Political Connection and Sell-Side Research Quality

Yangyang Chen — *ByteDance Ltd*

Kaizhao Guo * — *Adam Smith Business School, University of Glasgow*

Jingshu Wen — *Said Business School, University of Oxford*

August 13, 2022

Abstract

Using textual analysis to measure financial analysts' industry knowledge for A-share healthcare reports issued during 2012-2019, we find that unqualified analysts exist in the Chinese stock market. These unqualified analysts concentrate in state-owned (SOE) brokerages. They tend to plagiarize other analysts and follow salient news, and investors lose 12.01% on average following their recommendations. After an exogenous shock that severs many analysts' political ties, research quality improves at local SOEs and informational efficiency increases for firms most intensely covered by local SOEs relative to other brokerages. Our results suggest that some unqualified analysts obtain their positions through their connection with Chinese securities regulators. In return, brokerage directors who hire unqualified analysts receive promotions prior to their loss of political connection.

Keywords: Security analysts; Industry knowledge; Political connection; Rent seeking; Cronyism

JEL classification: G15; G24; D73; P26

*Corresponding author: Kaizhao Guo. Email: k.guo.1@research.gla.ac.uk. Address: Main Building, University of Glasgow, G12 8QQ, UK.

1. Introduction

Institutional investors consistently rank industry knowledge as the single most important quality of sell-side analysts (Bradley, Gokkaya, and Liu, 2017a; Brown, Call, Clement, and Sharp, 2015, 2016; Kadan, Madureira, Wang, and Zach, 2012). However, anecdotal evidence abounds that analyst reports lack strong supporting evidence and their recommendations lack investment value[1]. Besides conflicts of interests and behavioral biases, it is possible that some sell-side analysts lack basic understanding of the business model and technology of the industry that they cover. This is especially a concern in autocratic countries where political connection can be a deciding factor for resource allocation and employment opportunities so that unqualified analysts are hired. China is a regulated economy whose large brokerages are state-owned enterprises (SOEs). Because Chinese securities regulators have power over SOE brokerages, they may extract rents and transfer resources to their networks, imposing costs on public investors.

This paper investigates whether investors pay for the service of financial analysts with little industry knowledge due to their political connection. Given the importance of industry knowledge, we use it as a proxy for analyst qualification. Unqualified analysts who are linked to China Securities Regulatory Committee (CSRC) may obtain and keep their positions at SOE brokerages so long as their political connection is intact. This is possible because CSRC officials appoint the chairmen and CEOs of SOE brokerages (the "one-level-up" policy mentioned in Chen and Kung (2018)), and the salary of SOE directors is government-controlled so they may not necessarily maximize brokerage profits. Brokerage directors are unlikely to reject CSRC officials' requests due to their power distribution, and they may receive promotions or other government resources as quid pro quo.

We test our hypothesis on the healthcare industry, whose firms' clinical trials and products are publicly available, so our measure for industry knowledge is transparent and replicable. We measure industry knowledge as the frequency of industry-specific words in their reports. We use a top-down approach to build the healthcare industry dictionary. For each healthcare industry sector, we first collect product names and other technical terms from Chinese drug regulator's, listed firms' and third-party

[1] For example, articles in *Financial Times* and *The Economist* attribute analyst mistakes to their conflicts of interests or herding. See $https://www.ft.com/content/0609b1b4-ec51-11e6-ba01-119a44939bb6 and https://www.economist.com/finance-and-economics/2016/12/01/sell-side-share-analysis-is-wrong$

websites. Then we supplement the word list by training an algorithm that uses word embedding, a method from computational linguistics, to identify words with similar context as the industry-specific terms in annual reports and IPO prospectuses. Analysts are unlikely to forecast firms' cash flows meaningfully without knowing their products and technology, and these industry-specific words may be key to in-depth analyses on firms' profits and growth. Because we directly quantify industry knowledge at the report level, the greater granularity allows us to capture the cross-sectional variation of an analyst's expertise for different portfolio firms and the time variation of an analyst's expertise as they strengthen or lose their knowledge and skills.

We first validate our industry knowledge measure on 15,080 A-share healthcare industry reports issued during 2012-2019. We regress observable measures of analyst performance on their industry knowledge, controlling for plagiarism tendency to ensure report originality. We find that analysts who demonstrate industry expertise are more accurate, more insightful, and more likely to recommend firms with improving future performance. If analysts use one extra word from our industry knowledge dictionary, their recommendation profitability increases by 0.156% and 0.337% for all reports and revisions, respectively. Industry experts are also more likely to move to the buy-side, though we do not find a strong relationship between analysts' industry knowledge and their star status. The overall evidence supports that industry knowledge is the most important dimension of analyst qualification, consistent with practitioners' view.

Then, we examine whether unqualified analysts exist and estimate the costs that they impose on investors. China's anti-corruption campaign reached the financial sector in 2015. In 2016, the State Council of China changed the chairman and vice-chairmen of CSRC. The anti-corruption campaign and CSRC official turnover are likely to sever many analysts' political ties and reduce the cronyism in SOE brokerages, based on the effectiveness of the anti-corruption campaign in Chen and Kung (2018). We sort our sample analysts into quintiles based on their average industry knowledge. Over our sample period, the average recommendation of analysts in the top quintile generates a six-month abnormal return of 7.214%, while investors do not benefit from the recommendations of analysts in the bottom quintile. Analysts in the bottom quintile generate a one-year average abnormal return of -12.01% for investors before 2016, which is significant at the 1% level. We conduct t tests for each year and find that this pattern persists, especially before 2016. Analysts in the top quintile are more likely to become stars, but this relationship is only significant after 2016. Before 2016, analysts in the bottom quintile are significantly more likely to be promoted than those in the top quintile, while the relationship reverses after 2016. Because our industry knowledge measure is predefined and not reliant on analyst performance, the evidence suggests that many unqualified analysts exist in the A-share market, especially before 2016.

We show that these unqualified analysts are unlikely to be informed as they rely on salient news or other analysts for writing their reports. If the analysts in the bottom quintile issue Buy ratings randomly due to their lack of insight, they should generate zero abnormal return on average in a weakly efficient market. However, we find that these analysts are not issuing reports randomly. They tend to issue reports soon after earnings announcements and plagiarize other analysts' reports. Because they follow salient news or lead analysts, they tend to recommend overvalued stocks with recent runups, leading to the negative average abnormal return.

Then, we use the exogenous shock in 2016 to test whether these unqualified analysts work at SOE brokerages due to their political connection. In difference-in-difference (DID) tests, local SOEs' research quality, proxied by their analysts' average industry knowledge, significantly improves after 2016 relative to that of private brokerages, while central SOEs' research quality does not change significantly. After the exogenous shock in 2016 that severs many analysts' political ties, the average industry knowledge improves by 13.25 at local SOE brokerages relative to non-SOEs. Both the loss of previous political ties and China's clamp-down on corruption likely reduce the rent seeking at SOEs. However, research quality only improves at local SOE brokerages possibly because they are more profit-oriented than central SOEs, so they train the unqualified analysts, or fire them and hire industry experts after 2016. Our data shows that local SOE brokerages are competing with non-SOEs, while central SOEs are larger, less efficient and do not appear to maximize profits. We find that local SOE brokerages have significantly higher net profits after 2016, while central SOE brokerages' profits do not change.

As our proxy for political connection is noisy, we strengthen the causal interpretation by testing whether CSRC officials promote brokerage directors as a return to their favors. We hand-collect SOE directors' employment history and regress their likelihood of being promoted on their brokerages' research quality, controlling for director and brokerage characteristics. Directors at brokerages with lower research quality are significantly more likely to be promoted, which is only statistically significant before 2016 and driven by local SOEs. Within our SOE brokerages, 23 directors of 19 brokerages are accused of misconduct after 2016. Most of them were promoted two to three times before 2016. On average, these brokerages' industry knowledge score increased from 17.99 to 36.65, and their employee turnover rate increased from 9.92% to 11.06% after 2016. Among them, 13 directors are investigated and charged with corruption by the Central Commission for Discipline Inspection (CCDI), the highest anti-corruption government body in China.

The above results suggest that cronyism exists in SOE brokerages in China, but institutional investors may purchase sell-side service for the analysts' political connection and do not follow their research recommendations. To address this concern, we test the effects of analysts' loss of political ties on the A-share market. We find that informational efficiency improves for our sample firms on average after 2016. Our DID tests show that the informational efficiency improves after 2016 for firms that are more intensely covered by local SOE brokerages compared with non-SOE brokerages. The results suggest that investors do use analyst reports in China and the market benefits from the reduction in cronyism.

Our findings suggest that charlatan analysts exist due to their political connection and that investors eventually bear the costs of this form of corruption, which could be one of the costs that lead to the smaller capital market in countries of weak legal institutions. An alternative explanation is that the charlatan analysts exist due to information asymmetry. We dismiss this hypothesis as analysts' clients are sophisticated institutional investors and most of their reports are publicly available on financial websites. In addition, our simple bag-of-words measure can differentiate between analysts with high and low levels of industry knowledge, so the level of information asymmetry between analysts and investors is likely to be low.

