The Value of Social Media Anonymity: Evidence from the Stock Market*

Huasheng Gao
International School of Finance
Fudan University
huashenggao@fudan.edu.cn

Kai Li
Sauder School of Business
University of British Columbia
kai.li@sauder.ubc.ca

Yongxian Tan
School of Accounting, Economics and Finance
Curtin University
yongxian.tan@curtin.edu.au

Hailong Zhao
School of Accounting
Southwestern University of Finance and Economics
zhaohl@swufe.edu.cn

This version: June 2024

Abstract: This study examines the value implications of social media anonymity in a stock market context, exploiting China's regulatory change that removes anonymity and requires real-name registration on social media. We find that this policy has led to a significant decrease in firm value, especially when investors rely more on social media for information. Moreover, it has resulted in poorer information dissemination on stock message boards, lower stock price informativeness and liquidity, higher stock crash risk, deteriorated corporate disclosure, and increased earnings manipulation. Overall, our findings indicate that removing social media anonymity decreases firm value by worsening firms' information environment.

Key Words: Firm value; Social media anonymity; Stock message board; Information environment

JEL: G32; G38; K24

---

* We are grateful for the helpful comments by Sumit Agarwal, Simba Xin Chang, Jason Huafeng Chen, Zhiguo He, Sheng Huang, Feng Jiang, Ron Kaniel, Dan Li, Ting Luo, Yiqing Lv, Jun Pan, Jun Qian, Jiangping Qiu, Zhongzhi Song, Nancy Su, Dragon Tang, Jean Tirole, Jun Tu, Frank Yu, Renxuan Wang, Yuxi Wang, Qi Zhang, and the seminar participants from CEIBS, NYU Shanghai, Shanghai Jiaotong University, the First Fanhai-Haitong Finance Workshop, the Fifth Chinese Accounting Scholar Workshop, and the First Fudan-TSE Economics and Finance Workshop. An earlier version of this paper was circulated under the title of "Social Media Anonymity and Stock Price Crash Risk." All errors are our own.

1. Introduction

Over the past two decades, an intense debate has raged regarding anonymity on social media. One side argues that anonymity should be eliminated because it facilitates abuse, illegal activity, and the spread of fake news on the Internet. The other side believes that anonymity is vital for the protection of freedom of expression and helps to achieve better information dissemination (Lidsky and Cotter, 2007; Moore, 2009; Choi, 2012; Kaminski, 2012; Lee and Liu, 2016; Huang, Li and Markov, 2020). Despite the numerous pros and cons of Internet anonymity, quantitative evidence identifying its impact on society is lacking, possibly because it is difficult to quantify. This study examines the impact by focusing on stock markets, which provide a good setting because (1) richer data in this market help to quantify any value implications reasonably accurately, and (2) investors are increasingly relying on social media for financial information (Antweiler and Frank, 2004; Das and Chen, 2007; Chen et al., 2014; Ang et al., 2021).

Ex ante, the effect of social media anonymity on firm value can go either way. Anonymity may encourage the spread of false information because of the difficulty in holding individuals accountable for anonymous posts. False news can make firm environments opaque and subject investors to greater uncertainty, which could in turn decrease firm value. Conversely, people are more likely to share important but sensitive information when they are free of privacy concerns and when they have improved protection from detection, retribution, and embarrassment. Anonymous communication can facilitate the flow of information by protecting privacy and freedom of expression. Furthermore, free communication of information, especially negative information, on social media can prompt firm managers to increase their disclosure and prevent them from hiding and accumulating negative information (Xu and Zhang, 2013; Cade, 2018; Dube and Zhu, 2021). Thus, anonymity on social media can enrich firms' information environments for investors, enhancing firm value.

In this study, we identify the effect of social media anonymity on firm value based on the exogenous shock to social media anonymity generated by China's regulatory change toward real-name registration. In February 2015, the Cyberspace Administration of China (CAC) issued Provisions on the Administration of Account Names of Internet Users, under which Internet information service providers must ensure that their users register accounts after undergoing identity information authentication; this effectively imposes real-name policies on all Internet services in China. In this setting, this is the first empirical study on the value implications of social media anonymity by examining the impact of removing anonymity from Internet stock message boards on firm value.

We focus on China where regulatory changes provide us with a unique empirical setting for four reasons. First, the government sought to tighten its control of the Internet via strengthened real-name account registration regulations. This policy change did not aim to affect the stock market; therefore, any potential effect on firm value is likely an unintended consequence. Second, because all domestic Internet service providers are required to adopt real-name policies, Chinese users of Internet stock message boards cannot switch to other anonymous platforms to release sensitive information. Third, the Chinese stock market is characterized by high retail investor participation (Carpenter, Lu and Whitelaw, 2015; Liu, Stambaugh and Yuan, 2019). This enhances the power of our tests, because retail investors are more likely to rely on social media for information. Fourth, similar settings are not readily available in other countries. For example, the U.S. does not have a nation- or state-wide real-name registration regulation, and very few mainstream social media platforms have adopted a genuine real-name system. When platforms require such registration, users can easily switch to others.[1]

We conduct our analysis in a difference-in-differences (DiD) framework using posts data from Eastmoney Guba. As the first and most popular stock message board in China (e.g., Hong et al., 2014; Huang, Qiu and Wu, 2016; Ang et al., 2021), Guba hosts a firm-specific stock message board for each firm listed on the Shanghai and Shenzhen Stock Exchanges. Considering that firms with more pre-existing Guba posts should be more affected by the real-name policy, we use the number of Guba posts for each sample firm in 2014 (the year immediately before the policy change) as a measure of treatment intensity.

Using a panel of 14,770 firm-year observations from 2010 to 2019, we find that after the real-name policy, firms with higher treatment intensity experience significantly larger reductions in firm value (measured by Tobin's Q) than firms with lower treatment intensity. On average, an increase in treatment intensity by one standard deviation (SD) leads to a 5.02% reduction in firm value, which translates to RMB 667 million (approximately USD 95 million). Our inference is largely unchanged when we implement propensity score matching (PSM) based on observable firm characteristics, use an instrumental variable approach to address any possible omitted variables associated with a firm's Guba posts, or apply an event study around the policy change announcement.

We further investigate the channels through which social media anonymity affects firm value. First, we examine the cross-sectional variation in the treatment effect. If real-name registration

[1] For example, Facebook and Quora have real-name policies, but it is practically impossible for them to force all users to provide identification information. They enforce their policies by merely using computer algorithms to flag suspicious accounts. A user seeking to communicate sensitive information anonymously may still create accounts with a fake name and any email address. Moreover, a Facebook user can choose to release sensitive information about a stock anonymously on Twitter.

decreases firm value by worsening investors' information environments, the impact should be more pronounced in situations where investors rely more on social media for information. Consistent with this conjecture, we find that the treatment effect is more pronounced in firms with less analyst and newspaper coverage, lower institutional ownership, and greater local usage of mobile Internet.

Next, based on a textual analysis of each post in Guba, we show that real-name registration leads to a significant reduction in the number, length, originality, and quantifiability of Guba posts. This evidence indicates that the removal of user anonymity decreases both the quantity and quality of information exchange on social media, reinforcing support for the relevance condition of this policy change. We also observe a greater decrease in negative posts compared to positive posts, consistent with the view that sharing negative information is potentially riskier, and anonymity is particularly effective in protecting posters (Froomkin, 1995; Lidsky and Cotter, 2007). We find a decrease in the predictive power of Guba posts for negative earnings surprises, and an increase in the magnitude of negative stock market reactions at the announcement of such surprises.

Similarly, we discover that the treatment policy leads to lower stock price informativeness and liquidity, higher stock crash risk, poorer corporate disclosure, greater real earnings management, and more financial restatements. Taken together, these findings support the mechanism by which real name registration decreases firm value—worsening firms' information environment faced by investors.

We investigate two alternative channels for our main findings. The first is an overvaluation channel. Firms with more Guba posts in the pre-event period were overvalued and thus experience a greater decline in firm value in the post-event period. We show that our inference remains unchanged when we measure treatment intensity based on the number or fraction of negative posts (rather than the total number of posts). As firms with more negative posts are less likely to be overvalued (rather, undervalued), the overvaluation channel is unlikely to hold true. The second channel is investor attention: Real-name registration deters information exchange on social media, decreases investor attention paid to the stock, and thus leads to a decline in firm valuation. However, we find no significant decline in investor attention following the treatment, and this attention-based channel too is unlikely to explain our main results.

This study contributes to literature on online anonymity, a problem that can be traced back to the early days of the Internet. In November 1997, the American Association for the Advancement of Science organized a conference on anonymous Internet communication. While conference participants agreed that such communication is a constitutional right protected by the First Amendment, they acknowledged that online anonymity offers opportunities and risks related to how people exchange information. Existing evidence on the net effect of online anonymity on society is limited, possibly due to the challenge of quantifying these effects in various social contexts. We shed light on this issue by focusing on financial markets, in which stock information can be used to identify possible value implications. To our knowledge, this provides the first empirical evidence that the removal of social media anonymity imposes significant costs on the financial market by inhibiting free communication of information.

Furthermore, our study extends the growing field of research that examines the role of social media in financial markets. Previous studies have shown that social media influences investors' trading behavior (Hong et al., 2014), market volatility (Antweiler and Frank, 2004), future firm performance (Chen et al., 2014; Bartov, Faurel and Mohanram, 2018; Tang, 2018), and corporate information dissemination (Blankespoor, Miller and White, 2014; Miller and Skinner, 2015). Our study complements extant research by providing evidence that anonymity is crucial to social media for harnessing the wisdom of crowds to uncover valuable information.

2. Institutional Background

The Chinese government's earliest attempt to remove social media anonymity was made in 2002 when it ordered Internet cafés to collect customer identification information (Gao and Jiang, 2016)[2] to fight crime and protect minors. In the following decade, lawmakers, government agencies, and industry associations adopted more policy initiatives toward a real-name identification system. In 2004, the Ministry of Education required all university-affiliated Internet message boards to strictly enforce real-name registration policies (Feng, Pan and Qian, 2014).[3] In 2007, Internet companies in China signed the China Internet Industry Self-Discipline Act, which strongly encouraged bloggers' real-name identification.[4] In 2012, to comply with the Beijing Municipal Government's new regulatory requirements, China's major microblog companies, such as Sina, Sohu, NetEase, and Tencent, announced they would implement real-name policies.[5] The 2013 institutional reform plan of the State Council of China required the Ministry of Industry and Information Technology, CAC, and Ministry of Public Security to implement a real-name registration system (State Council of China, 2013).

---

[2] According to Article 23 of the "Regulations on the Administration of Internet Access Service Business Places" issued by the State Council of China in September 2002, Internet service providers are required to verify, register, and record the identification cards or other valid certificates of Internet access customers. In November 2002, all Internet cafés began to require customers to provide their ID cards for registration.

[3] On December 28, 2004, the Ministry of Education announced that it would strengthen campus network management in universities and require that the campus network be "real-name based" for users. In 2005, major Chinese universities, such as Tsinghua University, Peking University, and Fudan University, started to implement the real-name system for their on-campus Internet service.

[4] China Internet Industry Self-Discipline Act, https://www.isc.org.cn/article/15537.html.

[5] https://fzsd124.com/index/index/mob_show/id/8502.html.

After these initial preparations, the real-name policy was announced on February 4, 2015, and took effect on March 1, 2015.[6] The policy requires Internet information service providers to demand that users register their accounts after undergoing identity information authentication. If a user fails to provide identity information, the service is stopped. This has effectively imposed real-name policies on all Internet services in China, including Internet message boards. The CAC also initiated a series of "regulatory talks" to supervise online service providers and urge them to rectify their unlawful practices. Social media platforms validate users' identities by combining their social media accounts with their mobile phone numbers, as Chinese mobile phone users must provide personal IDs to telecommunications operators when purchasing a mobile phone SIM card. This authentication is further accompanied by a facial recognition process that matches individuals' faces with their ID photos so that a user cannot register under another ID. Once verified, all personal information, such as ID number, name, mobile number, and other details, is linked to the social media account.[7]

The 2015 regulatory change provides a unique opportunity to examine the effect of real-name registration on firm value because it has the following desirable features: Regulatory changes were initiated for reasons unrelated to stock market performance. The stated purposes of the new policy include "strengthening the administration of account names of Internet users and protecting the lawful rights and interests of citizens." Analyzing Chinese blog entries regarding real-name registration, Yang and Kang (2016) find freedom, privacy, and deterrence of cybercrime to be the three broad themes.

[6] The official announcement (in Chinese) can be found at https://www.cac.gov.cn/2015-02/04/c_1114246561.htm. This policy is also widely reported in the international media, such as CNN, BBC, Reuters, and Bloomberg (e.g., https://www.reuters.com/article/china-internet-censorship/china-to-ban-online-impersonation-accounts-enforce-real-name-registration-idINKBN0L811520150204/).

[7] Prior to the real-name policy, the traditional method to track an Internet user was to identify the geographic location of the computer based on the user's IP address. This is inadequate for precise user identification, for reasons such as network equipment shared between multiple users, dynamic IP addresses, proxy servers, and mobile devices.

We extensively searched for newspaper articles on real-name registration in China, but none cited financial markets as a major concern. Therefore, the real-name policy generates a plausible exogenous shock to the information exchange on social media for stock markets.

Moreover, as all Internet service providers in China must adopt real-name policies, a social media platform cannot choose whether to comply and Internet users cannot switch to other anonymous platforms. This condition ensures a more accurate assessment of the impact of real-name registration without confounding biases stemming from the voluntary adoption or evasion of the policy.

