variable for media coverage. As shown in model 7, *SentiNews* is constructed as a proxy for media tone. It is measured as the difference between the number of positive and negative news articles, scaled by the sum of positive and negative news articles. Similarly, we compare the coefficients of *ComIssue* for IPOs in the sample with *SentiNews* values above and below its median value.

$$\text{SentiNews} = (\text{Positive news} - \text{Negative news}) / (\text{Positive news} + \text{Negative news}) \quad (7)$$



Furthermore, as a robustness test, the number of forum postings (*Posting1*, *Posting2*, and *Posting3*) and amount of browsing (*Reading1*, *Reading2*, and *Reading3*) are substituted for the number of comments as independent variables. To analyze sensitivity, we also use an alternative sample of under- and overpricing that excludes IPOs with negative under- or overpricing observations. Additionally, the relationship between online discussions and media coverage is tested, and the difference-in-differences method is used to alleviate endogeneity concerns.

## 4.2. Variable definitions

### 4.2.1. IPO initial returns

Following Loughran and McDonald (2013), Bajo and Raimondo (2017), Song et al. (2014), and Gao (2010), initial return (*IR*) is defined as the percentage of change from an IPO's offer price to its first-day closing price, which is the difference between the first-day closing price and the offer price, divided by the offer price. We then use the following variables from the literature to test our hypotheses.

### 4.2.2. IPO underpricing and overpricing

Following Song et al. (2014), we calculate underpricing as the difference between intrinsic value and offer price, scaled by offer price. We compute overpricing as the difference between first-day closing price and intrinsic value, scaled by offer price. The sum of the two is the IPO's initial return.

### 4.2.3. Online forum discussions

We use the number of relevant online forum comments to measure forum discussion. *ComList* is the number of comments posted within 7 days (*ComList1*), 14 days (*ComList2*), or 60 days (*ComList3*) before the listing date. *ComIssue* is the number of comments posted within 7 days (*ComIssue1*), 14 days (*ComIssue2*), or 60 days (*ComIssue3*) before the offer price declaration date.

### 4.2.4. Control variables to capture IPO characteristics (*FirmIPO*)

*Gap* is the natural logarithm of the number of days between the issue and listing dates. Unlike the U.S. market in which the IPO offer price is set 1 day before listing, China's market has a significant time lag between an IPO's offering date and its listing date. This longer processing time gives investors more time to collect and digest information, thereby reducing information asymmetry.

*IssueSize* is the natural logarithm of the funds raised through an IPO. Beatty and Ritter (1986) argue that smaller issues are subject to more uncertainty and find that issue size is negatively correlated with an IPO's initial return. Small firms may attract less attention, which leads to more information asymmetry during an IPO.

*Underwriter* is a dummy variable that is coded as 1 if the IPO is underwritten by a top 10 underwriter, which are those that raise the most funds. Individual investor attention can influence underwriters' offer price adjustment behavior (Huang and Zhang, 2020). Carter and Manaster (1990) find that in the 1980 s, IPOs underwritten by high-quality underwriters were less underpriced. However, Beatty and Welch (1996) find an inverse relationship between the two.

*EPS* is the earnings per share reported in the annual report before an IPO. Following Loughran and McDonald (2013), we include *EPS* to control for the historical performance of the IPO company.

*Board* is a dummy variable that is coded as 1 if a firm lists in the small and medium-sized enterprise or ChiNext sector. Firms in the small and medium-sized enterprise sector are smaller and have greater growth uncertainty, so investors and analysts tend to be more cautious, which pushes them to collect more information regarding the target firm. The ChiNext sector has a strict regulation regarding information disclosure. The greater growth uncertainty of ChiNext firms makes investors and analysts more discreet.

### 4.2.5. Control variables to test rational theory (*Rational*)

*PriceRange* is the initial offer price range (Hanley, 1993). It is defined as the difference between the upper and lower price limits, scaled by the mean value of the upper and lower price limits (mid-range price). During book-building, upper and lower limits are set as the initial price range for the offer price, which is the basis for

the IPO's book-building. After book-building, a final offer price is set. A wider range indicates more uncertainty regarding the IPO's valuation. Hanley (1993) finds that a wider price range is associated with higher IPO initial returns.

*Revision* is the offer price adjustment measured as the difference between the offer price and mid-range price, scaled by the mid-range price. *Revision* is used as strong evidence to support the book-building theory developed by Benveniste and Spindt (1989).

*Prospectus* denotes the length of an IPO prospectus measured as the natural logarithm of the number of sentences, which captures the amount of official information released by the IPO firm. Loughran and McDonald (2013) find that the information in an IPO prospectus affects investors' ability to precisely assimilate value-relevant information and thus influences pricing.

*NewsListing* is the tone of media coverage before the listing date. *NewsIssue* is the tone of media coverage before the offer price declaration date. Media coverage acts as an information intermediary to reduce information asymmetry (Bushee et al., 2010) and shapes the attitudes of society (Shaw, 1979). A positive tone sends a strong signal to investors regarding the expected riskiness and valuation of an IPO in the book-building process, and in turn it reduces information asymmetry (Hanley and Hoberg, 2010).

#### 4.2.6. Control variables to test behavioral theory (Behavioral)

*Oversubscription* is the natural logarithm of the oversubscription ratio measured as the number of shares effectively subscribed, scaled by the number of shares issued. IPOs subject to strong individual investor demand have higher initial returns and suffer lower long-term returns, indicating that these IPOs are overvalued (Derrien, 2005; Cornelli et al., 2006).

*Volume* is the trading volume (in tens of millions) on the listing date. Cornelli et al. (2006) find that the aftermarket total IPO trading volume is positively correlated with individual investor sentiment, which leads to high IPO first-day prices and low long-run returns. Ofek and Richardson (2003) show that high initial returns occur when institutions sell IPO shares to retail investors on the first day. The total trading volume is an indicator of individual investor behavior, especially given that China's market is largely driven by individual investors.

Momentum is calculated in two ways. *MomentumList* is the market returns over the 30 days before the listing date. *MomentumIssue* is the market returns over the 30 days before the issue date. *MomentumIssue* is controlled in the underpricing model, whereas *MomentumList* is controlled in the overpricing model. Classic rational theory implies that public information, such as market momentum, should not affect IPO underpricing. The underwriter should fully adjust the offer price to eliminate the effect of public information. However, Loughran and Ritter (2002) use prospect theory to explain that underwriters only partially adjust the offer price for public information on market momentum, and IPOs in high-momentum markets are more underpriced.

### 4.3. Data and descriptive statistics

#### 4.3.1. Data collection

As the data for the initial pricing range are only available after November 2010, the sample starts from 2010. In 2014, IPO pricing in China was limited to a P/E ratio of less than 23, which limits the IPO sample before 2014. In fact, from November 2012 to December 2013, IPO activity in China stagnated under the CSRC rule. The final sample is from 2010 to 2012 and includes 430 IPOs. The data for the IPOs are obtained from WIND, a leading capital market information provider in China. Information about the online discussion forums, media coverage, and underwriters is retrieved from the Chinese Research Data Services (CNRDS) platform. The IPO prospectuses are obtained from the Cninfo website. The other market- and firm-level data are obtained from the China Stock Market and Accounting Research (CSMAR) database.

Specifically, the key independent variables of forum comments, posts, and readings are calculated using information from the CNRDS's subdatabase, Stocks Comments of Chinese Listed Companies. This database is a professional database of Internet financial and economic texts, providing text analysis and the quantitative statistics of forum comments and postings regarding listed firms in China. We first obtain the statistics for

each natural day from the database and then calculate the statistics for the required time windows (7, 14, and 60 days).

The sentiment information provided by the Stocks Comments of Chinese Listed Companies database is used to calculate *SentiPosts*. The database uses a supervised learning model to judge the emotion of a comment or posting. Sentiment predictions are divided into three categories according to the emotional tendency of the text: positive, negative, or neutral. After labeling training materials, a support-vector machine algorithm is used to train and obtain the classification model. The trained model is then used to label all of the texts. Table 2 shows examples of the sentiment classification, which are obtained from the CNRDS database specifications. We use an asterisk to hide the names of the stocks.

The sentiment information from the CNRDS subdatabase Financial News Database of Chinese Listed Companies is used to calculate *SentiNews*, *NewsListing*, and *NewsIssue*. Similarly, this database uses a support-vector machine algorithm to analyze the financial news of listed firms. The database includes information from more than 400 major online media outlets and more than 600 major newspapers. We first obtain the number of news articles in the three sentiment categories for each natural day and then calculate the variables.

In addition, one of the control variables, *Prospectus*, is constructed by collecting the IPO prospectuses of the firms in the sample and conducting textual analysis with Python. The IPO prospectuses are obtained from the Cninfo website, which is designated by the CSRC as an information disclosure website for listed firms in China. We use Python to automatically crawl and download the prospectuses and then calculate the number of sentences in each prospectus to measure its length.

