

Firm News and Market Views: The Informational Role of Official Newspapers in China

Joseph D. Piotroski*

Stanford Graduate School of Business

T.J. Wong

USC Marshall School of Business

Shubo Zhang

Chinese University of Hong Kong

Tianyu Zhang

Chinese University of Hong Kong

Abstract: Prior research documents that official “party” newspapers in China publish business news articles that are more positively biased and contain greater political content, and as a result, are relatively less informative than articles published by commercially-oriented, non-official newspapers. We posit and find that despite these political biases, official newspapers serve an important informational role by conveying *relatively more* value-relevant industry and market-wide information than non-official newspapers. We show that the strength of this information role varies based on the political proximity of the newspaper to the central government and intensifies during highly politicized time periods. Our results suggest that politically captured, state-owned newspapers can serve an important and potentially unique capital markets-oriented information role in developing economies.

Please Do Not Circulate Without Permission

25 October 2019

* Corresponding Author. The authors would like to thank Gwen Yu and workshop participants at the University of Southern California China Conference and the University of Melbourne for helpful suggestions and comments on draft version of this paper. The authors would like to thank the Stanford Graduate School of Business, USC Marshall School of Business, and the CUHK Business School for financial support.

1. Introduction

In an autocracy, it is unclear whether increasing (decreasing) state control of the media will uniformly decrease (increase) the value of the information supplied to the market. On the one hand, state-controlled media have strong incentives to bias corporate news for political purposes, with these biases adversely affecting the information value of their reports. On the other hand, state media outlets can serve as channels for the government to credibly convey information about economic and industrial policies to market participants. Relaxing the state's control of the media should allow affected outlets to reduce biases, to focus on providing information that readers demand, and to increase the overall value of their articles. However, a reduction in political dependency could also weaken these outlets connection to the government and jeopardize their ability to obtain information about the government's economic agendas and policies, thus actually reducing the value of their reports. Thus, the expected impact is dependent upon the type of information these news organizations supply to the market.

In this paper, we examine whether differences in the degree of state control over the media affect the type of information supplied to the market. Our paper addresses this issue by investigating corporate news articles published in China. China is a natural setting to examine articles published by state-owned media. Similar to most other developing economies, Chinese newspapers are directly or indirectly controlled by the government or party affiliated organizations, with editorial boards and journalists subject to political oversight and incentivized to cater to politicians (and, in this context, members of the Communist Party of China). Due to these political pressures, newspapers in China publish articles that are, on average, positively biased and frequently contain political slogans or support current political orthodoxy.¹ That said, China's newspapers must also serve the market economy. In response to demand for business news and entertainment, China sanctioned the creation of "non-official" newspapers.² These newspapers are still state-controlled, but have stronger commercial incentives and a clear mandate to

¹ See Stockmann (2012) for a comprehensive review.

² See Hu (201) for a summary and discussion of the rise of the business news media in China.

attract and serve a broader readership. This is in stark contrast to “official” newspapers, which are closely aligned to their sponsoring organization and are required to serve as the government’s mouthpiece.

Prior research establishes that, due to these structural differences, non-official newspapers publish articles that are less positively biased and contain less political rhetoric than official newspapers (Stockmann, 2012; Piotroski, Wong, and Zhang, 2017; You, Zhang, and Zhang, 2018; Qin, Stromberg, and Wu, 2018). More importantly, these differences impact the relative effectiveness of these newspapers at fulfilling their informational and governance roles. As information mechanisms, non-official newspapers publish more timely articles and convey incrementally more information than official newspapers, as measured by the average market reaction to the publication of a signed news article and the strength of the correlation between the article’s tone and the covered firm’s future performance. And, unlike official newspapers, non-official newspapers also seem to serve an important monitoring and corporate governance role, as captured by subsequent executive turnover at Chinese organizations following the publication of negative news stories (You, Zhang, and Zhang, 2018). Thus, prior research concludes that official newspapers, because of their coverage and reporting biases, convey less information to the capital markets than non-official newspapers and perform no governance role.

This extant empirical literature, however, fails to consider that these two types of newspapers may have incentives or strategic advantages to focus on different business topics, thus providing different information to market participants.³ We posit that official and non-official newspapers serve two different informational roles in the Chinese media, and thus their corporate news articles will potentially convey different information to the financial markets. Given their market orientation and commercial incentives,

³ It is important to recognize that the analyses and conclusions of You, Zhang, and Zhang (2018) primarily reflect only one dimension of article informativeness – the timely reporting of bad, firm-specific news. For example, to measure the timeliness of an article, they examine how long after the initiation of an accounting fraud does a newspaper report on the event. This event has two characteristics – it is firm-specific and it is a negative event. Prior research suggests that non-official newspapers have stronger incentives to write an article about this type of negative event, as compared to news articles about a positive event (e.g., IPO, opening of a new manufacturing facility) or an economic outcome that speaks to broader macro-economic conditions (e.g., higher than expected sales in the quarter). Similarly, the newspaper articles used in You, Zhang, and Zhang (2018) to examine the governance role of these newspapers are again firm-specific in nature and reflect the publication of negative information that leads to executive turnover.

non-official business newspapers are expected to focus on conveying credible information about the firm-specific activities of the covered firm and the decisions of its executives. In contrast, because of their closer proximity to the government and the party, official newspapers are in a position to gather and credibly convey information about government objectives, plans, and policy that would have broad implications to specific industries or the economy at large. As such, we expect official newspaper articles to be tilted more towards the coverage of these industry and market-wide topics when reporting on individual firms.

To investigate this prediction, we gathered a comprehensive sample of domestic corporate news articles written about Chinese listed firms over the period 2000 – 2017. Our main sample consists of 2.78 million articles drawn from 52 official and 48 non-official business newspapers across China. For each news article, we estimate the market reaction on its publication date, the fraction of the market reaction related to the incorporation of firm-specific versus industry and market-wide news, and the percent of the articles' textual content that focuses on firm-specific versus industry and market-wide topics.

First, we confirm that Chinese business newspaper articles have information content, as measured by the absolute stock market reaction to their publication. Consistent with You, Zhang, and Zhang (2018), we find that publication of a corporate news article by a non-official newspaper generates a significant market reaction, suggesting that these depoliticized newspapers' articles convey value relevant information to market participants. More importantly, we confirm that the articles published by official newspapers also generate a significant, albeit smaller, market reaction, suggesting that these politicized newspapers' articles also convey useful information to China's capital markets.

Given that both sets of articles have information content, we next focus on identifying the relative *type* of information conveyed by these corporate news articles. Following the methodology in Liu (2011), we decompose daily returns into three components – a market component, an industry component, and a firm-specific components – and use this decomposition to measure the relative amount of industry, market, and firm-level news influencing the stock price of a given firm on a given day. We find that, relative to the average trading day without news coverage, the portion of industry and market news versus

firm-specific news influencing stock prices shifts around the publication of a corporate news article about the covered firm. Specifically, a smaller fraction of the firm's daily stock return is related to industry or market news on the date these articles are published (i.e., a greater fraction of the firm's daily stock return relates to firm-specific information component), consistent with corporate news articles conveying, on average, incremental, value-relevant firm-specific information to the financial markets. This firm-specific information effect is actually larger than the shift observed on earnings announcement dates, suggesting that our business news articles (and other underlying information events) are preempting the release of firm-specific earnings news in China. Interestingly, the shift from industry and market news to firm-specific news is more (less) pronounced when non-official (official) newspapers publish the articles. This observed difference in the *relative* components of the market reaction to these news articles suggests that official and non-official newspapers are potentially conveying different types of information to the market.⁴

To test this conjecture, we next use textual analysis techniques to identify the topics contained in these news articles, and use this data to measure the proportion of each article's textual content that is related to industry and/or market related topics versus firm-specific topics. First, we find that articles published by official newspapers contain a greater proportion of industry and market-wide topics than articles published by non-official newspapers. Second, we find that the resultant empirical measures of article content are positively correlated with the components of the observed market reaction, namely, articles containing a greater (lesser) preponderance of market and industry-wide topics are associated with publication returns that have larger industry and market (firm-specific) return components, *ceteris paribus*. And, interestingly, this mapping of industry and market-wide news topics into these articles' publication returns' components is significantly stronger for official newspapers. Taken together, these stock return and text-based analyses establish that official and non-official newspapers convey different types of information, consistent with our main conjecture. This conclusion is then confirmed by

⁴ Our estimations suggesting that the average official newspaper article conveys 10% more (less) industry and market-level (firm-specific) news about the covered firm than comparable non-official business newspaper article.

examining the relation between stock return synchronicity and annual newspaper activity in China, where greater official (non-official) newspaper activity of the covered firm over the calendar year is found to be associated with an increase (decrease) in the stock return synchronicity over the same time period, similar to the relative informational effects observed for financial analyst coverage (executives' insider trades) in prior research (e.g., Piotroski and Roulstone, 2004).

The remainder of our paper examines whether and how political forces interact with this unique informational role of official party newspapers. Preliminary tests reveal four key findings. First, we find that official newspapers controlled by the central government convey relatively more industry and market level information than official newspapers controlled by local governments, consistent with central government newspapers having greater access to policy-relevant information, wider circulation (i.e., broader dissemination of the information), and greater perceived credibility. Second, all newspapers in China convey relatively more industry and market information during National Congress periods, consistent with the newspapers' in-depth (and perhaps disproportionate) coverage of those major political events. Third, we find that observed differences in the type of information conveyed by non-official and official newspapers widened following President Xi's visit to leading state media outlets on February 19th, 2016, with official (non-official) newspaper conveying relatively more industry and market-wide (firm-specific) information after this high profile "loyalty" event. Lastly, we show that industry information content of official newspaper's articles is incrementally stronger on days when the central government introduced new industrial economic policies.

Our paper contributes to the literature in several unique ways. First, we show that official and non-official newspapers serve different informational roles in China's capital markets. Specifically, the two newspapers focus on publishing different types of information, and that this difference in the mix of topics explains differences in the market reaction to the information. Non-official newspapers create and disseminate credible firm-specific information about the covered firm, thus fulfilling an important information intermediary role in China's capital markets. Official newspapers, despite their role as the Party's mouthpiece, also play an important role by delivering industry-level and market-wide information

to the capital markets, and that the information content of official newspapers' articles varies by the paper's nexus to the central government, contemporaneous political events, and in response to heightened political pressures. Given that economic growth and government policy is central to emerging market firms' investment opportunities and performance, official newspapers appear to be serving a critical, and complimentary, informational role in China's capital markets at this stage of country's development.

Second, we show how the publication activity of state-controlled newspapers is affecting the stock return synchronization of China's listed firms. We find that greater intensity of coverage by non-official newspapers is associated with a reduction in stock return synchronicity of the listed firm, consistent with these newspaper improving the flow firm-specific information into prices. More importantly, we show that greater intensity of coverage by official newspapers is associated with an incremental increase in stock return synchronicity. This evidence suggests that the current structure of the Chinese media, where a significant proportion of corporate news articles are produced by official newspapers, may be a source the unusually high level of price synchronization observed in the China economy (e.g., Morck, Yeung, and Yu, 2000).

Third, more generally, we contribute towards an understanding of the information environment supporting China's capital markets (see Piotroski and Wong (2012) for a summary). And, because China is the largest economy practicing state capitalization, our paper serves as an important first step towards understanding whether, and under what contextual conditions, China's state-controlled information institutions will be able to fulfill their market development roles within the country's broader institutional, political and economic framework.

Finally, our results need to be interpreted with caution. First, domestic news articles could reflect the dissemination of information already released by the firm or government bureau on the same date (e.g., management forecast, ad hoc report, dividend announcement, etc.). In such a case, the observed market reaction is confounded by the effects of the concurrent, omitted disclosure event, making it difficult to distinguish how much of the reaction is due to the newspaper's role as a dissemination vehicle. Prior research shows that these types of information events occur with a fairly low frequency, suggesting

that our results may be separable from such correlated events; future versions of the paper will attempt to control for these types of information events directly, similar to our current control for contemporaneous earnings announcements. Second, our measure of information content – i.e., the absolute market reaction to the publication of the news article – implicitly assumes that market prices efficiently impound new information. To the extent that market prices are inefficient, or that market participants in China are not efficient processors of information, our results instead reflect the market’s *perception* of an article’s relative informativeness as opposed to its actual information content. Regardless, our evidence remains relevant, as ours is the first meaningful study of the types of business topics covered in Chinese financial news media and how the market reacts to this information.

2 Background on Chinese Newspapers and Hypothesis Development

2.1 Information content of Chinese corporate news articles

Shortly after the founding of the People’s Republic of China in 1949, all published news articles were delivered through China’s state-owned and state-supported official newspapers. These official newspapers (so called “party papers”) traditionally received financial support from the government, have a well-defined and installed readership base, and serve the primary goal of conveying the government’s political narrative to its constituents. Following the introduction of economic reforms after 1979, the government commercialized the press by creating a new category of newspapers – non-official newspapers – that operate like stand-alone commercial units. These non-official newspapers receive no ongoing financial support from the government for day-to-day operations and target a specific audience to fill a particular market demand for information. As a result of these differences, non-official newspapers have greater freedom and financial incentives to respond to market demands for news in order to maximize advertising and subscription revenues, and accordingly, tend to publish news articles that are significantly less biased than similar articles published by official party newspapers (Stockmann, 2012). And, as a result of this liberalization of the media market, China experienced a ten-fold growth in the number of newspapers in the country over the period 1979 to 2009, suggesting that the overall

information environment in China may have potentially improved with the introduction of these less politicized news media outlets.

Focusing on corporate news reporting, the reopening of China's stock exchanges, and the impact of the Asian Financial Crisis, reinforced the need for a domestic, market-oriented business media to provide heightened transparency for the capital markets. These views were ultimately supported and endorsed by the Chinese government, and as a result, a number of high-profile, non-official newspapers emerged in the late 1990's and early 2000's.⁵ These non-official business publications are characterized by their focus on corporate and financial news and almost propaganda free reporting. And, consistent with earlier evidence on the general media in China, corporate news articles published in non-official business newspapers are less biased and less politically tainted than corporate news articles published in by official newspapers (Piotroski, Wong, and Zhang, 2017; You, Zhang, and Zhang, 2018; Qin, Stromberg, and Wu, 2018), suggesting that these newspapers are likely filling an important information role for capital market participants.

However, despite these reforms and prevailing commercial incentives of non-official newspapers, the government continues to have direct and indirect control of all media outlets through ownership and the appointment of editorial boards. These channels of control provide politicians and party official a conduit to shape editorial practices and article content for both types of newspapers. By design, official party papers are expected to publish articles that cater to politicians, the party, and/or the newspaper's sponsoring governmental or party unit, reducing the expected information content and credibility of official newspaper articles. These same political incentives could also influence the editorial decisions of non-official newspapers, especially if the newspaper's editorial board and journalists have concerns about the risk imposed upon the newspaper's license or their own careers by publishing news articles that "cross a political line." This political influence means that political incentives could also adversely affect the credibility and information content of news articles published by non-official newspapers. This concern is

⁵ Examples include *Caijing*, *21st Century Business Herald*, *Economic Observer*, *New Fortune Magazine*, and *Business Watch*.

reinforced by existing evidence showing that although the bias in articles published by non-official newspapers is smaller than the bias associated with official newspaper articles, the average bias is also non-zero (Stockmann, 2012; Piotroski, Wong and Zhang, 2017). Thus, to the extent that such biases swamp the information content of the news article, it is plausible that even business news articles published by non-official newspapers lack information content.

These arguments lead to our first set of null (alternative) hypotheses. First, focusing on the absolute information content of Chinese business news articles, we predict:

Hypothesis 1: The average business news article published in China does not provide (provides) incremental information to the capital markets.

Second, focusing on differences in the relative amount of information conveyed by the business news articles of official and non-official newspapers, we predict:

Hypothesis 2: The information content of business news articles published by official newspapers is equal to (not equal to) the information content of business news articles published by non-official newspapers.

Our empirical evidence on these two hypotheses is presented in Section 4.

2.2 Relative information content of China's business news articles

Prior research establishes that individual market participants have different relative informational advantages with respect to firm, industry, or market-level information, with their resultant information-based activities having a differential impact on the price formation process. For example, insiders have an advantage with respect to acquiring and disseminating firm-specific information, as evidenced by their trading behavior in advance of firm-specific announcements (Seyhun 1992, 1998; Meulbroek 1992; Damodoran and Liu 1993) and the net negative impact of their private trading activity on stock price synchronicity (Piotroski and Roulstone, 2004). Conversely, financial analysts have an advantage with respect to the production, interpretation, and dissemination of industry and market level information and trends, as evidenced by the net positive impact of their reporting activity on stock return synchronicity

(Piotroski and Roulstone, 2004; Crawford, Roulstone, and So, 2012) and subsequent revision of analysts' forecasts in response to the earnings announcements of other firms in the same industry (Ramnath, 2002).