The Chinese stock market is characterized by high levels of retail investor participation. According to the annual report of the Shanghai Stock Exchange in 2016,[8] retail investors comprised 99.8% of total investor accounts, held 42.87% of shares outstanding, and contributed to approximately 86.9% of the yearly trading volume.[9] This enhances the power of our test, because retail investors are more likely to rely on social media for information.

When we examine the stock market reaction toward the policy announcement made on February 4, 2015 (details in Subsection 5.5), we find that it causes a dramatically negative market reaction and that the cross-sectional variation of the firm-level stock reaction is related to firms' exposure to social media. These findings provide supporting evidence for the significance of the real-name policy for the stock market.

Anecdote evidence indicates that corporate managers indeed attempt to retaliate netizens for spreading negative information about their firms. For example, in 2022 the Cloud Live Technology

[8] https://www.sse.com.cn/aboutus/publication/yearly/documents/c/tjnj_2016.pdf.

[9] In comparison, Brav, Cain and Zytnick (2022) find that individual investors hold 26% of all shares outstanding in the U.S. stock market.

Group (a public firm listed in the Shenzhen Stock Exchange) sued three Guba users, alleging that they fabricated facts and damaged its reputation. The court ultimately ruled against the company's claims, because the three users simply expressed their disappointment about the firm's performance.[10] In 2017, Dr. Qindong Tan, a physician in China, posted on social media questioning the "toxicity" of Hongmao Liquor, a well-known Chinese medicinal product owned by Jinyu Group, a publicly listed company on the Shanghai Stock Exchange. This post caused an uproar on Chinese social media, leading the liquor company to file an infringement lawsuit against him. In January 2018, based on Dr. Tan's real-name-registered ID, the police arrested him on charges of "damaging the reputation of a company." After being detained for more than three months, the company eventually withdrew their complaint and the infringement lawsuit in May 2018, following a formal apology from Dr. Tan's wife.[11] These cases highlight the litigation risk faced by real-name-registered netizens from posting negative information about firms.

3. Hypothesis Development

Information plays a crucial role in allocating resources to capital market economies. All types of investors, from sophisticated to small, struggle to acquire and analyze information to evaluate the return potential of investment opportunities and monitor the use of their capital resources once they have been allocated. Explicit information processing costs and implicit opportunity costs affect investors' information choices, trade, and market outcomes (Blankespoor, deHaan and Marinovic, 2020). Owing to their lack of ability and resources to collect and analyze firm information, individual investors usually face information disadvantages compared to corporate insiders and professional

[10] https://stcn.com/article/detail/944236.html

[11] https://www.ft.com/content/4528da22-4237-11e8-803a-295c97e6fd0b

investors (La Porta et al., 2000; Bhattacharya et al., 2007; Malmendier and Shanthikumar, 2007). The growth of Internet stock message boards has enabled them to exchange information and express opinions effectively (Antweiler and Frank, 2004; Das and Chen, 2007). A priori, however, how the anonymity in stock message boards affects firm value can go either way.

Anonymity on the message board can increase firm value by improving the information environment for investors for three reasons. First, anonymity plays a vital role in facilitating information exchange on social media, irrespective of whether the information is positive or negative about a firm. As Miller and Skinner (2015) note, this phenomenon weakens the control firms typically wield over their information environments. Information is typically exchanged both interactively and collaboratively on social media platforms, in which the participation of individual contributors is crucial to gain the "wisdom of crowds" (Surowiecki, 2004). The provision of anonymity enables individuals to share information that they might otherwise hesitate to disclose under their real identity owing to social, professional, or cultural constraints. It helps lower participation barriers and attracts a larger and more diverse group of individuals, such as investors, employees, suppliers, and competitors, to stock message boards. Thus, anonymity can potentially increase the volume of discussion, foster a more active and engaged community, and mitigate biases in the information environment.

Second, anonymity helps reveal negative information by shielding whistleblowers or activists who want to expose wrongdoing or raise awareness of the negative aspects of firms without risking personal safety or retaliation. As corporate disclosures tend to be biased toward good news, the timely dissemination of potentially negative information is particularly valuable for individual investors (Verrecchia, 1983; Jorgensen and Kirschenheiter, 2003; Acharya, DeMarzo and Kremer, 2011). This view is broadly consistent with that of previous studies. For example, Lee and Kim (2020) show that employees are likely to be critical of their companies, especially on anonymous social media platforms. Huang, Li and Markov (2020) find that employee social media disclosures, which can strongly predict bad corporate events, are more important as a source of bad news than good news. According to Ang et al. (2021), social media criticism contains incremental value-relevant information that can predict if a potential buyer would withdraw their acquisition attempt.

Third, improved information exchanges enabled by anonymity can encourage managers to provide better corporate disclosures. Corporate managers generally have incentives to withhold unfavorable news, and their voluntary disclosure decisions are influenced by shareholders' access to information (Beyer et al., 2010). When there are limited ways for outside investors to access firm-specific information, self-interested managers are more inclined to release positive news promptly while obfuscating and withholding negative news (Healy and Palepu, 2001; Li, 2008; Hutton, Marcus and Tehranian, 2009; Kothari, Shu and Wysocki, 2009) because this strategy buys managers time for more favorable developments or allows them to adjust accounting measures (Dye, 1985, 2001). Active information sharing on social media limits the room for information manipulation and increases managers' risk of detection or litigation. This discourages managers from hoarding negative information or delaying disclosure (Skinner, 1994, 1997).

Collectively, therefore, anonymity fosters a more favorable information environment for investors, consequently reducing their required returns on investments and enhancing firm value, leading to our first hypothesis (the information environment hypothesis):

H1a. Elimination of anonymity in the stock message board decreases firm value.

However, anonymity in stock message boards comes with risks, mainly the potential dissemination of misinformation. If not promptly corrected by other participants, false information can mislead investors and significantly diminish a platform's overall reliability. Anonymous users of stock message boards can avoid reputational damage or legal responsibility by changing their online identities. Existing literature suggests that anonymous users may post deceptive messages to gain at the expense of other investors. Van Bommel (2003), Eren and Ozsoylev (2009) and Jiang, Mahoney and Mei (2005) find that anonymous fraudsters spread misleading information in stock message boards to "pump" up the price of a stock and then "dump" their shares at the inflated price. When fraudsters stop hyping a stock after dumping, its price typically falls, causing other investors to lose money. Kogan, Moskowitz and Niessner (2021) find that fraudulent articles on Seeking Alpha (SA) and The Motley Fool led to abnormal trading volumes accompanied by significant short-term increases in stock prices and subsequent reversals. Mitts (2020) documents how pseudonymous authors in SA's "short idea" category generate sharp price drops and reversals in target firms, a pattern associated with manipulative stock options trading. Dyer and Kim (2021) show that investors discount information in anonymous equity research compared to non-anonymous research.

Thus, anonymity on stock message boards may make it easier for individuals to spread false information about companies, which can be used to manipulate stock prices and mislead investors. Anticipating such risks, investors require higher returns when making an investment, resulting in lower firm value. This leads to our alternative hypothesis (manipulation hypothesis):

H1b. Elimination of anonymity in the stock message board increases firm value.

4. Sample and Descriptive Statistics

We obtain accounting information and stock return data from the China Stock Market & Accounting Research (CSMAR) database and Guba posts data from https://guba.eastmoney.com/

(more details in Appendix B). Our initial sample consists of all A-share stocks listed on the Shanghai and Shenzhen Stock Exchanges from 2010 to 2019. We ensure that for the sample firms, we have five years of data before and after the regulatory change to allow us to compare the same set of firms between the pre- and post-event periods, effectively using each firm as its own control. This condition ensures that our results are not driven by firms that went IPO around the policy change, as such firms usually have great exposure in social media and perform poorly in the stock market following the IPO (Jain and Kini, 1994; Brau, Couch and Sutton, 2012). Our final sample consists of 14,770 firm-year observations of 1,477 unique firms.

If real-name registration has any impact on firm value, firms previously exposed more in social media should be relatively more affected. Based on this rationale, we calculate the total number of posts about a firm on its firm-specific Guba message board in 2014 (the year immediately before the real-name policy) as a measure of the firm's treatment intensity.

We measure firm value using Tobin's Q, defined as the sum of the market value of equity and book value of long-term debt divided by total assets (Roll, Schwartz and Subrahmanyam, 2009; Gurun and Butler, 2012). Motivated by the literature (e.g., Gurun and Butler, 2012; Cremers and Ferrell, 2014), we control for a vector of firm characteristics that could potentially affect firm value, including firm size, operating performance (*EBIT*), sales growth, physical investment (*PPE*), capital expenditure (*Capex*), book leverage, number of geographic business segments, and past-year stock returns. To minimize the effect of outliers, we winsorize all continuous variables at the 1st and 99th percentiles. Appendix A provides the variable definitions.

Table 1 Panel A presents the summary statistics. On average, our sample firms have a *Tobin's Q* of 1.902, book value assets of RMB 15 billion (or approximately USD 2.15 billion), *EBIT* of 4.880%, *PPE* of 27.854%, *Capex* of 4.575%, book leverage ratio of 60.616%, sales growth of 19.700%, and 5.246 geographic business segments. The sample average market return in the preceding 12 months is 17.607%. In 2014, the number of Guba posts for each sample firm (*#Posts 2014*) has a mean of 7,240 and a SD of 5,821, indicating a substantial variation in message board posts across firms prior to the regulatory change.

Panel B of Table 1 presents the sample firms in two groups. The treatment (control) group consists of firms with a higher (lower) total number of posts on their firm-specific Guba message boards than the sample median in 2014. We compared the pre-event firm characteristics between the two groups. Compared to the control group, the treatment firms are larger, have higher sales growth and profitability, invest more in capital expenditure, and have more debt and segments. Roberts and Whited (2013) point out that it is preferable for treatment and control groups to be relatively similar across observable dimensions prior to treatment. However, when this similarity is not achieved, control variables can be directly incorporated into the regression, as in our later analysis.

5. Results

5.1. Virtual Illustration and Univariate Test

In Figure 1, we plot the mean cumulative excess returns of the treatment and control group firms from 60 months before the regulatory change (i.e., month 0) to 60 months after. Before the regulatory change, the 60-month cumulative excess returns of the two groups move in tandem: −17.46% for the treatment group and −16.80% for the control group, respectively. In the 24 months after the regulatory change, the treatment firms had substantially lower stock performance than control firms. This difference remains stable for 36 months. Over the entire 60 months after the regulatory change, the cumulative excess returns were −46.75% for the treatment group and −33.66% for the control group.

Table 1 Panel C reports the univariate test to examine the before-after effect of the change in firm value for the treatment group compared to the before-after effect in the control group. The average *Ln(Q)* of the treatment group decreases from 0.270 in the pre-event period to 0.162 in the post-event period, and this change is significantly different from zero at the 1% level. By contrast, the average *Ln(Q)* of the control firms is statistically indistinguishable between the pre- and post-event periods (0.391 vs. 0.389). Column (3) reports the DiD estimation of *Ln(Q)*, defined as the average difference in *Ln(Q)* change between the treatment and control groups. These two differences are significantly different at the 1% level.

Figure 1 and Panel C of Table 1 together indicate that, after the regulatory change, the treatment firms experience a significant decline in firm value relative to the control firms. These results are consistent with H1a, the information environment hypothesis, which suggests that eliminating anonymity from stock message boards decreases firm value.

5.2. Baseline Results

We perform our baseline regression analysis in a general DiD framework as follows.

$$Ln(Q)_{i,t+1} = \alpha + \beta Ln(\#Posts\ 2014)_i \times After_t + \gamma' Z_{i,t} + \delta_t + \varphi_i + \varepsilon_{i,t}. \qquad (1)$$

In Equation (1), i denotes firm, and t denotes year. $Ln(\#Posts\ 2014)_i$ is the natural logarithm of the total number of Guba posts on firm i in 2014 and is used to measure the treatment intensity for this firm. $After_t$ is an indicator variable that takes the value of 1 for the post-treatment period (2015–2019) and 0 for the pre-treatment period (2010–2014). The β coefficient is the DiD estimate that captures the causal effect of real-name registration on firm value. Z is a vector of the control variables described in Section 4. We include year fixed effects, δ_t, to account for time-specific shocks to a firm's value and firm fixed effects, φ_i, to absorb any time-invariant unobservable firm characteristics. In the presence of these fixed effects, we do not include *Ln(#Posts 2014)* or *After* in the regression to avoid collinearity issues. We cluster standard errors by firm in all the regressions.

In columns (1)–(3) of Table 2, we include only *Ln(#Posts 2014)×After,* firm, and year fixed effects in the regression and implement the estimation using 2-, 6-, and 10-year windows surrounding the regulatory change, respectively. In columns (4)–(6), we re-estimate the models in columns (1)–(3) by including a full set of control variables. In all the six columns, the coefficients of *Ln(#Posts 2014)×After* are negative and statistically significant at the 1% level. The economic magnitude of this effect is significant. Taking column (6) (the time window is five years each before and after the treatment), the coefficient on *Ln(#Posts 2014)×After* is –0.071, indicating that a one-SD increase in treatment intensity (0.690) is associated with a 5.02% $(= e^{0.071\times0.690} - 1)$ decrease in Tobin's Q. Considering that the sample average market capitalization is RMB 13.3 billion, this reduction in Tobin's Q can be translated into a decline in market capitalization by RMB 667 million (equivalent to USD 95 million).[12]

As the validity of the DiD estimate depends critically on the parallel trends assumption, we conduct a diagnostic test to check whether this assumption holds. We construct nine indicator variables for flag years relative to regulatory changes. *Before*^{n} *(n= 1, 2, 3, 4)* indicates the nth year prior to the regulatory change. Likewise, *After*^{n} *(n= 1, 2, 3, 4, 5)* indicates the nth year after the regulatory change. In column (7), we re-estimate Equation (1) after replacing the *After* indicator variable with these nine indicators (the fifth year before treatment served as the benchmark year). The coefficients of *Ln(#Posts*

[12] In untabulated analysis, we use Q or industry-median-adjusted Q as dependent variables and our inference is largely the same.