#### 4.3.2. Descriptive statistics

Table 3 presents the summary statistics of the sample. On average, more than 2000 (1000) comments are posted within 7 days before a listing (issue) date. The mean value of *IR* is 25.9%. The mean value of *Quiet* is 0.512, suggesting that approximately 50% of the observations are subject to the quiet period regulation. The mean values of *SentiPosts* (0.406, 0.426, and 0.427) are higher than those of *SentiNews* (0.333, 0.333, and 0.143), indicating that online forums produce more sentiment information than media news reports produce. On average, *NewsListing* (0.133) is higher than *NewsIssue* (0.084), implying that media coverage produces more positive news as the IPO listing date approaches.

Table 4 shows the correlation coefficients and their significance at the 1% confidence level. The Spearman correlation coefficients are on the upper right triangle, and the Pearson correlation coefficients are on the lower left triangle. Both the Pearson and Spearman correlation coefficients show that initial return (*IR*) is positively and significantly associated with *ComList1*, indicating that online forum comments are positively associated with information asymmetry or investor sentiment. In addition, *Underpricing* has a positive relationship with *ComList1* and *ComIssue1*, further confirming the conjecture that online forum discussions create noise. The coefficients between *Overpricing* and forum comments are not significant, suggesting that forum information does not further increase investors' enthusiasm. The results of the correlation analysis are consistent with hypotheses *H1b* and *H2b* concerning the negative influence of online forums.

Table 2  
Examples of sentiment classification.

Posting title	Sentiment classification
*, restructuring resolution passed, is about to soar	Positive
A broken stock, identification completed	Negative
Civilized Posting, put an end to abuse	Neutral
Let me tell you another joke today	Neutral
Ha (laugh), I yesterday at the end whole warehouse into *, today up! rich!	Positive
You should stop arguing about trifles, really not clear	Neutral
Never saw a stock as disgusting as * again	Negative
It will fall sideways!!! Quick out!!!	Negative
* is soft!!	Negative
!!!	Neutral
At last there was something to look forward to, and it began to rise gradually	Positive

Table 3  
Summary statistics.

Variable	Obs.	Mean	SD	P25	Median	P75
<i>IR</i>	430	0.259	0.444	0.004	0.171	0.371
<i>Underpricing</i>	430	−0.024	0.459	−0.270	−0.110	0.151
<i>Overpricing</i>	430	0.283	0.364	0.108	0.299	0.499
<i>ComList1</i>	430	2339	3320	860.0	1522	2791
<i>ComList2</i>	430	4585	6060	1864	3160	5151
<i>ComList3</i>	430	5418	10,609	2251	3486	5663
<i>ComIssue1</i>	430	1255	3687	187.0	420.0	1174
<i>ComIssue2</i>	430	1408	4296	219.0	468.5	1230
<i>ComIssue3</i>	430	1425	4330	224.0	470.5	1237
<i>Posting1</i>	430	333.1	382.2	149.0	237.5	400.0
<i>Posting2</i>	430	555.0	667.2	259.0	388.0	628.0
<i>Posting3</i>	430	633.8	1101	289.0	419.5	679.0
<i>Reading1</i> (millions)	430	0.501	0.773	0.210	0.345	0.590
<i>Reading2</i> (millions)	430	1.052	1.239	0.494	0.768	1.213
<i>Reading3</i> (millions)	430	1.226	1.893	0.580	0.874	1.312
<i>Gap</i>	430	2.286	0.216	2.079	2.197	2.398
<i>IssueSize</i>	430	20.319	0.689	19.846	20.229	20.682
<i>Underwriter</i>	430	0.474	0.500	0.000	0.000	1.000
<i>EPS</i>	430	0.872	0.502	0.560	0.780	1.010
<i>Board</i>	430	0.858	0.349	1.000	1.000	1.000
<i>PriceRange</i>	430	0.737	0.160	0.636	0.723	0.832
<i>Revision</i>	430	0.102	0.120	0.026	0.094	0.183
<i>Prospectus</i>	430	7.704	0.188	7.593	7.703	7.814
<i>NewsListing</i>	430	0.133	0.357	−0.136	0.128	0.385
<i>NewsIssue</i>	430	0.084	0.467	−0.259	0.042	0.385
<i>Oversubscription</i>	430	4.465	0.915	3.951	4.625	5.112
<i>Volume</i>	430	2.769	7.917	1.019	1.503	2.419
<i>MomentumList</i>	430	−0.018	0.056	−0.058	−0.029	0.025
<i>MomentumIssue</i>	430	−0.012	0.063	−0.059	−0.025	0.035
<i>Quiet</i>	430	0.512	0.500	0.000	1.000	1.000
<i>SentiPosts1</i>	430	0.406	0.255	0.333	0.456	0.560
<i>SentiPosts2</i>	430	0.426	0.212	0.357	0.465	0.561
<i>SentiPosts3</i>	430	0.427	0.206	0.366	0.463	0.555
<i>SentiNews1</i>	430	0.333	0.000	0.333	0.333	0.333
<i>SentiNews2</i>	430	0.333	0.000	0.333	0.333	0.333
<i>SentiNews3</i>	430	0.143	0.000	0.143	0.143	0.143

Notes. The table provides the summary statistics of the variables for the 2010–2012 period. *Reading1*, *Reading2*, and *Reading3* are in millions.

To understand the distribution of IPO first-day returns among the groups, Fig. 1 depicts the average first-day returns of each group according to the percentile rank of the number of comments. The low group contains IPO observations for which the number of comments is below the 20th percentile. The high group includes IPO observations for which the number of comments is above the 80th percentile.

Fig. 1 illustrates the power of online forum discussions in explaining differences in first-day returns without the IPO control variables. The figure plots IPO first-day returns sorted by the number of comments posted before the listing date. Each IPO observation in our sample is placed in one of five groups based on its number of comments. The figure shows a monotonic increase in initial returns from the group with least comments to the group with the most comments. For example, IPOs in the low 7-day comment (*ComList1*) group have average first-day returns of 13.35% compared with 45.11% for the high 7-day comment (*ComList1*) group, which is a difference of 31.76% between the extreme groups. This large difference indicates a positive relationship between online forum discussions and IPO first-day returns.

Table 4  
Correlation analysis.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. <i>IR</i>	1	0.468*	0.396*	0.262*	0.047	0.026	−0.354*	−0.016	−0.285*	0.046	0.108	−0.066	0.049	−0.035
2. <i>Underpricing</i>	0.676*	1	−0.539*	0.148*	0.033	−0.086	−0.429*	−0.040	−0.200*	0.049	0.117	0.013	0.073	−0.110
3. <i>Overpricing</i>	0.366*	−0.438*	1	0.098	−0.011	0.131*	0.115	0.026	−0.060	−0.019	−0.022	−0.090	−0.042	0.046
4. <i>ComList1</i>	0.245*	0.186*	0.064	1	0.491*	−0.246*	0.145*	0.005	−0.167*	−0.312*	0.116	0.079	0.148*	−0.132*
5. <i>ComIssue1</i>	0.136*	0.154*	−0.028	0.557*	1	0.038	0.229*	0.103	−0.049	−0.430*	0.089	0.150*	0.105	−0.204*
6. <i>Gap</i>	−0.005	−0.082	0.097	−0.194*	0.048	1	0.037	0.061	−0.027	−0.093	0.024	−0.065	0.012	0.000
7. <i>IssueSize</i>	−0.363*	−0.403*	0.065	0.243*	0.298*	0.086	1	0.098	0.131*	−0.426*	0.063	−0.026	−0.060	0.122
8. <i>Underwriter</i>	−0.015	−0.032	0.021	0.034	0.114	0.063	0.117	1	0.050	−0.108	0.003	0.056	−0.098	0.005
9. <i>EPS</i>	−0.182*	−0.182*	0.007	0.010	0.034	−0.030	0.187*	0.045	1	0.176*	−0.155*	0.008	−0.040	−0.074
10. <i>Board</i>	0.007	0.024	−0.023	−0.354*	−0.482*	−0.103	−0.496*	−0.108	0.115	1	−0.094	−0.028	−0.020	0.006
11. <i>PriceRange</i>	0.149*	0.177*	−0.042	0.101	0.085	0.033	0.033	0.018	−0.075	−0.055	1	0.177*	−0.027	0.001
12. <i>Revision</i>	−0.232*	−0.151*	−0.093	0.025	0.075	−0.090	0.002	0.063	−0.006	−0.001	0.171*	1	0.001	−0.190*
13. <i>Prospectus</i>	−0.029	0.005	−0.042	0.126*	0.103	0.043	−0.007	−0.088	−0.009	0.009	0.019	−0.005	1	−0.070
14. <i>NewsListing</i>	−0.114	−0.161*	0.065	−0.152*	−0.201*	0.004	0.099	0.010	−0.094	−0.001	−0.022	−0.144*	−0.084	1

Notes. This table presents the correlation matrix for the key variables. The Spearman correlation coefficients are on the upper right triangle, and the Pearson correlation coefficients are on the lower left triangle. \* denotes significance at the 1% confidence level.