Our paper is related to the studies of favor exchanges between politicians and business owners (Agarwal, Qian, Seru, and Zhang, 2020; Carvalho, 2014; Claessens, Feijen, and Laeven, 2008; Faccio and Hsu, 2017; Ferguson and Voth, 2008). Our paper is closest to Chen and Kung (2018), where local Chinese officials give favorable business deals to Politburo members' family for promotions in exchange. Firms' use of political power have social costs including investment efficiency distortion, insider trading and worker mortality increase (Duchin and Sosyura, 2012; Fisman and Wang, 2015; Jagolinzer, Larcker, Ormazabal, and Taylor, 2020). We show the losses that investors suffer from following the recommendations of unqualified analysts who obtain their positions unfairly.

Our study is the first to separate industry knowledge from professional connections. Kadan et al. (2012) indirectly measure industry expertise as the profitability of analyst recommendations, which can be explained by many confounding factors. In addition, Bradshaw (2012) view the across-industry recommendation profitability in Kadan et al. (2012) as being disconnected from the practical implication of industry knowledge, which is mainly within-industry rather than across-industry or macroeconomic knowledge. Bradley et al. (2017a) study the effect of industry knowledge on analysts' forecast accuracy, but their proxy for industry knowledge is pre-analyst work experience, which can lead to connection with insiders and private information. This is not a trivial concern, as access to management is still analysts' competitive advantage after Regulation Fair Disclosure (Green, Jame, Markov, and Subasi, 2014). Because knowledge is a necessary but insufficient condition for skill, our bag-of-words measure cannot differentiate among skilled and unskilled analysts.

We contribute to the literature on the cross-sectional variation of analyst performance (Asquith, Mikhail, and Au, 2005; Clement and Tse, 2003; Kadan et al., 2012; Stickel, 1992). We show that lack of industry knowledge could explain analyst underperformance, which is different from their optimism bias or shirking in Altinkilic, Balashov, and Hansen (2019). Loh and Stulz (2010) show that only a small group of skilled analysts issue influential reports persistently. Our study suggests that institutional factors may increase the variability in analyst skill. Li (2005) and Mikhail, Walther, and Willis (2004) show that analysts with superior past performance continue to outperform in future periods. Superior industry knowledge may be the source of this persistent outperformance.

We also contribute to the literature of textual analysis in finance and accounting. Asquith et al. (2005) study the strength of analysts' argument by summarizing the number of positive and negative words that are related to firm operations, risk and prospects in their reports. Huang, Zang, and Zheng (2014) show that the sentiment in analyst reports provides valuable signals conditional on their earnings forecasts and recommendations. Lehavy, Li, and Merkley (2011) find that analysts spend greater efforts and achieve lower forecast accuracy for firms with less readable annual reports, but it is probable that annual report readability reflect the complexity and riskiness of the business and the amount of industry knowledge required for analysis.

In addition, our study contributes to the topic of replacing or equipping financial analysts with artificial intelligence (AI) technology. Analysts could issue biased reports for brokerage commissions and investment banking fees (Groysberg, Healy, and Maber, 2011), but Robo-Analysts suffer less from conflicts of interests and give more profitable buy recommendations than traditional analysts (Coleman, Merkley, and Pacelli, 2020). To build more effective Robo-Analysts, we need to understand how financial analysts add value. Our study provides insight on how financial analysts gather and use their industry knowledge, which can be automated in part though quantitative analysis of public information. Ball and Ghysels (2018) use high-frequency data to forecast earnings and obtain higher accuracy than analyst forecasts in certain situations, but their approach doesn't shed light on the key drivers of forecast success. Cao, Jiang, Wang, and Yang (2021) build an AI analyst that can use public information to forecast stock prices, but their AI analyst does not perform well in industries with high percentage of intangibles, where industry knowledge is essential for understanding business models and firms' growth potential. Deeper understanding of industry knowledge could help algorithm designers reduce over-extrapolation and other modeling issues.

The remainder of this paper is organized as follows. Section 2 presents the background and our hypotheses. Section 3 describes the data and presents our methods. Section 4 presents our results and discussion. Section 5 reports our robustness tests and Section 6 concludes the paper.

2. Background and hypothesis development

2.1. Background

Studies on the US market show that analysts are important information intermediaries who can affect firm policies (Derrien and Kecskes, 2013; Guo, Perez-Castrillo, and Toldra-Simats, 2019). The recent study of Crane and Crotty (2020) suggests that the majority of sell-side analysts in the US market are skilled. Although there are some overlaps between them, knowledge is generally factual information that is the prerequisite to skill. A qualified analyst must possess adequate amount of industry knowledge to understand firms' business models, competitiveness and growth potential to forecast future cash flows. However, there is a great deal of uncertainty in long-term cash flows, which cannot be predicted based on knowledge of existing facts alone. In this sense, industry knowledge is a necessary, but insufficient condition for analyst skill.

Many papers show the importance of industry knowledge in investment. Industry knowledge could help venture capital firms select and nurture innovative startups (Chemmanur, Loutskina, and Tian, 2014). Martens and Sextron (2021) argue that analysts' industry knowledge can benefit firms' innovation as analysts know both the target firms and their competitors. Bradley, Gokkaya, Liu, and Xie (2017b) find that analysts with industry expertise are more effective at monitoring firms and reducing agency conflicts. Hutton, Lee, and Shu (2012) find that analysts can forecast earnings as accurately as managers and attribute this to their industry expertise.

Analysts can create value for investors by collecting private information or processing public information. Some studies argue that analysts' value is in their collection of private information (Clement and Tse, 2005; Frankel, Kothari, and Weber, 2006; Ivkovic and Jegadeesh, 2004), while others cast doubt on the information discovery role of analysts (Kim and Song, 2015; Livnat and Zhang, 2012). If industry knowledge can be applied independently from inside information, analysts should be more accurate in interpreting public information if they possess superior industry knowledge, in contrast to their information collection role in Ivkovic and Jegadeesh (2004).

Bradley et al. (2017a) show that brokerages sometimes allocate analysts without related work experiences to covered firms, but these inexperienced analysts are not necessarily unqualified because analysts can acquire their industry knowledge through prior work experience or self-learning after becoming an analyst. For high-technology industries, education in relevant fields may contribute to an analyst's industry knowledge. Our current paper focuses on their demonstrated level of knowledge rather than their learning process. A larger number of analysts covering an industry can improve information efficiency (Merkley, Michaely, and Pacelli, 2017).

As most analysts use simple and standardized valuation models, the content of their analysis or the textual argument in their reports is highly important for understanding their target prices. Sell-side analysts are more specialized than buy-side analysts by industry sectors (Brown et al., 2016), and generally cover fundamentally related firms or industry peers (Ali and Hirshleifer, 2020; Parsons, Sabbatucci, and Titman, 2020). Because industry knowledge includes understanding of the technology and innovation in an industry, an expert in one sector is likely to possess technological expertise as measured by Tan, Wang, and Yao (2019) in other sectors with similar patents.

If analysts with inadequate industry knowledge lead some investors to make suboptimal decisions, investors with industry expertise will exploit profit opportunities arising from such mistakes. If the market is efficient, investors that lose money due to unqualified investment recommendations will exit the market and the unqualified analysts will lose their client and their job. Even if unqualified analysts do not exhibit any systematic bias, their existence doesn't provide valuable market information or investment insight and imposes costs on investors. Unlike a patient-doctor relationship where information asymmetry is high, fund managers should be able to recognize an analyst's industry knowledge after reading their reports, speaking with them or using their recommendations. Rational and self-interested investors should recognize such costs and discontinue their services, so unqualified analysts will exit the market in the long run under these ideal assumptions.

If a country has weak legal institutions, it tends to have a smaller, less valuable and less efficient capital market (La Porta, Lopez-De-Silanes, Shleifer, and Vishny, 1997, 2002; Shleifer and Wolfenzon, 2002), where information intermediaries tend to be less specialized. The relatively weak legal institutions and investor protection in China lead to a less competitive and transparent capital market (Allen, Qian, and Qian, 2005; Cheung, Rau, and Stouraitis, 2006; Jiang, Lee, and Yue, 2010). In 2016, Bloomberg released an article criticizing Chinese sell-side analysts[2]. Online searches suggest that some financial analysts in China lack industry background. For example, Chinese securities regulator fined Wu Chaoze for her unsubstantiated reports, who is the head of research of China Securities Co., Ltd[3]. Wu Chaoze is also the head of the telecommunications research group, but she has neither related degree nor industry work experience. Bradley et al. (2017a) show that 73% of forecasts in the US are made by analysts with previous work experience and 37% of forecasts are made by analysts with industry-related experience during the period from 2008 to 2011. We search on LinkedIn and find a small percentage of our sample analysts' profiles. Among those with LinkedIn profiles, most lack pre-analyst experience related to the healthcare industry.

As most large brokerages in China are state-owned and government controlled, it is possible that some analysts have political connections that help them obtain employment opportunities and shield them from market competition. As China is a civil law country with an autocratic political structure, cronyism is arguably rampant in its government and state-owned enterprises. The Chinese market also has capital control and low institutional ownership, so the roles of institutional investors and financial analysts could be relatively limited[4]. These frictions may lead to a large variability in the industry expertise of Chinese sell-side analysts, which can be captured by a relatively simple measure.