In the context of China's newspapers, it is possible that official and non-official newspapers convey different information to the financial markets due to their relative advantages with respect to acquiring, interpreting, and disseminating firm-level versus industry and market-wide information. Because of their commercial incentives, non-official newspapers may be more likely to engage in firm-specific reporting than official newspapers, either to increase readership through investigative reports or to cater to the covered firms (through the publication of favorable news stories) in exchange for advertising revenue (or in the extreme, journalist compensation). In both cases, these reporting activities can lead to an increase in the supply of firm-specific information about the covered firm. Similarly, non-official newspapers may be more inclined to report on factors behind the recent stock price or earnings performance of individual firms, with this focus on idiosyncratic performance drivers also increasing the supply of firm-specific information about the covered firm.

In contrast, official newspapers enjoy a close relationship with their sponsoring governmental agencies and party units. Due to this relation, official newspapers may become conduits for conveying political opinions and policy-related information to the market. And, even in the absence any specific communication directives from their sponsoring entities, official newspapers are also likely to have greater access to economic data, more likely to cover political events, and more likely to frame corporate news in the broader political / economic policy context than non-official newspapers. Thus, despite their well-known, pro-government biases, official newspaper could actually convey timely information about current and future government policy, industry-level trends and regulation, and/or macro-economic data when reporting on individual companies, thus increasing the relative supply of industry and market-wide information available to capital market participants.

These arguments lead to our final set of null (alternative) hypotheses. First, focusing on differences in the relative type of topics covered by the business news articles of official and non-official newspapers, we predict:

Hypothesis 3: The relative amount of industry and market-wide topics contained in official newspaper articles is equal to (not equal to) the relative amount of industry and market-level topics contained in non-official newspapers' articles.

Second, focusing on markets' response to the relative type of information conveyed by Chinese business news articles, we predict:

Hypothesis 4: The relative amount of industry and market-wide information conveyed by official newspapers articles is equal to (not equal to) the relative amount of industry and market-wide information conveyed by non-official newspapers' articles.

Lastly, focusing on the impact of each type of newspaper's activity on Chinese listed firms' stock return synchronicity, we predict:

Hypothesis 5: Non-official newspaper coverage of the listed firm will have no impact on (have an impact on) stock return synchronicity over the year.

Hypothesis 6: Official newspaper coverage of the listed firm will have no impact on (have an impact on) stock return synchronicity over the year.

Our empirical evidence on these hypotheses is presented in Sections 4 and 5.

3. Research Design, Sample Construction and Descriptive Statistics

3.1. Newspaper Data Source

To perform our analyses, we utilize data from Wisenews, a commercial database of domestic Chinese language newspaper and magazine articles. Established in 1998, Wisenews archives all variety of newspaper and magazine articles, including corporate news articles published by China's official and non-official newspapers, allowing us to gather the most comprehensive set of financial news articles available about China's listed firms. Due to database limitations before 2000, we focus on domestic newspaper articles published about China's listed firms over the period 2000-2017.

To construct our sample of country-specific newspaper articles, we used a refinement of the methodology employed in Piotroski, Wong, and Zhang (2017). First, using the trading name of each listed

company, we develop an automated article-crawling robot to search the archives of the database to identify all Chinese language articles featuring the listed company. This process was performed for each firm listed on the Shanghai and Shenzhen stock exchanges during our sample period. Second, we attribute individual articles to a specific listed firm based upon the corporate name that appears most frequently in the specific article. To ensure that our sample consists of press-generated news articles, we specifically exclude articles identified as summary lists (for example, firms with the largest price change or trading activity) and articles that relate to regulation-mandated announcements or press releases made by the company (for example, quarterly earnings reports, annual reports, ad hoc filings). We also rely on machine learning processes to clean the crawled news articles which are not related to the company due to ambiguity in the firm's name. When a news article mentions more than one firm, we assign the article to the firm with the highest frequency in the article or to the firm whose name appears first in the article if all firms have the same frequency in the article. Finally, we only keep articles published in official newspapers and non-official business newspapers (i.e., non-official newspapers focusing on financial and economic news reporting). These procedures yield a final sample of 2,777,103 unique corporate news articles about 3,103 unique domestic listed companies, drawn from 100 unique newspapers.

3.2 Classification of newspapers

Because our research focuses on corporate news, our sample only includes articles published in official newspapers and non-official business newspapers.⁶ Newspapers are classified as *official newspapers* if they are under the direct control of the central and local Communist Party Committees (party newspapers) or government institutions (political organ papers) and serve as their mouthpiece. Business newspapers are classified as *non-official* if they have an orientation towards reporting business, economic and financial news and are not controlled by any of the preceding organizations. Most non-

⁶ We exclude all commercialized metro and evenings newspapers given orientation toward non-business news content. There could be potential under-identification of official papers as some of the evening papers claim to be official newspapers (e.g., *Hefei Evenings Newspaper*). However, this exclusion should not bias our results.

official papers are controlled by news groups (e.g. *21st Century Business Herald* by Nan Fang News Group in Guangzhou), commerce-related ministries (e.g. *Financial News* by the People's Bank of China) or have private individuals or companies as shareholders (e.g. *Economics Observer* by Shandong Sanlian Group Ltd). Because official and non-official newspapers are formed to fulfill a specific niche and licensed accordingly, these newspaper classifications are stable through time. Based on these definitions, our sample of 100 newspapers consists of 52 official and 48 non-official newspapers.

3.3 Measurement of the information content (i.e., market reaction) publication of the news article

3.2.1 Absolute market reaction

Following prior research on the information content of firm-specific information releases, we measure the incremental information content of a news article based on the stock market's reaction to its publication (e.g., Fama, Fisher, Jensen, and Roll, 1969). Specifically, we estimate two variants of the stock price reaction to the publication of a news article. The first variable, $Return_{it}$, is measured as the one day raw stock return for firm i on the date of newspaper article publication (i.e., day t). The second variable, $Mkt-Adj\ Return_{it}$, is measured as $Return_{it}$ minus the value-weighted market return for all stocks in China on day t . Because we are interested in the magnitude of the unsigned impact of newspaper coverage on the firm's stock price, we follow prior research and measure the information content as the absolute value of the stock price reaction on the publication date ($|Return_{it}|$ and $|Mkt-Adj\ Return_{it}|$, respectively), benchmarked against the firm's average, long-run daily volatility. All price and return data is gathered through CSMAR.

3.2.2 Market, industry, and firm-specific component of information

Prior research documents that firm-specific stock returns and earning news can be decomposed into three components: a market component, industry component, and firm-specific component (e.g., King, 1966; Ball and Brown, 1967; Roll, 1988). Following the methodology in Liu (2011), we disaggregate the daily stock return of a given firm into these three components, and estimate two

variables, *Industry Information_{it}* and *Industry & Market Information_{it}*, designed to measure the proportion of daily returns related to industry-level information and industry and market-level information, respectively, on day t .

The variable *Industry Information_{it}* captures the relative amount of industry wide news contained in firm i 's stock return on day t , following the decomposition in Liu (2011). *Industry Information_{it}* is defined as $\frac{|CR_{it}^I|}{|CR_{it}^F| + |CR_{it}^I| + |CR_{it}^M|}$, where CR_{it}^M , the component of market wide information embedded in firm i 's return on day t ; CR_{it}^I , the component of industry wide information embedded in firm i 's return on day t ; CR_{it}^F , the component of firm specific information embedded in firm i 's return on day t . To estimate these components, we annually estimate the following model for each firm to extract the market and industry beta of stock i , denoted β_i^M and β_i^I :

$$R_{it} = \alpha_i + \beta_i^M \times R_t^M + \beta_i^I \times (R_t^I - \beta_i^{IM} \times R_t^M) + \varepsilon_{it} \quad (1)$$

In these models, R_{it} is return of firm i in period t ; R_t^I is the value-weighted return on the one-digit industry portfolio (to which firm i is affiliated) in period of t ; R_t^M is the value-weighted market return in period t ; β_i^{IM} is the market beta of the industry to which the firm i is affiliated, estimated from the following market model regression in each calendar year:

$$R_t^I = \alpha_i + \beta_i^{IM} \times R_t^M + \varepsilon_{It} \quad (2)$$

CR_{it}^M , the component of market wide information among the firm i 's return during period t , is defined as $\hat{\beta}_i^M \times R_t^M$, where $\hat{\beta}_i^M$ is estimated from model (1); CR_{it}^I , the component of industry wide information embedded in firm i 's return during period t , is defined as $\hat{\beta}_i^I \times (R_t^I - \hat{\beta}_i^{IM} \times R_t^M)$, where $\hat{\beta}_i^{IM}$ is estimated from model (1) and $\hat{\beta}_i^{IM}$ is estimated from model (2); CR_{it}^F , the component of firm specific information embedded in firm i 's return during period t , is defined as $R_{it} - CR_{it}^M - CR_{it}^I$.⁷

⁷ In order to estimate these individual stock return components, we exclude firms that have fewer than 30 observations in a given year, and the daily return observations for those firm-years are not included in our main empirical tests.

Analogously, the variable *Industry & Market Information_{it}* captures the relative amount of industry and market wide information news contained in firm *i*'s stock return on day *t*. *Industry & Market Information_{it}* is measured as $\frac{|CR_{it}^I| + |CR_{it}^M|}{|CR_{it}^F| + |CR_{it}^I| + |CR_{it}^M|}$, where CR_{it}^M , the component of market wide information embedded in firm *i*'s return on day *t*; CR_{it}^I , the component of industry wide information embedded in firm *i*'s return at time *t*; CR_{it}^F , the component of firm specific information embedded in firm *i*'s return on day *t*.

3.4 Measurement of news articles' topical content

To better understand the topical content of corporate news articles published in China, we use textual analysis techniques to identify and categorize the individual topics contained in each news article. Specifically, we create the variable *Industry & Market Content_{it}*, which is the relative weight of industry and market-level topics covered in a specific news article about firm *i* on day *t*. To determine which topics are industry or market specific, we use first Latent Dirichlet Allocation (LDA) to identify and define 400 unique topics embedded in our sample of news articles.⁸⁹ Once these 400 topics have been identified and defined, we use the following methodology to categorize each of these topics as firm, industry, or market specific in nature. First, an article is characterized as covering a specific topic when the weight of the article on the topic is more than one percent of the article's total content. Second, we classify the topic as an industry topic if more than 20 percent of all articles covering firms within a two-digit SIC industry (over the full sample) are identified as containing this topic. This procedure identifies

⁸ We use Latent Dirichlet Allocation (LDA), one of the most popular topical-modeling techniques in textual analysis developed by Blei et al. (2003), to identify topics and their corresponding distributions within each news article. Past research has shown that LDA can meaningfully capture the topics of the textual content of analyst reports and 10-Ks (e.g., Bao and Datta, 2014; Dyer et al., 2017; Hoberg and Lewis, 2017; Huang et al., 2017). The advantage of LDA topic modeling is that it is an unsupervised learning algorithm and hence does not require any labeled data to generate topics. The LDA algorithm assumes that each document can be represented by a mixture of topics, and each topic also can be characterized by a probability distribution over the words. After specifying the number of topics, the algorithm will learn the probability distribution over all words for each topic. More importantly, for each document, the algorithm will also identify the distribution of topics within it. We rely on the document-topic distribution discovered by LDA to identify the content of each news article by instructing the LDA algorithm to generate 400 topics.

⁹ See Beineke, Hastie, Manning, and Vaithyanathan (2004), Lin and Hovey (1997), Njolstad, Hoysaeter, Wei and Gulla (2014), and Zhang, Zeng, Li, Wang, and Zou (2009) for specific details on the use of LDA.

64 industry-specific topics in our sample, or, in other words, 16% of all topics discussed in these news articles are classified industry-level topics. Second, we sort the remaining 336 topics according to the number of articles covering each specific topic. Those topics contained in the top quintile of this distribution (68 topics) are classified as market-wide topics. These market-wide topics account for 17% of the total topics identified in our sample. The remaining 268 topics are assumed to be primarily firm-specific in nature. The variable *Industry & Market Content_{it}* reflects the proportion of topics covered in a specific news article that are classified as industry or market topics using this methodology.

3.5 Descriptive statistics

3.5.1 Newspaper and news article distribution

Table 1 presents descriptive statistics on our sample of 100 newspapers and 2,777,103 unique corporate news articles. Focusing on the distribution of our sample of newspapers across regions, Panel A shows that 32 of our newspapers are controlled by the central government (consisting of 4 official and 27 non-official newspapers), while the remaining 68 newspapers are controlled by local government entities or local party organizations (consisting of 47 official and 21 non-official newspapers). In terms of newspaper articles, we see that 50.2% of the news articles in our sample are published by national newspapers. These articles account for 12.9% and 67.1% of the total number of articles published by official and non-official newspapers, respectively. Finally, in terms of local ownership, newspapers in Guangdong, Shanghai, Beijing, and Sichuan province accounted for 18.7%, 9.7%, 5.7%, and 3.3%, respectively, of all domestic corporate news articles published in China.¹⁰

Focusing on the distribution of our sample observations over time, Panel B shows that the number of newspapers included in the sample ranges from 37 to 87 newspapers per year. For the first half our sample, the number of newspapers and newspaper articles generally increased year over year, consistent

¹⁰ Panel A also highlights the existence of a potential selection bias in the Newswise database, as only 25 provinces are shown as having at least one official business newspaper included in the database over our sample period (out of a total of 32 provincial-level administration units, excluding the Special Administrative Regions of Hong Kong and Macau, in China), while 4 provincial-level administrative units have no newspaper coverage in our database.

with the underlying growth in China's capital markets and liberalization of the news media. Interestingly, the number of newspapers and news articles plateaued in 2008, and experienced a notable decline in the final years of the sample. This observed decline in sample newspapers and article coverage in the last half of our sample is due to both the closure of underperforming newspapers and the suspension of coverage of select newspapers, such as Qingdao Daily and Southern Daily, by Wisenews during the latter half of the sample period.

Finally, Table 1, Panel C presents evidence on the distribution of our news articles and newspaper coverage across two-digit SIC code industries in China. Consistent with the important roles that the manufacturing and financial sector plays in China's economy, these two industries account for 38.1% and 22.0%, respectively, of all corporate news articles written during our sample period.

3.5.2 Descriptive statistics

Table 2 presents descriptive statistics on the variables used in our main empirical analyses. These statistics reveal several interesting observations at this point. First, the average daily raw and market-adjusted return for our full sample of daily return observations is approximately zero, consistent with daily stock returns in China behaving like a random walk. Second, after decomposing daily stock returns into their market, industry, and firm-specific components, we observe that, on average, the industry information component (industry and market information components) explain 25% (52%) of daily stock returns in China. And, consistent with industry and market news playing an important role in daily stock returns, we also observe that firm-level daily raw returns move in the same direction as the firm's industry index return 69% of the time. These observations mirror the high level of stock return synchronization documented about the Chinese market in earlier research papers (e.g., Morck, Yeung, and Yu, 2000).

Focusing on the news articles themselves, we observe that 22% of individual firm-level trading days are associated with the publication of a corporate news article. This suggests that the average firm has one news article written about it each week (or approximately four articles per month) over our sample period. And, in terms of topical content, 38% (62%) of the topics covered the average news article

consists of industry or market-wide (firm-specific) topics. However, there is significant variation in this topical coverage, with the interquartile range of observations spanning from 18% to 45% of the article relating to industry and market level topics.

The remainder of the table presents data on the characteristics of the covered firms. Consistent with prior studies, the average listed firm is large and profitable (average net income scaled by assets of 3%), with significant growth options (average market-to-book ratio of 3.9). Finally, 53% of the daily return observations in our sample are associated with state-controlled listed firms and 1% of the trading days contain an earnings announcement.

4 Empirical analyses: Information content of Chinese business newspaper articles

In this section, we exploit our sample of domestic newspaper articles on China's listed firms to document the information content of these published news articles and to examine relative informational role played by China's two primary types of newspapers, referred to as "official" party papers and "non-official" newspapers. From a research design perspective, the existence of official and non-official newspapers in China allows us to observe how the relative intensity of political versus market-based incentives shape editorial decisions and the information content of the resultant articles. Following prior research on the media (Antweiler and Frank, 2004; Tetlock, 2007; 2011), we measure the incremental information content of a news article based on the stock market's reaction to its publication. Section 4.1 documents the absolute stock market reaction to the publication of these business news articles. Section 4.2 disaggregates publication stock returns into market, industry, and firm-specific components, and documents whether different types of newspaper contribute different types of information to the market.