2014)×Before¹, Ln(#Posts 2014)×Before², Ln(#Posts 2014)×Before³ and *Ln(#Posts 2014)×Before⁴* are especially important because their significance and magnitudes indicate any difference in the trend of firm value across firms with different treatment intensities prior to the regulatory change. Consistent with the pattern shown in Figure 1, we find that all these coefficients are statistically insignificant and trivial in magnitude, suggesting that the parallel trend assumption is not violated. By contrast, the coefficients for *Ln(#Posts 2014)×After¹, Ln(#Posts 2014)×After², Ln(#Posts 2014)×After³, Ln(#Posts 2014)×After⁴,* and *Ln(#Posts 2014)×After⁵* are significantly negative at or below the 5% level and are much larger in magnitude. For example, the coefficient of *Ln(#Posts 2014)×Before¹* is only −0.001, statistically indistinguishable from zero. By contrast, the coefficient of *Ln(#Posts 2014)×After⁵* is –0.063 and is statistically significant at the 1% level.

This lack of significant lead effects has three implications. First, it suggests that investors did not fully anticipate the real-name policy change. Consistent with this view, Subsection 5.5 shows that the stock market has a dramatically negative reaction to the policy announcement. Second, even if investors foresee a policy change, the actual information environment remains unaffected until the policy takes effect. Third, the effect of real-name registration on firm value cannot be attributed to policymakers reacting merely to past stock performance, which mitigates the concerns about reverse causality. This conclusion is also consistent with the discussion in Section 2 that the policy has been implemented to tighten control over the Internet rather than influence the stock market.

Overall, Table 2 shows that the real-name policy leads to a significant decrease in firm value, and that such an effect occurs after the policy change, suggesting a causal effect. These results support the information environment hypothesis.

5.3. Propensity Score Matching

A potential concern about our main findings is that firms with different treatment intensities may not be comparable owing to differences in firm characteristics (although we directly controlled for these characteristics in the regression). To alleviate this concern, we use a PSM algorithm to construct a matched sample of treatment and control firms. Specifically, we first follow the method used in Panel C of Table 1 to classify all sample firms into treatment and control groups: Firms whose total number of Guba posts in 2014 is above (below) the sample median in that year. We then estimate a probit model using all sample firms in 2014. The dependent variable is *Treat*, which equals 1 for treatment firms and 0 otherwise. We control for the full set of firm characteristics as in Equation (1), and for the average annual growth rates in Tobin's Q and Guba posts over the preceding three years (*Q 3-year growth* and *#Posts 3-year growth*, respectively) to ensure that the matched pairs have similar pre-existing trends in firm value and Guba posts. We include industry fixed effects in the probit model.

Column (1) of Table 3 presents the estimation results. The model captures a significant amount of variation in the choice variable, as indicated by a pseudo-R^2 of 21.5% and a p-value from the χ^2 test of the overall model fitness well below 0.1%. We match the treatment firms to the control firms using one-to-one nearest neighbor PSM without replacement, that is, we match each treatment firm to the control firm with the closest PSM score (i.e., the predicted probability) obtained from the estimation in column (1). We obtain 292 one-to-one pairs of matched firms (584 observations).

As the validity of PSM depends on finding closely matched firms, we conduct two diagnostic tests. In column (2) of Panel A, we re-estimate the probit model using a matched sample. All the coefficient estimates are statistically insignificant. Panel B reports the balance test results for the major characteristics of the matched pairs. We show that the pre-event characteristics of the treatment firms and their matched peers are indistinguishable. Overall, these diagnostic tests show that PSM eliminates meaningful observable differences in pre-event characteristics between treatment and control firms (other than the difference in Guba posts in 2014), and increases the likelihood of the observed difference in the change in firm value between the groups being caused by real-name registration.

We perform a DiD analysis using the matched sample, and report the estimation results in Panel C of column (1). The model in column (1) is the same as the baseline model in column (6) of Table 2 except that we replace *Ln(#Posts 2014)* with *Treat*. The coefficient of *Treat×After* is –0.096 and is statistically significant at the 1% level. This result indicates that the real-name policy leads to approximately 10% $(= e^{0.096} - 1)$ reduction in firm value for treatment firms relative to their propensity score-matched peers. The model in column (2) is similar to that in column (7) of Table 2, except that we replace *Ln(#Posts 2014)* with *Treat.* The estimation results indicate that the matched pairs have similar pre-treatment trends in firm value, and the effect of the real-name policy appears only after the treatment.

Taken together, the results in Table 3 are consistent with those in Table 2. Hence, analysis based on the PSM sample further alleviates the concern that our baseline results might be driven by differences in firm characteristics between the treatment and control groups.

5.4. Instrumental Variable Approach

In our baseline model, we use the number of Guba posts in 2014 to classify the treatment and control groups. However, it is possible that some omitted variables correlated with firms' Guba posts in that year drive their subsequent change in valuation. We apply a two-stage least squares (2SLS) regression with an instrumental variable (IV) to relate the exogenous variation in the number of Guba posts and thus help mitigate this concern.

We exploit the proportion of days with inclement weather in 2014 as the IV for the scale of the online discussion in that year. Weather information at the firm's headquarters is collected from the Chinese Research Data Services Platform (CNRDS). Inclement weather includes heavy rainfall, heavy snow, hail, sandstorm, violent wind (Beaufort scale greater than 7), and extreme temperature (>35 or <-25 degrees Celsius). It is well-documented that when the weather is bad, individuals are more likely to choose indoor (as opposed to outdoor) activities such as surfing the Internet and watching TV (Starr-McCluer, 2000; Gomez, Hansford and Krause, 2007). Additionally, individuals exhibit a strong preference for local stocks (Coval and Moskowitz, 1999; Ivkovic and Weisbenner, 2005). Therefore, the IV is likely to be positively associated with the number of Guba posts for a given local firm (satisfying the relevance condition). More importantly, the weather in 2014 is unlikely to affect the change in firm valuation during 2015–2019 (other than through social media exposure), thus possibly satisfying the exclusion condition.

In the first stage, we run an OLS regression with *Ln(#Posts 2014)* as the dependent variable. The independent variables include *Bad weather 2014* and the firm characteristics used in Equation (1). In the second stage, we re-estimate Equation (1) using the predicted value of *Ln(#Posts 2014)* obtained from the first stage. Column (1) of Table 4 presents the results of the first-stage regressions. The coefficient of IV is positive and significant at the 1% level, indicating that bad weather increases participation in online posts about local firms. The corresponding *F*-statistic on the IV is 12.16, exceeding 8.96, the critical value for Stock and Yogo's (2005) weak instrument test based on the 2SLS size. This result indicates that the instrument provides a significant degree of incremental explanatory power, and that our results are not susceptible to weak instrument bias.

Column (2) presents the second-stage regression results. Consistent with the findings of the OLS analysis, the coefficient estimate of *Ln(#Posts 2014)×After* is negative and significant at the 1%

level. Comparing the results obtained from the OLS regressions (column (6) of Table 2) with those obtained from the 2SLS regressions, we observe that the magnitudes of the 2SLS coefficient estimates (−0.164) are twice as large as those of the OLS estimate (−0.071), suggesting that OLS regressions bias the coefficient estimates downward. This finding suggests that some omitted variables (e.g., managerial talent) simultaneously result in higher firm valuation and more intensive discussions on social media by investors. Once we use the IV to clean up such correlations, the coefficient estimates become larger. Overall, the documented impact of the real-name policy on firm value is robust in addressing the endogeneity associated with a firm's exposure to Guba.

5.5. Event Study around the Policy Announcement

We examine whether the stock market reacts drastically to the real-name policy announcement and whether the cross-sectional variation in the firm-level stock reaction depends on the firm's exposure to Guba prior to the event. Stock price changes around that date reveal how investors expect policy changes to affect firms' prospects.

We compute the five-day cumulative stock returns (*AR[−2,+2]*) around the policy announcement for each sample firm. As shown in Panel A of Table 5, the average *AR[−2,+2]* for the treatment group is −2.497%, while that for the control group is −1.785%. This difference is statistically significant at the 1% level. By contrast, the average *AR[−2,+2]* around February 4 of the other nine sample years (i.e., February 4 of 2010–2014 and 2016–2019) is 1.068% for the treatment group and 1.193% for the control group;[13] the difference is not significant. In column (1) of Panel B, we regress *AR[−2,+2]* around the policy announcement on *Ln(#Posts 2014)*, and show that the coefficient of *Ln(#Posts 2014)*

[13] If February 4 of a certain year is a non-trading day, we use the closest trading day instead.

is negative and significant at the 1% level. In column (2), we re-estimate column (1) using *AR[−2,+2]* around February 4 for the other nine-year samples, and the coefficient of *Ln(#Posts 2014)* is close to zero in magnitude and is not significant.

The findings presented in Table 5 have three important implications. First, the real-name policy is unique and causes a dramatically negative market reaction. The cross-sectional variation in stock reactions is negatively related to firms' exposure to social media, suggesting that investors expect a greater negative impact of the policy change on treatment firms than on control firms. These results are consistent with our baseline regression, and support the information environment hypothesis. Second, considering that the post-treatment period in our baseline regression is five years, this event study helps mitigate concerns about confounding policies occurring during that period. Third, given the intangible nature of firms' information environments, the value of social media anonymity does not seem to have been fully incorporated into the stock price around the policy announcement date, but is reflected in Tobin's Q as the actual information is disseminated gradually over time.

6. Channel Test

6.1. Channel Test 1: Heterogenous Treatment Effect

To provide evidence of the mechanisms underlying our main findings, we examine the heterogeneous treatment effects across firms. Doing so can further mitigate the concern that some omitted variables drive our results because such omitted variables (if any) would need to be uncorrelated with the control variables in the regression, and would also need to explain the cross-sectional variation in the treatment effect (e.g., Claessens and Laeven, 2003; Raddatz, 2006).

If real-name registration decreases firm value by exacerbating the information environment faced by investors, the treatment effect should be stronger in situations in which investors rely more on social media for information. Therefore, we examine how the impact of real-name registration varies with the following four firm characteristics: analyst coverage, newspaper coverage, institutional ownership, and local usage of the mobile Internet.

Financial analysts are important traditional sources of information for investors. When a firm is covered by fewer analysts, investors have a greater need to obtain information via alternative sources such as social media (Hales, Moon and Swenson, 2018; Campbell, DeAngelis and Moon, 2019). Thus, we expect the treatment effect to be more pronounced in firms with lower analyst coverage. Similarly, when a firm is covered by fewer newspapers (other traditional information providers), investors tend to rely more heavily on social media for information (Fang and Peress, 2009; Bushee et al., 2010; You, Zhang and Zhang, 2018), and we expect the treatment effect to be more pronounced in firms with lower newspaper coverage. Unlike individual investors, institutional investors are less likely to rely on social media because they have their own research teams, subscribe to professional research services, and enjoy direct access to managers and equity analysts (Soltes, 2014; Green et al., 2014; Bushee, Jung and Miller, 2017). Hence, the treatment effect is expected to be stronger for firms with lower institutional holdings. Considering that mobile Internet is the most popular way for individuals to communicate and access timely news updates via social media (Federal Communications Commission, 2016), we expect a stronger treatment effect for firms in areas with greater local mobile Internet usage.

In columns (1) and (2) of Table 6, we partition our sample based on the amount of analyst coverage in 2014 (the year immediately before the policy change) to capture financial analysts' pre-existing level of coverage. We then re-estimate Equation (1) separately for terciles of firms with the highest and lowest analyst coverage. For the former group, the DiD estimate (i.e., the coefficient on *Ln(#Posts 2014)×After*) is −0.030 and is not statistically significant. By contrast, for firms with less analyst coverage, the DiD estimate is −0.074 and is significant at the 1% level. The corresponding Wald test indicates that these two coefficients differ significantly at the 5% level. This result indicates that the negative effect of real-name registration on firm value is more pronounced for firms with less analyst coverage. In columns (3) and (4), we partition our sample by newspaper coverage in 2014. The DiD estimate in the group of firms with lower newspaper coverage is −0.069, almost twice as large as that in the group of firms with higher newspaper coverage (−0.035), indicating a stronger treatment effect for firms with lower newspaper coverage.

Columns (5) and (6) partition the sample by institutional ownership in 2014. The DiD estimate for firms with smaller institutional ownership is −0.081, more than twice that for firms with larger institutional ownership (−0.036). The difference between these two coefficients is significant at the 5% level, indicating a stronger treatment effect for firms with more retail investors. We partition our sample by the local usage of mobile Internet in columns (7) and (8). In both columns, the DiD coefficients are significantly negative. However, the magnitudes of the estimates are much higher for firms with higher local mobile Internet usage than for other groups (–0.099 vs. –0.046), indicating a stronger treatment effect for them.[14] Taken together, Table 6 indicates that the impact of real-name registration on firm value is more pronounced when investors rely on social media for information.

6.2. Channel Test 2: Impact on Social Media

6.2.1. Number and Structure of Social Media Posts

[14] In untabulated analysis, we examine the heterogenous treatment effect between state-owned and non-state-owned firms, and find that the treatment effect is largely the same for these two groups. This result is understandable because, *ex ante*, it is unclear which group is more affected by the real-name policy. On the one hand, the treatment effect may be stronger for state-owned firms, because they may have greater resources to retaliate against the investors. On the other hand, the opposite may be true because state-owned firms are usually sufficiently covered by traditional information intermediaries like analysts and newspapers, making social media less important for investors to acquire information.