[2] $https://www.bloomberg.com/news/articles/2016-05-02/china-stock-analysts-were-among-world-s-worst-amid-surprise-rout$

[3] $https://news.stcn.com/sd/202012/t20201218_2640416.html$

[4] Institutional investors own only 18.7% of Chinese A-shares in 2021 and less than 10% in 2014 (Lin and Puchniak, 2021)

2.2. Hypotheses

Industry knowledge is not only regarded as the single most important characteristic of sell-side analysts, but also is the foundation for investment analysis and value investing. To identify the intrinsic value of a firm, the analyst needs to understand both the business model of the target firm and its competitors, customers and suppliers. Previous studies focus on analyst skills but overlooks a more fundamental issue: knowledge is the prerequisite of skills and some analysts may not even possess adequate industry knowledge. We hypothesize that analysts with inadequate industry knowledge provide less accurate forecasts and recommendations and are more likely to rely on salient public information. We directly test the value of industry knowledge to analysts and investors, which is also a way to validate our proxy for sell-side research quality.

Analyst Qualification Hypothesis: Industry knowledge is key to analysts' performance and the value of their recommendations.

From our analysis in the previous section, some financial analysts are likely to be unqualified due to their political connections, which might explain the cross-sectional variation in analysts' industry knowledge. We list the hypothesis as below.

Political Connection Hypothesis: Analysts with lower levels of industry knowledge are more likely to have political connections.

3. Methodology

3.1. Data

Although Kadan et al. (2012) study the industry expertise of both firm analysts and strategists, we only study firm analysts because strategists' analysis entails more macroeconomic than industry-specific knowledge. In addition, the bulk of sell-side service is within-industry investment consulting. Using Wind financial database, we find that around 70% of analyst reports are firm-level reports, while strategy reports make up 30% of the total number of A-share reports from 2006 to 2020.

We identify the healthcare industry according to $Wind$, which follows the Global Industry Classification Standard (GICS). Most Chinese brokerages started sell-side research service after 2000, so more recent samples are more reliable[5]. In addition, Crane and Crotty (2020) show that the proportion of skilled analysts is increasing over time, so our results may be more generalizable if we can identify unqualified analysts in a recent sample. Our sample consists of all the listed healthcare firms that receive analyst coverage in the Chinese A-share market from 2012/1/1 to 2019/12/31, as this period also includes both bull and bear markets. We collect financials and stock data from $Wind$ and the observation period for firms is from 2011/1/1 to 2020/12/31. For newly listed shares and firms with missing data, we use all the available data within the sample period. We scrape analyst reports from $Hexun.com$ and manually collect from $Huibo$ and $Wind$ the reports that are not listed on $Hexun$.

To test whether industry knowledge is an important dimension of analyst qualification, we study all analysts' ratings, including forecast revisions, initial coverages, and other events. Bradley, Clarke, and Zeng (2020) use the time stamp from IBES and a financial technology firm to study the value of early recommendation issuance. However, many large brokerage houses in China do not share their recommendations or reports on financial terminals or websites, and the report release dates on $Huibo$ are generally several days later than the day the reports are released to the brokers' paying clients. Therefore, we only use the report release date in the reports, which are in daily frequency. The details of our textual data cleaning are in Section B.1 of the Appendix.

3.2. *Industry knowledge measure*

Although star analysts and those in more prestigious brokerages are more likely to have rigorous training and research resources, most buy-side analysts care more about the actual experience of sell-side analysts than their star status or company size (Brown et al., 2016). Standard measures of brokerage prestige, such as size, may not accurately reflect research quality in China, where most large brokerage houses are state-owned enterprises with political goals. In our sample, 20.7% of star analysts are in central SOEs and 62.1% in local SOEs, although most central SOE brokerages are much larger than either local SOEs or private brokerages.

Expert analysts are likely to be well versed in the products, R&D, patents or services of the firm, and their reports are more likely to be specific and to-the-point, rather than general and vague. A qualified analyst should be familiar with the key drivers of firm operations and growth and center her reports

[5] The New Fortune magazine in China started ranking sell-side analysts in June, 2003.

on them. Products, patents, and R&D are highly relevant for pharmaceutical and biotech firms, while services are more important for Clinical Research Organizations (CROs) and hospitals. Unlike biotech firms which focus on innovation, active pharmaceutical ingredient (API) manufacturers care more about the costs of ingredients. Tan et al. (2019) find that analysts with technological expertise are more accurate. Knowledge of technological terms is a necessary condition for technological expertise. Industry knowledge is likely to be time-varying, as analysts will not be able to keep up with the latest technology and industry trend if they don't continuously learn.

Although we need context and meaning to evaluate the comprehensiveness and depth of analyst reports, a bag-of-words approach is likely to be suitable for measuring industry knowledge in our case. Given the institutional characteristics in China, some analysts may not even master the vocabulary or understand the business models in the industries, so word counts may differentiate the qualified from the unqualified analysts. We measure industry knowledge as the number of occurrences of terms from our industry knowledge dictionary in analyst reports. To ensure that our industry knowledge measure is unaffected by inside information, we rely on public sources to build our industry knowledge dictionary. First, we gather drug, medical device and equipment names, clinical service as well as drug targets from the websites of government, companies and third parties. These terms encompass the approved products and services of all the sectors in the healthcare industry. A list of our word sources is in Section B.2.1 of the Appendix.

Second, we add key terms from firm disclosures that are contextually similar to the jargons above. Hoberg and Phillips (2016) use 10-k business descriptions to classify firms' industries, because firms generally discuss their main products in annual fillings. Klein, Li, and Zhang (2020) show that healthcare analysts directly access information on US FDA (Food and Drug Administration) websites. Gibbons, Iliev, and Kalodimos (2021) find that analysts write more informative recommendation reports with stronger investor reaction when they directly access corporate disclosures via EDGAR. As reported by Brown et al. (2016), financial reports like 10-k filings are more important for buy-side analysts than conference calls or management earnings guidance. Annual reports and IPO prospectus are important for investors to understand the firms' current operations, competitive position and growth potential and they generally include firms' main products, R&D, patents and services. We scrape A-share healthcare firms' filings (including annual, semiannual, and quarterly reports, and IPO prospectus) during 2010-2020 from the official website CNINFO, which is the equivalent of EDGAR in China. We use word embedding, a method that is also used in Li, Mai, Shen, and Yan (2021), to find terms in disclosures that are contextually similar as our precompiled words above. We provide the technical details in Section B.2.2 of the Appendix.

Although top products may deserve a higher weight than minor products, we give the same weight to all the words in our industry knowledge dictionary. For this study, we simply assume that different products and product categories are equally important for firms. Future studies may explore potential word weighting schemes that better reflect the practical and investment value of different words in an industry.

3.3. Main tests

We use Equation (1) below to validate our industry knowledge measure and to test whether industry knowledge is the most important dimension of analyst qualification. The dependent variable Y_{it} is the value of analyst reports or analyst performance, and industry knowledge is measured at the report level for the former and at the analyst level for the latter. We estimate Equation (1) at both the report level and the analyst-year level. We measure recommendation profitability at the report level and then aggregate it to the analyst level. At the analyst level, measures for overall performance include their star status, overall recommendation profitability, target firm performance, tendency to piggyback, forecast boldness and employment change. We use logistic regressions to estimate the likelihood that analysts become star analysts in China. We use heteroskedasticity-robust standard errors that are clustered at the firm levels for regressions involving firm-level accuracy. We cluster robust standard errors at the analyst level for regressions on analysts' overall accuracy. Because an analyst's industry knowledge is likely to contain time-invariant components that are absorbed by analyst fixed effects, we do not include analyst fixed effects in Equation (1). For robustness check, we first use OLS estimates and then add year fixed effects.

$$Y_{it} = a + b \times Industry\,Knowledge_{it} + c \times Control_{it} + \delta_t + \varepsilon_{it} \tag{1}$$

To find unqualified analysts, we first sort our sample analysts into quintiles based on their industry knowledge and conduct *t* tests on the value of their recommendations. Then we show the average industry knowledge and recommendation profitability among analysts who are likely to be unqualified. We also estimate how much investors lose by following the recommendations of unqualified analysts.

Then we explore why unqualified analysts exist in the Chinese stock market through Equation (2). We test our political connection story at the analyst-year level/analyst level. From our analysis in previous sections, we proxy for analysts' political connection by their employers' ownership category, in which we differentiate local from central SOEs. The list of control variables is in the next section and in the Appendix.

$$Industry\,Knowledge_{it} = a + b \times Political\,Connection_{it} + c \times Control_{it} + \delta_t + \varepsilon_{it} \qquad (2)$$

To strengthen the causal inference of our test on political connection, we use the change of CSRC chairman in 2016 as an exogenous shock and conduct difference-in-difference (DID) test on analysts employed by SOEs and Non-SOEs[6]. Because the CSRC is the highest oversight committee for the securities and asset management industry in China, the change in CSRC chairman is likely to make some analysts lose their political connection. As cronyism and exchange of favors for employment are more rampant in government and SOEs, this shock is likely to mainly affect analysts in SOEs. We choose this event because three out of four vice-chairmen of CSRC also changed following the inauguration of Liu Shiyu, so many analysts associated with their predecessors may lose their access to private benefits and exit the firms or the industry afterwards. In addition, this event occurs shortly after China's anti-corruption campaign touched the financial sector in 2015[7], so the new CSRC and SOE officials are unlikely to abuse their powers as much as their predecessors.