4.1 Absolute market reaction to news article publication

Our first set of empirical analyses examine the incremental information content of business news articles published by domestic Chinese newspapers. Specifically, we estimate the following cross-sectional model using firm-level data:

$$\begin{aligned}
|Return_{it}| \text{ or } |Mkt-Adj Return_{it}| = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 News Coverage_{it} \\
& + \beta_2 News Coverage_{it} * Official_{it} + \beta_3 |Market Return_{it}| + \beta_4 Firm Size_{it} + \beta_5 Market-to-book_{it} \\
& + \beta_6 ROA_{it} + \beta_7 Leverage_{it} + \beta_8 SOE_{it} + \beta_9 Earnings Announcement_{it} + \varepsilon_{it}
\end{aligned} \tag{3}$$

In these estimations, the dependent variable $|Return_{it}|$ ($|Mkt-Adj Return_{it}|$) is the absolute value of the daily raw return (absolute value of the market-adjusted return) of firm i on day t . Our main variable of interest, $News Coverage_{it}$, is an indicator variable equal to one if there was a business news article about firm i published in any domestic Chinese newspaper on day t , zero otherwise. Similarly, the indicator variable $Official_{it}$ equals one if a business news article about firm i was published in an official newspaper on day t , zero otherwise. To the extent that Chinese business news articles convey meaningful information, we expect a positive coefficient on $News Coverage$ ($\beta_1 > 0$). However, as discussed earlier, news articles published by official newspapers tend to be more biased and contain greater political rhetoric than articles published by non-official newspapers. If these attributes reduce the relative information content of articles published by official newspapers, we expect a negative coefficient on the interaction term $News Coverage * Official$ ($\beta_2 < 0$).

In these estimations, we also include firm-level control variables ($Firm Size_{it}$, $Market-to-book_{it}$, ROA_{it} , $Leverage_{it}$, and SOE_{it}) to capture both fundamental characteristics expected to influence the tone or political slant of media coverage over the reporting period (e.g., Piotroski, Wong, and Zhang, 2017) and risk-related factors imbedded in realized returns (e.g., Fama and French, 1992). The indicator variable SOE_{it} is included to capture a positive political bias introduced because the covered firm is state-owned. Because dates associated with corporate earnings announcements induce stock price volatility, we include an indicator variable ($Earnings Announcement_{it}$) that equals one if the firm announced quarterly or annual earnings on day t , zero otherwise. We include year-month indicator variables to capture systematic news shocks and economic trends that are expected to vary over time. We include firm fixed effects to control for differences in fundamentals and the average riskiness and volatility of the firms in our sample, and we cluster all standard errors by firm. Finally, in all estimations utilizing raw returns as the dependent

variable, we include the absolute market return ($|Market Return_t|$) on day t to control for the impact of market-level news on individual firm stock prices. All variables are as defined in the Appendix.

Table 4 presents select coefficients (and t-statistics) from estimations of these models. Consistent with newspapers serving an information role for China's capital market participants, we find that the average absolute daily stock return for Chinese listed firms is significantly larger when a business news article is published about the company, as captured by the positive coefficient on *News Coverage* in equation (1) and (3). This average absolute increase of about 50 to 70 basis points is economically meaningful, as it reflects a nearly 17 to 24 percent increase over average daily return volatility. Interestingly, these news stories introduce more volatility than quarterly earnings announcements, consistent with certain media-related events pre-empting the news contained in historical financial reports

Focusing on estimations (2) and (4), we observe that articles published in both official and non-official newspaper articles have information content, but that the magnitude of the market reaction is significantly smaller for official newspapers than non-official newspapers, as captured by the significant negative coefficient on *News Coverage*Official*. The smaller market reaction to official newspaper articles is consistent with them conveying less information (e.g., stale information; less value relevant information) or being perceived as less credible (due to their stronger political bias).

4.2 Disaggregation of market reaction into market, industry, and firm-specific components

Our main empirical analyses will examine whether official and non-official newspapers convey different information to the capital markets. Prior research documents that firm-specific stock returns and earning news can be decomposed into three components: a market component, industry component, and firm-specific component (e.g., King, 1966; Ball and Brown, 1967; Roll, 1988). This decomposition of stock returns allows one to measure, on any given day, the relative amount of each type of new information impacting the stock price formation process (e.g., Liu, 2011). In this section, we use this decomposed stock price behavior around news article publication to examine whether official and non-

official newspapers convey different amounts of industry and market-wide vs. firm-specific information to the market.

Prior research shows that financial intermediaries and select market participants contribute different information into stock prices, and that the type of information contributed is related to their relative informational advantage. Piotroski and Roulstone (2004) show insider and institutional investors primarily reveal firm-specific information through their trades and changes in holdings, consistent with their private access to firm-specific information sources. In contrast, Piotroski and Roulstone (2004) show that the reporting activities of financial analysts contribute relatively more industry and market level information than firm-specific information into market prices, with the effect especially pronounced following the initiation of analyst coverage (Crawford, Roulstone, and So, 2012). Finally, focusing on the news media, Dang, Moshirian, and Zhang (2015) show that newspapers operating in stronger institutional environments write idiosyncratic articles, contribute more firm-specific information, and generate lower stock return synchronicity than news outlets operating in weak institutional settings.

In the China context, our preceding evidence shows that domestic business news articles about China's listed firms convey information. However, it is an open empirical question about the type of information domestic newspapers contribute to the market, and whether official and non-official newspaper create and disseminate the same type of information. Table 5 explores this issue. Specifically, we estimate the following cross-sectional models using firm-level data:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or } \text{Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} \\ & + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} \\ & + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned} \quad (4a)$$

$$\begin{aligned} \text{Comovement}_{it} = \text{Logit} (& \alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} \\ & + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} \\ & + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned} \quad (4b)$$

In these estimations, the dependent variables *Industry Information_{it}* (*Industry & Market Information_{it}*) equals the fraction of firm i's raw stock return that reflects industry information (industry and market

information) on day t . The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. For completeness, we also create a non-parametric dependent variable, $Comovement_{it}$, which is an indicator variable equal to one if firm i 's raw stock return moved in the same direction as the firm's relevant industry index on day t , zero otherwise. This version of the model (equation 4b) is estimated using a Logit regression.

Our main variable of interest, $News\ Coverage_{it}$, is an indicator variable equal to one if there was a business news article about firm i published in any domestic Chinese newspaper on day t , zero otherwise. Similarly, the indicator variable $Official_{it}$ equals one if a business news article about firm i was published in an official newspaper on day t , zero otherwise. To the extent that Chinese business news articles about Chinese listed firms primarily convey meaningful firm-specific (industry and market) information, we expect a negative (positive) coefficient on $News\ Coverage$ ($\beta_1 < 0$). However, as discussed earlier, we expect news articles published by official newspapers to convey *relatively more* industry and market-level information than articles published by non-official newspapers. If this conjecture is correct, we expect a positive coefficient on the interaction term $News\ Coverage * Official$ ($\beta_2 > 0$). Finally, to control for omitted determinants of the information content of these newspaper articles, we continue to include firm-level control variables ($Firm\ Size_{it}$, $Market-to-book_{it}$, ROA_{it} , $Leverage_{it}$, and SOE_{it}), year-month indicator variables, and firm (industry and province) fixed effects in all OLS (Logit) estimations, and all standard errors are clustered by firm.

Table 5 presents select coefficients (and t-statistics) from estimations of these models. First, focusing on estimations (1) and (3), we find that the average proportion of daily returns explained by industry and market factors is significantly lower on days associated with the publication of a corporate news article about a specific firm than the proportion observed on days without such articles, as reflected by the negative coefficient on $News\ Coverage$. This negative relation is consistent with these newspaper articles generally conveying firm-specific information about the firm's being covered. This reduction

represents an approximate 10% reduction in the relative amount of industry and market-level information (vs. firm-specific information) contained in daily returns on these publication dates. This main result is confirmed in estimation (6), where we see that the likelihood of an individual firm's stock price moving in the same direction as its industry index is approximately 3% lower on dates associated with a published news article.

Focusing on differences between official and non-official newspapers, estimations (2) and (4) reveal that the relative amount of firm-specific vs. industry and market-wide information conveyed by official newspaper articles is significantly larger than the proportion conveyed by non-official newspapers, as reflected by the positive coefficient on the term $News\ Coverage_{it} * Official_{it}$. In other words, official newspapers convey relatively more (less) industry and market level (firm-specific) information than non-official newspapers, *ceteris paribus*. These differences are confirmed in estimation (6), where the observed reduction in the likelihood of co-movement with industry returns around newspaper article publication dates is marginally smaller when the business news article is published by an official newspaper. Thus, although both newspapers' articles primarily convey firm-specific information about the covered firm, the degree to which the article conveys other, macro information seems to be correlated with the specific newspapers' political role, with official newspapers seemingly able to credibly convey information about government policies toward and/or overall economic performance of specific industries or the broader economy.

5 Empirical Analysis: Textual Content of Chinese Business News Articles

The results in Tables 4 and 5 show that both official and non-official Chinese newspaper convey information to the capital markets. However, consistent with differences in the political and market-oriented roles, we find that they both convey different absolute amounts of information and that they do not convey the same proportion of firm-specific vs. industry and market-level news to the market. This section explores these latter issue directly by examining the topical content contained in these news

articles, and examining whether the topics covered by these articles map into the observed information content / stock price relations observed in Table 5.

5.1 Analysis of news articles' textual content

In this section, we directly examine whether articles published by official and non-official newspapers focus on different topics. Specifically, following the methodology outlined in section 3, we measure the relative importance of industry and market level topics contained in a specific news article. The resultant variable, *Industry & Market Content_{it}*, is measured as the sum of the weight of industry and market-wide topics covered in a specific news article about firm *i* on day *t*. Our first objective is to determine whether articles published by official (non-official) newspapers contain relatively more (less) industry and market specific topics. Table 6 provides this evidence.

Table 6, panel A provides simple descriptive evidence on the textual content of Chinese financial news articles. Based on our textual analysis, we find that about 32 percent (30 percent) of the topics covered in the average (median) news article about an individual Chinese listed company relates to industry or market-wide topics. However, there is considerable cross-sectional variation in the amount of industry and market topics contained in the articles, with the interquartile range of *Industry and Market Content_{it}* spanning from 18% to 45% of covered topics.

Partitioning on newspaper type, we observe that the mean news article published in an official newspaper contains a greater relative amount of industry and market topics than an article published in a non-official newspaper (34.8% vs. 31.4%, respectively). However, because these newspapers may choose to write articles on different firms who have different exposure to industry and market forces, or write about these firms at different times, we next examine this potential differences by estimating the following cross-sectional model using firm-article data:

$$\begin{aligned} \text{Industry \& Market Content}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Official}_{it} + \beta_2 \text{Firm Size}_{it} \\ & + \beta_3 \text{Market-to-book}_{it} + \beta_4 \text{ROA}_{it} + \beta_5 \text{Leverage}_{it} + \beta_6 \text{SOE}_{it} + \beta_7 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned} \quad (5)$$

In this estimation, the dependent variable is *Industry & Market Content_{it}*, which measures the relative weight of industry and market-level topics covered in a news article about firm *i* on day *t*. Our main variable of interest, *Official_{it}*, equals one if a business news article about firm *i* was published in an official newspaper on day *t*, zero otherwise. If news articles published by official newspapers convey *relatively* more industry and market-level information than articles published by non-official newspapers, we would predict a positive coefficient on the *Official_{it}* ($\beta_1 > 0$). We include an array of firm-level control variables (*Firm Size_{it}*, *ROA_{it}*, *Market-to-book_{it}*, *Leverage_{it}*, and *SOE_{it}*) potentially correlated with the relative mixture of industry, market, and firm-specific topics covered in a financial news article. The model also includes year-month indicator variables and firm fixed effects; all standard errors are clustered by firm.

Table 6, panel B presents select coefficients and t-statistics (in parenthesis) from our estimation of this model. First, focusing on firm-level determinants of textual content, we find that news articles about larger firms are more likely to contain industry and market information, consistent with large firms serving as bell weather stocks for overall economic conditions. Second, we observe a weak positive association for SOEs, again consistent with the fact SOE performance is highly dependent on government policy. Third, we see that news articles published on the day of a firm's earnings announcement discuss fewer industry and market-wide topics, consistent with those news events traditionally focusing on individual, firm-specific outcomes. Finally, and most importantly, we find that after controlling for these attributes, the average article published by official newspapers contains a greater portion of industry and market level topics than articles published by non-official newspaper about the same firm.

5.2 Link between articles' content and disaggregated stock market reaction.

The results in Table 6 confirm that business news articles written by official newspapers contain a disproportionate amount of industry and market-related topics relative to news articles written by non-official newspapers about the same firm. To the extent that the market reaction results documented in Table 5 reflect these differences in article content, we would expect the impact of a published news article

on the components of the stock price response to the article (i.e., *Industry & Market Info*) to be positively correlated with the relative amount of industry specific topics contained in the article (i.e., *Industry & Market Content*). Moreover, to the extent that official newspapers have greater credibility with respect to disseminating industry and market information, we would expect the effect to be stronger for their articles.

To test these arguments, we estimation variations of the following empirical model using firm-article level data:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or } \text{Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} \\ & + \beta_1 \text{Industry \& Market Content}_{it} + \beta_2 \text{Industry and Market Content}_{it} * \text{Official}_{it} \\ & + \beta_3 \text{Official}_{it} + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} \\ & + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned} \quad (6a)$$

$$\begin{aligned} \text{Comovement}_{it} = \text{Logit} (& \alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Provincial FE} \\ & + \beta_1 \text{Industry \& Market Content}_{it} + \beta_2 \text{Industry and Market Content}_{it} * \text{Official}_{it} + \beta_3 \text{Official}_{it} \\ & + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} \\ & + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned} \quad (6b)$$

In these estimations, the dependent variables *Industry Information_{it}* (*Industry and Market Information_{it}*) equals the fraction of firm i's raw stock return that reflects industry information (industry and market-wide information) on day t, as defined earlier. For completeness, we continue to use the dependent variable, *Comovement_{it}*, which is an indicator variable equal to one if firm i's raw stock return moved in the same direction as the firm's relevant industry index on day t, zero otherwise, and estimate this version of the model (equation 6b) using a Logit regression.

Our main variable of interest, *Industry & Market Content_{it}*, is the relative weight of industry and market-level topics covered in a news article about firm i on day t. The indicator variable *Official_{it}* equals one if a business news article about firm i was published in an official newspaper on day t, zero otherwise. To the extent that the relative components of the market reaction to published business news articles reflect differences in the relative amount of industry and market information versus firm-specific information conveyed by the news article, we expect a positive coefficient on *Industry & Market Content_{it}*

($\beta_1 > 0$). Additionally, our earlier tests show that news articles published by official newspapers convey *relatively* more industry and market-level information than articles published by non-official newspapers. If the market views this information as having greater credibility (or more value relevance) than similar articles published by non-official sources, we expect a positive coefficient on the interaction term *Industry & Market Content_{it} * Official_{it}* ($\beta_2 > 0$). Finally, to control for omitted determinants of the information content of these newspaper articles, we continue to include firm-level control variables (*Firm Size_{it}, Market-to-book_{it}, ROA_{it}, Leverage_{it}, and SOE_{it}*), year-month indicator variables, and firm (industry and province) fixed effects in all OLS (Logit) estimations, and all standard errors are clustered by firm.

Table 7 presents select coefficients and t-statistics (in parenthesis) from our estimation of these models. Because our goal is to understand the mapping between article topic content and market reaction, this particular model is estimated using only our subsample of article-level observations (unlike Table 5, which includes stock return data from dates with and without news articles). Consistent with the market reaction to news articles differing by the articles' content, we find that news articles containing a greater preponderance of industry and market level (firm-specific) topics are associated with publication returns that have larger industry and market (firm-specific) return components. Moreover, we observe that the positive relationship between topics and return components is incrementally stronger for articles published in official newspapers. This incremental effect, which is nearly one to two times larger than the base relation observed for non-official newspapers, suggests that industry and market news conveyed by official newspapers has a *stronger relative impact* on prices than similar industry and market level news conveyed by non-official newspapers.