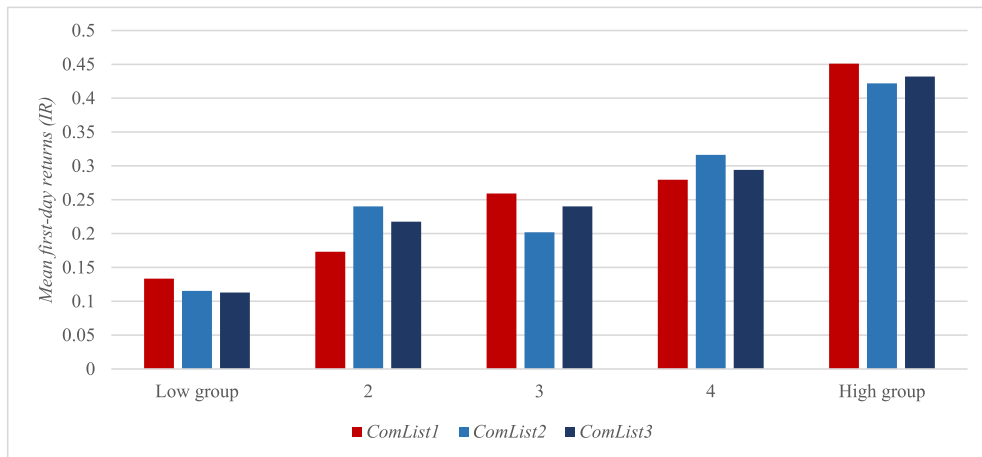


Fig. 1. Mean IPO first-day returns sorted by the number of comments posted before the listing date.

## 5. Empirical results

### 5.1. Online forum discussions and initial returns

Table 5 presents the regression results of model 1. Columns (1), (3), and (5) report the regression results with the traditional IPO control variables, whereas columns (2), (4), and (6) include individual investor over-subscription, trading volume, and momentum, respectively, which support the behavioral theory explanation.

In all of the columns, *ComList1*, *ComList2*, and *ComList3* are positively and significantly related to *IR* ( $p$ -value  $< 0.01$ ), which indicates that online forums produce noise that increases information asymmetry, produces sentiment that affects investors' trading behavior, or both. As IPO initial returns are not an appropriate proxy for underpricing in the Chinese market, it is unclear which effect dominates. Models 4 and 5 isolate under- and overpricing, which should provide a clearer picture.

The adjusted  $R^2$  values in columns (1), (3), and (5) are 38.4%, 40.0%, and 40.5%, respectively. When over-subscription, trading volume, and market momentum are added in columns (2), (4), and (6), the adjusted  $R^2$  values all increase to approximately 45%, indicating the additional explanatory power of the variables in the behavioral framework. Therefore, the empirical results do not entirely support the rational underpricing argument in China's IPO market. The complexity of the Chinese market inspires us to decompose first-day returns into the components of underpricing by issuers or underwriters and overpricing by investors, which contributes to the high initial returns in the Chinese market.

The control variables are also of interest. *IssueSize* is negatively associated with *IR*, which is consistent with the idea that smaller issues are subject to more uncertainty (Beatty and Ritter, 1986). In line with the findings of Hanley (1993), the coefficients of *PriceRange* are positive and significant, suggesting that a wider price range indicates more information asymmetry. The coefficients of *Prospectus* are all negative and significant, which we interpret as implying a positive role of IPO prospectuses in reducing information asymmetry. *NewsListing* also has negative coefficients. Similarly, this shows that the positive signal of media coverage increases information transparency. *MomentumList* is positively and significantly correlated with *IR*. Rational theory implies that underwriters fully adjust the offer price using market information, such as momentum. However, Loughran and Ritter (2002) explain that underwriters only partially adjust the offer price. IPOs in high-momentum markets are significantly underpriced. The results suggest that both rational and behavioral variables have explanatory power. Further regressions are performed on the separated components to provide a clearer picture.



Table 5  
Online forum discussions and initial returns.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
<i>ComList1</i>	0.170*** (4.74)	0.158*** (3.72)				
<i>ComList2</i>			0.198*** (3.53)	0.201*** (2.87)		
<i>ComList3</i>					0.212*** (3.39)	0.214*** (2.73)
<i>Gap</i>	0.140 (1.43)	0.097 (0.93)	0.046 (0.46)	0.021 (0.21)	−0.081 (−0.64)	−0.104 (−0.78)
<i>IssueSize</i>	−0.405*** (−4.36)	−0.437*** (−3.54)	−0.398*** (−4.48)	−0.442*** (−3.49)	−0.409*** (−4.52)	−0.452*** (−3.48)
<i>Underwriter</i>	0.059 (1.08)	0.048 (0.88)	0.043 (0.86)	0.035 (0.69)	0.043 (0.87)	0.035 (0.69)
<i>EPS</i>	0.001 (0.02)	0.035 (1.16)	−0.003 (−0.10)	0.022 (0.66)	−0.004 (−0.14)	0.021 (0.60)
<i>Board</i>	−0.217*** (−3.19)	−0.262*** (−3.96)	−0.187*** (−2.83)	−0.219*** (−3.38)	−0.188*** (−2.86)	−0.227*** (−3.56)
<i>PriceRange</i>	0.495** (2.13)	0.533** (2.30)	0.478** (2.23)	0.498** (2.45)	0.476** (2.26)	0.499** (2.50)
<i>Revision</i>	−0.813 (−1.55)	−0.778 (−1.52)	−0.829* (−1.68)	−0.771 (−1.63)	−0.805* (−1.68)	−0.754 (−1.65)
<i>Prospectus</i>	−0.236* (−1.88)	−0.217* (−1.81)	−0.242* (−1.87)	−0.225* (−1.90)	−0.238* (−1.84)	−0.212* (−1.82)
<i>NewsListing</i>	−0.135** (−2.16)	−0.144** (−2.59)	−0.125** (−2.23)	−0.136*** (−2.66)	−0.117** (−2.15)	−0.129*** (−2.60)
<i>Oversubscription</i>		0.005 (0.15)		−0.015 (−0.34)		−0.013 (−0.29)
<i>Volume</i>		0.004 (1.37)		0.004 (1.41)		0.003 (1.27)
<i>MomentumList</i>		1.456*** (2.98)		1.766*** (4.13)		1.649*** (3.74)
<i>MomentumIssue</i>		0.549 (0.96)		0.345 (0.64)		0.483 (0.84)
<i>Constant</i>	8.914*** (3.67)	9.525*** (3.16)	8.567*** (3.91)	9.432*** (3.28)	8.927*** (4.00)	9.674*** (3.31)
<i>Observations</i>	430	430	430	430	430	430
<i>Adj. R<sup>2</sup></i>	0.384	0.436	0.400	0.456	0.405	0.461

Notes. This table shows the results for model (1). The regressions include industry and calendar year dummies. *ComList* is in log form. The *t* statistics are reported in parentheses and are based on robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

## 5.2. Online forum discussions and IPO valuation components

As under- and overpricing coexist on the first day of an IPO, we examine the link between online forum comments and each pricing component. Table 6 reports the regression results for models (4) and (5). According to rational theory, IPOs are deliberately underpriced to compensate investors for information asymmetry (Rock, 1986; Benveniste and Spindt, 1989). Thus, if forum comments play a positive role in information transparency, its coefficients should exhibit negative signs. Table 6 indicates a positive relationship between forum comments and *Underpricing*, but there is no significant relationship between comments and *Overpricing*. *ComIssue1*, *ComIssue2*, and *ComIssue3* in columns (1) to (3) are all positively and significantly related to *IR* ( $p$ -value < 0.05), suggesting that forum comments before the issue date play a negative role in information symmetry. However, the coefficients of *ComList1*, *ComList2*, and *ComList3* in columns (4) to (6) are insignificant, which we interpret as indicating that Chinese investors' IPO enthusiasm cannot be further increased by forum information given its already high level. As a belief in undefeated new shares is deeply rooted in Chinese investors' ideas, additional forum information has little influence on this belief.

Table 6  
Online forum discussions and IPO valuation components.