Our DID design is in Equation (3). Our treatment group is SOEs and control group is non-SOEs, with the exogenous shock occurring in the February of 2016. The coefficient on $SOE_i \times Post_t$ is the DID coefficient and it captures the effect of losing political connection on the industry knowledge of analysts at SOEs relative to those at non-SOEs.

$$\begin{aligned} Industry\,Knowledge_{it} = a + b \times SOE_i + c \times Post_t + d \times SOE_i \times Post_t \\ + e \times Control_{it} + \delta_t + \varepsilon_{it} \end{aligned} \qquad (3)$$

[6]On February 20, 2016, Liu Shiyu replaced Xiao Gang as the Chairman of the China Securities Regulatory Commission.

[7]$https://www.ft.com/content/e50b1036-ab73-11e4-8070-00144feab7de$

3.4. Variables

3.4.1. Dependent variables

This section reports the dependent variables for Equation (1), which are different dimensions for analyst performance. New Fortune's Best Analyst ranking in China is the equivalent of Institutional Investor All Star Analyst ranking, where analysts recognized as Best analysts earn substantially more than non-star analysts. We acquire star analysts' ranking from New Fortune's website, and use a dummy variable that equals one if an analyst is ranked by New Fortune in each year and zero otherwise.

It is ideal to use different dimensions to measure analysts' accuracy, but we cannot obtain Chinese analysts' forecast data directly from data vendors, so we use the investment recommendations directly extracted from analyst reports[8]. Because the issuance time of our sample reports is at daily frequency, we cannot differentiate among reports issued before, during or after trading hours. In addition, most A-share investors are individuals without access to analyst reports, so announcement day abnormal return is a highly noisy measure for the price impact of analyst reports. Therefore, we do not study the short-term influence of analysts' research but use a longer time horizon. We study the investment profitability of analysts' ratings by trading on their recommendations at report issuance date T with a holding period of 180 days. We follow the literature and use buy-and-hold abnormal return to measure analysts' recommendation profitability (Crane and Crotty, 2020; Jegadeesh and Kim, 2010). The abnormal return ABR_i for recommendation revision i is as follows:

$$ABR_i(T) = D_i \times [\Pi_{t=T}^{T+180}(1 + r_{i,t}) - \Pi_{t=T}^{T+180}(1 + r_{m,t})] \tag{4}$$

Where $r_{i,t}$ is the return on the target stock in revision i, $r_{m,t}$ is the market return, and D_i is equal to 1 and -1 for upgrades and downgrades, respectively.

In addition to using revisions, we also test analysts' accuracy by trading on their recommendations in all reports. This allows us to use the whole sample of reports with investment recommendations, including initial coverages and other non-revisions. We buy the target stock at market price if the stock receives a Buy recommendation (including Strong Buy and Buy), do not trade for Hold ratings, and sell

[8] In our sample analyst reports, recommendations are reported in separate sections while stock price forecasts and earnings forecasts are reported together with other numbers. Because it is more error-prone to use algorithms to extract forecast earnings and target stock prices, we only extract investment recommendations by algorithms and then manually check their accuracy.

the stock for Sell ratings. Then we recalculate the ABR for all the analyst recommendations in our sample. We aggregate recommendation profitability at the report level to the analyst level by averaging each measure for each analyst. In addition to ABR, we define another variable $BHAR$ as the buy-and-hold abnormal return following analyst recommendations with a holding period of one year, which measures relatively longer horizon stock performance of covered firms.

Lee and So (2017) find that analysts in the US tend to cover underpriced firms with improving fundamental performance. The profitability of analysts' recommendations is the performance of firms' stock prices after receiving analyst ratings as negative ratings are extremely rare in China[9]. In addition to stock performance, we measure firms' accounting performance using $FScore$, a composite index that measures firms' profitability, leverage and liquidity, and operating efficiency (Lee and So, 2017; Piotroski, 2000; Piotroski and So, 2012). We first test the relationship on the whole sample and then on the subsample of firms with "Buy" ratings.

Industry experts are more likely to provide new information to investors (Li, Ramesh, Shen, and Wu, 2015; Luo and Nagarajan, 2015), rather than piggyback on financial news without providing new insight or investment value as reported by Altinkilic and Hansen (2009) as well as Loh and Stulz (2010). Because a piggybacking analyst issues reports after salient and positive corporate disclosure, the average pre-issuance return is likely to be high for the analyst. We measure an analyst's piggybacking tendency as his or her average pre-recommendation returns, which is related to the recommendation screening approach in Loh and Stulz (2010). Another sign of piggybacking is the tendency to issue revisions around salient corporate information release such as earnings announcements. Ivkovic and Jegadeesh (2004) find that analysts are less accurate immediately after earnings announcements compared with other periods. Coleman et al. (2020) find that Robo-Analysts are less reliant on earnings announcements for their investment recommendation revisions. Similar to Coleman et al. (2020), we define $EARevisions$ as an analyst's percentage of recommendation revisions that are issued within 5 days of covered firms' earnings announcements at the firm-year level.

Besides the performance measures above, we also study the likelihood that an analyst moves to a higher ranked brokerage or the buy-side. We define $Employment$ as a dummy variable that is equal to 1 if an analyst has a promotion, or moves to a higher ranked brokerage or the buy-side during the year, and to 0 otherwise. Because all sell-side analysts are required to register their profiles on the Securities

[9]Only three out of our 15136 sample reports have a Sell rating.

Association of China (SAC), we directly obtain their sell-side employment history on SAC website. For all the analysts who have left the sell side within our sample period, we search online to find their next employer. Our sources include financial websites like $Hexun.com$ and $Eastmoney.com$, as well as the websites of asset management firms.

3.4.2. *Control variables*

Our control variables include plagiarism likelihood, the beta of each industry sector, brokerage size, ranking, and mutual fund affiliation, analyst education, experience and portfolio complexity, firm size, stock price momentum, ownership category, trading volume and R&D intensity.

Due to the relatively weak protection on intellectual property rights in China (Fang, Lerner, and Wu, 2017), some financial analysts may directly copy the reports of other analysts. If a report is mainly plagiarized, the issuing analyst may not possess the industry knowledge in the report, so we need to control for the likelihood of plagiarism. We measure the likelihood of plagiarism as the maximum cosine similarity between a report and all the reports issued within seven days before, whose details are reported in Section B.2.3 of the Appendix.

Firm analyst reports include both broad industry-level analysis and firm-level information. Liu (2011) finds that analysts provide more industry information than firm information when the industry is riskier or less mature. To address this confounding effect, we control for the beta of each sector. Because the IPO regulatory process is heavily controlled by the Chinese government, the age of a sector in the stock market may not reflect its age in the real economy, so we do not control for sector age.

Clement (1999) show that analysts' experience, brokerage size and portfolio complexity affect analysts' forecast accuracy. Mikhail, Walther, and Willis (1997, 2003) find that analysts tend to become more accurate as their experience covering a firm increases. Analysts could accumulate industry knowledge as their work experience increases. If work experience also proxies for inherent ability or efforts, it should be positively correlated with the amount of industry knowledge that analysts have. Besides practical experience, healthcare analysts may also benefit from the analytical tools provided by an advanced degree or relevant majors such as pharmacy or medicine. Due to data availability, we could only control for analysts' highest degree and number of years working as a security analyst. We control for the resources that analysts have, proxied by brokerage size and ranking, because more prestigious brokerages provide more research assistants, administrative support and funding. These benefits could affect analysts' access to management and forecast accuracy. We acquire analyst education level, experience in the sell-side, brokerage ranking, revenue and profit data from the Securities Association of China (SAC). We also control for portfolio complexity, as busy analysts are likely to devote less time to each portfolio firm, which negatively affect their accuracy but do not affect their industry knowledge directly.

Due to the large percentage of retail investors in China, some brokerages may specialize in retail, rather than institutional business, so the total revenue of a brokerage may not reflect the scale of its research service. We also measure brokerage size as brokerage profit, as the profit margin for institutional services is higher than that for retail services. Mutual fund affiliation could cause analysts to issue biased reports on stocks held by affiliated fund (Mola and Guidolin, 2009), so we add it to our control variables.

Following Bradley et al. (2017a), we control for firms' size and momentum, as these could reflect firms' informational transparency. Additionally, we control for healthcare firms' ownership category and trading volume, as the trading restrictions and investors for SOE and non-SOEs in China are quite different, and speculative tendency is generally higher for firms with higher trading volume. Finally, we control for firms' R&D intensity, as innovation is key to the success of healthcare firms and R&D spending is directly observable.

We include the definition of all the variables in Table 12 of the Appendix A.

4. Empirical results and discussion

4.1. Summary statistics

We have scraped 12,924 reports from $Hexun.com$ and downloaded 2,156 reports from $Huibo$ and $Wind$. The pharmaceuticals sector accounts for 34.72% of total coverage, the largest among all healthcare sectors. The second popular sector is the traditional Chinese medicine sector, accounting for 29.20%, which slightly outnumbers the biotechnology sector (28.81%).