5.3 Link between official and non-official newspaper activity and stock return synchronicity

Prior research shows that the level of stock price synchronization observed in an economy reflects the amount of firm-specific information being incorporated in prices, with the degree of synchronization affected by both the strength of country-level institutions and firm-level mechanisms and incentives to

promote transparency (e.g., Morck, Yeung, and Yu, 2000; Durnev, Morck, Yeung, and Zarowin, 2003; Jin and Myers, 2006; Fernandes and Ferreira, 2008). Due to the country's historically weak institutional arrangements and prevailing incentives for opacity, China's stock prices exhibit an unusually high level of stock return synchronization, with the effect more pronounced for state controlled firms, firms with low quality auditors, and minimal foreign shareholders (Gul, Kim, and Qiu, 2010). Given the important role that the state media plays in shaping the flow of information in this economy, China's state-controlled newspapers can also have a net, incremental impact on the observed level of synchronicity in the economy. Specifically, we expect that the different information roles by played official and non-official newspapers will differentially affect the degree of stock price synchronization we observe for China's listed firms, with a greater intensity of articles published by official (non-official) newspapers leading to more (less) stock price synchronization, *ceteris paribus*.

To examine this issue, we estimate the following empirical model using our full sample of listed firm-year observations:

$$\begin{aligned}
Synchronicity_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Number of Articles}_{it} \\
& + \beta_2 \text{Number of Official Articles}_{it} + \beta_3 \text{Number of Non-Official Articles}_{it} + \beta_4 \text{Relative Official}_{it} \\
& + \beta_5 \text{Fundamental Correlation}_{it} + \beta_6 \text{Herfindahl Index}_{it} + \beta_7 \text{Std Dev (ROA)}_{it} \\
& + \beta_8 \text{Log(\# of Industry Firms)}_{it} + \beta_9 \text{Firm Size}_{it} + \beta_{10} \text{Market-to-book}_{it} + \beta_{11} \text{ROA}_{it} \\
& + \beta_{12} \text{Leverage}_{it} + \beta_{13} \text{SOE}_{it} + \varepsilon_{it}
\end{aligned} \tag{7}$$

In this model, the dependent variable, *Synchronicity_{it}*, captures the annual comovement of the firm's stock return with the overall market return. *Synchronicity_{it}* is measured as $\log(\frac{R^2}{1-R^2})$, where R^2 is estimated from the market model for firm *i* in year *t* using daily return data. Our measure of overall newspaper coverage intensity, *Number of Articles_{it}*, is measured as the natural logarithm of one plus the number of news articles covering firm *i* in year *t*. Our measure of official newspaper coverage intensity, *Number of Official Articles_{it}*, is measured as the natural logarithm of one plus the total number of news articles about firm *i* published in official newspapers in year *t*. Our measure of non-official newspaper

coverage intensity, *Number of Non-Official Articles_{it}*, is measured as the natural logarithm of one plus the total number of news articles about firm *i* published in non-official newspapers in year *t*. Our measure of the relative intensity of media coverage by official newspapers, *Relative Official Coverage_{it}*, is measured as the ratio of one plus *Number of Official Articles_{it}* to one plus *Number of Non-Official Articles_{it}*.

These estimations confirm and reinforce the interpretation of our earlier findings. Namely, the number of articles published by official (non-official) newspaper are positively (negatively) associated with stock return synchronicity, consistent with these newspapers primarily contributing industry level (firm specific information) into the price formation process. The results for official newspaper are strikingly similar to the effects observed for financial analysts in earlier research, who also have a competitive advantage with respect to gathering, analyzing, and disseminating information about industry and market-wide trends (Piotroski and Roulstone, 2004; Crawford, Roulstone, and So, 2012).

Finally, relative to the extant literature on stock return synchronicity and the media, we are finding these differential information effects across newspapers after holding country-level institutional characteristics constant, again demonstrating the importance that newspaper-level incentives are having on these two classes of newspapers in China. One means of interpreting this incremental relation is that official newspapers have more credibility when communicating industry and macro information than non-official newspapers (and vice versa). This credibility can arise due to these official newspapers proximity to its sponsoring governmental agency or party unit. We test this argument, along with the broader role of political forces, in the next section.

6. Political variation in the information content of China's state-controlled newspapers

Our remaining tests search for political variation in the information content of China's state-controlled newspapers.

6.1 Relative influence of central government vs. local government official newspapers

The preceding analyses show that official newspapers contribute a greater amount of industry and market-wide information into security prices than non-official newspapers. Given that official newspapers

are sponsored by governmental agencies or specific communist party units, this information effect seems plausible. First, official newspapers' editors and journalists may have preferential access to policy information due to their relations with these controlling entities, which they can then incorporate into reporting activities. Second, official newspaper business articles tend to contain more political content and rhetoric than non-official newspapers (Piotroski, Wong, and Zhang, 2017). Due to the newspapers own political incentives, they may be voluntarily including political content in their articles that provides insights to the market about current or future government policies and long-term objectives. Lastly, official newspapers are frequently used as the mouthpiece of these organizations, and may be actively used as instruments to convey government policy-related information to the capital markets.

To the extent that proximity to policy-relevant information is behind the relative strength of the informativeness of official newspapers with respect to industry and market level topics, we would expect newspapers affiliated with central government agencies (i.e., national newspapers) to produce stronger information effects than official newspapers affiliated with local governments and agencies. Similarly, if national newspapers are perceived to possess greater credibility, we would document similar differences in the information content of the articles. To test these arguments, first we categorize our official newspapers based on the locality of their sponsoring organization. This process identifies five official newspapers as being controlled by units of the central government, and 47 official newspapers being controlled by local governmental units, with these newspapers having published 111,010 and 752,095 business news articles about individual listed firms, respectively. Next, we estimate variants of the following cross-sectional models:

$$\begin{aligned}
\text{Industry Information}_{it} \text{ or } \text{Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} \\
& + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Central Official}_{it} + \beta_3 \text{News Coverage}_{it} * \text{Local Official}_{it} \\
& + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} \\
& + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}
\end{aligned} \tag{8a}$$

$$\begin{aligned}
\text{Comovement}_{it} = & \text{Logit} (\alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} \\
& + \beta_2 \text{News Coverage}_{it} * \text{Central Official}_{it} + \beta_3 \text{News Coverage}_{it} * \text{Local Official}_{it} \\
& + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it}
\end{aligned}$$

$$+ \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \quad (8b)$$

These models are extensions of equations (4a) and (4b), with the indicator variable *Official_{it}* being replaced by two separate indicator variables – *Central Official_{it}* and *Local Official_{it}* – reflecting the locality of the official newspaper. Specifically, *Central Official_{it}* is an indicator variable equal one if an official newspaper controlled by the central government published a corporate news article about firm *i* on day *t*, zero otherwise, and *Local Official_{it}* is an indicator variable equal one if an official newspaper controlled by a local government agency or party unit published a corporate news article about firm *i* on day *t*, zero otherwise. To the extent that national newspapers convey more (or more credible) information about industry and market factors, we would expect the coefficient on *Central Official_{it}* to be larger than the coefficient on *Local Official_{it}* ($\beta_2 > \beta_3$).

Table 9 presents select coefficient and t-statistics (in parentheses) for these estimations. Consistent with our earlier arguments, we find that the positive incremental impact of official newspaper article publication on our two stock return based measures of industry and market information is larger for national newspapers than local official newspapers, with the difference in coefficients significantly different from each other at the one percent level of significance.¹¹

6.2 Impact of CPC National Congress on the informativeness of official vs. non-official newspapers

Meetings of the CPC National Congress correspond to a highly politicized period in China. During the National Congress itself, the CPC publicly unveils its leadership team for the following five years and begins the process of outlining the Party's goals and objective (i.e., so-called Five Year Plan). Prior to the Congress, the party is focused on developing and building consensus for its next set of five year objectives, with those plans, goals, and related policy items announced during the Congress and at subsequent events and meetings in the months thereafter. As such, during National Congress years, it is

¹¹ Untabulated analyses reveal minimal, incremental effect among non-official newspapers with national versus regional circulation (and only in certain empirical specifications), suggesting that differences in circulation are likely not behind the observed differences in information effects of articles published by local government vs. central government official newspapers.

natural to expect that newspapers in China will convey a disproportionate amount of industry and macro-economic information. And, to the extent that official newspapers are the primary channel by which this information is credibly disseminated to the market, we would expect the industry and market role of official newspapers to be incrementally stronger during these periods.¹²

To examine this issue directly, we separately re-estimate variants of equations (4a) and (4b) after partitioning the sample based on whether or not a specific firm-day observation or newspaper article relates to a national congress year. Our eighteen years' sample period spans four National Congress periods (2002, 2007, 2012, 2017), providing sufficient opportunity to test for these effects (i.e., we have approximately 2.9 million National Congress observations, or 23% of the total daily return sample). Select coefficients and t-statistics from these estimations are presented in Table 10.

These estimations reveal two key findings. First, we find that all corporate news articles in China convey relatively more industry and market-wide information during National Congress periods, consistent with these newspapers' in-depth (and perhaps disproportionate) coverage of those major political events. Second, and quite surprisingly, we find little evidence to suggest that this temporal shift is larger for official newspapers than non-official newspapers. If anything, select specifications suggest that the effect may actually be more pronounced for non-official newspapers who, during non-Congress periods, are less likely to have conveyed material amounts of industry or market-level information, whereas official newspapers are already fulfilling that role.

6.3 Impact of Xi Jinping Media Visit on the informativeness of official vs. non-official newspapers

On February 19th, 2016, President and Chairman Xi Jinping made an extraordinary, high profile visit to the country's three most important state-run media outlets, the party newspaper People's Daily, state-run news agency Xinhua, and state broadcaster China Central Television (CCTV). During these visits, President Xi told editors and reporters they must pledge "absolute loyalty" to the Communist party

¹² Prior research already documents that Chinese newspapers tilt their coverage and reporting decisions around these political events to facilitate the suppression of negative corporate news (e.g., Piotroski, Wong, and Zhang, 2015).

and closely follow its leadership in “thought, politics and action.” In particular, he noted that “the media run by the party and the government are the propaganda fronts and must have the party as their family name.” And, when speaking about their relationship with the government and party, “all the work by the party’s media must reflect the party’s will, safeguard the party’s authority, and safeguard the party’s unity” and “they must love the party, protect the party, and closely align themselves with the party leadership in thought, politics and action.”

Anecdotal evidence suggests that the reporting incentives and activities of Chinese newspapers shifted in response to this powerful signal from the party.¹³ We exploit this event to identify whether the tightening of this relationship between the party and official newspapers affected the relative information content of their reports. To examine this issue directly, we separately re-estimate variants of equation (4) after partitioning the sample based on whether a given firm-day observation or newspaper article occurred before or after February 19th, 2016. Although the event occurs near the end of our sample, we have approximately 2.0 million post-visit observations, which should provide a sufficient opportunity to test for these effects (approximately 15% of the total daily return sample). Select coefficients and t-statistics from these estimations are presented in Table 11.

These estimations reveal two pronounced effects. First, we document a dramatic reduction in the amount of industry and market-level information being conveyed by non-official newspapers following Xi’s policy statement, as reflected by shift in the main *News Coverage* effect in these estimations. Second, we document a significant increase in the incremental amount of industry and market information conveyed by official newspapers following President Xi’s media visit. These divergent effects parallel changes that occurred among official and non-official newspapers during the conglomeration reform of the 2000s, where conglomeration allowed the two different newspapers types to focus on their respective market and political roles (Piotroski, Wong, and Zhang, 2017). More importantly, these current results

¹³ Untabulated analyses documented a small, but statistically significant, reduction in the overall information content of news articles published by both official and non-official newspapers following Xi’s media visit. This overall reduction in the informational role of the business news media is consistent with anecdotal evidence suggesting that the loyalty event has had a chilling effect on China’s information environment.

reinforce our interpretation that official and non-official newspapers also serve different informational roles for China's capital markets. Specifically, following President Xi's request of party loyalty and his explicit reminder of the party media's role as their propaganda mouthpiece, it seems that official newspapers began publishing news articles that convey a greater amount of industry and macroeconomic information likely correlated with government and party policy and objectives, which aligns well with the overall political and propaganda role they play in the economy. In contrast, non-official newspapers are publishing articles that seem to be deemphasizing industry and market information, suggesting that they are potentially deemphasizing their political role and focusing on their commercial incentives and generating articles that primarily provide firm-specific information about the covered firms.¹⁴

6.4 Political variation in textual content

Our final set of analyses examine the mix of topics covered in Chinese business news articles conditional upon these political attributes. Specifically, we re-estimate variants of equation (5) after partitioning the sample on the centrality of the official newspapers, whether articles were published during National Congress years, and whether articles were published following President Xi's high profile media visit. Select coefficients from these estimations are presented in Table 12. First, consistent with our returns-based tests of information content (i.e., Table 9), we observe that articles published by official newspapers owned by the central government contain a significantly greater mix of industry and market-wide topics than other, locally-owned official newspapers. Second, also consistent with our returns-based information content results (i.e., Tables 10 and 11), we observe that articles published by official newspapers contain a significantly greater mix of industry and market-wide topics during both National Congress periods and following President Xi's media visit. Together, the observed variation in textual topics across these three tests lends support to our conjecture that the informational role of official newspapers is affected by these organizations proximity to the government, the party, and their policies.

¹⁴ The results in this section are robust to restricting the sample to only articles written during the Xi Jinping era (i.e., 2012-2017) and during the one year period before and after Xi's visit to China's media outlets.

7 Additional analyses

7.1 Articles published around corporate earnings announcements

The preceding analyses find that official and non-official newspapers convey different relative amounts of firm-specific vs. industry and market-wide news to the market. However, this effect could be the result of these two types of newspaper writing corporate news reports in response to fundamentally different economic events. As such, the observed information content is endogenous to the decision to write an article. To address this potential concern, we employ a secondary research design that examines the attributes of financial news articles written immediately following corporate earnings announcements. Earnings announcements have two features that make for a powerful setting to study media behavior. First, earnings announcements are well-defined, value relevant events during which newspapers have salient market-based incentives to report and interpret the financial performance of the listed firm. Second, because multiple newspapers have economic and political incentives to report on the performance of the announcing firm, earnings announcements are a setting where newspaper professionals (journalists and editors) across China are concurrently performing the same task (e.g., publishing a story) in response to the same underlying economic event and using similar inputs (i.e., the earnings announcement press release and related financial disclosures). Thus, articles written immediately following the earnings announcement relate to a common event, but individually reflect the reporting preferences, editorial direction, and comparative advantage of each publishing newspaper.

To perform this analysis, we first identify quarterly, semi-annual and annual earnings announcement dates for all Chinese listed firms over our sample time period through CSMAR. We then match our full sample of newspaper articles against these earnings announcement event dates, and retain all news articles published on or within one day of the listed firm's respective earnings announcement date (days -1 to +1). This selection procedure results in a sample of 136,130 corporate news articles published by our sample of 100 newspapers, and covering approximately 93,230 unique earnings announcement events over our sample period.

Untabulated estimations of equations (4), (5), and (6) using this sub-sample of daily return observations confirm our main findings in Tables 5 through 7. Namely, articles published by official newspapers during these earnings announcement windows convey relatively more industry and market news information – as measured by both topical content and disaggregated market returns - that similar articles published by non-official newspapers. Moreover, the official newspaper effects with respect to differences in topical coverage are substantially larger during these earnings announcements than in the full sample.

7.2 Restricted sample of newspapers

Our descriptive evidence suggests that our newspaper database has selection biases, both across time and across provinces. The systematic absence of local official and/or non-official business newspapers in these databases has the potential to induce a selection bias into our analysis if these database coverage decisions are non-random. To validate that our results are not driven by such these coverage biases, we re-estimate our main analysis using two restricted samples: (1) articles published in newspapers that appear throughout our entire sample period and (2) articles published only newspapers domiciled in China's three tier 1 cities (Beijing, Guangzhou, and Shanghai). The constant sample will help mitigate selection biases over time, while the focus on Tier 1 cities will mitigation concerns about regional biases. We focus on Tier 1 cities because these three cities capture China's leading political and business centers and because our primary data source, Newswise, seems to provide comprehensive coverage of leading local newspapers in these three cities. Untabulated estimations of equations (4), (5), and (6) using these restricted samples confirm our main results.

7.3 Alternative measure of industry content

Our measures of textual industry content are based on the identification of topics via LDA techniques and our industry, market, and firm-specific topic classification methodology. Both aspects of our research design introduce noise into the identification of industry topics and measurement of an

article's industry content. To address these concerns, we adopt an alternative approach whereby we search for the presence of unambiguous, industry-specific words or phrase in each article. Specifically, we count the number of times an article mentions industry names. Our dictionary of industry names are defined based on the Guidelines for Industry Classification of Listed Companies (2013 revision) as published by the China Securities Regulatory Commission (CSRC). We assume that if a newspaper mentions the related industry name when covering certain firms, the article is likely to contain industry wide. The resultant variable, *Industry Count_{it}*, is the number of times an industry name is mentioned in an article, deflated by the number of sentences in an article.