Variable	<i>Underpricing</i>			<i>Overpricing</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>ComIssue1</i>	0.098** (2.32)					
<i>ComIssue2</i>		0.104** (2.22)				
<i>ComIssue3</i>			0.109** (2.34)			
<i>ComList1</i>				0.007 (0.27)		
<i>ComList2</i>					0.025 (0.77)	
<i>ComList3</i>						0.016 (0.47)
<i>Gap</i>	−0.269*** (−2.61)	−0.274*** (−2.65)	−0.272*** (−2.65)	0.220*** (2.73)	0.225*** (2.99)	0.211*** (2.68)
<i>IssueSize</i>	−0.415*** (−4.11)	−0.418*** (−4.09)	−0.419*** (−4.13)	−0.016 (−0.29)	−0.022 (−0.40)	−0.019 (−0.34)
<i>Underwriter</i>	−0.012 (−0.24)	−0.014 (−0.29)	−0.015 (−0.30)	0.047 (1.24)	0.047 (1.23)	0.046 (1.22)
<i>EPS</i>	−0.123*** (−2.74)	−0.121*** (−2.71)	−0.122*** (−2.76)	0.171*** (5.57)	0.166*** (5.32)	0.169*** (5.38)
<i>Board</i>	−0.080 (−1.00)	−0.072 (−0.88)	−0.065 (−0.81)	−0.155** (−2.29)	−0.142** (−2.07)	−0.149** (−2.16)
<i>PriceRange</i>	0.574*** (3.16)	0.576*** (3.20)	0.573*** (3.19)	−0.019 (−0.15)	−0.030 (−0.23)	−0.025 (−0.19)
<i>Revision</i>	−1.036** (−2.36)	−1.057** (−2.39)	−1.061** (−2.41)	0.204 (0.61)	0.211 (0.64)	0.208 (0.63)
<i>Prospectus</i>	−0.157 (−1.32)	−0.161 (−1.34)	−0.156 (−1.32)	−0.035 (−0.31)	−0.037 (−0.32)	−0.035 (−0.31)
<i>Oversubscription</i>	−0.106*** (−2.63)	−0.105*** (−2.59)	−0.107*** (−2.62)	0.140*** (3.91)	0.131*** (3.70)	0.136*** (3.81)
<i>Volume</i>	0.006* (1.82)	0.006* (1.81)	0.005* (1.76)	0.000 (0.13)	0.000 (0.02)	0.000 (0.07)
<i>NewsListing</i>				−0.050 (−1.00)	−0.046 (−0.92)	−0.048 (−0.96)
<i>MomentumList</i>				1.182*** (3.04)	1.221*** (3.16)	1.199*** (3.10)
<i>NewsIssue</i>	−0.046 (−1.09)	−0.045 (−1.07)	−0.045 (−1.06)			
<i>MomentumIssue</i>	0.804** (2.10)	0.855** (2.16)	0.862** (2.18)			
<i>Constant</i>	9.600*** (4.01)	9.630*** (4.01)	9.589*** (4.06)	0.076 (0.06)	0.093 (0.08)	0.097 (0.08)
<i>Observations</i>	430	430	430	430	430	430
<i>Adj. R<sup>2</sup></i>	0.377	0.377	0.381	0.275	0.277	0.276

Notes. This table shows the results for models (4) and (5). The regressions include industry and calendar year dummies. *ComList* and *ComIssue* are in log form. The *t* statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

Regarding the control variables, *EPS* is negatively associated with underpricing, whereas it is positively related to overpricing, implying that good firm performance signals positive information that reduces information asymmetry but also triggers overvaluation. *PriceRange* and *Revision* are only significant in the underpricing model, which is consistent with our conjecture that *PriceRange* and *Revision* have explanatory power within the rational framework. *Oversubscription* is positively related to overpricing, whereas it is negatively associated with underpricing. Greater individual investor demand leads to higher initial returns and overval-

uation (Derrien, 2005; Cornelli et al., 2006). The variables of market momentum are also significant and exhibit the expected signs.

### 5.3. Additional analyses

#### 5.3.1. Online forum discussions during the quiet period and underpricing

Table 7 presents the results for the split-sample design. The dependent variable is *Underpricing*. *Quiet* is a dummy variable that is coded as 1 if the IPO observation is subject to the quiet period regulation. The coefficients of *ComIssue1*, *ComIssue2*, and *ComIssue3* are only positive and significant ( $p$ -value < 0.05) in columns (1), (3), and (5), respectively. This implies that the impact of stock forums is more significant on firms that are subject to the quiet period regulation.

This finding provides evidence that the quiet period regulation contravenes its goal. The regulation's aim is to protect investors from noisy information released by IPO firms. However, our results show that although issuers are not allowed to disclose promotional information during the quiet period, online forum discussions create noise that increases information asymmetry. It is natural for investors to seek information through

Table 7  
Online forum discussions during the quiet period and underpricing.

Variable	(1) <i>Quiet</i> = 1	(2) <i>Quiet</i> = 0	(3) <i>Quiet</i> = 1	(4) <i>Quiet</i> = 0	(5) <i>Quiet</i> = 1	(6) <i>Quiet</i> = 0
<i>ComIssue1</i>	0.112** (2.55)	0.033 (1.04)				
<i>ComIssue2</i>			0.120** (2.36)	0.036 (1.10)		
<i>ComIssue3</i>					0.129** (2.57)	0.038 (1.16)
<i>Gap</i>	−0.019 (−0.13)	−0.341*** (−3.03)	−0.041 (−0.28)	−0.341*** (−3.03)	−0.032 (−0.21)	−0.342*** (−3.03)
<i>IssueSize</i>	−0.583*** (−4.29)	−0.344*** (−6.22)	−0.589*** (−4.25)	−0.345*** (−6.26)	−0.590*** (−4.31)	−0.347*** (−6.29)
<i>Underwriter</i>	−0.008 (−0.11)	0.029 (0.58)	−0.007 (−0.10)	0.028 (0.56)	−0.005 (−0.06)	0.027 (0.55)
<i>EPS</i>	−0.092 (−1.47)	−0.076 (−0.97)	−0.092 (−1.45)	−0.076 (−0.98)	−0.093 (−1.51)	−0.077 (−0.99)
<i>Board</i>	0.293 (0.50)	−0.073 (−0.88)	0.246 (0.44)	−0.068 (−0.81)	0.257 (0.46)	−0.064 (−0.77)
<i>PriceRange</i>	0.762*** (3.09)	0.274* (1.72)	0.772*** (3.12)	0.273* (1.71)	0.764*** (3.12)	0.272* (1.71)
<i>Revision</i>	−1.356*** (−2.61)	−0.246 (−1.15)	−1.395*** (−2.66)	−0.255 (−1.18)	−1.408*** (−2.70)	−0.256 (−1.19)
<i>Prospectus</i>	−0.410 (−1.64)	0.115 (0.95)	−0.404 (−1.59)	0.118 (0.99)	−0.391 (−1.55)	0.120 (1.01)
<i>NewsIssue</i>	−0.107 (−1.34)	−0.026 (−0.49)	−0.104 (−1.31)	−0.025 (−0.47)	−0.103 (−1.31)	−0.025 (−0.46)
<i>Oversubscription</i>	−0.142** (−2.13)	−0.115*** (−3.49)	−0.137** (−2.06)	−0.116*** (−3.60)	−0.140** (−2.08)	−0.117*** (−3.65)
<i>Volume</i>	0.032 (1.21)	0.008 (3.62)	0.029 (1.10)	0.008** (3.59)	0.029 (1.09)	0.007*** (3.57)
<i>MomentumIssue</i>	1.297 (1.46)	0.634* (1.68)	1.335 (1.48)	0.649* (1.72)	1.348 (1.50)	0.653* (1.73)
<i>Constant</i>	14.027*** (4.12)	7.494*** (5.16)	14.112*** (4.09)	7.487*** (5.17)	13.942*** (4.15)	7.500*** (5.20)
<i>Observations</i>	220	210	220	210	220	210
<i>Adj. R<sup>2</sup></i>	0.483	0.306	0.481	0.307	0.487	0.308

Notes. This table demonstrates the results of the split-sample design. The regressions include industry and calendar year dummies. *ComIssue* is in log form. The  $t$  statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

online discussion forums when no official information is available. This was probably not anticipated and has perhaps gone unnoticed by regulators.

Information asymmetry during the quiet period is likely to be high, so an endogeneity concern is that firms with high information asymmetry may drive investors to online forums to gather information. Thus, it is possible that investors choose to comment online because of the quiet period regulation that restricts official information supply. The significance of *ComIssue* in the split-sample design may be the result of this underlying information asymmetry instead of online discussions.

However, when testing the regression coefficient of *Quiet* with *Underpricing* as the dependent variable and with the other variables controlled, we find insignificance. This result indicates that the quiet period regulation does not increase information asymmetry. Investors are more influenced by online forums during the quiet period purely because there is less information from the firm, even if the information restricted by the regulation is also confusing and deceptive. Thus, the insignificance of *Quiet* implies that investors are confused either by the IPO firm or by the online forum. Even if the regulators stop IPO firms from excessively promoting themselves with false information, investors turn to the Internet and receive noisy information anyway.

Table 8  
Online forum discussions and underpricing with different forum sentiments.