We have 86 brokerages, 411 analysts and 250 healthcare firms. After excluding analysts who do not issue reports with ratings and those with missing observations, we have 265 analysts with observations. Each analyst has been issuing forecasts for 4.92 years and covers 6.23 firms on average over our sample period. Only 3 reports have "Sell" ratings, while 92.22% of all reports give positive ratings, ranging from "Hold-outperform" to "Strong Buy". Out of all revision reports, 61% are upgrades.

Table 1 reports the summary statistics for our sample. We collapse report-level observations to the analyst-year level. The average number of industry-specific words used by each analyst each year has a mean of 25.42, median of 16.53 and standard deviation of 28.75. The first and third quartile of industry knowledge show large variability. If our measure of industry knowledge is a good proxy for research quality, the results corroborate our previous analysis based on China's economic institutions.

[Insert Table 1 here]

We report the pairwise correlations between our key variables in Table 2. The correlation of our industry knowledge measure with analyst and brokerage characteristics is relatively low, suggesting that there is heterogeneity even among the expertise of analysts with similar background. If the low correlation is driven by time-series variation, analysts' research caliber changes over time. In our sample, the exogenous shock in 2016 likely changes the research quality at SOEs.

[Insert Table 2 here]

4.2. The value of industry knowledge

Table 3 reports the estimation results on Equation (1). After adding control variables, analysts' recommendation profitability is significantly positively associated with the industry knowledge demonstrated in each report. Across different performance dimensions, we find qualitatively similar results: the higher the industry knowledge, the higher the forecast accuracy. If analysts use one extra word from our industry knowledge dictionary, their ABR_All, ABR_Revision, and BHAR increase by 0.156%, 0.337%, and 0.151%, respectively. The results for other dimensions of analyst performance are not statistically significant, possibly because China's sell-side industry is not fully competitive so that more accurate analysts are not necessarily recognized. The insignificant results for the regression on analysts' star status may also be due to the small number of stars relative to the total number of analysts[10].

[Insert Table 3 here]

Overall, we find that our industry knowledge measure does reflect the value of analysts' research for investors, as analysts with higher industry knowledge provide more profitable recommendations.

[10]It is also possible that the New Fortune ranking is not reflective of analysts' investment accuracy but more like a popularity contest.

Our results also survive changes in control variables, which are omitted for the sake of brevity. The results support the validity of our bag-of-words measure for the Chinese stock market. As knowledge of vocabulary or jargons is a minimum requirement for financial analysts to conduct fundamental analysis, this bag-of-words approach could differentiate qualified from unqualified analysts in markets where the variability in research quality is high.

4.3. The existence of unqualified analysts

We sort analysts into quintiles based on their industry knowledge, and use t tests to compare the performance of analysts in the top and lowest quintile. Table 4 reports the results. Analysts in the top quintile use significantly more industry-related words than analysts in the bottom quintile, and the discrepancy widens after 2016. Our t-test on ABR for the whole sample shows that analysts in the top quintile generates a 7.214% positive abnormal return for investors on average, which is statistically significant. In contrast, investors do not benefit from the recommendations of analysts in the bottom quintile. These abnormal returns are before deducting transaction fees, so investors' actual abnormal returns following their recommendations are lower. Before 2016, analysts in the bottom quintile generate a BHAR of -12.01% for investors, which is significant at the 1% level, while the BHAR for analysts in the top quintile is not statistically significant. Based on the value of their recommendations, analysts in the bottom quintile are likely to be unqualified and their existence imposes net costs on investors.

To see whether the qualified and unqualified analysts receive fair returns, we compare their employment outcome and star status. Analysts in the top quintile are more likely to become stars, but this relationship is only significant after 2016. Before 2016, analysts in the bottom quintile are significantly more likely to be promoted than those in the top quintile, which is quite surprising. However, after 2016, analysts in the top quintile are significantly more likely to move to better positions. The results suggest that unqualified analysts are likely to be shielded from market competition before 2016, and their more favorable career outcome may be a result of their political connections. Some brokerages may recruit more competent analysts after the unqualified analysts lose their political connection but not all brokerages fire the unqualified analysts, which would explain the larger discrepancy between the two groups after 2016.

[Insert Table 4 here]

These results suggest that charlatans and experts coexist in the A-share sell-side industry. To ensure the robustness of our findings, we conduct t tests for each year separately and find that the pattern persists: analysts in the top quintile outperform those in the bottom quintile in terms of their accuracy. However, experts start to receive market recognition for their skills only after 2016. For brevity, we do not report these results.

A challenge to our argument is that uninformed analysts should only generate zero abnormal return on average, if they cannot conduct fundamental analysis and issue Buy ratings randomly, as long as the A-share market is at least weakly efficient during our sample period. However, uninformed analysts might be able to generate non-zero abnormal returns, if they follow certain rules to form investment opinions and these rules introduce systematic bias to their recommendations. It is natural to suspect that unqualified analysts will try to hide that they lack the industry knowledge to gather and interpret information by copying the content of news or other analysts' reports. This strategy takes a minimum amount of efforts and is not risky in a country of weak protection on intellectual property rights[11]. By doing so, they issue reports after the release of salient news or lead analysts' reports, so they tend to recommend overvalued stocks, leading to the average negative abnormal return.

To understand why these unqualified analysts generate negative, rather than zero, abnormal returns before 2016, we regress analysts' trend-chasing and plagiarism tendency on their industry knowledge. Table 5 reports the results. Industry knowledge is negatively correlated with Plagiarism and EARevisions, which are significant at the 1% and 5% level, respectively. Analysts with higher industry knowledge are also less likely to piggyback on recent stock price runups, though the relationship is not statistically significant. The evidence supports that unqualified analysts tend to issue reports soon after earnings announcements and to plagiarize other analysts' reports.

[Insert Table 5 here]

[11] China receives very low score on IPR protection. For example, see $https://www.gtipa.org/publications/2021/11/30/release-2021-international-property-rights-index$

4.4. *Why do investors pay for unqualified analysts?*

4.4.1. *Analysts' political connection*

In this section, we test our hypothesis on unqualified analysts' political connection. Table 6 presents the results of t-tests on analysts' performance for subsamples created based on brokerage ownership category. Analysts at local SOEs have the highest average industry knowledge and the best overall performance. Analysts at local SOEs are more likely to be stars and they outperform other analysts in terms of the accounting performance of the firms that they recommend.

Analysts at central SOEs have the lowest average industry knowledge and the worst overall performance, but the pattern is not so clear in terms of their recommendation profitability. However, analysts at central SOEs are more likely to move to higher ranked brokerages or the buy-side.

The market-based and government-controlled systems coexist in China, possibly leading to the large variability in financial intermediaries' qualification. It seems that both local SOEs and private enterprises are competing for profit-maximization, while central SOEs care less about their brokerages' service quality and profitability.

[Insert Table 6 here]

We find similar results after changing the univariate tests to multiple regressions at the analyst-year level. From Table 7, for our whole sample, the average industry knowledge of local SOE analysts is higher by 12.35 than that of their non-SOE counterparts, which is statistically significant after controlling for confounding variables that may affect their performance. After 2016, this difference is 25.97 and is significant at the 1% level. However, the research quality at local SOE brokerages is not significantly different from that at private brokerages before 2016. And the research quality at central SOE brokerages is not significantly different from that at private brokerages either before or after 2016, despite their largest size among Chinese brokerage houses. After changing our control variables, we find qualitatively similar results.

The results show that the research quality at local SOEs improves after 2016 relative to other types of brokerages. Central SOEs seem to be less efficient and do not display scale advantage, as analysts working at large central SOEs do not have higher industry expertise than those at smaller brokerages. Because the CSRC directly appoints the directors at SOE brokerages, unqualified analysts may obtain their positions if they have connection with the CSRC or brokerage directors. In this way, the CSRC officials and brokerage directors transfer government-controlled market resource to their personal networks. Due to their political connections, these analysts are also less motivated to increase their industry knowledge. It is likely that both central and local SOEs hired unqualified analysts before 2016, but local SOEs either trained these unqualified analysts or fired them after 2016.

[Insert Table 7 here]

To mitigate endogeneity issues, we use the DID design in Equation (3) to test analysts' political connection. We first use non-SOEs as the control group and SOEs as the treatment group. Then we further divide SOEs into local and central SOEs. Table 8 reports our DID estimation results. Among the DID terms, only Local $SOE \times Post$ is statistically significant. After the exogenous shock in 2016 that severs many analysts' political ties, the average industry knowledge improves by 13.25 at local SOE brokerages relative to non-SOEs. Both the loss of previous political ties and China's clamp-down on corruption likely reduce the rent seeking at SOEs. As cronyism lessens after 2016 at local SOEs, research quality improves. A new set of control variables gives us qualitatively similar results.

[Insert Table 8 here]

While the research quality at central SOEs doesn't change relative to non-SOEs after 2016, both analysts at local and central SOEs likely lose their political connections after the CSRC official turnover and anti-corruption campaign. Why is it that research quality only improves at local SOEs? A likely explanation is that the ownership concentration of local SOE brokerages is higher and local government are more profit-oriented than the central government, the latter of which focuses more on political goals. So the local SOEs are more likely to train the unqualified analysts, or fire them and hire industry experts after 2016 than central SOEs, because the local government has more to gain if research quality improves and revenue increases. Qin, Stromberg, and Wu (2018) find that local governments are more profit-oriented while the central government cares more about political goals in China.