Table 13 presents select coefficients from re-estimations of equations (5), (6a) and (6b) using *Industry Count_{it}* in lieu of our LDA-derived measure of industry topical content (i.e., *Industry & Market Content_{it}*). These estimations confirm our main results, namely, official newspaper more frequently reference the covered firm's CSRC industry in their articles than non-official newspapers, and that these references by official newspapers are more strongly correlated with the observe stock market reaction to the publication of the articles.

7.4 Examination around macro and industry-level economic policy shocks

One of the concerns with our research design pertains to how to interpret documented differences in the informational effects of official vs. non-official newspaper articles. To address this limitation, we exploit time-series variation in these newspapers access to value-relevant industry information and their incentives to promote such a narrative. Specifically, we examine the relative informational properties of newspaper articles published around the introduction of new macro-level or industry-level economic policies by the Chinese government. To the extent our results reflect differences in official newspaper's access to policy level data and incentives to report on such issues, we would expect our results for official newspapers to be stronger during these policy-sensitive periods, *ceteris paribus*.

To test this argument, we collect information on all economic policies issued by Chinese governmental units from the China Industrial Policy website (<http://zc.wefore.com>) over our sample

period.¹⁵ Given the objective of our tests, we retained the subsample of policy introductions that were introduced by central government agencies, include all Ministry level organizations and the State Council, as these policy changes would have national implications and warrant coverage by all business media outlets. For each policy introduction, we identify the calendar date the policy was formally announced by the government, and determine whether the policy applies to a specific set of CSRC industries, or the economy as a whole.

To examine whether the proximity of an article's publication date to the unveiling of a key industry a policy affects the relative informativeness of the news article, we re-estimate equation (4a) conditional upon whether or not policy shock affected a specific firm on that day. These estimations are presented in Table 14. The first set of columns consider the impact of macro-level policy shocks on all firms on the relevant announcement days; the second set of columns consider the impact of an industry-level shock to the covered firm's industry on that day. These estimations confirm that the perceived value of an official newspaper article is increasing in expected market and industry policy relevance of the article.

7.5 Robustness tests: Focus on innovative articles

An additional concern with our media analysis is that, by examining all corporate news articles in the Newswise database, we are capturing many articles that are disseminating stale news. As such, the relative information content of a newspaper article is likely to be impacted by the timing of its publication relative to similar news articles. To the extent that articles in official or non-official newspapers are systematically published sooner or later, or that the publication of a news article by one newspaper has the potential to shape the reporting behavior of other, follower newspapers, our market-based results may reflect those timing differences.

¹⁵ This search yields approximately 52,000 new policies that were introduced by various municipal, provincial, and central government-level organizations over the period 2000 to 2017.

To alleviate this concern, we re-estimate our main analysis, equations (4a) and (4b), after retaining only those corporate news articles that we classify as “innovative,” based on the textual content of the article. Specifically, we categorize an article as innovative if there are no corporate news articles with “similar” content published about the covered firm in the prior fourteen-day period. We define articles as being “similar” if the cosine similarity score between a pair of news articles is greater than 0.4. Estimations using these innovative articles yield similar inferences, and marginally stronger results, than those reported in Table 4.

8 Conclusion

State-controlled media in developing economies typically perform both market and political functions. Through the creation of new information and the broad dissemination of public information, the media can serve important information and governance roles. When effective, an independent, market-oriented media improves transparency, reduces informational frictions, and strengthens monitoring mechanisms, which together should lead to deeper and more efficient markets.¹⁶ However, the reporting objectives of state-controlled media are frequently not aligned with transparency and resultant market-oriented benefits. In these settings, editorial decisions can be shaped by the incentives of politically captured editors and journalists, with published articles shaped by the objectives and preferences of politicians, bureaucrats, regulators, and their cronies. In the extreme, media organizations become a propaganda mouthpiece of the government, communicating the government’s narrative by biasing news stories and engaging in censorship.

Prior research argues and finds that corporate news articles published by official newspapers in China are biased, provide less information to the capital markets, and fail to serve an important corporate governance role. Although all three conclusions are valid, they also overlook that official newspapers,

¹⁶ Prior research shows that media outlets create and disseminate financial information (Zingales, 2000; Bushman, Piotroski, and Smith 2004; Bushee et al., 2010; Soltes, 2010) and serve as a watchdog in the market (Dyck and Zingales, 2002; Miller, 2006; Dyck, Volkchkova, and Zingales, 2008; Liu and McConnell, 2013), even helping autocrats to monitor their bureaucrats (Egorov, Goriev, and Sonin, 2009; Lorentzen, 2014).

because of their close proximity to the government, can still play an important role in communicating industry and market-wide news and information to the capital markets. Our paper provides the first evidence to suggest that China's official newspapers may serve a different and complimentary informational role than that played by non-official newspapers in China's market economy.

Finally, our main assumption throughout the paper is the stock price reaction to the publication of a news article represents the incremental, value-relevant information being conveyed by the article. To the extent that the markets in China are informationally inefficient, our stock return based measures would instead be capturing the *perceived* information content of the news article. Regardless, our evidence remains relevant, as ours is the first comprehensive study to document the different types of business topics covered in Chinese financial news media and how the market reacts to this information. Future research can investigate whether the market corrects for known biases embedded in the content of these newspapers' articles, especially as it relates to the industry and market-wide information conveyed by China's politicized official newspapers.

References

- Antweiler, W., and M. 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.
- Ball, R., and P. Brown. 1967. Some preliminary findings on the association between the earnings of a firm, its industry and the economy. *Journal of Accounting Research* 5 (Supplement): 55-77.
- Bao Y., and A. Datta. 2014. Simultaneously Discovering and Quantifying Risk Types from Textual Risk Disclosures. *Management Science* 60(6): 1371-1391.
- Beineke, P., T. Hastie, C. Manning, and S. Vaithyanathan. 2004. Exploring sentiment summarization. In *Proceedings of the AAAI spring symposium on exploring attitude and affect in text: theories and applications* (Vol. 39). Palo Alto, CA: The AAAI Press.
- Blei, D., A. Ng, and M. Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research* 3: 993-1022.
- Bushee, B., J. Core, W. Guay, and S. Hamm. 2010. The Role of the Business Press as an Information Intermediary. *Journal of Accounting Research* 48: 1-19.
- Bushman, R., J. Piotroski, and A. Smith. 2004. What Determine Corporate Transparency? *Journal of Accounting Research* 42: 207-52.
- Crawford, S., D. Roulstone, and E. So. 2012. Analyst Initiations of Coverage and Stock Return Synchronicity. *The Accounting Review* 87 (5): 1527-1553.
- Dang, T., F. Moshirian, and B. Zhang, B. 2015. Commonality in news around the world. *Journal of Financial Economics* 116 (1): 82-110.
- Djankov, S., C. McLiesh, T. Nenova, and A. Shleifer. 2003. Who Owns the Media? *Journal of Law and Economics* 46: 341-81.
- Damodoran, A., and C. H. Liu. 1993. Insider trading as a signal of private information. *Review of Financial Studies* 6 (1): 79-119.
- Dyck, A., N. Volchkova, and L. Zingales. 2008. The Corporate Governance Role of the Media: Evidence from Russia. *The Journal of Finance* 63: 1093-135.
- Dyck, A. and L. Zingales. 2002. The Corporate Governance Role of the Media. 107-140 in *The Right to Tell: The Role of Mass Media in Economic Development*, edited by R. Islam. Washington D.C.: World Bank.
- Dyer, T., M. Lang, and L. Stice-Lawrence. 2017. The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation. *Journal of Accounting and Economics* 64 (2): 221-245.
- Durnev, A., R. Morck, B. Yeung, and P. Zarowin. 2003. Does greater firm-specific return variation mean more or less informed stock pricing? *Journal of Accounting Research* 41 (5): 797-836.
- Egorov, G., S. Guriev, and K. Sonin. 2009. Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence from Panel Data. *American Political Science Review* 103: 645-68.

- Fama, E., L. Fisher, M. Jensen, and R. Roll. 1969. The Adjustment of Stock Prices to New Information. *International Economic Review* 10 (1): 1-21.
- Fama, E., and K. French. 1992. The Cross-Section of Expected Stock Returns. *The Journal of Finance* 47 (2): 427-465.
- Fernandes, N., and M. Ferreira. 2008. Does International Cross-Listing Improve the Information Environment. *Journal of Financial Economics* 88: 216-244
- Gul, F., J.B. Kim, and A. Qiu. 2010. Ownership Concentration, Foreign Shareholdings, Audit Quality, and Stock Price Synchronicity: Evidence from China. *Journal of Financial Economics* 95, 425-442.
- Hoberg, G., and C. Lewis. 2017. Do Fraudulent Firms Produce Abnormal Disclosure? *Journal of Corporate Finance* 43: 58-85.
- Huang, A., R. Lehavy, A. Zang, and R. Zheng. 2017. Analyst Information Discovery and Interpretation Roles: A topic modeling approach. *Management Science* 64 (6): 2833-2855.
- Hu, Shuli. 2010. The Rise of the Business Media in China. 77-90 in *Changing Media, Changing China*, edited by Susan L. Shirk. New York: Oxford University Press.
- Jin, L., and S. Myers. 2006. R2 Around the World: New theory and new tests. *Journal of Financial Economics* 79: 257-292.
- King, B. 1966. Market and Industry Factors in Stock Price Behavior. *Journal of Business* 39: 139–190.
- Lin, C. Y., and E. Hovy. 1997. Identifying Topics by Position. In *Proceedings of the fifth conference on Applied natural language processing* (pp. 283-290). Association for Computational Linguistics.
- Liu, M. H. 2011. Analysts' Incentives to Produce Industry-level versus Firm-specific Information. *Journal of Financial and Quantitative Analysis*, 46 (3): 757-784.
- Liu, B. and J. McConnell. 2013. The Role of Media in Corporate Governance: Do the Media Influence Managers' Capital Allocation Decisions. *Journal of Financial Economics* 110: 1-17.
- Lorentzen, P. 2014. China's Strategic Censorship. *American Journal of Political Science* 58: 402-14.
- Meulbroek, L. K. 1992. An Empirical Analysis of Illegal Insider Trading. *Journal of Finance* 47 (5): 1661–1699.
- Miller, G. 2006. The Press as a Watchdog for Accounting Fraud. *Journal of Accounting Research* 44: 1001-33.
- Morck, R., B. Yeung, and W. 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.
- Njølstad, P. C. S., L.S. Høysæter, W. Wei, and J.A. Gulla. (2014, August). Evaluating feature sets and classifiers for sentiment analysis of financial news. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014 IEEE/WIC/ACM International Joint Conferences on* (Vol. 2, pp. 71-78). IEEE.

- Piotroski, J., and D. Roulstone. 2004. The Influence of Analysts, Institutional Investors, and Insiders on the Incorporation of Market, Industry, and Firm-Specific Information into Stock Prices. *The Accounting Review* 79 (4): 1119-1151.
- Piotroski, J., and T.J. Wong. 2012. Institutions and the Information Environment of Chinese Listed Firms. 201-242 in *Capitalizing China*, edited by Joseph Fan and Randall Morck. Chicago: University of Chicago Press.
- Piotroski, J., Wong, T.J. and T. Zhang. 2015. Political Incentives to Suppress Negative Financial Information: Evidence from Chinese Listed Firms. *Journal of Accounting Research* 53 (2): 405-459.
- Piotroski, J., T.J. Wong, and T. Zhang. 2017. Political Bias of Corporate News: Role of Conglomeration Reform in China. *Journal of Law and Economics* 60 (1): 173-207.
- Qin, B., D. Stromberg, and Y. Wu. Media Bias in China. *American Economic Review* 108 (9): 2442-2476.
- Ramnath, S. 2002. Investor and Analyst Reactions to Earnings Announcements of Related Firms: An empirical analysis. *Journal of Accounting Research* 40 (December): 1351–1376.
- Roll, R. 1988. R2. *Journal of Finance* 43 (July): 541–566.
- Seyhun, H. N. 1992. Why does aggregate insider trading predict future stock returns? *Quarterly Journal of Economics* 107 (4): 1303–1331.
- Seyhun, H. N. 1998. *Investment Intelligence From Insider Trading*. Cambridge, MA: MIT Press.
- Soltes, E. 2010. “Disseminating Firm Disclosures.” Working Papers. Harvard Business School.
- Stockmann, D. 2012. *Media Commercialization and Authoritarian Rule in China*. New York. Cambridge University Press.
- Tetlock, P. C. 2007. Giving Content to Investor Sentiment: The role of the media in the stock market. *The Journal of Finance* 62 (3): 1139-1168.
- Tetlock, P. C. 2011. All the News That's Fit to Reprint: Do investors react to stale information? *The Review of Financial Studies*, 24(5), 1481-1512.
- You, L., Zhang, B., and Zhang, L. 2018. Who Captures the Power of the Pen? *Review of Finance Studies* 31(1), 43-96.
- Zhang, C., Zeng, D., Li, J., Wang, F. Y., & Zuo, W. (2009). Sentiment analysis of Chinese documents: From sentence to document level. *Journal of the American Society for Information Science and Technology* 60 (12): 2474-2487.
- Zingales, L. 2000. “In Search of New Foundations.” *The Journal of Finance* 55: 1623-53.

Appendix

Variable Definitions

| | |
|--|--|
| $Return_{it}$ | Daily stock return for firm i on day t . |
| $Absolute\ Return_{it}$ | The absolute value of $Return_{it}$. |
| $Mkt-Adj\ Return_{it}$ | The value weighted market-adjusted return for firm i on day t . |
| $Absolute\ Mkt-Adj\ Return_{it}$ | The absolute value of $Mkt-Adj\ Return_{it}$. |
| $Market\ Return_t$ | The value-weighted market return on day t . |
| $ Market\ Return_t $ | The absolute value of $Market\ Return_t$. |
| $Industry\ Information_{it}$ | <p>This is a measure of the relative amount of industry wide information news contained in firm i's stock return on day t. It is defined as $\frac{ CR_{it}^I }{ CR_{it}^F + CR_{it}^I + CR_{it}^M }$, where CR_{it}^M, the component of market wide information among the firm i's return on day t; CR_{it}^I, the component of industry wide information among the firm i's return on day t; CR_{it}^F, the component of firm specific information among the firm i's return on day t, following the decomposition by Liu (2011).</p> <p>Specifically, in each calendar year from 2000 to 2017, we run the following regression to estimate the market and industry beta of stock n, β_i^M and β_i^I:</p> $R_{it} = \alpha_i + \beta_i^M \times R_t^M + \beta_i^I \times (R_t^I - \beta_i^{IM} \times R_t^M) + \varepsilon_{it} \quad (1)$ <p>Where, R_{it} is return of firm n in period t; R_t^I is the value-weighted return on the one-digit industry portfolio, in which firm n is affiliated, in period of t; R_t^M is the value-weighted market return in period t; β_i^{IM} is the market beta of the industry, in which the firm n is affiliated, estimated from the following regression in each calendar year:</p> $R_t^I = \alpha_i + \beta_i^{IM} \times R_t^M + \varepsilon_{it} \quad (2)$ <p>CR_{it}^M, the component of market wide information among the firm n's return during period t, is defined as $\beta_i^M \times R_t^M$, while β_i^M is estimated from model (1); CR_{it}^I, the component of industry wide information among the firm n's return during period t, is defined as $\beta_i^I \times (R_t^I - \beta_i^{IM} \times R_t^M)$, while β_i^I is estimated from model (1) and β_i^{IM} is estimated from model (2); CR_{it}^F, the component of firm specific information among the firm i's return during period t, is defined as $R_{it} - CR_{it}^M - CR_{it}^I$.</p> |
| $Industry\ \&\ Market\ Information_{it}$ | <p>This is a measure of the relative amount of industry and market information news contained in firm i's stock return on day t. It is defined as $\frac{ CR_{it}^I + CR_{it}^M }{ CR_{it}^F + CR_{it}^I + CR_{it}^M }$, where CR_{it}^M, the component of market wide information among the firm i's return on day t; CR_{it}^I, the component of industry wide information among the firm i's return at time t; CR_{it}^F, the component of firm specific information among the firm i's return on day t, following the decomposition by Liu (2011).</p> |
| $Comovement_{it}$ | The indicator variable that equals one if the firm's stock price moves in the same direction as that of the firm's industry index on day t , and zero otherwise. |
| $Industry\ \&\ Market\ Content$ | <p>The weight of industry- and market- wide topics in a news article. We use Latent Dirichlet Allocation (LDA) to define 400 topics for all news articles. $IND\ \&\ MKT\ Content$ is the sum of the weight of industry- and market- wide topics in the news article. Method to create the market, industry and firm spec topics is as follow: First, we define an article as covering a topic when the weight of article on this topic is more than one percent (Ideally, an article can be defined as 100 topics at most). Second, we classify the topic as industry topic (<i>Industry-wide Topics</i>) when the articles covering the topic take more than 20 percent all the articles covering the firms within a two-digit CSRC industry. This procedure yields 64 industry-specific topics. Finally, we sort the remaining 336 topics according to the number of articles covering each of them and the top quintile (68 topics) out of 336 topics are defined as market-wide topics.</p> |
| $Synchronicity_{it}$ | <p>This is the synchronicity of firm return with market and marker return. It is measured as $\log(\frac{R^2}{1-R^2})$, where R^2 is estimated from the market model as following Morck, Yeung and Yu (2000) for the firm in each calendar year.</p> $RET_{i,t} = \alpha + \beta_1 * MARET_t + \beta_2 * INDRET_{i,t} + \varepsilon_{i,t}$ <p>Where $RET_{i,t}$ is the return in day t for firm i, $MARET_t$ is the value-weighted market return on day t, and $INDRET_{i,t}$ is the value-weighted one-digit industry return on day t.</p> |
| $News\ Coverage_{it}$ | Indicator variable for newspaper coverage, which equals one if there is a news article covering the firm on the day, and zero otherwise. |
| $Official_{it}$ | Indicator variable for coverage by an official newspaper, which equals one if there is an official paper article covering the firm on the day, and zero otherwise. |
| $Non-Official_{it}$ | Indicator variable for coverage by a non-official newspaper, which equals one if there is a non-official paper article covering the firm on the day, and zero otherwise. |
| $Number\ of\ Articles_{it}$ | This is the coverage intensity for a firm in a year, which is measured as the natural logarithm value of one plus number of news articles covering the firm in a year. |