In this subsection, we examine the impact of the real-name policy on the quantity and structure of information exchanged on social media. In column (1) of Table 7 Panel A, the dependent variable *Ln(#Posts)* measures the total number of Guba posts about a given firm. The coefficient on *Ln(#Posts 2014)×After* is –0.176 and significant at the 1% level, indicating that a one-SD increase in treatment intensity (0.690) decreases the number of Guba posts by 12.9% ($= \mathrm{e}^{0.690\times0.176} - 1$). Similarly, column (2) shows that the number of words contained in each post significantly decreases following the treatment policy. These findings indicate that the quantity of information exchanged on social media is reduced when anonymity is removed.

Next, we define a rehashed post as one that predominantly comprises quotes sourced from public materials (e.g., corporate announcements, analyst reports, newspaper reports, and other social media posts) and has less than 10 words excluding quotes. As Chen, Magdy and Wolters (2020) note, such rehashed posts typically contain little original information. Column (3) shows that the treatment policy leads to a significant increase in the proportion of rehashed posts among all posts (i.e., a decline in the originality of Guba posts). Likewise, we define a first-person post as one that uses first-person expressions. Column (4) shows that the proportion of such posts decreases significantly after the treatment policy, indicating that personal opinions are now less likely to be revealed on social media.

We further define a quantitative post as the one that contains at least 10 words and a number in the following cases: (i) the number is preceded by a dollar/Chinese Yuan sign ("$/￥"); (ii) the number is followed by the words "million"/"billion"/"trillion"; (iii) the number is followed by a percentage sign ("%") or by the word "percent"; (iv) the number is followed by an unit symbol for kilometer (km), kilogram (kg), and kilowatt (kW), square meters (m^2), and so on. Siano and Wysocki (2018) show that such quantitative information is more likely to make the text objective, concise, precise, verifiable, and free of rhetoric, and is thus associated with a higher quality of information exchange. Column (5) shows that the proportion of quantitative posts significantly decreases after the treatment policy.

The variable *#Comments per post* in column (6) measures the number of comments for each post. The coefficients of *Ln(#Posts 2014)×After* are significantly negative, implying that the posts receive fewer comments following the policy change. To the extent that *#Comments per post* reflects investors interacting to share opinions (Chen et al., 2014), this result is consistent with the view that removing online anonymity discourages freedom of expression and information exchange in social media.

Next, we examine how real-name registration affects the flow of negative information on Guba. Negative information is particularly important to retail investors because they often hold undiversified portfolios, have lower risk tolerance (Goetzmann and Kumar, 2008), and pay more attention to bad news because of loss aversion and endowment effects (Kahneman, 1979). Furthermore, they are more sensitive to bad news because it spreads faster and more widely via the Internet than good news (Hornik et al., 2015). In column (7), the dependent variable is the number of negative posts minus the number of positive posts normalized by the total number of posts (*Net proportion of negative posts*), which measures the extent of negative messages communicated on social media. Following Antweiler and Frank (2004), a post is classified as negative, positive, or neutral using the Naive Bayes Classification method based on a manually labeled training sample with a Chinese sentiment dictionary. The coefficient on *Ln(#Posts 2014)×After* is −0.382 and statistically significant at the 1% level, indicating that a one-SD increase in treatment intensity reduces the proportion of negative information in Guba by more than 26 percentage points (=0.690×0.382), relative to the sample mean of −18%. This shows that real-name registration discourages netizens from revealing negative information on social media.

Following Antweiler and Frank (2004), we construct the contributors' agreement index as $1 - \sqrt{1 - B^2}$, where B is (number of posts with a negative tone − the number of posts with a positive tone)/(number of posts with a negative tone + number of posts with a positive tone). The index ranges from 0 to 1, with a higher value indicating greater homogeneity (i.e., a lack of different perspectives). When all posts are homogeneous in tone (i.e., they are all positive or negative), the agreement index is 1. By contrast, when opinions are equally divided between positive and negative, the index is 0. Column (8) shows that the agreement index increases significantly following the treatment policy, indicating that removing anonymity makes people less likely to express disagreement or different perspectives.

6.2.2. Predicting Earnings Surprise Based on Social Media Posts

If real-name registration impedes the flow of information (particularly, negative information) on social media, we should observe a decline in the predictive power of social media for earnings surprises after the regulatory change. We examine the extent to which negative (positive) Guba posts in a quarter predict negative (positive) quarterly earnings surprises. As analysts in China rarely issue quarterly forecasts (Huyghebaert and Xu, 2016; Ke and Zhang, 2021), we measure earnings surprise as earnings per share (EPS) in quarter q minus EPS in q-4 (i.e., the same quarter in the previous year), scaled by the absolute value of EPS in q-4 (Bamber and Cheon, 1995; Bartov, Radhakrishnan and Krinsky, 2000).

Column (1) of Panel B in Table 7 focuses on our sample's firm-quarter observations of negative earnings surprises. The dependent variable is quarterly earnings surprises, and we include firm and quarter fixed effects and the same set of control variables as in column (6) of Table 2. The key variable of interest is *Proportion of negative posts*, defined as the ratio of negative posts to all Guba posts created during the quarter. We also include an additional interaction term with *After* to capture the incremental predictive power of *Proportion of negative posts* in the post-treatment period.

The coefficient of *Proportion of negative posts* is –0.036 and statistically significant at the 1% level, indicating that negative information communicated on anonymous social media can predict poor firm performance. This finding is broadly consistent with Hales, Moon and Swenson (2018), and Huang, Li and Markov (2020), who show that social media in the U.S. conveys important information for predicting corporate earnings. However, the coefficient of the interaction term *Proportion of negative posts* ×*After* is positive (0.028) and statistically significant at the 5% level. These results indicate that negative posts on social media can predict poor firm performance mainly when social media is anonymous, but such predictability shrinks by 78% (=0.028/0.036) once anonymity is removed.

Column (2) reports the results estimated using the firm-quarter observations of positive earnings surprises. The model specifications are the same as those in column (1), except that we replace *Proportion of negative posts* with *Proportion of positive posts,* defined as the ratio of positive posts to all Guba posts created during the quarter. The coefficient of *Proportion of positive posts* is 0.021 and is significant at the 5% level, and the coefficient of the interaction term *Proportion of positive posts* × *After* is –0.002 and is not statistically significant. This result means that a favorable message on social media predicts positive earnings surprises, and such predictability is largely unchanged irrespective of the real-name policy.

Overall, Panel B of Table 7 indicates that while positive messages on social media can predict good firm performance irrespective of anonymity, negative messages can predict poor firm performance only when social media is anonymous. These results provide supporting evidence that the real-name policy deters the dissemination of negative but valuable information on social media.

6.2.3. Stock Reaction toward Earnings Surprise

If real-name registration discourages the disclosure and exchange of negative information, investors should be surprised and react more strongly when a firm's earnings fall short of expectations (Kothari, Shu and Wysocki, 2009). In Panel C, we follow Andrei, Friedman and Ozel (2023) to compute the dependent variable *CAR[−2,+2]* as the five-day size-adjusted cumulative abnormal return (firm's stock return minus the value-weighted size matched portfolios return) around the announcement date of quarterly earnings.[15] Column (1) estimates the model using the firm-quarter observations of negative earnings surprises. We include the same control variables as in Panel B, and additionally control for the magnitude of earnings surprises. The coefficient of *Ln(#Posts 2014)×After* is negative and statistically significant at the 1% level, indicating that investors have stronger negative reactions to earnings surprises when social media anonymity is eliminated.

In column (2), we repeat the same analysis focusing on firm-quarter observations with positive earnings surprises. The coefficient of *Ln(#Posts 2014)×After* is negligible and statistically indistinguishable from zero. The difference between these two DiD estimates is significant at the 5% level, which indicates that the real-name policy leads to stronger negative market reactions to negative earnings surprises, whereas investor reactions to positive earnings surprises are unaffected. Overall, the findings in Panel C align closely with those in Panel B, confirming that negative information is suppressed with the removal of anonymity from social media platforms.

[15] Our main results are robust to various ways of defining excess daily returns, such as firm-specific returns minus market returns.

6.2.4. Placebo Test on Newspaper

A potential concern is that our findings might be driven by a change in China's overall media environment rather than just the removal of anonymity in social media. We investigate this possibility by examining the change in the coverage and tone in China's traditional media. China's newspaper organizations consist of a set of central party-state media outlets (e.g., the People's Daily and Xinhua News Agency) and regional official outlets (e.g., Shanghai Daily and Beijing Daily). All are state-owned, strictly controlled by the Propaganda Department of the Chinese Communist Party (Qin, Strömberg and Wu, 2018), and key to the government's goal of maintaining "correct public opinion guidance." Newspapers, never anonymous, were thus unaffected by the real-name policy.

In Panel D of Table 7, we collect information on the newspaper coverage of each stock from the CNRDS dataset. We re-estimate our baseline regression using newspaper coverage and tone as dependent variables in columns (1) and (2), respectively, in which the coefficients of *Ln(#Posts 2014)×After* are negligible in magnitude and not significantly different from zero, indicating that newspapers are unaffected by the real-name policy. Overall, the results show that the policy affects only social media and has no effect on traditional paper-based media, indicating that our findings are likely due to the change in policy rather than in China's overall media environment.

Taken together, Table 7 shows that the real-name policy decreases the quantity and quality of value-relevant information (particularly negative information) on social media, supporting the information environment hypothesis.

6.3. Channel Test 3: Impact on Firm Information Environment

Next, we provide further evidence that real-name registration worsens the information environment for investors. Column (1) of Table 8 focuses on firms' stock price informativeness.

Higher stock price informativeness indicates more firm-specific information conveyed in firms' stock prices, and thus a better information environment for investors (Morck, Yeung and Yu, 2000; Hutton, Marcus and Tehranian, 2009). Following this strand of the literature, we first compute the R^2 obtained by estimating the expanded index model detailed in Appendix A, and then define a firm's stock price informativeness (*Stock informativeness*) as $Log(\frac{1-R^2}{R^2})$, where a higher value indicates greater informativeness. The coefficient of *Ln(#Posts 2014)×After* is −0.037, statistically significant at the 5% level, which suggests that the policy leads to a significant decrease in firms' stock price informativeness.

In column (2), we focus on stock liquidity because increased disclosure and transparent information environments improve liquidity (Leuz and Verrecchia, 2000; Kelly and Ljungqvist, 2012; Balakrishnan et al., 2014). If real-name registration hinders information dissemination on social media, it probably decreases a firm's stock liquidity. The variable *Bid-ask spread* is constructed as follows. We first compute the average daily bid-ask spread, and then take the natural logarithm to correct for non-normality in the distribution (Fang, Noe and Tice, 2009).[16] A higher value of *Bid-ask spread* indicates lower stock liquidity. The coefficient of *Ln(#Posts 2014)×After* is 0.055 and is statistically significant at the 5% level, suggesting that the real-name policy indeed reduces stock liquidity.

Informationally opaque firms tend to have higher stock price crash risk (Jin and Myers, 2006; Hutton, Marcus and Tehranian, 2009). If negative firm-specific information is not reflected in stock prices in a timely manner, a stock price crash is possible when the accumulated negative firm-specific information is revealed. Therefore, we use a firm's stock price crash risk as another proxy for investors'

[16] In untabulated tests, we use Armihud's (2002) illiquidity measure and obtain similar inferences.

information environments. In column (3), we measure *Crash risk* as the negative of the third moment of firm-specific weekly returns in a sample year divided by the SD of firm-specific weekly returns raised to a third power (Chen, Hong and Stein, 2001; Kim, Li and Zhang, 2011; Xu, Xuan and Zheng, 2021).[17] The coefficient of *Ln(#Posts 2014)×After* is positive and statistically significant at the 1% level, indicating that firms with a higher treatment intensity experience a significantly larger increase in the likelihood of stock price crashes after the treatment policy.

Column (4) focuses on voluntary corporate disclosures, which helps reduce information asymmetry between corporate insiders and outside investors. Following Bryan (1997), we measure voluntary disclosure using the logarithm of the number of words about future discussions in the Management Discussion and Analysis (MD&A) section of a company's annual report.[18] The coefficient of *Ln(#Posts 2014)×After* is significantly negative, indicating that the real-name policy leads to a significant decrease in voluntary corporate disclosure. The literature suggests that corporate managers' voluntary disclosure decisions depend on shareholders' access to information (Healy and Palepu, 2001; Hutton, Miller and Skinner, 2003; Li, 2008; Kothari, Shu and Wysocki, 2009). When investors have greater access to information from external sources (e.g., active information sharing on anonymous social media), managers tend to provide more voluntary disclosures to reduce detection or

[17] In untabulated tests, we also use the down-to-up volatility as an alternative proxy for the stock price crash risk and our inference is largely the same.