Variable	(1) <i>SentiPosts1</i> <i>Above</i>	(2) <i>SentiPosts1</i> <i>Below</i>	(3) <i>SentiPosts2</i> <i>Above</i>	(4) <i>SentiPosts2</i> <i>Below</i>	(5) <i>SentiPosts3</i> <i>Above</i>	(6) <i>SentiPosts3</i> <i>Below</i>
<i>ComIssue1</i>	0.042 (1.35)	0.129** (2.50)				
<i>ComIssue2</i>			0.028 (0.84)	0.135*** (2.61)		
<i>ComIssue3</i>					0.021 (0.60)	0.153*** (2.92)
<i>Gap</i>	−0.199 (−1.56)	−0.231 (−1.63)	−0.155 (−1.10)	−0.249* (−1.78)	−0.092 (−0.66)	−0.263* (−1.94)
<i>IssueSize</i>	−0.372*** (−4.14)	−0.526*** (−4.32)	−0.422*** (−4.79)	−0.460*** (−3.81)	−0.398*** (−4.71)	−0.511*** (−3.87)
<i>Underwriter</i>	−0.065 (−1.02)	−0.007 (−0.09)	−0.102 (−1.57)	0.036 (0.40)	−0.085 (−1.35)	0.037 (0.43)
<i>EPS</i>	−0.053 (−0.40)	−0.122** (−2.09)	−0.111 (−0.95)	−0.092* (−1.75)	−0.085 (−0.69)	−0.097* (−1.81)
<i>Board</i>	0.142 (1.11)	−0.001 (−0.01)	0.064 (0.44)	0.066 (0.46)	0.110 (0.78)	0.002 (0.01)
<i>PriceRange</i>	0.369** (1.98)	0.689*** (2.75)	0.331 (1.60)	0.738*** (2.99)	0.318 (1.61)	0.693*** (2.88)
<i>Revision</i>	−0.205 (−0.83)	−1.476*** (−2.75)	0.097 (0.36)	−1.616*** (−3.03)	0.050 (0.19)	−1.524*** (−2.92)
<i>Prospectus</i>	−0.115 (−0.57)	−0.353** (−2.10)	−0.215 (−1.09)	−0.290 (−1.37)	−0.094 (−0.49)	−0.297 (−1.57)
<i>NewsIssue</i>	−0.032 (−0.54)	0.021 (0.28)	−0.027 (−0.51)	−0.017 (−0.21)	−0.037 (−0.69)	0.031 (0.38)
<i>Oversubscription</i>	−0.152*** (−3.04)	−0.119** (−2.05)	−0.187*** (−3.04)	−0.080 (−1.41)	−0.183*** (−3.34)	−0.104* (−1.66)
<i>Volume</i>	0.054** (2.44)	0.006 (0.46)	0.052*** (2.72)	0.007 (1.53)	0.053*** (2.92)	0.006 (1.45)
<i>MomentumIssue</i>	0.346 (0.72)	0.897 (1.31)	0.606 (1.40)	0.780 (1.06)	0.637 (1.41)	0.775 (0.98)
<i>Constant</i>	8.704*** (2.93)	13.371*** (4.90)	10.816*** (3.77)	11.224*** (3.91)	9.231*** (3.36)	12.483*** (4.50)
<i>Observations</i>	215	215	215	215	215	215
<i>Adj. R<sup>2</sup></i>	0.304	0.502	0.293	0.497	0.202	0.537

Notes. This table demonstrates the results of the split-sample design. The regressions include industry and calendar year dummies. *ComIssue* is in log form. The *t* statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

The significance of *ComIssue* in the split-sample design is not because of underlying information asymmetry but because investors rely more on information from online forums during the quiet period.

### 5.3.2. Effect of tone in online forum posts and news coverage

As investors' comments and opinions can have different effects on abnormal returns (Huang et al., 2018; Yang et al., 2020), we examine the influence of online comments with different sentiments using a split-sample test. To capture the sentiment information from the online forums, we construct *SentiPosts* as a proxy for forum tone, which is the difference between the number of positive and negative postings, scaled by the sum of positive and negative postings. Table 8 reports the coefficients of *ComIssue* for the observations with a *SentiPosts* value above and below its median value. To further support the conjecture regarding tone, we also construct a tone variable for media coverage, which is the difference between the number of positive and negative news articles scaled by the sum of positive and negative news articles. Table 9 reports the coefficients of *ComIssue* for the observations with a *SentiNews* value above and below its median value.

Table 9  
Online forum discussions and underpricing with different news sentiments.

Variable	(1) <i>SentiNews1</i> Above	(2) <i>SentiNews1</i> Below	(3) <i>SentiNews2</i> Above	(4) <i>SentiNews2</i> Below	(5) <i>SentiNews3</i> Above	(6) <i>SentiNews3</i> Below
<i>ComIssue1</i>	0.037 (1.39)	0.133** (2.05)				
<i>ComIssue2</i>			0.038 (1.27)	0.168** (2.18)		
<i>ComIssue3</i>					0.035 (1.08)	0.147** (2.13)
<i>Gap</i>	−0.093 (−0.99)	−0.328* (−1.83)	−0.109 (−1.07)	−0.335** (−2.07)	−0.148 (−1.18)	−0.341** (−2.16)
<i>IssueSize</i>	−0.405*** (−6.45)	−0.487*** (−2.84)	−0.414*** (−6.27)	−0.517*** (−3.17)	−0.372*** (−5.70)	−0.542*** (−3.47)
<i>Underwriter</i>	−0.069 (−1.39)	0.036 (0.45)	−0.061 (−1.16)	0.055 (0.64)	−0.047 (−0.91)	0.008 (0.10)
<i>EPS</i>	−0.000 (−0.01)	−0.211** (−2.18)	−0.052 (−1.13)	−0.162** (−1.98)	−0.062 (−1.16)	−0.099 (−1.31)
<i>Board</i>	0.101 (0.89)	−0.111 (−0.67)	0.050 (0.43)	−0.115 (−0.74)	0.165 (1.48)	−0.122 (−0.83)
<i>PriceRange</i>	0.254 (1.61)	0.982*** (3.04)	0.202 (1.15)	1.076*** (3.27)	0.419** (2.28)	0.713*** (2.73)
<i>Revision</i>	−0.338* (−1.66)	−1.423** (−2.35)	−0.365 (−1.55)	−1.460*** (−2.78)	−0.356 (−1.48)	−1.313** (−2.54)
<i>Prospectus</i>	−0.064 (−0.55)	−0.252 (−1.21)	−0.137 (−1.03)	−0.262 (−1.36)	−0.206 (−1.46)	−0.173 (−0.91)
<i>NewsIssue</i>	0.021 (0.43)	0.000 (0.00)	0.019 (0.36)	0.017 (0.15)	0.067 (0.98)	−0.082 (−0.63)
<i>Oversubscription</i>	−0.121*** (−3.06)	−0.168** (−2.00)	−0.126*** (−3.04)	−0.160** (−2.11)	−0.137*** (−3.54)	−0.146* (−1.91)
<i>Volume</i>	0.064*** (4.02)	−0.003 (−0.65)	0.053*** (3.46)	−0.004 (−0.87)	0.071*** (3.84)	−0.003 (−0.60)
<i>MomentumIssue</i>	0.468 (1.25)	1.731 (1.65)	0.428 (1.16)	1.615* (1.68)	0.267 (0.73)	1.356* (1.68)
<i>Constant</i>	8.759*** (5.48)	11.793*** (2.88)	9.673*** (5.21)	12.124*** (3.42)	9.157*** (4.90)	12.229*** (3.56)
<i>Observations</i>	228	202	220	210	218	212
<i>Adj. R<sup>2</sup></i>	0.358	0.439	0.336	0.450	0.340	0.458

Notes. This table shows the results for the split-sample design. The regressions include industry and calendar year dummies. *ComIssue* is in log form. The *t* statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

As shown in Tables 8 and 9, the coefficients of *ComIssue1*, *ComIssue2*, and *ComIssue3* are only positive and significant ( $p$ -value  $< 0.05$  or  $p$ -value  $< 0.01$ ) in columns (1), (3), and (5), respectively. This implies that the impact of stock forums is more significant for firms that are exposed to negative sentiment. Compared with the effects of news coverage, which is supposed to be objective, forum sentiment tends to be subjective. The results show that both objective and subjective sentiment environments show a negative impact from online forums under negative sentiment, whereas there is no significant effect under positive sentiment. In the literature, Veronesi (1999) and Epstein and Schneider (2008) show that the responses to positive and negative news are not necessarily symmetric. As investors' belief in undefeated new shares is deeply rooted, additional information has little influence. However, negative information ferments and spreads to a greater extent, giving online comments broader impact.

Table 10

Online forum discussions and initial returns using alternative independent variables.