| | |
|---|---|
| <i>Number of Official Articles_{it}</i> | This is the coverage intensity by official newspaper for a firm in a year, which is measured as the natural logarithm value of one plus the number of news articles from official newspapers covering the firm in a year. |
| <i>Number of Non-Official Articles_{it}</i> | This is the coverage intensity by non-official newspaper for a firm in a year, which is measured as the natural logarithm value of one plus the number of news articles from non-official newspapers covering the firm in a year. |
| <i>Relative Official Coverage_{it}</i> | This is the relative coverage intensity of party newspapers over non-party newspapers, which is measured as the natural logarithm value of one plus the number of news articles from official newspapers to one plus that from non-official newspapers, $(1 + OFFICIAL_NUM) / (1 + NON-OFFICIAL_NUM)$. |
| <i>Firm Size_{it}</i> | Firm size, which is measured as the natural logarithm value of total market value of the firm at the end of the year. |
| <i>Market-to-book_{it}</i> | Market-to-book equity ratio, which is the ratio of total market value of equity to the book value of equity at the end of the year. |
| <i>ROA_{it}</i> | Return on assets, which is calculated as the net income divided by total assets of the firm at the end of the year. |
| <i>Leverage_{it}</i> | Financial leverage of the firms, which is measured as the ratio of total long-term liabilities to total assets at the end of the year. |
| <i>SOE_{it}</i> | Indicator variable for state-owned enterprises (SOEs), which equals one if the firm is ultimately controlled by state, and zero otherwise. |
| <i>Earnings Announcement_{it}</i> | Indicator variable for announcement date of earnings report, which equals one for the day when the firm make the announcement for quarterly, semi-annual and annual report, and zero otherwise. |
| Fundamental Correlation _{it} | The logarithmic transformation of the R from a regression of the firm i's quarterly return on assets on a value-weighted industry index of ROA in year t. The variable is estimated using three years of historical data (i.e., 12 quarterly observations in each firm-specific regression). |
| Herfindahl Index _{it} | A revenue-based Herfindahl index of firm's i's industry-level concentration |
| Std Dev (ROA _{it}) | Standard deviation of firm i's quarterly return on assets realizations, measured over the three years including and preceding year t. Each quarter's return on assets is measured as income before extraordinary items scaled by average total assets. |
| Log (# of Industry Firms _{it}) | Log of the number of firms in firm i's two-digit SIC industry in year t. |
| | |

Table 1
Distribution of Sample Newspapers and Sample News Articles

This table presents the distribution of newspapers, newspaper articles, and firm-trading days by region (panel A), by year (panel B), and by industry (panel C). Official refers to newspapers that are affiliated with government agencies, mainly serving a propaganda role for the government and the Chinese communist party (also referred to as “party papers”). Non-Official refers to newspapers other than Official newspapers. Non-official newspapers activities are confined to newspapers with a focus on financial and economic news.

Panel A: Distribution by region

| Province | Newspapers | | | Newspaper Articles | | |
|--------------|------------|----------|--------------|--------------------|----------|--------------|
| | All | Official | Non-Official | All | Official | Non-Official |
| National | 32 | 5 | 27 | 1,395,326 | 111,010 | 1,284,316 |
| Shanghai | 7 | 3 | 4 | 268,266 | 99,335 | 168,931 |
| Yunnan | 2 | 2 | - | 6,207 | 6,207 | - |
| Beijing | 5 | 2 | 3 | 158,383 | 68,458 | 89,925 |
| Sichuan | 6 | 4 | 2 | 100,549 | 28,223 | 72,326 |
| Tianjin | 2 | 1 | 1 | 14,392 | 11,734 | 2,658 |
| Ningxia | 2 | 2 | - | 389 | 389 | - |
| Anhui | 3 | 2 | 1 | 11,155 | 7,061 | 4,094 |
| Shandong | 3 | 3 | - | 26,871 | 26,871 | - |
| Shanxi | 1 | 1 | - | 2,093 | 2,093 | - |
| Guangdong | 10 | 6 | 4 | 520,013 | 281,889 | 238,124 |
| Guangxi | 1 | 1 | - | 14,598 | 14,598 | - |
| Jiangsu | 2 | 2 | - | 14,135 | 14,135 | - |
| Jiangxi | 1 | 1 | - | 15,460 | 15,460 | - |
| Hebei | 1 | 1 | - | 431 | 431 | - |
| Henan | 3 | 2 | 1 | 8,729 | 5,671 | 3,058 |
| Zhejiang | 1 | 1 | - | 20,836 | 20,836 | - |
| Hainan | 1 | 1 | - | 5,512 | 5,512 | - |
| Hubei | 4 | 2 | 2 | 76,347 | 56,510 | 19,837 |
| Hunan | 1 | 1 | - | 39,403 | 39,403 | - |
| Gansu | 3 | 2 | 1 | 45,664 | 19,721 | 25,943 |
| Fujian | 1 | 1 | - | 18,008 | 18,008 | - |
| Liaoning | 4 | 3 | 1 | 6,919 | 6,264 | 655 |
| Chongqing | 1 | - | 1 | 4,131 | - | 4,131 |
| Shaanxi | 1 | 1 | - | 1,370 | 1,370 | - |
| Qinghai | 1 | 1 | - | 611 | 611 | - |
| Heilongjiang | 1 | 1 | - | 1,305 | 1,305 | - |
| Total | 100 | 52 | 48 | 2,777,103 | 863,105 | 1,913,998 |

Panel B: Distribution by year

| Year | Newspapers | | | Newspaper Articles | | | Firm Day Observations | | |
|-------|------------|----------|--------------|--------------------|----------|--------------|-----------------------|---------------------|-------------------|
| | All | Official | Non-Official | All | Official | Non-Official | All | With Media Coverage | No Media Coverage |
| 2000 | 60 | 39 | 21 | 35,404 | 18,171 | 17,233 | 342,345 | 28,514 | 306,914 |
| 2001 | 65 | 40 | 25 | 55,900 | 27,209 | 28,781 | 392,216 | 41,205 | 336,226 |
| 2002 | 58 | 36 | 22 | 66,633 | 36,365 | 30,268 | 425,955 | 46,919 | 359,322 |
| 2003 | 83 | 45 | 38 | 95,428 | 40,602 | 54,826 | 466,264 | 57,294 | 370,836 |
| 2004 | 87 | 47 | 40 | 103,384 | 40,183 | 63,201 | 490,379 | 59,169 | 386,995 |
| 2005 | 86 | 46 | 40 | 168,633 | 40,909 | 127,724 | 563,755 | 79,730 | 395,122 |
| 2006 | 68 | 34 | 34 | 151,815 | 35,212 | 116,603 | 542,734 | 83,218 | 390,919 |
| 2007 | 68 | 35 | 33 | 197,699 | 49,803 | 147,896 | 598,320 | 93,655 | 400,621 |
| 2008 | 83 | 48 | 35 | 267,506 | 71,179 | 196,327 | 695,545 | 114,984 | 428,039 |
| 2009 | 60 | 37 | 23 | 208,410 | 63,100 | 145,310 | 669,694 | 94,245 | 461,284 |
| 2010 | 40 | 23 | 17 | 184,776 | 62,873 | 121,903 | 678,907 | 89,720 | 494,131 |
| 2011 | 47 | 26 | 21 | 199,052 | 68,746 | 130,306 | 772,253 | 98,705 | 573,201 |
| 2012 | 47 | 25 | 22 | 233,405 | 70,154 | 163,251 | 961,099 | 113,747 | 727,694 |
| 2013 | 50 | 26 | 24 | 236,039 | 63,493 | 172,546 | 1,023,431 | 113,959 | 787,392 |
| 2014 | 49 | 25 | 24 | 184,336 | 55,656 | 128,080 | 993,367 | 93,157 | 809,031 |
| 2015 | 43 | 20 | 23 | 149,211 | 42,523 | 106,688 | 1,009,052 | 82,639 | 859,841 |
| 2016 | 39 | 18 | 21 | 133,699 | 41,915 | 91,784 | 1,079,573 | 78,436 | 945,874 |
| 2017 | 37 | 17 | 20 | 105,683 | 35,012 | 70,671 | 1,080,153 | 28,514 | 974,470 |
| Total | | | | 2,777,103 | 863,105 | 1,913,998 | 12,785,042 | 1,369,296 | 10,007,939 |

Table 1 (continued)
Sample: Newspaper and News Article Distributions

Panel C: Distribution by industry

| Industry | Newspapers | | | Newspaper Articles | | | Firm Day Observations | | |
|----------------------|------------|----------|--------------|--------------------|----------|--------------|-----------------------|---------------------|-------------------|
| | All | Official | Non-Official | All | Official | Non-Official | All | With Media Coverage | No Media Coverage |
| Agriculture | 94 | 49 | 45 | 22,495 | 5,214 | 17,281 | 182,388 | 16,952 | 159,893 |
| Mining | 100 | 52 | 48 | 131,410 | 35,207 | 96,203 | 413,529 | 50,089 | 282,119 |
| Manufacturing | 100 | 52 | 48 | 1,058,766 | 300,806 | 757,960 | 6,947,044 | 684,063 | 5,888,278 |
| Energy & Utilities | 100 | 52 | 48 | 73,728 | 16,946 | 56,782 | 547,237 | 53,212 | 473,509 |
| Construction | 96 | 50 | 46 | 60,267 | 21,964 | 38,303 | 324,200 | 35,804 | 263,933 |
| Retail & Wholesale | 100 | 52 | 48 | 147,043 | 54,198 | 92,845 | 846,047 | 89,735 | 699,004 |
| Transportation | 100 | 52 | 48 | 166,697 | 76,050 | 90,647 | 506,822 | 77,307 | 340,125 |
| Travel & Tourism | 92 | 49 | 43 | 12,588 | 4,560 | 8,028 | 61,289 | 8,466 | 48,701 |
| Info Svc & Software | 100 | 52 | 48 | 115,108 | 33,265 | 81,843 | 581,200 | 59,472 | 466,092 |
| Financial | 100 | 52 | 48 | 610,551 | 183,027 | 427,524 | 743,778 | 106,499 | 133,227 |
| Real Estate | 100 | 52 | 48 | 198,905 | 60,460 | 138,445 | 856,898 | 93,042 | 657,993 |
| Leasing & Commere. | 100 | 52 | 48 | 59,235 | 24,157 | 35,078 | 185,603 | 27,172 | 126,368 |
| Scientific Research | 81 | 42 | 39 | 4,460 | 1,692 | 2,768 | 51,792 | 3,075 | 47,332 |
| Environmental | 96 | 51 | 45 | 47,970 | 21,911 | 26,059 | 150,691 | 22,041 | 102,721 |
| Education | 60 | 30 | 30 | 935 | 155 | 780 | 15,656 | 751 | 12,721 |
| Health; Social work | 78 | 38 | 40 | 3,824 | 908 | 2,916 | 30,653 | 3,246 | 26,829 |
| Sport; Entertainment | 94 | 50 | 44 | 39,692 | 16,216 | 23,476 | 179,203 | 22,655 | 139,511 |
| Diversified | 99 | 51 | 48 | 23,429 | 6,369 | 17,060 | 163,012 | 15,715 | 139,583 |
| Total | | | | 2,777,103 | 863,105 | 1,913,998 | 12,785,042 | 1,369,296 | 10,007,939 |

Table 2
Descriptive Statistics

This table presents descriptive statistics on the variable used in our main empirical analyses. All variables are defined in the Appendix. All continuous variables are winsorized by 1% on both tails.

| | Number of Observations | Mean | 25 th Percentile | Median | 75 th Percentile | Standard Deviation |
|-------------------------------|---------------------------|-------|--------------------------------|--------|--------------------------------|-----------------------|
| Return | 12,785,042 | 0.00 | -0.01 | 0.00 | 0.02 | 0.04 |
| Return | 12,785,042 | 0.03 | 0.01 | 0.02 | 0.03 | 0.03 |
| Mkt-Adj Return | 12,785,042 | 0.00 | -0.01 | 0.00 | 0.01 | 0.03 |
| Mkt-Adj Return | 12,785,042 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 |
| Market Return | 12,785,042 | 0.00 | -0.01 | 0.00 | 0.01 | 0.02 |
| Market Return | 12,785,042 | 0.01 | 0.00 | 0.01 | 0.02 | 0.01 |
| Industry Information | 12,780,776 | 0.25 | 0.05 | 0.17 | 0.41 | 0.25 |
| Industry & Market Information | 12,780,776 | 0.52 | 0.33 | 0.53 | 0.72 | 0.25 |
| Comovement | 12,785,042 | 0.69 | 0.00 | 1.00 | 1.00 | 0.47 |
| News Coverage | 12,785,042 | 0.22 | 0.00 | 0.00 | 0.00 | 0.41 |
| Industry & Market Content | 2,770,718 | 0.32 | 0.18 | 0.30 | 0.45 | 0.18 |
| Firm Size | 12,785,042 | 15.58 | 14.62 | 15.36 | 16.22 | 1.42 |
| Market-to-book | 12,785,042 | 3.90 | 1.73 | 2.75 | 4.57 | 3.92 |
| ROA | 12,785,042 | 0.03 | 0.01 | 0.03 | 0.06 | 0.06 |
| Leverage | 12,785,042 | 0.49 | 0.32 | 0.49 | 0.64 | 0.22 |
| SOE | 12,785,042 | 0.53 | 0.00 | 1.00 | 1.00 | 0.50 |
| Earnings Announcement | 12,785,042 | 0.01 | 0.00 | 0.00 | 0.00 | 0.11 |

Table 3**Univariate evidence on the information content of Chinese newspapers' business articles**

This table presents univariate evidence on the information content of newspaper articles published in official and non-official newspapers in China. Panel A presents average stock return characteristics on days with or without a published newspaper article about the listed firm. Panel B presents average one-day stock return characteristics on the first trading day related to the publication of a particular news article. All variables are defined in the Appendix.