[18] Although the inclusion of an MD&A section in financial reports is mandatory, managers' discretion drives its content (Li, 2008; Loughran and McDonald, 2014). Leuz and Schrand (2009) and Bourveau et al. (2018) use its length as proxy for the level of value-relevant voluntary disclosure. In general, firms are required to disclose both retrospective and prospective analyses in MD&A. Bryan (1997) finds prospective disclosures (share of content related to the future) to be especially helpful in assessing firms' prospects. Such evidence suggests that our measurement is a reasonable proxy for voluntary disclosure.

litigation risks (Skinner, 1994, 1997; Beyer et al., 2010). Once the removal of anonymity reduces investors' access to information, managers face less pressure to voluntarily disclose it.[19]

Finally, in poor information environments, shareholders often lack sufficient information to monitor the actions of managers. This may induce managers to engage in earnings management activities (Schipper, 1989; Warfield, Wild and Wild, 1995). We test this conjecture in columns (5) and (6). In column (5), the dependent variable measures firms' real earnings management, defined as abnormal production costs minus abnormal cash flows from operations and abnormal discretionary expenses (Kim, Park and Wier, 2012). The coefficient of *Ln(#Posts 2014)×After* is significantly positive at the 5% level, suggesting that the policy leads to an increase in real earnings management. In column (6), the dependent variable *Restatement* is an indicator variable that takes the value of 1 for firms announcing the correction of misstatements in a subsequent reporting period, and 0 otherwise. The coefficient on *Ln(#Posts 2014)×After* is 0.017 and statistically significant at the 5% level, indicating that the likelihood of financial restatement increases after the treatment policy.[20] The magnitude of the coefficient is economically significant: A one-SD increase in treatment intensity results in a higher likelihood of restatement by 1.17 (=0.690×0.017) percentage points compared to the unconditional probability of 8.10 percentage points. These results indicate that managers tend to engage in more earnings management after the policy change.

[19] *Ex ante*, managers may voluntarily enhance firm disclosure as an attempt to offset the negative impact of real-name registration on firms' information environment. However, our findings do not support this view.

[20] In column (6), we use OLS regression rather than the logit/probit regression because of the "incidental parameters problem" associated with fixed effects. Greene (2004) and Arellano and Hahn (2007) suggest that nonlinear maximum likelihood models can be inconsistent and biased when fixed effects are included. Ai and Norton (2003) point out that the coefficient on the interaction term does not represent the marginal effect in logit or probit DiD specifications. Using linear models with discrete outcome variables is common practice (e.g., Gillette, Samuels and Zhou, 2020; Jennings et al., 2023).

Taken together, Table 8 provides supporting evidence that real-name registration decreases firm value by worsening the information environment for shareholders: It decreases stock price informativeness and liquidity, increases stock price crash risk, reduces voluntary corporate disclosure, and boosts earnings manipulation.

6.4. Alternative Channels: Overvaluation and/or Attention

An alternative explanation for our baseline results is that firms with higher treatment intensities are overvalued during the pre-event period. Consequently, these firms attract more Guba posts in the pre-event period and experience a greater decline in firm value in the post-event period. This explanation is unlikely to hold because firms with higher treatment intensities have lower Tobin's Q in the pre-event period. As shown in Panel C of Table 1, in the pre-event period, the average *Ln(Q)* of firms with high treatment intensities is only 0.270, which is smaller than that of firms with low treatment intensities (0.391). Figure 1 shows that the stock returns of these two groups are indistinguishable during the 60-month period prior to the treatment policy.

To further investigate this overvaluation-based explanation, we re-estimate the baseline model in column (6) of Table 2 using three alternative measures of treatment intensity: *Ln(#Negative posts 2014)*, *Ln(#Non-positive posts 2014)*, and *Net proportion of negative posts 2014*. The first two variables measure the total number of negative (non-positive) Guba posts on a firm in 2014 (the year immediately prior to the regulatory change). The third is *Net proportion of negative posts* measured in 2014. A higher value for these three variables indicates a more negative tone toward the firm. As real-name registration discourages negative posts more than positive ones, firms with relatively more negative pre-existing Guba posts should be more affected. A key advantage of using these alternative measures of treatment intensity is that negative posts are unlikely to be correlated with stock overvaluation and are more likely to be correlated with stock undervaluation (Chen et al., 2014). Panel A of Table 9 reports the estimation results. In all three columns, the DiD estimates are negative and significant at or below the 5% level, indicating that firms with more negative posts in the pre-event period experience larger declines in value after the treatment. These results, combined with evidence from Panel C of Table 1 and Figure 1, show our main findings are unlikely to be driven by an overvaluation channel.

Another possible alternative is investor attention. Real-name registration deters information exchange on social media and thus decreases investor attention paid to the stock. Considering that investor attention is usually positively associated with stock valuation (Barber and Odean, 2008; Hirshleifer, Lim and Teoh, 2009; Aboody, Lehavy and Trueman, 2010), a drop in investor attention may lead to a decline in firm valuation. To examine this explanation, we use the Baidu Searching Index, the Chinese equivalent of Searching Volume Index of Google Trends, to measure stock-level retail investor attention (Da, Engelberg and Gao, 2011; DeHaan, Shevlin and Thornock 2015; Choi, Gao and Jiang, 2020). The Baidu Searching Index for stocks covers all Internet searches on Baidu with keywords related to the stock, and the keywords must include the stock code, current/historical stock name, or current/historical name of the listed company.[21] In column (1) of Panel B of Table 9, we measure investor attention using the *Overall index*, which measures the search volumes from all devices. In columns (2) and (3), we measure investor attention using *Mobile index* and *PC index*, which represent searches from mobile phones and personal computers, respectively. We re-estimate our baseline regression model using these attention proxies as dependent variables. In all three columns,

[21] Information on the Baidu Searching Index is obtained from MARK datasets (https://www.macrodatas.cn/article/1147467078). As this dataset starts in 2011, our sample period for this analysis is from 2011 to 2019.

the coefficients of $Ln(\#Posts\ 2014) \times After$ are not significantly different from zero, indicating that real-name registration does not alter the investor attention paid to firms. Thus, our main findings cannot be attributed to changes in investor attention.

7. Conclusion

Despite numerous arguments regarding Internet anonymity in the past two decades, quantitatively assessing its societal impact has been challenging. After weighing the potential cost and benefits, we ask what the net benefit is of Internet anonymity. We explore this issue by examining the effects of Internet anonymity on firm value in the stock market, especially because of the growing investor reliance on social media for financial information, and using the rich stock market data for a quantitative evaluation.

Our study utilizes China's transition to a real-name registration system for all users of social media in 2015. We observe a substantial decline in firm value for firms more affected by this policy than for those less affected. This finding is robust to various time windows, PSM method, 2SLS approach, and event studies. Through a battery of tests to examine the underlying mechanism, we show that (1) stronger treatment effects exist for firms whose investors rely more heavily on social media for information; (2) the treatment policy reduces both the quantity and quality of information exchange on social media platforms; and (3) the policy leads to lower stock price informativeness and liquidity, higher stock crash risk, poorer corporate disclosure, and more earnings manipulation. Collectively, these results support the information environment hypothesis that real-name registration on social media decreases firm value by worsening investors' information environment.

Several governments have considered or passed laws restricting Internet users' rights to maintain online anonymity. On July 31, 2023, President Putin signed a legal amendment requiring all Russian online platforms to verify the identities of new users before providing access to services. In 2021, British politicians called for an end to online anonymity in the UK's Online Safety Bill, arguing that such anonymity contributed to greater toxicity in public discourse; the proposal was later not included in the draft Bill. In the U.S., Senator John Kennedy, Republican of Louisiana, announced in 2021 a plan to introduce legislation requiring social media users to verify their legal identities (critics argue this may be unconstitutional, but Kennedy has yet to introduce the bill). This is the first quantitative analysis of this debate in the context of stock markets. Our results suggest that eliminating Internet anonymity does more harm than good, with implications for public policy everywhere.

Finally, we acknowledge that caution should be exercised in generalizing our results to broader social domains because the nature of information exchange and participant engagement may vary significantly across different socio-political contexts. Future research should explore the impact of Internet anonymity regulations on diverse social landscapes.

REFERENCES

Aboody, David, Reuven Lehavy, and Brett Trueman. 2010. Limited attention and the earnings announcement returns of past stock market winners. *Review of Accounting Studies* 15: 317-344.

Acharya, Viral V., Peter DeMarzo, and Ilan Kremer. 2011. Endogenous information flows and the clustering of announcements. *American Economic Review* 101 (7): 2955-2979.

Ai, Chunrong, and Edward C. Norton. 2003. Interaction terms in logit and probit models. *Economics Letters* 80 (1): 123-129.

Amihud, Yakov. 2002. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets* 5 (1): 31-56.

Andrei, Daniel, Henry Friedman, and N. Bugra Ozel. 2023. Economic uncertainty and investor attention. *Journal of Financial Economics* 149 (2): 179-217.

Ang, James S., Charles Hsu, Di Tang, and Chaopeng Wu. 2021. The role of social media in corporate governance. *The Accounting Review* 96 (2): 1-32.

Antweiler, Werner, and Murray Z. Frank. 2004. Is all that talk just noise? The information content of Internet stock message boards. *The Journal of Finance* 59 (3): 1259-1294.

Arellano, Manuel, and Jinyong Hahn. 2007. Understanding bias in nonlinear panel models: Some recent developments. *Econometric Society Monographs* 43: 381.

Balakrishnan, K., Billings, M. B., Kelly, B., and Ljungqvist, A. 2014. Shaping liquidity: On the causal effects of voluntary disclosure. *The Journal of Finance* 69 (5), 2237-2278.

Bamber, Linda Smith, and Youngsoon Susan Cheon. 1995. Differential price and volume reactions to accounting earnings announcements. *The Accounting Review* 70 (3): 417-441.

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

Bartov, Eli, Lucile Faurel, and Partha S. Mohanram. 2018. Can Twitter help predict firm-level earnings and stock returns? *The Accounting Review* 93 (3): 25-57.

Bartov, Eli, Suresh Radhakrishnan, and Itzhak Krinsky. 2000. Investor sophistication and patterns in stock returns after earnings announcements. *The Accounting Review* 75 (1): 43-63.

Beyer, Anne, Daniel A. Cohen, Thomas Z. Lys and Beverly R. Walther. 2010. The financial reporting environment: Review of the recent literature. *Journal of Accounting and Economics* 50 (2-3), 296-343.

Bhattacharya, Nilabhra, Black, Erv, Christensen, Ted, and Mergenthaler, Rick. 2007. Who trades on pro forma earnings information? *The Accounting Review* 82 (3): 581-619.

Blankespoor, Elizabeth, Ed deHaan, and Ivan Marinovic. 2020. Disclosure processing costs, investors' information choice, and equity market outcomes: A review. *Journal of Accounting and Economics* 70 (2-3): 1013-1044.

Blankespoor, Elizabeth, Gregory S. Miller, and Hal D. White. 2014. The role of dissemination in market liquidity: Evidence from firms' use of Twitter. *The Accounting Review* 89 (1): 79–112.

Bourveau, Thomas, Yun Lou, and Rencheng Wang. 2018. Shareholder litigation and corporate disclosure: Evidence from derivative lawsuits. *Journal of Accounting Research* 56 (3): 797-842.

Brau, James C., Robert B. Couch, and Ninon K. Sutton. 2012. The desire to acquire and IPO long-run underperformance. *Journal of Financial and Quantitative Analysis* 47 (3): 493-510.

Brav, Alon, Matthew Cain, and Jonathon Zytnick. 2022. Retail shareholder participation in the proxy process: Monitoring, engagement, and voting. *Journal of Financial Economics* 144 (2): 492-522.

Bryan, Stephen H. 1997. Incremental information content of required disclosures contained in management discussion and analysis. *The Accounting Review,* 285-301.

Bushee, Brian J., John E. Core, Wayne Guay and Sophia J.W. Hamm. 2010. The role of the business press as an information intermediary. *Journal of Accounting Research* 48(1): 1-19.

Bushee, Brian J., Michael J. Jung, and Gregory S. Miller. 2017. Do investors benefit from selective access to management? *Journal of Financial Reporting* 2 (1): 31-61.

Cade, Nicole L. 2018. Corporate social media: How two-way disclosure channels influence investors. *Accounting, Organizations and Society* 68-69: 63-79.

Campbell, John L., Matthew D. DeAngelis, and James R. Moon. 2019. Skin in the game: Personal stock holdings and investors' response to stock analysis on social media. *Review of Accounting Studies* 24: 731-779.

Carpenter, Jennifer N., Fangzhou Lu, and Robert F. Whitelaw. 2015. The real value of China's stock market. *Journal of Financial Economics* 139 (3): 679-696.

Chen, Hailiang, Prabuddha De, Yu (Jeffrey) Hu, and Byoung-Hyoun Hwang. 2014. Wisdom of crowds: The value of stock opinions transmitted through social media. *The Review of Financial Studies* 27 (5): 1367-1403.

Chen, Joseph, Harrison Hong, and Jeremy C. Stein. 2001. Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. *Journal of Financial Economics* 61 (3): 345-381.

Chen, Lucia Lushi, Walid Magdy, and Maria K. Wolters. 2020. The effect of user psychology on the content of social media posts: Originality and transitions matter. *Frontiers in Psychology* 11: 1-9.

Choi, Bryan H. 2012. The anonymous Internet. *Maryland Law Review* 72: 501-570.

Choi, Darwin, Zhenyu Gao, and Wenxi Jiang. 2020. Attention to global warming. *The Review of Financial Studies* 33 (3): 1112-1145.

Claessens, Stijn, and Luc Laeven. 2003. Financial development, property rights, and growth. *The Journal of Finance* 58 (6): 2401-2436.

Coval, Joshua, and Tobias Moskowitz, 1999. Home bias at home: Local equity preference in domestic portfolios, *Journal of Finance* 54 (6): 2045-2073.

Cremers, Martijn, and Allen Ferrell. 2014. Thirty years of shareholder rights and firm value. *The Journal of Finance* 69 (3): 1167-1196.

Da, Zhi, Joseph Engelberg, and Pengjie Gao. 2011. In search of attention. *The Journal of Finance* 66 (5): 1461- 1499.

Das, Sanjiv R., and Mike Y. Chen. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science* 53 (9): 1375-1388.