Variable	(1)	(2)	(3)	(4)	(5)	(6)
<i>Posting1</i>	0.186*** (3.96)					
<i>Posting2</i>		0.222*** (3.71)				
<i>Posting3</i>			0.248*** (3.46)			
<i>Reading1</i>				0.189*** (3.80)		
<i>Reading2</i>					0.235*** (2.89)	
<i>Reading3</i>						0.271*** (2.80)
<i>Gap</i>	0.125 (1.21)	−0.013 (−0.11)	−0.124 (−0.88)	0.151 (1.51)	0.051 (0.52)	−0.116 (−0.85)
<i>IssueSize</i>	−0.436*** (−3.57)	−0.443*** (−3.58)	−0.450*** (−3.58)	−0.432*** (−3.53)	−0.438*** (−3.46)	−0.450*** (−3.49)
<i>Underwriter</i>	0.045 (0.84)	0.041 (0.79)	0.041 (0.78)	0.047 (0.87)	0.033 (0.65)	0.035 (0.69)
<i>EPS</i>	0.038 (1.24)	0.027 (0.85)	0.021 (0.64)	0.038 (1.18)	0.013 (0.34)	0.004 (0.09)
<i>Board</i>	−0.249*** (−3.70)	−0.208*** (−3.14)	−0.205*** (−3.13)	−0.238*** (−3.58)	−0.174** (−2.48)	−0.161** (−2.26)
<i>PriceRange</i>	0.517** (2.26)	0.485** (2.23)	0.465** (2.26)	0.547** (2.33)	0.516** (2.44)	0.509** (2.51)
<i>Revision</i>	−0.758 (−1.50)	−0.779 (−1.56)	−0.747 (−1.55)	−0.789 (−1.53)	−0.783 (−1.62)	−0.750 (−1.64)
<i>Prospectus</i>	−0.210* (−1.80)	−0.221* (−1.92)	−0.213* (−1.87)	−0.233* (−1.92)	−0.242** (−1.98)	−0.235* (−1.95)
<i>NewsListing</i>	−0.134** (−2.45)	−0.125** (−2.39)	−0.118** (−2.33)	−0.146** (−2.57)	−0.135** (−2.59)	−0.125** (−2.48)
<i>Oversubscription</i>	0.020 (0.68)	0.011 (0.37)	0.012 (0.37)	−0.003 (−0.08)	−0.032 (−0.65)	−0.040 (−0.75)
<i>Volume</i>	0.004 (1.25)	0.004 (1.35)	0.003 (1.15)	0.003 (0.87)	0.003 (1.05)	0.002 (0.85)
<i>MomentumList</i>	1.409*** (2.84)	1.648*** (3.57)	1.568*** (3.36)	1.279** (2.51)	1.688*** (3.86)	1.535*** (3.37)
<i>MomentumIssue</i>	0.554 (0.97)	0.303 (0.58)	0.413 (0.76)	0.638 (1.08)	0.357 (0.66)	0.540 (0.92)
<i>Constant</i>	9.473*** (3.17)	9.660*** (3.23)	9.817*** (3.28)	8.181*** (2.97)	7.885*** (3.20)	8.002*** (3.31)
<i>Observations</i>	430	430	430	430	430	430
<i>Adj. R<sup>2</sup></i>	0.444	0.452	0.459	0.435	0.450	0.462

Notes. This table shows the robustness test results for model (1). The regressions include industry and calendar year dummies. *Posting* and *Reading* are in log form. The  $t$  statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.



#### 5.4. Robustness tests

##### 5.4.1. Alternative proxies for online forum information

*Posting* and *Reading* are used as key independent variables to test the robustness of our results. *Posting* (*Posting1*, *Posting2*, and *Posting3*) is the number of forum articles posted within 7, 14, and 60 days before the listing date, respectively. *Reading* (*Reading1*, *Reading2*, and *Reading3*) is the number of times articles

Table 11

Online forum discussions and IPO valuation components using an alternative sample.

Variable	<i>Underpricing</i>			<i>Overpricing</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>ComIssue1</i>	0.153*** (2.65)					
<i>ComIssue2</i>		0.155** (2.61)				
<i>ComIssue3</i>			0.157*** (2.72)			
<i>ComList1</i>				0.025 (1.12)		
<i>ComList2</i>					0.046* (1.73)	
<i>ComList3</i>						0.043 (1.51)
<i>Gap</i>	−0.267 (−1.47)	−0.278 (−1.53)	−0.275 (−1.53)	0.195*** (2.84)	0.188*** (2.95)	0.161** (2.44)
<i>IssueSize</i>	−0.446** (−2.54)	−0.459** (−2.57)	−0.454** (−2.57)	−0.052 (−1.33)	−0.061 (−1.52)	−0.059 (−1.47)
<i>Underwriter</i>	0.096 (0.98)	0.095 (0.98)	0.095 (0.98)	0.008 (0.30)	0.006 (0.22)	0.007 (0.24)
<i>EPS</i>	0.006 (0.03)	0.012 (0.07)	0.008 (0.05)	0.104*** (4.11)	0.099*** (3.89)	0.100*** (3.92)
<i>Board</i>	−0.122 (−0.59)	−0.127 (−0.62)	−0.128 (−0.63)	−0.135** (−2.20)	−0.121* (−1.95)	−0.125** (−1.99)
<i>PriceRange</i>	1.035*** (2.82)	1.019*** (2.81)	1.033*** (2.83)	0.061 (0.52)	0.043 (0.38)	0.046 (0.40)
<i>Revision</i>	−1.710*** (−2.89)	−1.706*** (−2.88)	−1.708*** (−2.88)	−0.030 (−0.17)	−0.027 (−0.16)	−0.026 (−0.16)
<i>Prospectus</i>	−0.608** (−2.39)	−0.593** (−2.32)	−0.581** (−2.27)	−0.039 (−0.46)	−0.039 (−0.47)	−0.039 (−0.46)
<i>Oversubscription</i>	−0.129 (−1.47)	−0.132 (−1.48)	−0.131 (−1.48)	0.071*** (2.92)	0.059** (2.35)	0.063** (2.53)
<i>Volume</i>	0.001 (0.11)	0.001 (0.14)	0.000 (0.09)	−0.003 (−1.52)	−0.003* (−1.78)	−0.003* (−1.77)
<i>NewsListing</i>				−0.042 (−1.02)	−0.037 (−0.89)	−0.036 (−0.87)
<i>MomentumList</i>				0.962*** (3.62)	1.023*** (3.82)	1.010*** (3.79)
<i>NewsIssue</i>	−0.171* (−1.89)	−0.160* (−1.80)	−0.162* (−1.81)			
<i>MomentumIssue</i>	1.494 (1.58)	1.483 (1.55)	1.466 (1.55)			
<i>Constant</i>	13.870*** (2.89)	14.013*** (2.91)	13.797*** (2.91)	1.139 (1.18)	1.182 (1.25)	1.210 (1.27)
<i>Observations</i>	162	162	162	358	358	358
<i>Adj. R<sup>2</sup></i>	0.457	0.455	0.460	0.251	0.259	0.257

Notes. This table shows the results of the robustness test using models (4) and (5). The regressions include industry and calendar year dummies. *ComList* and *ComIssue* are in log form. The *t* statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

are read within 7, 14, and 60 days before the listing date, respectively. In Table 10, the coefficients of *Posting* and *Reading* are all positive and significant ( $p$ -value < 0.01), which is consistent with the results in Table 5.

#### 5.4.2. Alternative sample selection

To further ensure robustness and to analyze sensitivity, we use an alternative sample of under- and overpricing that excludes observations with negative under- or overpricing. Specifically, observations with a negative value for *Underpricing* or *Overpricing* are excluded. This reduces the sample size, but the results are consistent with our major findings. The regression results in Tables 11 and 12 are consistent with the results in Tables 6 to 9, confirming the robustness of our findings.

#### 5.4.3. Reverse causality

The empirical results imply that forums create noise and exacerbate information asymmetry during IPOs. An endogeneity concern for this study is reverse causality. Instead of the noise effect, the positive link between underpricing and forum discussions may also be explained by the following logic. Firms with less public information and greater information asymmetry may drive investors to online forums to gather information and comment on IPOs, leading to the positive coefficients of *ComIssue*. Therefore, we use media coverage as a proxy for the supply of public information and test the correlation between online discussions and media coverage in the same period. If a lack of public information induces more forum comments, postings, or readings,

Table 12  
Online forum discussions and underpricing using an alternative sample.

Variable	(1) <i>Quiet</i> = 1	(2) <i>Quiet</i> = 0	(3) <i>SentiPostsI</i> Above	(4) <i>SentiPostsI</i> Below	(5) <i>SentiNewsI</i> Above	(6) <i>SentiNewsI</i> Below
<i>ComIssueI</i>	0.145** (2.62)	0.064 (1.08)	−0.029 (−0.62)	0.209** (2.15)	0.012 (0.25)	0.240*** (3.11)
<i>Gap</i>	0.256 (1.28)	−0.394 (−1.66)	−0.180 (−0.73)	−0.267 (−1.09)	−0.236 (−1.47)	−0.622 (−1.26)
<i>IssueSize</i>	−0.839*** (−4.06)	−0.153 (−1.04)	−0.122 (−0.72)	−0.465 (−1.64)	−0.163 (−0.85)	−0.708*** (−3.46)
<i>Underwriter</i>	−0.021 (−0.18)	−0.085 (−0.83)	−0.132 (−1.37)	0.064 (0.34)	−0.014 (−0.18)	0.079 (0.46)
<i>EPS</i>	0.248 (1.25)	0.036 (0.15)	0.213 (0.88)	−0.134 (−0.55)	0.165 (0.84)	0.109 (0.39)
<i>Board</i>	1.173* (1.91)	0.026 (0.11)	−0.083 (−0.30)	0.107 (0.26)	0.257 (1.03)	−0.414 (−1.14)
<i>PriceRange</i>	0.746* (1.90)	0.457 (1.11)	−0.014 (−0.04)	1.505*** (3.01)	0.334 (0.60)	1.247 (1.58)
<i>Revision</i>	−0.682 (−1.12)	−0.523 (−1.27)	0.342 (1.05)	−2.850*** (−3.50)	0.147 (0.36)	−2.706*** (−2.98)
<i>Prospectus</i>	−0.897** (−2.06)	−0.282 (−0.86)	−0.181 (−0.62)	−0.608 (−1.42)	−0.175 (−0.41)	−0.835* (−2.04)
<i>NewsIssue</i>	−0.226* (−1.75)	−0.033 (−0.37)	0.011 (0.15)	−0.314* (−1.84)	−0.022 (−0.18)	−0.052 (−0.19)
<i>Oversubscription</i>	−0.169* (−1.82)	−0.007 (−0.08)	−0.075 (−1.25)	−0.197 (−1.14)	−0.049 (−0.56)	−0.290** (−2.11)
<i>Volume</i>	0.180*** (2.96)	0.004 (0.88)	0.020 (0.80)	0.002 (0.28)	0.041* (1.87)	0.001 (0.21)
<i>MomentumIssue</i>	3.167** (2.22)	0.164 (0.22)	−0.241 (−0.29)	2.511 (1.42)	−0.376 (−0.41)	4.218** (2.31)
<i>Constant</i>	20.995*** (4.36)	5.898 (1.20)	6.651 (1.35)	14.175* (1.74)	5.043 (0.88)	21.417*** (2.94)
<i>Observations</i>	94	68	78	84	83	79
<i>Adj. R<sup>2</sup></i>	0.703	0.276	0.389	0.618	−0.019	0.666