Panel A: Daily stock return characteristics conditional on the existence of news coverage

| Information Content Metric | Media Coverage (n=2,777,103) | No Media Coverage (n=10,007,939) | Difference (Cover – No Cover) | t-stat |
|---|---------------------------------|-------------------------------------|----------------------------------|---------|
| Absolute Return _{it} | 0.032 | 0.029 | 0.003*** | 120.78 |
| Absolute Mkt-Adj. Return _{it} | 0.021 | 0.019 | 0.002*** | 103.73 |
| Industry Information _{it} | 0.263 | 0.251 | 0.012*** | 73.48 |
| Industry & Market Information _{it} | 0.513 | 0.526 | -0.010*** | -70.17 |
| Industry Comovement _{it} | 0.654 | 0.693 | -0.039*** | -126.61 |

Panel B: Daily stock return characteristics conditional on newspaper outlet

| Information Content Metric | Official Newspaper (n=863,105) | Non-Official Newspaper (n=1,913,998) | Difference (Official – Non-Official) | t-stat |
|---|--------------------------------------|--|---|---------|
| Absolute Return _{it} | 0.030 | 0.034 | -0.004*** | -138.62 |
| Absolute Mkt-Adj. Return _{it} | 0.018 | 0.022 | -0.004*** | -154.29 |
| Industry Information _{it} | 0.268 | 0.261 | 0.007*** | 19.81 |
| Industry & Market Information _{it} | 0.348 | 0.314 | 0.034*** | 144.32 |
| Industry Comovement _{it} | 0.663 | 0.650 | 0.013*** | 21.15 |

Table 4
Information content of Chinese corporate news articles (average absolute stock price reaction)

This table presents evidence on the information content of Chinese newspaper business articles, specifically, select coefficients from an estimation of the following model:

$$|Return_{it}| \text{ or } |Mkt-Adj Return_{it}| = \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 |Market Return_{it}| + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}$$

In these estimations, the dependent variable $|Return_{it}|$ ($|Mkt-adj Return_{it}|$) is the absolute value of firm i 's raw stock return (market adjusted stock return) on day t . The indicator variable $News Coverage_{it}$ equals one if a domestic Chinese newspaper published a business news article about firm i on day t , zero otherwise. The indicator variable $Official_{it}$ equals one if an official newspaper published a business news article about firm i on day t , zero otherwise. All other variables are defined in the appendix. All reported t -statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. All models include an array of year-month and firm fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. $N=12,785,042$ observations.

| Dependent Variable: | $ Return_{it} $ | | $ Mkt-Adj Return_{it} $ | |
|--|----------------------|-----------------------|-------------------------|-----------------------|
| | (1) | (2) | (3) | (4) |
| News Coverage _{it} | 0.005*** (44.85) | 0.006*** (47.23) | 0.006*** (46.107) | 0.007*** (48.28) |
| News Coverage _{it} * Official _{it} | | -0.002*** (-13.97) | | -0.002*** (-13.92) |
| $ Market Return_{it} $ | 0.666*** (152.74) | 0.666*** (152.73) | | |
| Firm Size _{it} | 0.001*** (5.09) | 0.001*** (5.03) | 0.001*** (5.96) | 0.001*** (5.89) |
| Market-to-book _{it} | 0.000*** (14.53) | 0.000*** (14.50) | 0.001*** (17.76) | 0.001*** (17.72) |
| ROA _{it} | -0.005*** (-3.61) | -0.005*** (-3.60) | -0.001 (-0.84) | -0.001 (-0.83) |
| Leverage _{it} | -0.002** (-2.40) | -0.002** (-2.38) | -0.001 (-1.54) | -0.001 (-1.52) |
| SOE _{it} | -0.001** (-2.19) | -0.001** (-2.20) | -0.001** (-2.42) | -0.001** (-2.43) |
| Earnings Announcement _{it} | 0.003*** (26.54) | 0.003*** (26.07) | 0.004*** (30.76) | 0.004*** (30.43) |
| Intercept | 0.001 (0.43) | 0.001 (0.50) | -0.000 (-0.01) | 0.000 (0.07) |
| Sum of $\beta_1 + \beta_2$ | | 0.004*** (75.13) | | 0.005*** (71.00) |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.250 | 0.250 | 0.096 | 0.096 |

Table 5**Information content of Chinese corporate news articles: Firm-specific, industry or market news?**

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, and presents select coefficients from an estimation of the following empirical models:

$$\text{Industry Information}_{it} \text{ or Industry \& Market Information}_{it} = \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it}$$

$$\text{Comovement}_{it} = \text{Logit} (\alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it})$$

In the first set of estimations, the dependent variable *Industry Information_{it}* equals the fraction of firm i's raw stock return that reflects industry information on day t. In the second set of estimations, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day t. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm i's raw stock return moved in the same direction as the firm's relevant industry index on day t, zero otherwise. The indicator variable *News Coverage_{it}* equals one if a domestic Chinese newspaper published a business news article about firm i on day t, zero otherwise. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm i on day t, zero otherwise. All other variables are defined in the appendix. An OLS model is applied to estimations (1) – (4), and a Logit model is applied to estimations (5) – (6). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively.

| Dependent Variable: | Industry Information (n=12,780,711) | | Industry & Market Information (n=12,780,711) | | Comovement _{it} (n=12,785,042) | |
|---|--|-----------------------|---|-----------------------|--|------------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| News Coverage _{it} | -0.023*** (-24.89) | -0.025*** (-26.08) | -0.049*** (-52.56) | -0.053*** (-54.34) | -0.118*** (-17.441) | -0.125*** (-17.746) |
| News Coverage _{it} *Official _{it} | | 0.006*** (5.59) | | 0.013*** (12.89) | | 0.024*** (3.612) |
| Firm Size _{it} | 0.000 (0.16) | 0.000 (0.17) | -0.010*** (-5.20) | -0.009*** (-5.16) | -0.017*** (-3.289) | -0.017*** (-3.326) |
| Market-to-book _{it} | -0.001*** (-4.25) | -0.001*** (-4.24) | -0.005*** (-17.99) | -0.005*** (-17.96) | -0.014*** (-18.938) | -0.014*** (-18.894) |
| ROA _{it} | 0.022 (1.15) | 0.022 (1.15) | -0.011 (-0.79) | -0.011 (-0.80) | 0.243*** (4.559) | 0.243*** (4.561) |
| Leverage _{it} | 0.005 (0.57) | 0.005 (0.57) | 0.005 (0.85) | 0.005 (0.83) | 0.055*** (3.328) | 0.054*** (3.297) |
| SOE _{it} | -0.000 (-0.028) | -0.000 (-0.028) | 0.003 (0.80) | 0.003 (0.80) | 0.024*** (3.156) | 0.023*** (3.154) |
| Earnings Announcement _{it} | -0.024*** (-20.71) | -0.024*** (-20.51) | -0.049*** (-38.36) | -0.048*** (-38.04) | -0.182*** (-14.994) | -0.181*** (-14.950) |
| Intercept | 0.161*** (4.24) | 0.160*** (4.22) | 0.755*** (26.70) | 0.753*** (26.65) | 1.397*** (17.210) | 1.400*** (17.230) |
| Sum of $\beta_1 + \beta_2$ | | -0.019*** (21.33) | | -0.040*** (12.23) | | -0.101*** (14.87) |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | No | No |
| Adj. / Pseudo R-squared | 0.055 | 0.055 | 0.040 | 0.040 | 0.010 | 0.010 |

Table 6
Analysis of Corporate News Articles: Topical Content

This table documents the relative amount of industry and market topics versus firm-specific content contained in business news articles published by Chinese newspapers over the period 2000-2017. Panel A presents descriptive statistics and univariate evidence. Panel B presents select coefficients from an estimation of the following model:

$$\text{Industry \& Market Content}_{it} = \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Official}_{it} + \beta_2 \text{Firm Size}_{it} + \beta_3 \text{Market-to-book}_{it} + \beta_4 \text{ROA}_{it} + \beta_5 \text{Leverage}_{it} + \beta_6 \text{SOE}_{it} + \beta_7 \text{Earnings Announcement}_{it} + \varepsilon_{it}$$

In this model, the dependent variable *Industry & Market Content_{it}* is the relative weight of industry and market-level topics covered in a news article about firm *i* on day *t*. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm *i* on day *t*, zero otherwise. All other variables are defined in the appendix. All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. All models include an array of year-month and firm fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. N=2,770,718.

Panel A: Descriptive evidence

| MI | Mean All Articles (N=2,770,718) | Official Newspaper (n=863,105) | Non-Official Newspaper (n=1,913,998) | Difference Official – Non-Official (t-statistic) |
|---------------------------|---------------------------------------|--------------------------------------|--|--|
| Industry & Market Content | 0.324 | 0.348 | 0.314 | 0.034*** (144.32) |

Panel B: Multivariate estimation

| Dependent variable: | Industry & Market Content |
|-------------------------------------|---------------------------|
| Official _{it} | 0.016*** (7.86) |
| Firm Size _{it} | 0.010*** (5.269) |
| Market-to-book _{it} | -0.000 (-0.484) |
| ROA _{it} | 0.001 (0.08) |
| Leverage _{it} | 0.002 (0.23) |
| SOE _{it} | 0.009* (2.84) |
| Earnings Announcement _{it} | -0.068*** (-28.80) |
| Intercept | 0.201*** (6.24) |
| Year-Month Fixed Effect | Yes |
| Firm Fixed Effect | Yes |
| Adjusted R-squared | 0.015 |

Table 7
Influence of Corporate News Topical Content on their Information Content

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, and presents select coefficients from an estimation of the following empirical models:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or Industry \& Market Information}_{it} &= \alpha + \text{Year-Month FE} + \text{Firm FE} \\ &+ \beta_1 \text{Industry \& Market Content}_{it} + \beta_2 \text{Industry and Market Content}_{it} * \text{Official}_{it} + \beta_3 \text{Official}_{it} + \beta_4 \text{Firm Size}_{it} \\ &+ \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it} \\ \\ \text{Comovement}_{it} &= \text{Logit} (\alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Industry \& Market Content}_{it} \\ &+ \beta_2 \text{Industry and Market Content}_{it} * \text{Official}_{it} + \beta_3 \text{Official}_{it} + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} \\ &+ \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned}$$

In the first set of estimations, the dependent variable *Industry Information_{it}* equals the fraction of firm i's raw stock return that reflects industry information on day t. In the second set of estimations, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day t. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm i's raw stock return moved in the same direction as the firm's relevant industry index on day t, zero otherwise. The independent variable *Industry \& Market Content_{it}* is the relative weight of industry and market-level topics covered in a news article about firm i on day t. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm i on day t, zero otherwise. All other variables are defined in the appendix. An OLS model is applied to estimations (1) – (4), and a Logit model is applied to estimations (5) – (6). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively.

| Dependent Variable: | Industry Information | Industry \& Market Information | Comovement |
|---|-------------------------|-----------------------------------|---------------------|
| Industry \& Market Content _{it} | 0.009* (1.70) | 0.007*** (3.59) | 0.032* (1.887) |
| Industry \& Market Content _{it} * Official _{it} | 0.012*** (5.08) | 0.024*** (11.18) | 0.043*** (2.590) |
| Official _{it} | -0.001 (-0.81) | 0.001 (0.78) | 0.004 (0.596) |
| Control Variables | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes |
| Firm-Fixed Effects | Yes | Yes | No |
| Adjusted R-squared | 0.047 | 0.031 | 0.012 |
| N | 2,768,021 | 2,768,021 | 2,770,718 |

Table 8
State owned media and stock return synchronicity

This table presents select coefficients from various estimations of the following empirical model:

$$\begin{aligned} \text{Synchronicity}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Number of Articles}_{it} + \beta_2 \text{Number of Official Articles}_{it} \\ & + \beta_3 \text{Number of Non-Official Articles}_{it} + \beta_4 \text{Relative Official}_{it} + \beta_5 \text{Fundamental Correlation}_{it} \\ & + \beta_6 \text{Herfindahl Index}_{it} + \beta_7 \text{Std Dev (ROA)}_{it} + \beta_8 \text{Log(\# of Industry Firms)}_{it} + \beta_9 \text{Firm Size}_{it} \\ & + \beta_{10} \text{Market-to-book}_{it} + \beta_{11} \text{ROA}_{it} + \beta_{12} \text{Leverage}_{it} + \beta_{13} \text{SOE}_{it} + \varepsilon_{it} \end{aligned}$$

The dependent variable, *Synchronicity_{it}*, captures the comovement of the firm's stock return with the overall market return. Following Morck, Yeung, and Yu (2000), *Synchronicity_{it}* is measured as $\log(\frac{R^2}{1-R^2})$, where R^2 is estimated from the market model for firm *i* in year *t* using daily return data. Our measure of overall newspaper coverage intensity, *Number of Articles_{it}*, is measured as the natural logarithm of one plus the number of news articles covering firm *i* in year *t*. Our measure of official newspaper coverage intensity, *Number of Official Articles_{it}*, is measured as the natural logarithm of one plus the number of news articles about firm *i* published in official newspapers in year *t*. Our measure of non-official newspaper coverage intensity, *Number of Non-Official Articles_{it}*, is measured as the natural logarithm of one plus the number of news articles about firm *i* published in non-official newspapers in year *t*. Our measure of the relative intensity of media coverage by official newspapers, *Relative Official Coverage_{it}*, is measured as the ratio of one plus *Number of Official Articles_{it}* to one plus *Number of Non-Official Articles_{it}*. All other variables are defined in the appendix. All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. All models include an array of year and firm fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. N=32,815 firm-year observations.

| Estimation: | (1) | (2) | (3) |
|---|------------------------|------------------------|------------------------|
| Number of Articles _{it} | -0.193*** (-20.620) | | |
| Number of Official Articles _{it} | | 0.035*** (4.506) | |
| Number of Non-Official Articles _{it} | | -0.232*** (-24.466) | |
| Relative Official Coverage _{it} | | | 0.200*** (7.724) |
| Fundamental Correlation | 0.009*** (3.427) | 0.009*** (3.222) | 0.009*** (3.125) |
| Herfindhal Index | -0.856*** (-5.708) | -0.872*** (-5.858) | -0.828*** (-5.445) |
| Std Dev (ROA) | -1.162*** (-6.571) | -1.128*** (-6.392) | -1.353*** (-7.632) |
| Log(# of Industry Firms) | 0.003 (0.085) | -0.000 (-0.005) | 0.018 (0.455) |
| Firm Size _{it} | -0.073*** (-4.795) | -0.069*** (-4.521) | -0.185*** (-13.215) |
| Market-to-book _{it} | -0.040*** (-18.476) | -0.039*** (-18.165) | -0.041*** (-18.933) |
| ROA _{it} | -0.286** (-2.566) | -0.286*** (-2.582) | -0.230** (-2.060) |
| Leverage _{it} | -0.028 (-0.533) | -0.023 (-0.439) | -0.099* (-1.901) |
| SOE _{it} | 0.046* (1.820) | 0.041 (1.624) | 0.069*** (2.818) |
| Intercept | 1.998*** (7.078) | 1.836*** (6.542) | 2.813*** (10.315) |
| Year Fixed Effects | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes |
| Adjusted R-squared | 0.332 | 0.336 | 0.316 |

Table 9
Information Content of Official Corporate News Articles conditional upon Newspaper Centrality

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, and presents select coefficients from an estimation of the following empirical models:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} \\ & + \beta_2 \text{News Coverage}_{it} * \text{Central Official}_{it} + \beta_3 \text{News Coverage}_{it} * \text{Local Official}_{it} + \beta_4 \text{Firm Size}_{it} \\ & + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned}$$

$$\begin{aligned} \text{Comovement}_{it} = & \text{Logit} (\alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} \\ & + \beta_2 \text{News Coverage}_{it} * \text{Central Official}_{it} + \beta_3 \text{News Coverage}_{it} * \text{Local Official}_{it} + \beta_4 \text{Firm Size}_{it} \\ & + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned}$$

In the first set of estimations, the dependent variable *Industry Information_{it}* equals the fraction of firm *i*'s raw stock return that reflects industry information on day *t*. In the second set of estimations, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day *t*. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm *i*'s raw stock return moved in the same direction as the firm's relevant industry index on day *t*, zero otherwise. The indicator variable *News Coverage_{it}* equals one if a domestic Chinese newspaper published a business news article about firm *i* on day *t*, zero otherwise. The indicator variable *Central Official_{it}* equals one if an official newspaper controlled by the central government published a business news article about firm *i* on day *t*, zero otherwise. The indicator variable *Local Official_{it}* equals one if an official newspaper controlled by a local government agency or party unit published a business news article about firm *i* on day *t*, zero otherwise. All other variables are defined in the appendix. An OLS model is applied to estimations (1) and (2), and a Logit model is applied to estimation (3). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively.