DeHaan, Ed, Terry Shevlin, and Jacob Thornock. 2015. Market (in) attention and the strategic scheduling and timing of earnings announcements. *Journal of Accounting and Economics*, *60* (1): 36-55.

Dube, Svenja, and Chenqi Zhu. 2021. The disciplinary effect of social media: Evidence from firms' responses to Glassdoor reviews. *Journal of Accounting Research* 59(5): 1783-1825.

Dye, Ronald A. 1985. Disclosure of nonproprietary information. *Journal of Accounting Research*: 123-145.

Dye, Ronald A. 2001. An evaluation of “essays on disclosure” and the disclosure literature in accounting. *Journal of Accounting and Economics* 32 (1-3): 181-235.

Dyer, Travis, and Eunjee Kim. 2021. Anonymous equity research. *Journal of Accounting Research* 59 (2): 575-611.

Eren, Nevzat, and Han N. Ozsoylev. 2009. Hype and dump manipulation, Working Paper (University of Minnesota).

Fang, Lily, and Joel Peress. 2009. Media coverage and the cross - section of stock returns. *The Journal of Finance* 64 (5): 2023-2052.

Fang, Vivian W., Thomas H. Noe, and Sheri Tice. 2009. Stock market liquidity and firm value. *Journal of Financial Economics* 94 (1): 150-169.

Federal Communications Commission, 2016. 2016 Broadband Progress Report.

Feng, Zhiyue, Congping Pan, and Yibin Qian. 2014. Scientific transformation, path and strategy of university BBS development. *Advertising Panorama* (in Chinese) 12: 87-92.

Froomkin, A. Michael. 1995. Flood control on the information ocean: Living with anonymity, digital cash, and distributed databases. *Journal of Law and Commerce* 15: 395-507.

Gao, Yifei, and Lian Jiang, 2016. The development and regulation of the Internet real-name system. *Guangxi Social Sciences* (in Chinese) 2: 102-108.

Gillette, Jacquelyn R., Delphine Samuels, and Frank S. Zhou. 2020. The effect of credit ratings on disclosure: Evidence from the recalibration of Moody's municipal ratings. *Journal of Accounting Research* 58 (3): 693-739.

Goetzmann, William N., and Alok Kumar. 2008. Equity portfolio diversification. *Review of Finance* 12 (3): 433-463.

Gomez, Brad T., Thomas G. Hansford, and George A. Krause. 2007. The Republicans should pray for rain: Weather, turnout, and voting in US presidential elections. *The Journal of Politics* 69 (3): 649-663.

Green, T. Clifto, Russell Jame, Stanimir Markov, and Musa Subasi. 2014. Access to management and the informativeness of analyst research. *Journal of Financial Economics* 114 (2): 239-255.

Greene, William. 2004. The behaviour of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects. *The Econometrics Journal* 7 (1): 98-119.

Gurun, Umit G., and Alexander W. Butler. 2012. Don't believe the hype: Local media slant, local advertising, and firm value. *The Journal of Finance* 67 (2): 561-598.

Hales, Jeffrey, James R. Moon Jr, and Laura A. Swenson. 2018. A new era of voluntary disclosure? Empirical evidence on how employee postings on social media relate to future corporate disclosures. *Accounting, Organizations and Society* 68: 88-108.

Healy, Paul M., and Krishna G. Palepu. 2001. Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics* 31 (1-3): 405-440.

Hirshleifer, David, Sonya Seongyeon Lim, and Siew Hong Teoh. 2009. Driven to distraction: Extraneous events and underreaction to earnings news. *The Journal of Finance* 64 (5): 2289-2325.

Hong, Harrison, Wenxi Jiang, Na Wang, and Bin Zhao. 2014. Trading for status. *Review of Financial Studies* 27 (11): 3171–3212.

Hornik, Jacob, Rinat Shaanan Satchi, Ludovica Cesareo, and Alberto Pastore. 2015. Information dissemination via electronic word-of-mouth: Good news travels fast, bad news travels faster! *Computers in Human Behavior* 45: 273-280.

Huang, Kelly, Meng Li, and Stanimir Markov. 2020. What do employees know? Evidence from a social media platform. *The Accounting Review* 95 (2): 199-226.

Huang, Yuqin, Huiyan Qiu, and Zhiguo Wu. 2016. Local bias in investor attention: Evidence from China's Internet stock message boards. *Journal of Empirical Finance* 38: 338–354

Hutton, Amy P., Alan J. Marcus, and Hassan Tehranian. 2009. Opaque financial reports, R^2, and crash risk. *Journal of Financial Economics* 94 (1): 67-86.

Hutton, Amy P., Gregory S. Miller, and Douglas J. Skinner. 2003. The role of supplementary statements with management earnings forecasts. *Journal of Accounting Research* 41 (5): 867-890.

Huyghebaert, Nancy, and Weidong Xu. 2016. Bias in the post-IPO earnings forecasts of affiliated analysts: Evidence from a Chinese natural experiment. *Journal of Accounting and Economics* 61 (2-3): 486-505.

Ivkovic, Zoran, and Scott, Weisbenner. 2005. Local does as local is: Information content of the geography of individual investors' common stock investments. *The Journal of Finance* 60 (1): 267-306.

Jain, Bharat A., and Omesh Kini. 1994. The post‐issue operating performance of IPO firms. *The Journal of Finance* 49 (5): 1699-1726.

Jennings, Jared, Jung Min Kim, Joshua Lee, and Daniel Taylor. 2023. Measurement error, fixed effects, and false positives in accounting research. *Review of Accounting Studies* 29*:* 1-37.

Jiang, Guolin, Paul G. Mahoney, and Jianping Mei. 2005. Market manipulation: A comprehensive study of stock pools. *Journal of Financial Economics* 77 (1): 147-170.

Jin, Li, and Stewart C. Myers. 2006. R^2 around the world: New theory and new tests. *Journal of Financial Economics* 79 (2): 257-292.

Jorgensen, Bjorn N., and Michael T. Kirschenheiter. 2003. Discretionary risk disclosures. *The Accounting Review* 78 (2): 449-469.

Kahneman, Daniel. 1979. Prospect theory: An analysis of decisions under risk. *Econometrica* 47 (2): 363-391.

Kaminski, Margot. 2012. Real masks and real name policies: Applying anti-mask case law to anonymous online speech. *Fordham Intellectual Property, Media & Entertainment Law Journal* 23: 815-896.

Ke, Bin, and Xiaojun Zhang. 2021. Does public enforcement work in weak investor protection countries? Evidence from China. *Contemporary Accounting Research* 38 (2): 1231-1273.

Kelly, Bryan, and Alexander Ljungqvist. 2012. Testing asymmetric-information asset pricing models. *The Review of Financial Studies* 25 (5): 1366-1413.

Kim, Jeong-Bon, Yinghua Li, and Liandong Zhang. 2011. Corporate tax avoidance and stock price crash risk: Firm-level analysis. *Journal of Financial Economics* 100 (3): 639-662.

Kim, Yongtae, Myung Seok Park, and Benson Wier. 2012. Is earnings quality associated with corporate social responsibility? *The Accounting Review* 87 (3): 761-796.

Kogan, Shimon, Tobias J. Moskowitz, and Marina Niessner. 2021. Social media and financial news manipulation. Available at SSRN 3237763.

Kothari, Sabino P., Susan Shu, and Peter D. Wysocki. 2009. Do managers withhold bad news? *Journal of Accounting Research* 47 (1): 241-276.

La Porta, Rafael, Lorencio Lopez ‐ de ‐ Silanes, Andrei Shleifer, and Robert Vishny. 2002. Investor protection and corporate valuation. *The Journal of Finance* 57(3): 1147-1170.

Lee, Jyh-An, and Ching-Yi Liu. 2016. Real-name registration rules and the fading digital anonymity in China. *Washington International Law Journal* 25 (1): 1-34.

Lee, Yeunjae, and Katie Haejung Kim. 2020. De-motivating employees’ negative communication behaviors on anonymous social media: The role of public relations. *Public Relations Review* 46 (4): 101955.

Leuz, Christian, and Catherine Schrand. 2009. Disclosure and the cost of capital: Evidence from firms' responses to the Enron shock. No. w14897. *National Bureau of Economic Research*.

Leuz, Christian, and Robert E. Verrecchia. 2000. The economic consequences of increased disclosure. *Journal of Accounting Research* 38: 91-124.

Li, Feng. 2008. Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics* 45 (2-3): 221-247.

Li, Xiaodong 2015. CAC: Microblog and posts bars to adopt real name registration this year. *The Beijing Times*.

Lidsky, Lyrissa Barnett, and Thomas F. Cotter. 2007. Authorship, audiences, and anonymous speech. *Notre Dame Law Review* 82: 1537-1603.

Liu, Jianan, Robert F. Stambaugh, and Yu Yuan. 2019. Size and value in China. *Journal of Financial Economics* 134 (1): 48-69.

Loughran, Tim, and Bill McDonald. 2011. Measuring readability in financial disclosures. *The Journal of Finance* 69 (4): 1643-1671.

Malmendier, Ulrike, and Devin Shanthikumar. 2007. Are small investors naive about incentives? *Journal of Financial Economics* 85 (2): 457-489.

Miller, Gregory S., and Douglas J. Skinner. 2015. The evolving disclosure landscape: How changes in technology, the media, and capital markets are affecting disclosure. *Journal of Accounting Research* 53 (2): 221-239.

Mitts, Joshua. 2020. Short and distort. *The Journal of Legal Studies* 49 (2): 287-334.

Moore, Susanna. 2009. The challenge of Internet anonymity: Protecting John Doe on the Internet. *UIC John Marshall Journal of Information Technology & Privacy Law* 26 (4): 469-486.

Morck, Randall, Bernard Yeung, and Wayne Yu. 2000. The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics* 58 (1-2): 215-260.

Qin, Bei, David Strömberg, and Yanhui Wu. 2018. Media bias in China. *American Economic Review* 108 (9): 2442-2476.

Raddatz, Claudio. 2006. Liquidity needs and vulnerability to financial underdevelopment. *Journal of Financial Economics* 80 (3): 677-722.

Roberts, Michael R., and Toni M. Whited. 2013. Endogeneity in empirical corporate finance. *Handbook of the Economics of Finance* 2: 493-572.

Roll, Richard, Eduardo Schwartz, and Avanidhar Subrahmanyam. 2009. Options trading activity and firm valuation. *Journal of Financial Economics* 94 (3): 345-360.

Roychowdhury, Sugata. 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42 (3): 335-370.

Schipper, Katherine. 1989. Earnings management. *Accounting Horizons* 3 (4): 91-102.

Siano, Federico, and Peter D. Wysocki. 2018. The primacy of numbers in financial and accounting disclosures: Implications for textual analysis research. Available at SSRN 3223757.

Skinner, Douglas J. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research* 32 (1): 38-60.

Skinner, Douglas J. 1997. Earnings disclosures and stockholder lawsuits. *Journal of Accounting and Economics* 23 (3): 249-282.

Soltes, Eugene. 2014. Private interaction between firm management and sell-side analysts. *Journal of Accounting Research* 52 (1): 245-272.

Starr-McCluer, Martha. 2000. The effects of weather on retail sales. Working Paper, Federal Reserve Board.

State Council of China, 2013. State Council Institutional Reform and Functional Transformation Plan. (in Chinese).

Stock, James H., and Motohiro Yogo. 2005. Testing for weak instruments in Linear IV regression. *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*. Cambridge University Press. 80-108.

Surowiecki, James. 2004. The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business. economies, societies and nations. New York: Doubleday.

Tang, Vicki Wei. 2018. Wisdom of crowds: Cross-sectional variation in the informativeness of third-party generated product information on Twitter. *Journal of Accounting Research* 56 (3): 989–1034.

Van Bommel, Jos. 2003. Rumors. *The Journal of Finance* 58 (4): 1499-1520.

Verrecchia, Robert E. 1983. Discretionary disclosure. *Journal of Accounting and Economics* 5: 179-194.

Warfield, Terry D., John J. Wild, and Kenneth L. Wild. 1995. Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics* 20 (1): 61-91.

Xu, Sean Xin, and Xiaoquan Zhang. 2013. Impact of Wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quarterly* 37 (4):1043-1068.

Xu, Yongxin, Yuhao Xuan, and Gaoping Zheng. 2021. Internet searching and stock price crash risk: Evidence from a quasi-natural experiment. *Journal of Financial Economics* 141 (1): 255-275.

Yang, Kenneth CC, and Yowei Kang. 2016. Real-name registration regulation in China: An examination of Chinese netizens' discussions about censorship, privacy, and political freedom. *In Political Discourse in Emergent, Fragile, and Failed Democracies* IGI Global: 61-87.

You, Jiaxing, Bohui Zhang, and Le Zhang. 2018. Who captures the power of the pen? *The Review of Financial Studies* 31 (1): 43-96.

Table 1. Summary Statistics

The full sample consists of 14,770 firm-year observations from 2010 to 2019. Panel A reports summary statistics for the entire sample. Panel B compares the firm characteristics of treatment and control groups in the pre-treatment period. Treatment (control) firms have a number of Guba posts above (below) the sample median in 2014. Panel C reports the univariate tests that examine the impact of the real-name policy on firm value. Appendix A provides the variable definitions. All continuous variables are winsorized at the 1^{st} and 99^{th} percentiles.