Notes. This table shows the results of the robustness test with the split-sample. The regressions include industry and calendar year dummies. *ComIssueI* is in log form. The  $t$  statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

the positive link between underpricing and forum discussions may be a result of reverse causality. Otherwise, the endogeneity concern is alleviated.

Table 13 reports the regression results for the relation between online discussions and media coverage. The number of news articles within 7, 14, and 60 days (*News1*, *News2*, and *News3*, respectively) before the listing date are positively and significantly associated with online forum comments, postings, and readings (*ComList*, *Posting*, and *Reading*, respectively) in the same period. The positive relationship between media coverage and forum discussions implies that less public information does not drive investors to online forums and thus alleviates the reverse causality concern.

#### 5.4.4. Difference-in-differences method

The results in Table 7 indicate that the effect of online forums on underpricing is more significant during the quiet period. Although regulators stop IPO firms from excessively promoting themselves with confusing and deceptive information, investors turn to the Internet and receive noisy information anyway. Hence, investors

Table 13  
Online forum discussions and media coverage.

Variable	(1) <i>ComList1</i>	(2) <i>ComList2</i>	(3) <i>ComList3</i>	(4) <i>Posting1</i>	(5) <i>Posting2</i>	(6) <i>Posting3</i>	(7) <i>Reading1</i>	(8) <i>Reading2</i>	(9) <i>Reading3</i>
<i>News1</i>	0.187*** (4.06)			0.182*** (4.24)			0.156*** (3.97)		
<i>News2</i>		0.261*** (4.38)			0.270*** (5.58)			0.233*** (4.83)	
<i>News3</i>			0.258*** (4.40)			0.268*** (5.95)			0.221*** (4.79)
<i>Gap</i>	−1.095*** (−6.39)	−0.309* (−1.95)	0.052 (0.35)	−1.072*** (−7.07)	−0.100 (−0.70)	0.116 (0.97)	−1.202*** (−7.84)	−0.383*** (−2.79)	0.081 (0.67)
<i>IssueSize</i>	0.451*** (5.22)	0.362*** (3.95)	0.393*** (4.35)	0.377*** (4.65)	0.329*** (4.34)	0.331*** (4.62)	0.353*** (5.07)	0.288*** (4.06)	0.305*** (4.44)
<i>Underwriter</i>	−0.119 (−1.63)	−0.034 (−0.49)	−0.027 (−0.40)	−0.086 (−1.29)	−0.060 (−0.99)	−0.048 (−0.83)	−0.097 (−1.60)	−0.020 (−0.36)	−0.022 (−0.41)
<i>EPS</i>	0.260*** (3.21)	0.257*** (3.76)	0.247*** (3.73)	0.198*** (2.77)	0.210*** (3.22)	0.209*** (3.30)	0.203*** (3.71)	0.260*** (5.13)	0.258*** (5.41)
<i>Board</i>	−0.556*** (−4.10)	−0.647*** (−5.07)	−0.526*** (−4.26)	−0.538*** (−4.11)	−0.628*** (−5.25)	−0.529*** (−4.59)	−0.593*** (−5.40)	−0.742*** (−7.27)	−0.653*** (−6.69)
<i>PriceRange</i>	0.352 (1.45)	0.402* (1.77)	0.386* (1.73)	0.369* (1.75)	0.403** (2.15)	0.444** (2.53)	0.220 (1.10)	0.259 (1.39)	0.256 (1.45)
<i>Revision</i>	−0.636** (−2.15)	−0.475 (−1.39)	−0.608* (−1.75)	−0.658** (−2.42)	−0.399 (−1.54)	−0.574** (−2.35)	−0.472* (−1.89)	−0.355 (−1.30)	−0.503* (−1.82)
<i>Prospectus</i>	0.026 (0.12)	0.067 (0.33)	−0.055 (−0.28)	−0.015 (−0.07)	0.039 (0.23)	−0.056 (−0.34)	0.106 (0.53)	0.126 (0.75)	0.036 (0.23)
<i>NewsListing</i>	−0.133 (−1.28)	−0.111 (−1.13)	−0.110 (−1.24)	−0.156* (−1.75)	−0.133 (−1.59)	−0.119 (−1.54)	−0.105 (−1.23)	−0.094 (−1.15)	−0.096 (−1.29)
<i>Oversubscription</i>	0.531*** (8.73)	0.508*** (8.82)	0.472*** (8.45)	0.372*** (6.88)	0.338*** (6.90)	0.305*** (6.49)	0.486*** (9.87)	0.505*** (11.43)	0.471*** (11.07)
<i>Volume</i>	0.023*** (5.63)	0.018*** (5.57)	0.019*** (5.87)	0.021*** (6.03)	0.014*** (5.31)	0.016*** (6.13)	0.027*** (8.30)	0.019*** (6.97)	0.019*** (6.92)
<i>MomentumList</i>	0.275 (0.32)	−1.182 (−1.52)	−0.581 (−0.78)	0.486 (0.58)	−0.521 (−0.71)	−0.157 (−0.23)	1.169 (1.61)	−0.669 (−1.07)	−0.030 (−0.05)
<i>MomentumIssue</i>	−2.412*** (−2.99)	−0.996 (−1.36)	−1.776** (−2.53)	−2.044*** (−2.90)	−0.708 (−1.17)	−1.282** (−2.29)	−2.491*** (−3.73)	−0.901 (−1.53)	−1.623*** (−2.92)
<i>Constant</i>	−2.558 (−1.02)	−2.189 (−0.85)	−2.607 (−1.05)	−1.892 (−0.82)	−3.090 (−1.47)	−2.837 (−1.45)	4.977** (2.28)	4.682** (2.23)	4.103** (2.04)
<i>Observations</i>	430	430	430	430	430	430	430	430	430
<i>Adj. R<sup>2</sup></i>	0.472	0.491	0.502	0.463	0.498	0.541	0.558	0.591	0.607

Notes. This table shows the results for the relationship between online discussions and media coverage. The regressions include industry and calendar year dummies. *News*, *ComList*, *Posting*, and *Reading* are in log form. The *t* statistics are reported in parentheses and are based on the robust standard errors. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

Table 14  
Difference-in-differences estimation results.

Panel A <i>ComIssue</i>					
Groups		<i>Underpricing</i>	S. Err.	t	P >  t
Before	Control	9.875			
	Treated	9.949			
	Diff (T-C)	0.074	0.059	1.24	0.214
After	Control	9.785			
	Treated	9.999			
	Diff (T-C)	0.215	0.052	4.10	0.000***
	Diff-in-Diff	0.141	0.077	1.84	0.067*
Panel B <i>Reading</i>					
Groups		<i>Underpricing</i>	S. Err.	t	P >  t
Before	Control	9.504			
	Treated	9.555			
	Diff (T-C)	0.051	0.060	0.84	0.401
After	Control	9.411			
	Treated	9.622			
	Diff (T-C)	0.211	0.053	4.00	0.000***
	Diff-in-Diff	0.160	0.077	2.09	0.038**

Notes. This table demonstrates the results of the difference-in-differences estimation. \*\*\*, \*\*, and \* denote significance at the 1%, 5%, and 10% confidence levels, respectively.

are more influenced by online forums during the quiet period due to the lack of information from IPO firms. Therefore, investors are confused either by IPO firms or by online forums. As the quiet period tends to induce greater dependence on unofficial information sources, such as online forums, we view the CSRC quiet period regulation as a shock to the effect of forum discussions on underpricing.

Table 14 presents the results of our difference-in-differences estimation. In Panel A, the treated group consists of observations with a *ComIssue* value (number of comments within 60 days before the offer price declaration date) greater than its median, and the untreated group consists of observations with a *ComIssue* value less than or equal to its median. In Panel B, the treated and untreated groups are constructed in the same way using *Reading* (number of times articles are read within 60 days before the offer price declaration date).