While it is possible, we do not hold the view that central SOEs are under greater scrutiny so they experience less cronyism before 2016. From the subsample t tests in Table 9, we can see that local SOEs have significantly higher net profits after 2016, while the change in central SOEs is insignificantly different from zero. From our previous results, central SOEs are larger, less efficient and do not seem to maximize profits. It is likely that the central SOEs are operated more for political motives, and that they do not fire the unqualified analysts as soon as the local SOEs which are more profit-driven.

[Insert Table 9 here]

4.4.2. Brokerage director promotion

The above findings only tell one side of the cronyism story. The CSRC officials benefit from the positions given to their social network, which lower the research quality and potentially commissions at the brokerages. As SOE directors' salaries are government-controlled, the directors may not lose much personally after they hire inept employees, but it still doesn't explain why they provide the favors to CSRC officials. Part of the reason could be power dynamics, as China is an autocratic country with pyramidal government structure and CSRC officials are their superiors, so the directors incur costs if they turn down the requests. Another reason could be the promotion prospects after giving the favors to their superiors. Promotion is a quite likely form of quid pro quo, as it gives the brokerage directors more power and resources while imposes minimal personal costs on CSRC officials, some of whom can appoint brokerage directors.

To test this form of quid pro quo, we hand collect the employment history of SOE directors, including the CEOs and board members, on company websites and Shanghai and Shenzhen Stock Exchange, as none of the brokerage directors post their profiles on $LinkedIn$. We do find that many of the directors at brokerages with low research quality move to higher ranked positions in brokerages and securities exchanges before 2016, but it is less common after 2016.

Then we regress directors' likelihood of being promoted on their brokerages' research quality, controlling for their gender and their brokerages' characteristics. We define director promotion as a dummy variable that is equal to one if the brokerage director moves to a higher ranked position at the current brokerage, other SOE brokerages or mutual funds, or stock exchanges.

Table 10 reports our estimation results. Directors at brokerages with lower research quality, proxied by their analysts' average industry knowledge, are significantly more likely to be promoted, which is only statistically significant before 2016 and driven by local SOEs. We find that the directors of central SOEs with low research quality are likely to move to higher ranked positions at brokerages or security exchanges prior to 2016, but this relationship is insignificant after 2016.

[Insert Table 10 here]

Besides the indirect evidence above, we find direct evidence of brokerage director corruption. Within our sample SOE brokerages, 23 directors of 19 brokerages (15 local SOEs and 4 central SOEs) are accused of misconduct and punished after 2016. Most of them were promoted to highly ranked positions at brokerages, mutual funds and stock exchanges before 2016. These brokerages' industry knowledge score increased from 17.99 to 36.65, and their employee turnover rate increased from 9.92% to 11.06% on average, from the period before to that after 2016. Among them, 13 directors are investigated and charged with corruption by the Central Commission for Discipline Inspection (CCDI), the highest anti-corruption government body in China. The additional evidence suggests that securities regulators promote brokerage directors in their exchange of favors, and these directors tend to engage in various forms of rent-seeking.

None of the directors above are charged with giving positions to unqualified analysts, though. The cronyism we report is a relatively mild form of corruption, and is quite indirect and difficult to uncover. In contrast to the exchange of favors between princeling firms and local officials in Chen and Kung (2018), CSRC officials are much lower ranked than the supreme rulers in the Politburo, and employment opportunities are a much less valuable form of bribe than the cheap land given to the princelings. So the analysts in our story are not necessarily the family members of the CSRC officials. The more distant relationships and the covert "bribe" make the cronyism less detectable. However, corruption-prone directors are likely to misuse their power in many ways, and the improvement in their firms' research quality after 2016 supports our cronyism story.

4.5. The impact on market informational efficiency

Although our findings support the existence of cronyism in SOE brokerages, A-share investors do not bear the cost of inept investment recommendations if they never really use analyst reports. Institutional investors may not follow analysts' research recommendations or even consult the information or logic in their reports, and they may purchase sell-side service for the analysts' political connection. To address this concern, we test the real effects of analysts' loss of political ties on the A-share market. If A-share investors actually use the information in analyst reports to make investment decisions, then the informativeness of their trades should improve after local SOE brokerages hired more industry experts in 2016.

To test this real effect, we conduct full-sample and split-sample tests on the change in firms' information asymmetry, which is proxied by Amihud's stock illiquidity measure (Amihud, 2002). To create subsamples, we sort healthcare firms into groups based on the ownership category of the brokerages that issue the largest total number of reports on each firm over our sample period. For the few firms that receive the same number of reports from central, local SOEs, and non-SOEs, we classify the firm by the ownership category in which the largest number of brokerages cover the firm. We follow Harford, Jiang, Wang, and Xie (2018) and control for factors that could affect firms' information asymmetry, including firm size, trading volume, institutional holding, book to market, leverage, past return, $\ln StockPrice$, ROA and volatility. We also use firm and year fixed effects to control for time invariant or macroeconomic confounders.

Using the whole sample, we find that informational efficiency improves for A-share healthcare firms on average after 2016. Table 11 reports the results of subsample regressions and DID tests. From Column 1 to 4, Amihud's stock illiquidity decreases significantly for stocks most intensely covered by both central and local SOEs, but not for stocks that receive the most coverage from non-SOEs. From Column 5-6, the coefficient of $SOE \times Post$ and Local $SOE \times Post$ are significantly negative, but Central $SOE \times Post$ is not statistically significant. Our DID tests show that the informational efficiency improves after 2016 for firms that are most intensely covered by local SOE brokerages compared with those most intensely covered by non-SOE brokerages. Our previous results show the improvement in research quality after 2016, and the results in this section suggest that investors do use analyst reports in China and the market benefits from the reduction in cronyism. Because the CSRC didn't issue no new rules on information disclosure between 2016 to 2018, our results are unlikely driven by changes in listed firms' disclosure.

[Insert Table 11 here]

5. Robustness tests

Our results are robust to various variable specifications. To address the concern that report length could confound our results, we normalized industry knowledge by dividing it with the total number of words in analyst reports (after excluding the appendix sections). We use different holding periods for ABR and BHAR, and different horizons for momentum. We replace brokerage ranking by their research department ranking published annually on SAC website. We substitute healthcare firms' size to their total assets. We add healthcare firms' R&D intensity to the list of controls, which is calculated as R&D expense/total assets and its annual growth rate. Our main results remain qualitatively similar.

6. Conclusions

Using a novel measure for analysts' industry knowledge, we show that unqualified analysts at state-owned brokerages impose costs on Chinese A-share investors. Many of them may obtain employment through connection with securities regulators and brokerage directors. Brokerage directors who provide favors to securities regulators are likely to be promoted in return. However, political connection is only one of the possible explanations for the existence of unqualified professionals in a state-regulated industry. In addition, we only investigate one form of the benefits that brokerage directors potentially receive. It is possible that some brokerage directors give positions to analysts affiliated with themselves or other directors to benefit their social and professional networks. And as not all CSRC officials have the power to appoint personnel, some SOE directors may receive other forms of benefits in return, such as business advantages or nonpublic information.

Despite the widespread use of computers and the decreased need for analysts' information production (Altinkilic, Hansen, and Ye, 2016), our study shows the value of industry knowledge, which suggests that industry experts may still have an edge in interpreting public information in this age. Future studies may explore the heterogeneity in investors' industry knowledge and its role in asset pricing. Although our industry word list comes from public sources, the words are meaningless without context, so we cannot directly test the value of background knowledge in investment analysis, which could be a direction for future research.

References

Agarwal, S., Qian, W., Seru, A., Zhang, J., 2020. Disguised Corruption: Evidence from Consumer Credit in China. Journal of Financial Economics 137, 430–450.

Ali, U., Hirshleifer, D., 2020. Shared Analyst Coverage: Unifying Momentum Spillover Effects. Journal of Financial Economics 136, 649–675.

Allen, F., Qian, J., Qian, M., 2005. Law, Finance, and Economic Growth in China. Journal of Financial Economics 77, 57–116.

Altinkilic, O., Balashov, V. S., Hansen, R. S., 2019. Investment Bank Monitoring and Bonding of Security Analysts' Research. Journal of Accounting and Economics 67, 98–119.

Altinkilic, O., Hansen, R. S., 2009. On the Information Role of Stock Recommendation Revisions. Journal of Accounting and Economics 48, 17–36.

Altinkilic, O., Hansen, R. S., Ye, L., 2016. Can Analysts Pick Stocks for the Long-Run? Journal of Financial Economics 119, 371–398.

Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. Journal of Financial Markets 5, 31–56.

Asquith, P., Mikhail, M. B., Au, A. S., 2005. Information Content of Equity Analyst Reports. Journal of Financial Economics 75, 245–282.

Ball, R. T., Ghysels, E., 2018. Automated Earnings Forecasts: Beat Analysts or Combine and Conquer? Management Science 64, 4936–4952.

Bradley, D., Clarke, J., Zeng, L., 2020. The Speed of Information and the Sell-Side Research Industry. Journal of Financial and Quantitative Analysis 55, 1467–1490.

Bradley, D., Gokkaya, S., Liu, X., 2017a. Before an Analyst Becomes an Analyst: Does Industry Experience Matter? Journal of Finance 72, 751–792.