Panel A. Summary Statistics of the Full Sample

Variable	N	Mean	P25	Median	P75	SD
Tobin's Q	14,770	1.902	0.767	1.277	2.223	1.930
Ln(Q)	14,770	0.303	-0.265	0.244	0.799	0.789
#Posts 2014	14,770	7,240	3,469	5,606	9,008	5,821
Ln(#Posts 2014)	14,770	1.731	1.244	1.724	2.198	0.690
Asset (in million RMB)	14,770	15055	1853	4308	11178	37303
Ln(Asset)	14,770	22.29	21.34	22.18	23.14	1.453
EBIT	14,770	4.880%	2.473%	4.668%	7.747%	6.820%
PPE	14,770	27.854%	11.279%	23.673%	40.684%	21.078%
Capex	14,770	4.575%	1.140%	3.124%	6.381%	4.735%
Book leverage	14,770	60.616%	38.463%	56.843%	75.412%	35.756%
Sales growth	14,770	19.700%	-4.806%	8.926%	25.276%	61.801%
Number of segments	14,770	5.246	2.000	3.000	7.000	5.400
Past stock return	14,770	17.607%	-20.008%	9.160%	48.001%	51.818%

Panel B. Firm Characteristics Between the Treatment and Control Groups in the Pre-treatment Period

	Treatment		Control		Test of difference	
	Mean	Median	Mean	Median	T-test	Wilcoxon test
	(1)	(2)	(3)	(4)	(1)-(3)	(2)-(4)
Tobin's Q	1.803	1.226	1.986	1.394	-0.183***	-0.168***
Ln(Q)	0.270	0.204	0.391	0.332	-0.121***	-0.128***
#Posts 2014	11,012	9,012	3,473	3,469	7,539	5,543
Ln(#Posts 2014)	2.284	2.199	1.179	1.244	1.105***	0.955***
Asset (in million RMB)	16120	4356	5246	2163	10754***	2193***
Ln(Asset)	22.362	22.195	21.544	21.495	0.818***	0.700***
EBIT	5.825%	5.273%	5.171%	4.956%	0.653%***	0.317%***
PPE	29.466%	25.143%	29.709%	26.023%	-0.243%	-0.880%
Capex	5.647%	4.183%	5.122%	3.595%	0.524%***	0.588%***
Book leverage	63.349%	60.249%	62.046%	57.941%	1.304%	2.308%***
Sales growth	21.356%	12.127%	19.999%	9.734%	1.358%	2.393%***
Number of segments	4.909	3.000	3.360	2.000	1.549***	1.000
Past stock return	23.104%	12.118%	23.046%	12.553%	0.058%	-0.435%

Panel C. Univariate Tests

	Treatment	Control	DiD
	(1)	(2)	(1)-(2)
Pre-event *Ln(Q)* (a)	0.270	0.391	
Post-event *Ln(Q)* (b)	0.162	0.389	
Difference=(b)-(a)	-0.108***	-0.002	-0.106***
	(0.000)	(0.899)	(0.000)

Table 2. Baseline Regression

Table 2 reports the general DiD tests that examine the effect of real-name policy on firm value. The dependent variable *Ln (Q)* is the natural logarithm of Tobin's Q. *Ln(#Posts 2014)* is the natural logarithm of Guba posts number in 2014. The indicator *After* equals 1 for the post-treatment period (2015–2019) and 0 for the pre-treatment period (2010–2014). The indicator $Before^n$ ($After^n$) equals 1 if it is *n* year before (after) the policy change and 0 otherwise. Appendix A provides the variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard error clustered by firm are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

DV=Ln(Q)	(1) (-1,+1)	(2) (-3,+3)	(3) (-5,+5)	(4) (-1,+1)	(5) (-3,+3)	(6) (-5,+5)	(7) (-5,+5)
Ln(#Posts 2014) ×After	-0.111***	-0.083***	-0.066***	-0.109***	-0.082***	-0.071***	
	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	
Ln(#Posts 2014) ×$Before^4$							-0.010
							(0.449)
Ln(#Posts 2014) ×$Before^3$							-0.007
							(0.613)
Ln(#Posts 2014) ×$Before^2$							-0.002
							(0.918)
Ln(#Posts 2014) ×$Before^1$							-0.001
							(0.948)
Ln(#Posts 2014) ×$After^1$							-0.121***
							(0.000)
Ln(#Posts 2014) ×$After^2$							-0.087***
							(0.000)
Ln(#Posts 2014) ×$After^3$							-0.044**
							(0.034)
Ln(#Posts 2014) ×$After^4$							-0.059***
							(0.004)
Ln(#Posts 2014) ×$After^5$							-0.063***
							(0.004)
Ln(Asset)				0.005	-0.295***	-0.351***	-0.351***
				(0.951)	(0.000)	(0.000)	(0.000)
EBIT				-0.240	0.511***	0.657***	0.659***
				(0.270)	(0.000)	(0.000)	(0.000)
PPE				-0.055	0.099*	0.047	0.047
				(0.627)	(0.057)	(0.284)	(0.292)
Capex				-0.009	0.193	0.328***	0.327***
				(0.977)	(0.157)	(0.003)	(0.003)
Book leverage				-0.014	-0.048**	-0.057***	-0.057***
				(0.866)	(0.050)	(0.005)	(0.005)
Sales growth				-0.003	0.013	0.022***	0.022***
				(0.870)	(0.158)	(0.004)	(0.004)
Number of segments				-0.003	0.001	0.003	0.003
				(0.561)	(0.742)	(0.159)	(0.168)
Past stock return				-0.032	0.189***	0.221***	0.222***
				(0.139)	(0.000)	(0.000)	(0.000)
Constant	0.656***	0.444***	0.360***	0.595	6.945***	8.113***	8.099***
	(0.000)	(0.000)	(0.000)	(0.756)	(0.000)	(0.000)	(0.000)
Firm_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,954	8,862	14,770	2,954	8,862	14,770	14,770
R-squared	0.877	0.811	0.766	0.876	0.846	0.830	0.830

Table 3. Propensity Score Matching

Table 3 examines the effect of real-name policy on firm value based on the propensity score matching method. Panel A reports the parameter estimates from the probit model used to estimate the propensity scores of the treatment and control groups. The dependent variable in the probit model is the *Treat* dummy, which equals 1 for firms with a number of Guba posts above the sample median in 2014, and 0 otherwise. Column (1) contains the probit model parameter estimates using the sample prior to matching. These estimates are then used to generate propensity scores for matching treatment and control firms. Column (2) contains the parameter estimates of the probit model using a subsample of matched treatment-control pairs. Panel B reports the balance test results for pairs of treatment and control firms after matching. Columns (1) and (2) of Panel C re-estimate columns (6) and (7) of Table 2 based on the propensity score-matched sample. Appendix A provides the variable definitions. P-values based on robust standard errors clustered by firm are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Propensity Score Matching and Diagnostic Regression

	Pre-matching (1)	Post-matching (2)
Ln(Asset)	0.364***	-0.041
	(0.000)	(0.536)
EBIT	-1.435**	0.129
	(0.024)	(0.890)
PPE	-0.210	-0.143
	(0.362)	(0.665)
Capex	-0.998	1.169
	(0.267)	(0.395)
Book leverage	-0.377***	-0.080
	(0.004)	(0.657)
Sales growth	0.064	0.072
	(0.393)	(0.561)
Number of segments	0.017*	-0.010
	(0.074)	(0.565)
Past stock return	-0.111	-0.066
	(0.359)	(0.716)
Q 3-year growth	0.090	0.010
	(0.271)	(0.935)
#Posts 3-year growth	0.220***	-0.008
	(0.000)	(0.783)
Constant	-7.797***	0.667
	(0.000)	(0.642)
Industry_FE	Yes	Yes
Observations	1,477	584
Pseudo R-Square	0.215	0.015
P-value	0.000	0.989

Panel B. Balance Tests

	Treatment	Control	Diff.	T-test	P-value
Ln(Asset)	21.917	21.950	-0.033	0.380	0.703
EBIT	4.880%	4.946%	-0.066%	0.130	0.898
PPE	28.824%	29.382%	-0.558%	0.320	0.746
Capex	4.659%	4.534%	0.124%	0.340	0.733
Book leverage	60.023%	60.461%	-0.438%	0.150	0.884
Sales growth	13.052%	11.711%	1.341%	0.320	0.749
Number of segments	3.887	4.065	-0.178	0.620	0.536
Past stock return	18.688%	20.693%	-2.005%	0.690	0.490
Q 3-year growth	11.292%	10.703%	0.588%	0.140	0.886
#Posts 3-year growth	-0.677%	-0.133%	-0.545%	0.320	0.749

Panel C. DiD Analysis Based on Matched Sample

	(1)	(2)
Treat×After	-0.096***	
	(0.001)	
Treat×Before^{4}		0.028
		(0.361)
Treat×Before^{3}		0.025
		(0.422)
Treat×Before^{2}		-0.021
		(0.542)
Treat×Before^{1}		-0.045
		(0.234)
Treat×After^{1}		-0.073**
		(0.047)
Treat×After^{2}		-0.097**
		(0.034)
Treat×After^{3}		-0.092**
		(0.039)
Treat×After^{4}		-0.096**
		(0.033)
Treat×After^{5}		-0.134***
		(0.005)
Controls	Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes
Year_FE	Yes	Yes
Observations	5,840	5,840
R-squared	0.813	0.813

Table 4. Two-stage Least Squares Regressions

Table 4 reports the two-stage least squares regression that examines the effect of the real-name policy on firm value. Column (1) reports the first-stage regression with *Ln(#Posts 2014)* as the dependent variable, and the province-level *Bad weather 2014* as the instrumental variable. *Bad weather 2014* is the average proportion of days with bad weather across all cities in a firm's headquarters' province in 2014, where bad weather refers to heavy rainfall, heavy snow, hail, sandstorm, violent wind (Beaufort scale greater than 7), and extreme temperature (>35 or <-25 degrees Celsius). Appendix A provides the variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard errors clustered by firm are reported in parentheses. ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

	First stage predicting *Ln(#Posts 2014)*	Second stage
DV=	*Ln(#Posts 2014)*	*Ln(Q)*
	(1)	(2)
Bad weather 2014	1.336***	
	(0.000)	
Ln(#Posts 2014)×After		-0.164***
(instrumented)		(0.000)
Controls	Same as those in Table 2 column (6)	
F-statistics	12.16***	
Industry_FE	Yes	No
Firm_FE	No	Yes
Year_FE	No	Yes
Observations	1,477	14,770
R-squared	0.206	0.832

Table 5. Market Reaction around the Announcement Date of the Real-name Policy

Panel A of column (1) reports the average accumulated stock return in the five-day window (*AR[-2,+2]*) around the announcement date of the real-name policy (February 4, 2015). Column (2) of Panel A reports the average *AR[-2,+2]* around February 4 of the other nine years (i.e., 2010–2014 and 2016–2019). Columns (1) and (2) of Panel B report the regression examining the cross-sectional variation in *AR[-2,+2]* around February 4, 2015, and February 4 in the other nine years, respectively. Appendix A provides the variable definitions. P-values based on robust standard errors clustered by firm are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Accumulative Stock Return around Policy Announcement Date

	AR [-2,+2] February 4, 2015	*AR [-2,+2]* February 4 of 2010–2014 and 2016–2019	Difference
	(1)	(2)	(1)-(2)
Treatment (a)	-2.497%***	1.068%***	-3.565%***
Control (b)	-1.785%***	1.193%***	-2.978%***
Difference=(b)-(a)	-0.712%***	-0.125%	-0.587%***
	(0.009)	(0.211)	(0.008)

Panel B. Cross-sectional Regression

	AR [-2,+2] February 4, 2015	*AR [-2,+2]* February 4 of 2010–2014 and 2016–2019
	(1)	(2)
Ln(#Posts 2014)	-0.006***	0.000
	(0.004)	(0.492)
Controls	Same as those in Table 2 column (6)	
Industry_FE	Yes	Yes
Year_FE	Yes	Yes
Observations	1,477	13,293
R-squared	0.153	0.013

Table 6. Channel Tests 1: Heterogeneous Treatment Effects Based on Reliance on Social Media

Table 6 re-estimates column (6) of Table 2 based on various subsamples. The subsamples are based on upper (High group) and lower (Low group) terciles of each of the following four variables. *Analyst coverage* is the number of annual earnings forecasts made by analysts on a given firm. *#Newspaper* is the number of newspapers reports covering the firm. *Institutional ownership* is the percentage of common shares outstanding held by institutional investors. *Mobile usage* is the number of local mobile Internet users normalized by the total population. All these four variables are measured in 2014. The Wald test examines the equivalence of the coefficients for *Ln(#Posts 2014)×After* between the high and low groups. Variable definitions are provided in Appendix A. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard errors clustered by firm are reported in parentheses. The superscripts ***, **, and * indicate significance at the 1%, 5% and 10% levels, respectively.