After the shock from *Quiet*, the treated groups have significantly higher values than the untreated groups, which is consistent with the finding that forum discussions increase the likelihood of underpricing. Table 14 shows that the difference-in-differences values are positive (0.067 and 0.038) and significant ( $p$ -value < 0.1 and  $p$ -value < 0.05, respectively). We interpret the results to imply that the CSRC's quiet period regulation induces greater dependence on online forums and that forum discussions create noise that increases information asymmetry and exacerbates IPO underpricing.

## 6. Conclusion

In this study, we use an IPO setting to examine the effect of online forums on information asymmetry. It is difficult to test whether online forums support information transparency under normal circumstances, but the effect of forum postings during IPOs can be easily tested because the level of IPO underpricing can be considered a proxy for information asymmetry. Hence, it is easier to test the influence of online discussions in an IPO setting, especially in China where underpricing was once fairly high.

Using both rational and behavioral frameworks, IPO initial returns are decomposed into their under- and overpricing components to test the influence of online stock forum discussions on IPO pricing. We find a positive relationship between the number of forum comments and underpricing, indicating that online forums produce noise and exacerbate information asymmetry during IPOs. The quiet period regulation causes inves-

tors to rely on forums to gather and exchange information, which amplifies the effect of online forums on information asymmetry. In addition, we examine the influence of online comments with different sentiment levels using split-sample tests. The results indicate a more significant negative impact of online forums under negative sentiment.

The findings of this study have important theoretical and policy implications. First, we extend the literature on the relationship between Internet information and IPO pricing. We provide a clearer picture of this relationship by separating under- and overpricing in a Chinese setting, which is ideal due to the prevalence of relatively high initial returns for Chinese IPOs. Second, the evidence of noise from forums advises investors not to put too much faith in stock forum discussions. According to the evidence, we find that in an IPO setting, online forums negatively affect information asymmetry. Third, the implication for the regulatory authority is that channels for firms to honestly voice against rumors and fake news during IPOs are necessary to alleviate information asymmetry.

## Funding

This work was supported by the “111” project funded by the Ministry of Education of China and the State Administration of Foreign Experts Affairs of China [grant number B18043].

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

- Antweiler, W., Frank, M.Z., 2004. Is all that talk just noise? The information content of Internet stock message boards. *J. Finance* 59 (3), 1259–1294.
- Bajo, E., Raimondo, C., 2017. Media sentiment and IPO underpricing. *J. Corp. Finance* 46, 139–153.
- Beatty, R.P., Ritter, J.R., 1986. Investment banking, reputation, and the underpricing of initial public offerings. *J. Financ. Econ.* 15 (1–2), 213–232.
- Beatty, R.P., Welch, I., 1996. Issuer expenses and legal liability in initial public offerings. *J. Law Econ.* 39 (2), 545–602.
- Benveniste, L.M., Spindt, P.A., 1989. How investment bankers determine the offer price and allocation of new issues. *J. Financ. Econ.* 24 (2), 343–361.
- Benveniste, L.M., Wilhelm, W.J., 1990. A comparative analysis of IPO proceeds under alternative regulatory environments. *J. Financ. Econ.* 28 (1–2), 173–207.
- Bushee, B.J., Core, J.E., Guay, W., Hamm, S.J., 2010. The role of the business press as an information intermediary. *J. Account. Res.* 48 (1), 1–19.
- Carter, R., Manaster, S., 1990. Initial public offerings and underwriter reputation. *J. Finance* 45 (4), 1045–1067.
- Chen, G., Firth, M., Kim, J.B., 2004. IPO underpricing in China's new stock markets. *J. Multinational Financ. Manage.* 14 (3), 283–302.
- Cornelli, F., Goldreich, D., Ljungqvist, A., 2006. Investor sentiment and pre-IPO markets. *J. Finance* 61 (3), 1187–1216.
- Das, S.R., Sisk, J., 2005. Financial communities. *J. Portfolio Manage.* 31 (4), 112–123.
- Delort, J.Y., Arunasalam, B., Milosavljevic, M., Leung, H., 2011. The impact of manipulation in Internet stock message boards. *Int. J. Bank. Finance* 8 (4), 1–18.
- Derrien, F., 2005. IPO pricing in “hot” market conditions: Who leaves money on the table? *J. Finance* 60 (1), 487–521.
- Dong, D., Wu, K., 2019. Investor attention is a risk pricing factor? Evidence from Chinese investors for self-selected stocks. *China Finance Rev. Int.* 10 (1), 95–112.
- Dorn, D., 2009. Does sentiment drive the retail demand for IPOs? *J. Financ. Quant. Anal.* 44 (1), 85–108.
- Epstein, L.G., Schneider, M., 2008. Ambiguity, information quality, and asset pricing. *J. Finance* 63 (1), 197–228.
- Gao, Y., 2010. What comprises IPO initial returns: evidence from the Chinese market. *Pacific-Basin Finance J.* 18 (1), 77–89.
- Hanley, K.W., 1993. The underpricing of initial public offerings and the partial adjustment phenomenon. *J. Financ. Econ.* 34 (2), 231–250.
- Hanley, K.W., Hoberg, G., 2010. The information content of IPO prospectuses. *Rev. Financ. Stud.* 23 (7), 2821–2864.
- Huang, H., Li, Y., Zhang, Y., 2018. Investors' attention and overpricing of IPO: An empirical study on China's growth enterprise market. *IseB* 16 (4), 761–774.
- Huang, Y., Zhang, H., 2020. Does individual investors' attention influence underwriters' IPO pricing? *Appl. Econ.* 52 (34), 3680–3687.
- Jiang, X., Akbar, A., 2018. Does increased representation of female executives improve corporate environmental investment? Evidence from China. *Sustainability* 10 (12), 4750.

- Jiang, F., Cai, W., Wang, X., Zhu, B., 2018. Multiple large shareholders and corporate investment: Evidence from China. *J. Corp. Finance* 50, 66–83.
- Loughran, T., McDonald, B., 2013. IPO first-day returns, offer price revisions, volatility, and form S-1 language. *J. Financ. Econ.* 109 (2), 307–326.
- Loughran, T., Ritter, J.R., 1995. The new issues puzzle. *J. Finance* 50 (1), 23–51.
- Loughran, T., Ritter, J.R., 2002. Why don't issuers get upset about leaving money on the table in IPOs? *Rev. Financ. Stud.* 15 (2), 413–444.
- Loughran, T., Ritter, J.R., Rydqvist, K., 1994. Initial public offerings: International insights. *Pacific-Basin Finance J.* 2 (2–3), 165–199.
- Mezhoud, M., Boubaker, A., 2011. Determinants of the components of IPO initial returns: Paris Stock Exchange. *Int. J. Account. Financ. Rep.* 1 (1), 190–211.
- Miller, E.M., 1977. Risk, uncertainty, and divergence of opinion. *J. Finance* 32 (4), 1151–1168.
- Mumtaz, M.Z., Smith, Z.A., Maqsood, A., 2016. An examination of short-run performance of IPOs using extreme bounds analysis. *Estudios de Economía* 43 (1), 71–95.
- Ofek, E., Richardson, M., 2003. Dotcom mania: the rise and fall of Internet stock prices. *J. Finance* 58 (3), 1113–1137.
- Purnanandam, A.K., Swaminathan, B., 2004. Are IPOs really underpriced? *Rev. Financ. Stud.* 17 (3), 811–848.
- Ritter, J.R., 1991. The long-run performance of initial public offerings. *J. Finance* 46 (1), 3–27.
- Ritter, J.R., Welch, I., 2002. A review of IPO activity, pricing, and allocations. *J. Finance* 57 (4), 1795–1828.
- Rock, K., 1986. Why new issues are underpriced. *J. Financ. Econ.* 15 (1–2), 187–212.
- Shaw, E.F., 1979. Agenda-setting and mass communication theory. *Gazette (Leiden, Netherlands)* 25 (2), 96–105.
- Shen, Z., Coakley, J., Instefford, N., 2014. Earnings management and IPO anomalies in China. *Rev. Quant. Financ. Acc.* 42 (1), 69–93.
- Song, S., Tan, J., Yi, Y., 2014. IPO initial returns in China: underpricing or overvaluation? *China J. Account. Res.* 7 (1), 31–49.
- Spiegel, U., Tavor, T., Templeman, J., 2010. The effects of rumours on financial market efficiency. *Appl. Econ. Lett.* 17 (15), 1461–1464.
- Tang, G., Li, L., 2013. Structure and distribution characteristics of corporate environmental investment: empirical evidence from A-share listed companies. *J. Audit Econ.* 28, 94–103 (In Chinese).
- Veronesi, P., 1999. Stock market overreactions to bad news in good times: a rational expectations equilibrium model. *Rev. Financ. Stud.* 12 (5), 975–1007.
- Wysocki, P.D., 1999. Cheap talk on the web: The determinants of postings on stock message boards. Working paper. University of Michigan.
- Yang, X., Zhu, Y., Cheng, T.Y., 2020. How the individual investors took on big data: The effect of panic from the Internet stock message boards on stock price crash. *Pacific-Basin Finance J.* 59 101245.
- Zheng, S.X., 2007. Are IPOs really overpriced? *J. Empirical Finance* 14 (3), 287–309.