Bradley, D., Gokkaya, S., Liu, X., Xie, F., 2017b. Are All Analysts Created Equal? Industry Expertise and Monitoring Effectiveness of Financial Analysts. Journal of Accounting and Economics 63, 179–206.

Bradshaw, M. T., 2012. Discussion of "Analysts' Industry Expertise". Journal of Accounting and Economics 54, 121–131.

Brown, L. D., Call, A. C., Clement, M. B., Sharp, N. Y., 2015. Inside the "Black Box" of Sell-Side Financial Analysts. Journal of Accounting Research 53, 1–47.

Brown, L. D., Call, A. C., Clement, M. B., Sharp, N. Y., 2016. The Activities of Buy-Side Analysts and the Determinants of Their Stock Recommendations. Journal of Accounting and Economics 62, 139–156.

Cao, S., Jiang, W., Wang, J. L., Yang, B., 2021. From Man vs. Machine to Man + Machine: The Art and AI of Stock Analyses. Working Paper 28800, National Bureau of Economic Research.

Carvalho, D., 2014. The Real Effects of Government-Owned Banks: Evidence from an Emerging Market. Journal of Finance 69, 577–609.

Chemmanur, T. J., Loutskina, E., Tian, X., 2014. Corporate Venture Capital, Value Creation, and Innovation. Review of Financial Studies 27, 2434–2473.

Chen, T., Kung, J., 2018. Busting the "Princelings": The Campaign Against Corruption in China's Primary Land Market. Quarterly Journal of Economics 134, 185–226.

Cheung, Y. L., Rau, P. R., Stouraitis, A., 2006. Tunneling, Propping, and Expropriation: Evidence from Connected Party Transactions in Hong Kong. Journal of Financial Economics 82, 343–386.

Claessens, S., Feijen, E., Laeven, L., 2008. Political Connections and Preferential Access to Finance: The Role of Campaign Contributions. Journal of Financial Economics 88, 554–580.

Clement, M. B., 1999. Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter? Journal of Accounting and Economics 27, 285–303.

Clement, M. B., Tse, S. Y., 2003. Do Investors Respond to Analysts' Forecast Revisions as if Forecast Accuracy is All That Matters? The Accounting Review 78, 227–249.

Clement, M. B., Tse, S. Y., 2005. Financial Analyst Characteristics and Herding Behavior in Forecasting. Journal of Finance 60, 307–341.

Coleman, B., Merkley, K. J., Pacelli, J., 2020. Man Versus Machine: A Comparison of Robo-Analyst and Traditional Research Analyst Investment Recommendations. Available at SSRN 3514879 .

Crane, A., Crotty, K., 2020. How Skilled Are Security Analysts? Journal of Finance 75, 1629–1675.

Derrien, F., Kecskes, A., 2013. The Real Effects of Financial Shocks: Evidence from Exogenous Changes in Analyst Coverage. Journal of Finance 68, 1407–1440.

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2018. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805 .

Duchin, R., Sosyura, D., 2012. The Politics of Government Investment. Journal of Financial Economics 106, 24–48.

Faccio, M., Hsu, H., 2017. Politically Connected Private Equity and Employment. Journal of Finance 72, 539–573.

Fang, L. H., Lerner, J., Wu, C., 2017. Intellectual Property Rights Protection, Ownership, and Innovation: Evidence from China. Review of Financial Studies 30, 2446–2477.

Ferguson, T., Voth, H. J., 2008. Betting on Hitler—The Value of Political Connections in Nazi Germany. Quarterly Journal of Economics 123, 101–137.

Fisman, R., Wang, Y., 2015. The Mortality Cost of Political Connections. Review of Economic Studies 82, 1346–1382.

Frankel, R., Kothari, S., Weber, J., 2006. Determinants of the Informativeness of Analyst Research. Journal of Accounting and Economics 41, 29–54.

Gibbons, B., Iliev, P., Kalodimos, J., 2021. Analyst Information Acquisition via EDGAR. Management Science 67, 769–793.

Green, T. C., Jame, R., Markov, S., Subasi, M., 2014. Access to Management and the Informativeness of Analyst Research. Journal of Financial Economics 114, 239–255.

Groysberg, B., Healy, P. M., Maber, D. A., 2011. What Drives Sell-Side Analyst Compensation at High-Status Investment Banks? Journal of Accounting Research 49, 969–1000.

Guo, B., Perez-Castrillo, D., Toldra-Simats, A., 2019. Firms' Innovation Strategy Under the Shadow of Analyst Coverage. Journal of Financial Economics 131, 456–483.

Harford, J., Jiang, F., Wang, R., Xie, F., 2018. Analyst Career Concerns, Effort Allocation, and Firms' Information Environment. Review of Financial Studies 32, 2179–2224.

Hoberg, G., Phillips, G., 2016. Text-Based Network Industries and Endogenous Product Differentiation. Journal of Political Economy 124, 1423–1465.

Huang, A. H., Zang, A. Y., Zheng, R., 2014. Evidence on the Information Content of Text in Analyst Reports. The Accounting Review 89, 2151–2180.

Hutton, A. P., Lee, L. F., Shu, S. Z., 2012. Do Managers Always Know Better? The Relative Accuracy of Management and Analyst Forecasts. Journal of Accounting Research 50, 1217–1244.

Ivkovic, Z., Jegadeesh, N., 2004. The Timing and Value of Forecast and Recommendation Revisions. Journal of Financial Economics 73, 433–463.

Jagolinzer, A. D., Larcker, D. F., Ormazabal, G., Taylor, D. J., 2020. Political Connections and the Informativeness of Insider Trades. Journal of Finance 75, 1833–1876.

Jegadeesh, N., Kim, W., 2010. Do Analysts Herd? An Analysis of Recommendations and Market Reactions. Review of Financial Studies 23, 901–937.

Jiang, G., Lee, C. M., Yue, H., 2010. Tunneling Through Intercorporate Loans: The China Experience. Journal of Financial Economics 98, 1–20.

Kadan, O., Madureira, L., Wang, R., Zach, T., 2012. Analysts' Industry Expertise. Journal of Accounting and Economics 54, 95–120.

Kim, Y., Song, M., 2015. Management Earnings Forecasts and Value of Analyst Forecast Revisions. Management Science 61, 1663–1683.

Klein, A., Li, T., Zhang, B., 2020. Seeking Out Non-Public Information: Sell-Side Analysts and the Freedom of Information Act. The Accounting Review 95, 233–257.

Kwon, O. W., Lee, J. H., 2003. Text Categorization Based on K-Nearest Neighbor Approach For Web Site Classification. Information Processing & Management 39, 25–44.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., Vishny, R., 1997. Legal Determinants of External Finance. Journal of Finance 52, 1131–1150.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., Vishny, R., 2002. Investor Protection and Corporate Valuation. Journal of Finance 57, 1147–1170.

Lee, C. M., So, E. C., 2017. Uncovering Expected Returns: Information in Analyst Coverage Proxies. Journal of Financial Economics 124, 331–348.

Lehavy, R., Li, F., Merkley, K., 2011. The Effect of Annual Report Readability on Analyst Following and the Properties of Their Earnings Forecasts. The Accounting Review 86, 1087–1115.

Li, E. X., Ramesh, K., Shen, M., Wu, J. S., 2015. Do Analyst Stock Recommendations Piggyback on Recent Corporate News? An Analysis of Regular-Hour and After-Hours Revisions. Journal of Accounting Research 53, 821–861.

Li, K., Mai, F., Shen, R., Yan, X., 2021. Measuring Corporate Culture Using Machine Learning. Review of Financial Studies 34, 3265–3315.

Li, X., 2005. The Persistence of Relative Performance in Stock Recommendations of Sell-Side Financial Analysts. Journal of Accounting and Economics 40, 129–152.

Lin, L., Puchniak, D. W., 2021. Institutional Investors in China: Corporate Governance and Policy Channeling in the Market Within the State. Available at SSRN 3858348 .

Liu, M. H., 2011. Analysts' Incentives to Produce Industry-Level versus Firm-Specific Information. Journal of Financial and Quantitative Analysis 46, 757–784.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V., 2019. Roberta: A Robustly Optimized Bert Pretraining Approach. arXiv preprint arXiv:1907.11692 .

Livnat, J., Zhang, Y., 2012. Information Interpretation or Information Discovery: Which Role of Analysts do Investors Value More? Review of Accounting Studies 17, 612–641.

Loh, R. K., Stulz, R. M., 2010. When Are Analyst Recommendation Changes Influential? Review of Financial Studies 24, 593–627.

Luo, R., Xu, J., Zhang, Y., Ren, X., Sun, X., 2019. Pkuseg: A Toolkit For Multi-Domain Chinese Word Segmentation. Working paper, Cornell University.

Luo, S., Nagarajan, N. J., 2015. Information Complementarities and Supply Chain Analysts. The Accounting Review 90, 1995–2029.

Martens, T., Sextron, C. J., 2021. Analyst Coverage Overlaps and Interfirm Information Spillovers. Journal of Accounting Research 59, 1425–1480.

Merkley, K., Michaely, R., Pacelli, J., 2017. Does the Scope of the Sell-Side Analyst Industry Matter? An Examination of Bias, Accuracy, and Information Content of Analyst Reports. Journal of Finance 72, 1285–1334.