DV=Ln(Q)	*Analyst coverage*		*#Newspaper*		*Institutional ownership*		*Mobile usage*	
	High	Low	High	Low	High	Low	High	Low
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(#Posts 2014) ×After	-0.030	-0.074***	-0.035	-0.069**	-0.036	-0.081***	-0.099***	-0.046**
	(0.103)	(0.003)	(0.104)	(0.015)	(0.103)	(0.001)	(0.000)	(0.042)
Wald Test	5.31**		2.96*		4.96**		8.10**	
P-value	0.021		0.085		0.026		0.004	
Controls	Same as those in Table 2 column (6)							
Firm_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,820	5,410	4,880	4,940	4,920	4,930	4,060	5,050
R-squared	0.863	0.812	0.849	0.810	0.847	0.825	0.839	0.835

Table 7. Channel Test 2: Impact on Social Media

Panel A examines the effect of the real-name policy on the number and structure of social media posts; the unit of observation is the firm-year. Panel B examines the predictive power of social media posts on firms' quarterly earnings; the unit of observation is the firm-quarter level. Panel C examines stock reactions toward quarterly earnings announcements; the unit of observation is at the firm-quarter level. Panel D reports the placebo test that examines the effect of the real-name policy on newspaper reports; the unit of observation is at the firm-year level. Appendix A provides the variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard errors clustered by firms are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Number and Structure of Social Media Posts

	Ln(#Posts)	Ln(#words each post)	Proportion of rehashed posts	Proportion of first-person posts	Proportion of quantitative posts	#Comments per post	Net proportion of negative posts	Agreement
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ln(#Posts 2014)×After	-0.176***	-0.043***	0.269***	-0.141**	-0.211**	-0.076***	-0.382***	0.308***
	(0.000)	(0.000)	(0.000)	(0.027)	(0.036)	(0.004)	(0.004)	(0.005)
Controls	Same as those in Table 2 column (6)							
Firm_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,770	14,770	14,770	14,770	14,770	14,770	14,770	14,770
R-squared	0.524	0.411	0.548	0.608	0.769	0.718	0.429	0.459

Panel B. Predicting Earnings Surprise based on Social Media Posts

	Negative surprises	Positive surprises
	(1)	(2)
Proportion of negative posts	-0.036***	
	(0.007)	
Proportion of negative posts×After	0.028**	
	(0.037)	
Proportion of positive posts		0.021**
		(0.023)
Proportion of positive posts×After		-0.002
		(0.898)
Controls	Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes
Quarter_FE	Yes	Yes
Observations	31,357	29,787
R-squared	0.111	0.069

Panel C. Stock Reaction toward Earnings Surprise

	CAR[-2,+2]	
	Negative surprises	Positive surprises
	(1)	(2)
Ln(#Posts 2014)×After	-0.003***	0.001
	(0.007)	(0.327)
Wald Test	4.77**	
P-value	0.029	
EarningSurprise	0.022***	0.032***
	(0.000)	(0.000)
Controls	Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes
Quarter_FE	Yes	Yes
Observations	29,549	25,916
R-squared	0.034	0.048

Panel D. Placebo Tests on Newspaper

	#Newspaper	Net proportion of negative newspaper reports
	(1)	(2)
Ln(#Posts 2014)×After	-0.002	-0.002
	(0.943)	(0.821)
Controls	Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes
Quarter_FE	Yes	Yes
Observations	14,770	14,770
R-squared	0.788	0.226

Table 8. Channel Tests 3: Effect on Firms' Information Environments

Table 8 examines the effect of the real-name policy on firms' information environments. The regression specification follows that in column (6) of Table 2. Appendix A provides the variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard errors clustered by firm are reported in parentheses. The ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

	Stock informativeness	Bid-ask spread	Crash risk	Voluntary disclosure	Earnings management	Restatement
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(#Posts 2014) ×After	-0.037**	0.055***	0.047***	-0.037**	0.012**	0.017**
	(0.025)	(0.000)	(0.008)	(0.044)	(0.047)	(0.026)
Controls	Same as those in Table 2 column (6)					
Firm_FE	Yes	Yes	Yes	Yes	Yes	Yes
Year_FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,769	14,770	14,765	13,824	13,754	14,770
R-squared	0.572	0.680	0.096	0.467	0.352	0.093

Table 9. Alternative Channels: Overvaluation and/or Attention

Panel A re-estimates column (6) of Table 2 using alternative methods to define the treatment intensity. *#Negative posts 2014* is the number of Guba posts with a negative tone in 2014. *#Non-positive posts 2014* is the number of Guba posts with non-positive tones (i.e., negative tones and neutral tones) in 2014. *Net proportion of negative posts 2014* is the difference between the number of posts with a negative tone and the number of posts with a positive tone scaled by the total number of Guba posts in 2014. Panel B re-estimates column (6) of Table 2 using proxies for investor attention as the dependent variable. The *Overall Index, Mobile Index*, and *PC Index* measure the Baidu search volume for a given stock on all devices, mobile phones, and personal computers. Appendix A provides the variable definitions. All continuous variables are winsorized at the 1st and 99th percentiles. P-values based on robust standard errors clustered by firms are reported in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Panel A. Posts with Negative Tones as Alternative Measures of Treatment Intensity

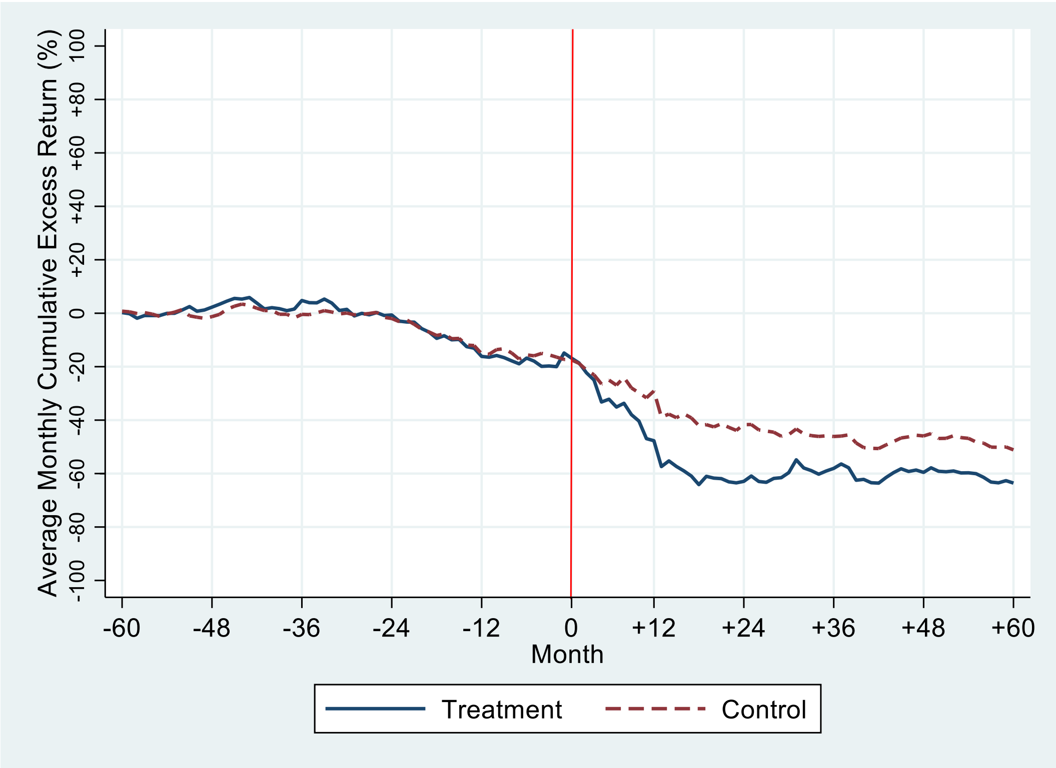
DV=Ln(Q)	(1)	(2)	(3)
Ln(#Negative posts 2014)×After	-0.069***		
	(0.000)		
Ln(#Non-positive posts 2014)×After		-0.071***	
		(0.000)	
Net proportion of negative posts 2014×After			-0.440**
			(0.022)
Controls		Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes	Yes
Quarter_FE	Yes	Yes	Yes
Observations	14,770	14,770	14,770
R-squared	0.830	0.830	0.829

Panel B. Investor Attention around the Policy Change

	Overall Index	*Mobile Index*	*PC Index*
	(1)	(2)	(3)
Ln(#Posts 2014)×After	0.022	-0.033	0.011
	(0.771)	(0.649)	(0.900)
Controls		Same as those in Table 2 column (6)	
Firm_FE	Yes	Yes	Yes
Quarter_FE	Yes	Yes	Yes
Observations	13,293	13,293	13,293
R-squared	0.928	0.922	0.899

Figure 1. Accumulative Stock Returns around the Real-name Policy

Figure 1 shows the average monthly cumulative excess returns of the treatment and control firms. The monthly cumulative excess return is calculated as the sum of the monthly returns in excess of the market return (measured by the Shanghai and Shenzhen A-share market indices). Treatment (control) firms have a number of Guba posts above (below) the sample median in 2014.



Appendix A. Variable Definitions

Variable	Definition
#Comments per post	Number of reply comments for each post.
#Newspaper	Number of newspaper reports covering a given firm.
#Posts	Number of Guba posts about a given firm.
#Posts 3-year growth	Three-year average annual growth rate in the number of Guba posts.
#words each post	Number of words contained in each post.
After	Indicator variable that equals 1 for the post-treatment period (2015–2019) and 0 for the pre-treatment period (2010–2014).
Agreement	Contributors' agreement level, calculated as $1-\sqrt{1-B^2}$. B=(number of posts with a negative tone – number of posts with a positive tone)/(number of posts with a negative tone +number of posts with a positive tone).
AR[-2,+2]	Five-day accumulated stock return.
Analyst coverage	Number of annual earnings forecasts made by analysts on a given firm.
Asset	Book value of total assets.
Bad weather 2014	Ratio of number of days with bad weather across all cities in a firm's headquarter province in 2014 normalized by 365, where bad weather refers to heavy rainfall, heavy snow, hail, sandstorm, violent wind (Beaufort scale greater than 7), and extreme temperature (>35 or <–25 degrees Celsius).
Bid-ask spread	We first compute the average daily bid-ask spread, and then we take natural logarithm to correct for non-normality in the distribution (Fang, Noe and Tice, 2009).
Book leverage	Book value of total debt normalized by book value of total assets.
Capex	Capital expenditures normalized by book value of total assets.
CAR[-2,+2]	Five-day accumulative stock return minus the value-weighted size matched portfolios return (Andrei, Friedman and Ozel, 2023).
Crash risk	Negative of the third moment of firm-specific weekly returns for each year, divided by the SD of firm-specific weekly returns raised to the third power.
Earnings management	Abnormal production costs minus abnormal cash flow from operations and abnormal discretionary expenses normalized by book value of total assets.
Earnings surprise	Difference between quarterly earnings per share (EPS) and EPS in the same quarter of the previous year normalized by the absolute value of the latter.
EBIT	Earnings before interest and taxes normalized by book value of total assets.
Institutional ownership	Percent of common shares outstanding held by institutional investors.
Mobile usage	Number of local mobile Internet users normalized by the total population.
Net proportion of negative newspaper reports	(Number of newspaper reports with a negative tone minus number of newspaper reports with a positive tone) / total number of newspaper reports.
Net proportion of negative posts	*Proportion of negative posts* minus *Proportion of positive posts*.
Number of segments	Number of a firm's geographic segments.
Past stock return	Stock return in the previous 12 months of the fiscal yearend.
PPE	Property, plant, and equipment normalized by book value of total assets.
Proportion of first-person posts	Number of Guba posts with first-person expressions normalized by the total number of Guba posts.
Proportion of negative posts	Number of Guba posts with a negative tone normalized by the total number of Guba posts.
Proportion of positive posts	Number of Guba posts with a positive tone normalized by the total number of Guba posts.
Proportion of quantitative posts	Number of quantitative posts normalized by the total number of posts. A quantitative post is the one that contains at least 10 words and a number in the following cases: (i) the number is preceded by a dollar/Chinese Yuan sign ("$/￥"); (ii) the number is followed by the words "million"/"billion"/"trillion"; (iii) the number is followed by a percentage sign ("%") or by the word "percent"; (iv) the number is followed by an unit symbol for kilometer (km), kilogram (kg), and kilowatt (kW), square meters (m^2), etc.
Proportion of rehashed posts	Number of rehashed posts normalized by the total number of posts. A rehashed post is the one that predominantly comprises quotes sourced from public materials (e.g., corporate financial reports, announcements, analysts' reports, newspaper reports, and other social media posts) and has less than 10 words excluding the quotes.
Q 3-year growth	Three-year average annual growth rate in Tobin's Q.

Restatement	Indicator variable that equals 1 if the firm announces the correction of misstatement in a subsequent reporting period by the end of 2022.
Sales growth	Percentage change in sales compared to the previous year.
Stock informativeness	Logit transformation of (1-R^2), where R^2 is obtained from the expanded value-weighted market and value-weighted industry index model using daily stock returns.
Tobin's Q	Market value of equity and book value of long-term debt divided by book value of total assets.
Treat	An indicator variable equals 1 for the firms with the number of Guba posts above the sample median in 2014 and 0 otherwise.
Voluntary disclosure	Natural logarithm of the number of words in MD&A about future.

Appendix B. Data from Social Media

We collect the Guba data from Eastmoney (http://guba.eastmoney.com), one of the largest and most comprehensive social media platforms focusing on the stock market in China. The platform includes all stocks traded in the market, with independent subforums named according to the stock ticker or name of each company. We develop a web crawler to download all the posts (including posts and replies) from the discussion boards of each company.

To construct the measurement of post tone, we utilize machine learning techniques to measure the tone of Guba posts using Chinese sentiment dictionary from Dalian University of Technology, Tsinghua University, CNKI, and Chinese translation of the dictionary of Loughran and McDonald. We label the tones of 5,000 posts as the training set and then classify all the posts as positive, negative, or neutral using the Naive Bayes Classification (NBC). As social media users often use emojis to express their opinions, we also consider emojis in our model to define the tone. The in-sample accuracy reaches 85.4%, higher than that in Antweiler and Frank (2004) (72.3%) using NBC for Yahoo! Finance messages. The out-of-sample validation using 1,000 randomly selected posts shows that the accuracy rate of our model is above 78% in labeling the tone of the posts. More importantly, the rate of false classification (i.e., positive posts classified as negative, and vice versa) is low (0.4% and 0.2%, respectively). Our final sample includes approximately 125.81 million posts and 399.30 million comments from 2010 to 2019.