| Dependent Variable: | Industry Information | Industry & Market Information | Comovement |
|--|-----------------------|-------------------------------|------------------------|
| News Coverage _{it} | -0.025*** (-26.13) | -0.053 (-54.44) | -0.125*** (-17.752) |
| News Coverage _{it} * Central Official _{it} | 0.012*** (7.15) | 0.020*** (15.53) | 0.030*** (3.176) |
| News Coverage _{it} * Local Official _{it} | 0.005*** (4.39) | 0.012*** (11.62) | 0.023*** (3.253) |
| Test of Difference (β ₂ – β ₃) (Chi-squared statistic) | 0.007*** (10.90) | 0.008*** (54.93) | 0.007 (1.45) |
| Control Variables | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes |
| Firm-Fixed Effects | Yes | Yes | No |
| Adjusted R-squared | 0.047 | 0.031 | 0.010 |
| N | 12,780,711 | 12,780,711 | 12,785,042 |

Table 10
Information Content of Corporate News Articles conditional upon Prevailing Political Forces

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, condition upon the presence of a National Congress Meeting of the CPC during the year. Specifically, the table presents select coefficients from an estimation of the following empirical models:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} \\ & + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} \\ & + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned}$$

$$\begin{aligned} \text{Comovement}_{it} = & \text{Logit} (\alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} \\ & + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned}$$

In the first set of estimations, the dependent variable *Industry Information_{it}* equals the fraction of firm i's raw stock return that reflects industry information on day t. In the second set of estimations, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day t. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm i's raw stock return moved in the same direction as the firm's relevant industry index on day t, zero otherwise. The indicator variable *News Coverage_{it}* equals one if a domestic Chinese newspaper published a business news article about firm i on day t, zero otherwise. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm i on day t, zero otherwise. All other variables are defined in the appendix. An OLS model is applied to estimations (1) – (4), and a Logit model is applied to estimations (5) – (6). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively

| Dependent Variable: | Industry Information | | Industry & Market Information | | Comovement _{it} | |
|--|-----------------------|-----------------------|-------------------------------|-----------------------|--------------------------|-----------------------|
| Estimation: | (1) | (2) | (3) | (4) | (5) | (6) |
| | No Congress | Congress Period | No Congress | Congress Period | No Congress | Congress Period |
| News Coverage _{it} | -0.025*** (-25.79) | -0.023*** (-16.94) | -0.054*** (-53.47) | -0.047*** (-36.52) | -0.133*** (-18.161) | -0.092*** (-9.451) |
| Test of Difference (C-NC) (Chi-Squared Statistic) | | 0.002*** (6.10) | | 0.007*** (179.91) | | 0.041*** (282.62) |
| News Coverage _{it} *Official _{it} | 0.005*** (4.79) | 0.009*** (5.28) | 0.013*** (12.49) | 0.014*** (8.91) | 0.021*** (2.755) | 0.032*** (3.444) |
| Test of Difference (C-NC) (Chi-Squared Statistic) | | 0.004** (26.91) | | 0.001 (2.52) | | 0.011*** (58.60) |
| Control Variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | No | No |
| Adjusted R-squared | 0.060 | 0.042 | 0.041 | 0.040 | 0.018 | 0.016 |
| N | 9,886,911 | 2,893,800 | 9,886,911 | 2,893,800 | 9,889,388 | 2,895,654 |

Table 11
Information Content of Corporate News Articles conditional upon Xi Jinping Visit to State Media

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, condition upon whether the article was written before or after Xi Jinping's visit to the State Media. Specifically, the table presents select coefficients from an estimation of the following empirical models:

$$\begin{aligned} \text{Industry Information}_{it} \text{ or Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} \\ & + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} \\ & + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned}$$

$$\begin{aligned} \text{Comovement}_{it} = & \text{Logit} (\alpha + \text{Year-Month FE} + \text{Industry FE} + \text{Province FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} \\ & + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned}$$

In the first set of estimations, the dependent variable *Industry Information_{it}* equals the fraction of firm *i*'s raw stock return that reflects industry information on day *t*. In the second set of estimations, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day *t*. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm *i*'s raw stock return moved in the same direction as the firm's relevant industry index on day *t*, zero otherwise. The indicator variable *News Coverage_{it}* equals one if a domestic Chinese newspaper published a business news article about firm *i* on day *t*, zero otherwise. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm *i* on day *t*, zero otherwise. All other variables are defined in the appendix. An OLS model is applied to estimations (1) – (4), and a Logit model is applied to estimations (5) – (6). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively

| Dependent Variable: | Industry Information | | Industry & Market Information | | Comovement _{it} | |
|---|-----------------------|-----------------------|-------------------------------|-----------------------|--------------------------|-----------------------|
| Estimation: | (1) | (2) | (3) | (4) | (5) | (6) |
| | Before | After | Before | After | Before | After |
| | Visit | Visit | Visit | Visit | Visit | Visit |
| News Coverage _{it} | -0.025*** (-26.53) | -0.026*** (-21.56) | -0.052*** (-54.30) | -0.063*** (-41.46) | -0.126*** (-18.014) | -0.160*** (-8.837) |
| Test of Difference (A-B) (Chi-Squared Statistic) | | -0.001 (0.04) | | -0.012*** (148.37) | | -0.034*** (82.06) |
| News Coverage _{it} *Official _{it} | 0.006*** (4.90) | 0.012*** (8.53) | 0.012*** (11.77) | 0.022*** (12.96) | 0.019*** (2.744) | 0.091*** (6.319) |
| Test of Difference (A-B) (Chi-Squared Statistic) | | 0.006** (97.29) | | 0.010*** (221.32) | | 0.072*** (18.45) |
| Control Variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | No | No |
| Adjusted R-squared | 0.059 | 0.008 | 0.038 | 0.024 | 0.018 | 0.012 |
| N | 10,765,779 | 2,014,932 | 10,765,779 | 2,014,932 | 10,769,023 | 2,016,019 |

Table 12
Analysis of Corporate News Articles Topical Content conditional upon Prevailing Political Forces

This table documents the relative amount of industry and market topics versus firm-specific content contained in business news articles published by Chinese newspapers over the period 2000-2017. Panel A presents descriptive statistics and univariate evidence. Panel B presents select coefficients from an estimation of the following model:

$$\begin{aligned} \text{Industry \& Market Content}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Official}_{it} + \beta_2 \text{Official}_{it} * \text{National Official}_{it} \\ & + \beta_3 \text{National Congress}_{it} + \beta_4 \text{Official}_{it} * \text{National Congress}_{it} + \beta_5 \text{Post Visit}_{it} + \beta_6 \text{Official}_{it} * \text{Post Visit}_{it} \\ & + \beta_7 \text{Firm Size}_{it} + \beta_8 \text{Market-to-book}_{it} + \beta_9 \text{ROA}_{it} + \beta_{10} \text{Leverage}_{it} + \beta_{11} \text{SOE}_{it} \\ & + \beta_{12} \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned}$$

In this model, the dependent variable *Industry & Market Content_{it}* is the relative weight of industry and market-level topics covered in a news article about firm *i* on day *t*. The indicator variable *Central Official_{it}* equals one if an official newspaper publishing a business news article about firm *i* on day *t* is controlled by the Central Government, zero otherwise. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm *i* on day *t*, zero otherwise. All other variables are defined in the appendix. All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. All models include an array of year-month and firm fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. N=2,770,718.

| Dependent variable: Political Attribute: | Industry & Market Content | | |
|---|----------------------------|-----------------------------|------------------------------------|
| | Centrality of Newspaper | National Congress Period | Post Xi Jinping Media Statement |
| Official _{it} | 0.009*** (3.326) | 0.017*** (7.623) | 0.018*** (8.007) |
| Official _{it} * Central Official _{it} | 0.008*** (3.823) | | |
| National Congress _{it} | | -0.006*** (-5.722) | |
| Official _{it} * National Congress | | 0.008*** (5.605) | |
| Post Visit _{it} | | | 0.007*** (2.657) |
| Official _{it} * Post Visit _{it} | | | 0.010*** (2.636) |
| Firm Size _{it} | 0.010*** (5.271) | 0.000 (0.321) | -0.001 (-0.771) |
| Market-to-book _{it} | -0.000 (-0.490) | 0.000 (1.388) | 0.000* (1.731) |
| ROA _{it} | 0.001 (0.065) | 0.012 (0.845) | 0.017 (1.221) |
| Leverage _{it} | 0.002 (0.229) | -0.018** (-2.181) | -0.018** (-2.215) |
| SOE _{it} | 0.009*** (2.828) | 0.010*** (3.041) | 0.011*** (3.532) |
| Earnings Announcement _{it} | -0.068*** (-28.779) | -0.070*** (-27.020) | -0.070*** (-27.068) |
| Intercept | 0.202*** (6.283) | 0.318*** (15.416) | 0.338*** (14.589) |
| Year-Month Fixed Effect | Yes | Yes | Yes |
| Firm Fixed Effect | Yes | Yes | Yes |
| Adjusted R-squared | 0.015 | 0.008 | 0.008 |

Table 13
Robustness Tests using an Alternative Measure of Industry Content

This table documents the relative amount of industry and market topics versus firm-specific content contained in business news articles published by Chinese newspapers over the period 2000-2017. Panel A (B) presents select coefficients from an estimation of the following empirical models:

$$\text{Industry \& Market Content}_{it} (\text{Industry Count}) = \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Official}_{it} + \beta_2 \text{Firm Size}_{it} + \beta_3 \text{Market-to-book}_{it} + \beta_4 \text{ROA}_{it} + \beta_5 \text{Leverage}_{it} + \beta_6 \text{SOE}_{it} + \beta_7 \text{Earnings Announcement}_{it} + \varepsilon_{it}$$

$$\begin{aligned} \text{Industry Information}_{it} \text{ or } \text{Industry \& Market Information}_{it} = & \alpha + \text{Year-Month FE} + \text{Firm FE} \\ & + \beta_1 \text{Industry Content Metric}_{it} + \beta_2 \text{Industry Content Metric}_{it} * \text{Official}_{it} + \beta_3 \text{Official}_{it} + \beta_4 \text{Firm Size}_{it} \\ & + \beta_5 \text{Market-to-book}_{it} + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it} \end{aligned}$$

$$\begin{aligned} \text{Comovement}_{it} = \text{Logit} (& \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{Industry Content Metric}_{it} \\ & + \beta_2 \text{Industry Content Metric}_{it} * \text{Official}_{it} + \beta_3 \text{Official}_{it} + \beta_4 \text{Firm Size}_{it} + \beta_5 \text{Market-to-book}_{it} \\ & + \beta_6 \text{ROA}_{it} + \beta_7 \text{Leverage}_{it} + \beta_8 \text{SOE}_{it} + \beta_9 \text{Earnings Announcement}_{it} + \varepsilon_{it}) \end{aligned}$$

In the first model, the dependent variable *Industry & Market Content_{it}* is the relative weight of industry and market-level topics covered in a news article about firm *i* on day *t*, and the dependent variable *Industry Count_{it}* is the frequency of industry-name mentioned in a news article about firm *i* on day *t*, deflated by the total number of sentences in an article. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm *i* on day *t*, zero otherwise. All other variables are defined in the appendix. All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. All estimations of the first model include an array of year-month and firm fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. In the second model, the dependent variable *Industry Information_{it}* equals the fraction of firm *i*'s raw stock return that reflects industry information on day *t*, the dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day *t*. In the third model, the dependent variable *Comovement_{it}* is an indicator variable equal to one if firm *i*'s raw stock return moved in the same direction as the firm's relevant industry index on day *t*, zero otherwise. The independent variable *Industry & Market Content Metric_{it}* is one of our two measures of article content, *Industry & Market Content_{it}* or *Industry Count_{it}*. An OLS model is applied to estimations (1) – (4), and a Logit model is applied to estimations (5) – (6). All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively. N=2,770,718.

Panel A: Multivariate estimation of News Article's Content

| Dependent variable: | Industry & Market Content | Industry Count |
|-------------------------------------|---------------------------|------------------------|
| Official _{it} | 0.016*** (7.86) | 0.005** (2.574) |
| Firm Size _{it} | 0.010*** (5.269) | 0.000 (0.221) |
| Market-to-book _{it} | -0.000 (-0.484) | -0.001*** (-5.162) |
| ROA _{it} | 0.001 (0.08) | 0.009 (0.634) |
| Leverage _{it} | 0.002 (0.23) | -0.009 (-1.280) |
| SOE _{it} | 0.009* (2.84) | 0.001 (0.352) |
| Earnings Announcement _{it} | -0.068*** (-28.80) | -0.023*** (-16.084) |
| Intercept | 0.201*** (6.24) | 1.068*** (38.809) |
| Year-Month Fixed Effect | Yes | Yes |
| Firm Fixed Effect | Yes | Yes |
| Adjusted R-squared | 0.015 | 0.016 |

Table 13 (Continued)
Robustness Tests using an Alternative Measure of Industry Content

Panel B: Mapping of Article Content into Market Information

| Dependent Variable: | Industry Information | | Industry & Market Information | | Comovement | |
|--|---------------------------|---------------------|-------------------------------|-----------------------|---------------------------|------------------------|
| Industry Content Metric: | Industry & Market Content | Industry Count | Industry & Market Content | Industry Count | Industry & Market Content | Industry Count |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Industry Content Metric _{it} | 0.009* (1.70) | 0.005*** (8.421) | 0.007*** (3.59) | 0.007*** (12.500) | 0.032* (1.887) | 0.006*** (6.526) |
| Industry Content Metric _{it} * Official _{it} | 0.012*** (5.08) | 0.007*** (4.640) | 0.024*** (11.18) | 0.022*** (12.286) | 0.043*** (2.590) | 0.241*** (18.844) |
| Official _{it} | -0.001 (-0.81) | -0.003* (-1.798) | 0.001 (0.78) | -0.013*** (-7.035) | 0.004 (0.596) | -0.252*** (-19.238) |
| Control Variables | Yes | Yes | Yes | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm-Fixed Effects | Yes | Yes | Yes | Yes | No | No |
| Adjusted R-squared | 0.047 | 0.038 | 0.031 | 0.017 | 0.012 | 0.018 |
| N | 2,768,021 | 2,768,021 | 2,768,021 | 2,768,021 | 2,770,718 | 2,770,718 |

Table 14
Information Content of Corporate News Articles conditional upon Policy Shock

This table presents evidence on whether the information content of Chinese business articles reflects firm specific, industry and/or market news, condition upon whether there is a new policy for a certain industry or the economy at large introduced on that day. Specifically, the table presents select coefficients from an estimation of the following empirical models:

$$\text{Industry \& Market Information}_{it} = \alpha + \text{Year-Month FE} + \text{Firm FE} + \beta_1 \text{News Coverage}_{it} + \beta_2 \text{News Coverage}_{it} * \text{Official}_{it} + \beta_3 \text{Firm Size}_{it} + \beta_4 \text{Market-to-book}_{it} + \beta_5 \text{ROA}_{it} + \beta_6 \text{Leverage}_{it} + \beta_7 \text{SOE}_{it} + \beta_8 \text{Earnings Announcement}_{it} + \varepsilon_{it}$$

The dependent variable *Industry and Market Information_{it}* equals the fraction of the firm's daily return that reflects industry-level and market-level information on day t. The disaggregation of daily returns into firm specific, industry-level, and market-level information components follows the methodology of Liu (2011), and as defined in the Appendix. In the third set of estimations, the dependent variable The indicator variable *News Coverage_{it}* equals one if a domestic Chinese newspaper published a business news article about firm i on day t, zero otherwise. The indicator variable *Official_{it}* equals one if an official newspaper published a business news article about firm i on day t, zero otherwise. All reported t-statistics (in parentheses) are adjusted for the clustering of standard errors by listed firm. OLS models include an array of year-month and firm fixed effects. Logistic models only include year-month fixed effects. ***, **, * denotes significance at the 1%, 5%, and 10% level, respectively

| Policy Shock: | Macro-Level Shock | | Industry-Level Shock | |
|---|------------------------|------------------------|------------------------|------------------------|
| Presence of Policy Shock | Policy Introduction | No Policy Introduction | Policy Introduction | No Policy Introduction |
| News Coverage _{it} | -0.051*** (-55.088) | -0.053*** (-45.280) | -0.053*** (-51.862) | -0.053*** (-47.722) |
| Test of Difference (P-O) (Chi-Squared Statistic) | 0.002*** (27.16) | | 0.000*** (49.61) | |
| News Coverage _{it} *Official _{it} | 0.013*** (12.981) | 0.012*** (10.666) | 0.014*** (9.635) | 0.012*** (11.884) |
| Test of Difference (P-O) (Chi-Squared Statistic) | 0.001*** (8.61) | | 0.002*** (20.74) | |
| Control Variables | Yes | Yes | Yes | Yes |
| Year-Month Fixed Effects | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes |
| Adjusted R-squared | 0.022 | 0.025 | 0.028 | 0.023 |
| N | 6,674,349 | 6,106,362 | 2,766,308 | 10,014,403 |