Mikhail, M. B., Walther, B. R., Willis, R. H., 1997. Do Security Analysts Improve Their Performance with Experience? Journal of Accounting Research 35, 131–157.

Mikhail, M. B., Walther, B. R., Willis, R. H., 2003. The Effect of Experience on Security Analyst Underreaction. Journal of Accounting and Economics 35, 101–116.

Mikhail, M. B., Walther, B. R., Willis, R. H., 2004. Do Security Analysts Exhibit Persistent Differences in Stock Picking Ability? Journal of Financial Economics 74, 67–91.

Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013a. Efficient Estimation of Word Representations in Vector Space. Working paper, Cornell University.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., Dean, J., 2013b. Distributed Representations of Words and Phrases and Their Compositionality. In: *Advances in Neural Information Processing Systems*, pp. 3111–3119.

Mola, S., Guidolin, M., 2009. Affiliated Mutual Funds and Analyst Optimism. Journal of Financial Economics 93, 108–137.

Parsons, C. A., Sabbatucci, R., Titman, S., 2020. Geographic Lead-Lag Effects. Review of Financial Studies 33, 4721–4770.

Piotroski, J. D., 2000. Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Journal of Accounting Research 38, 1–41.

Piotroski, J. D., So, E. C., 2012. Identifying Expectation Errors in Value/Glamour Strategies: A Fundamental Analysis Approach. Review of Financial Studies 25, 2841–2875.

Qin, B., Stromberg, D., Wu, Y., 2018. Media Bias in China. American Economic Review 108, 2442–76.

Shleifer, A., Wolfenzon, D., 2002. Investor Protection and Equity Markets. Journal of Financial Economics 66, 3–27.

Stickel, S. E., 1992. Reputation and Performance Among Security Analysts. Journal of Finance 47, 1811–1836.

Sun, X., Wang, H., Li, W., 2012. Fast Online Training With Frequency-Adaptive Learning Rates for Chinese Word Segmentation and New Word Detection. In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics*, pp. 253–262.

Tan, H., Wang, J., Yao, L., 2019. Analysts' Technological Expertise. Available at SSRN 2772363 .

Appendix A. Tables

Table 1. Summary statistics

This table reports the summary statistics of the variables in this paper. We report the mean, median, standard deviation (SD), 25th percentile (P25), 75th percentile (P75), and number of observations (N). All variables in this table are at the analyst-year level. The variables are defined in Table 12 of the Appendix A.

Variables	Mean	Median	SD	P25	P75	N
Star Analyst	0.04	0.18	0.00	0.00	0.00	821
Industry Knowledge	25.42	28.75	8.67	16.53	30.5	821
ABR_All	3.37	20.03	-7.21	3.22	13.91	809
ABR_Revision	0.22	23.09	-14.28	0.00	11.21	128
BHAR	2.80	34.77	-13.78	0.67	17.77	809
Piggyback	0.00	0.01	-0.00	0.00	0.00	807
Momentum	7.65	22.42	-4.00	6.04	17.58	800
EARevisions	0.15	0.24	0.00	0.00	0.24	821
First Mover	0.28	0.36	0.00	0.12	0.50	821
FScore	5.26	1.13	4.64	5.19	6.00	821
Plagiarism	0.36	0.07	0.33	0.36	0.39	821
Sector Beta	0.86	0.10	0.78	0.86	0.92	799
Brokerage Size	9690.26	9879.71	3168.00	5773.38	13353.21	488
Brokerage Ranking	36.11	31.23	12.00	25.00	53.00	821
Portfolio Complexity	6.35	8.06	1.00	3.00	8.00	821
Firm Size	4650.47	5069.21	1564.53	3183.87	5825.29	821
Broker Fund Affiliation	0.64	0.48	0.00	1.00	1.00	821
Broker Ownership Category	0.59	0.49	0.00	1.00	1.00	821
Analyst Education	2.01	0.47	2.00	2.00	2.00	808
Analyst Experience	5.05	3.01	3.00	4.00	7.00	808
Employment	0.12	0.32	0.00	0.00	0.00	702

Table 2. Correlations

This table displays the pairwise correlations between the key variables in this paper. All variables in this table are at the analyst-year level, and are defined in Table 12 of the Appendix A

	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1)	Star Analyst	1.00															
(2)	Industry Knowledge	0.03	1.00														
(3)	ABR_All	0.01	0.10	1.00													
(4)	BHAR	0.00	0.11	0.65	1.00												
(5)	Piggyback	0.01	0.04	0.08	0.10	1.00											
(6)	EARevisions	-0.03	-0.10	-0.06	0.00	-0.06	1.00										
(7)	First Mover	-0.07	-0.08	-0.01	-0.06	0.06	0.10	1.00									
(8)	FScore	0.00	0.03	0.14	0.13	0.08	0.03	0.10	1.00								
(9)	Plagiarism	0.05	-0.14	0.01	-0.01	-0.06	0.10	-0.03	0.05	1.00							
(10)	Brokerage Size	0.16	0.01	0.00	0.02	0.06	0.09	0.09	-0.05	0.13	1.00						
(11)	Portfolio Complexity	0.12	-0.02	-0.02	-0.02	-0.05	0.12	-0.20	-0.05	0.19	0.09	1.00					
(12)	Broker Fund Affiliation	-0.10	-0.02	-0.04	-0.01	-0.03	0.00	0.08	0.04	0.12	0.04	0.04	1.00				
(13)	Broker Ownership Category	0.09	0.02	0.07	0.05	-0.01	-0.03	-0.02	0.07	0.09	0.10	0.05	-0.14	1.00			
(14)	Analyst Education	0.01	-0.01	0.00	0.05	0.01	-0.01	0.05	0.08	0.02	0.12	0.05	0.10	0.06	1.00		
(15)	Analyst Experience	-0.02	0.01	0.04	0.03	0.04	-0.03	0.14	0.07	0.01	0.01	0.07	0.04	0.01	-0.12	1.00	
(16)	Employment	0.04	-0.02	-0.14	-0.12	0.01	-0.04	-0.03	-0.06	0.01	0.05	0.00	0.07	-0.02	0.04	-0.17	1.00

Table 3. The value of industry knowledge: Analyst-level regressions

Cross-sectional regressions of analyst performance on their industry knowledge. Measures for analyst performance include the following: (1) Star Analyst, the percentage of years that New Fortune magazine ranks an analyst as a star in our sample period; (2) ABR_All, the buy-and-hold abnormal return following analyst recommendations with a holding period of 180 days based on all reports; (3) ABR_Revision, the buy-and-hold abnormal return following analyst recommendations with a holding period of 180 days based on revisions; (4) BHAR, the buy-and-hold abnormal return following analyst recommendations with a holding period of one year; (5) First Mover, the percentage of an analyst's reports that are the first among all reports issued after earnings announcements; (6) EARevisions, the percentage of reports issued within 5 days after earnings announcements; (7) Piggyback, the average daily return for seven days before analyst report issuance; (8) FScore, a composite index of changes in firms' accounting performance; (9) Employment, the percentage of years during which an analyst has a promotion. Control variables include plagiarism likelihood, analyst education, experience and portfolio complexity, brokerage size and ranking, the beta of industry sectors under analyst coverage, firm size, momentum, R&D expense, log trading volume, firm ownership category, and brokerage's mutual fund affiliation. All control variables are at the analyst level and defined in Table 12 of the Appendix A. All regressions are OLS. Standard errors are in parenthesis. * $p < 0.1$, ** $p < 0.05$, and *** $p < 0.01$ denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Star analyst	(2) ABR_All	(3) ABR_Aevision	(4) BHAR	(5) First Mover	(6) EARevisions	(7) Piggyback	(8) FScore	(9) Employment
Industry Knowledge	0.011	0.156***	0.337*	0.151**	-0.007	0.001	0.001	0.002	0.000
	(0.010)	(0.046)	(0.187)	(0.074)	(0.006)	(0.001)	(0.002)	(0.003)	(0.001)
Plagiarism	1.722	0.595	76.730	18.919	0.529	0.534**	-0.881	1.000	0.193
	(5.342)	(18.528)	(80.314)	(29.610)	(2.494)	(0.220)	(0.700)	(1.058)	(0.270)
Portfolio Complexity	0.147***	0.214	0.678	0.128	0.229***	0.006***	-0.012*	-0.011	-0.003
	(0.032)	(0.176)	(0.424)	(0.281)	(0.053)	(0.002)	(0.007)	(0.010)	(0.003)
Sector Beta	-6.586	35.749**	85.850	114.378***	2.473	0.011	-0.378	2.757***	-0.373*
	(4.551)	(14.178)	(75.080)	(22.657)	(2.004)	(0.169)	(0.535)	(0.809)	(0.207)
Analyst Education	-0.216	1.785	-13.392	3.032	0.766**	-0.035	0.075	0.178	0.008
	(0.584)	(2.264)	(11.501)	(3.618)	(0.333)	(0.027)	(0.085)	(0.129)	(0.033)
Analyst Experience	0.091	-0.208	0.606	-0.525	0.094	-0.004	0.003	0.009	-0.010*
	(0.083)	(0.370)	(1.363)	(0.591)	(0.058)	(0.004)	(0.014)	(0.021)	(0.005)
Brokerage Size	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Brokerage Ranking	-0.029	0.026	0.295	0.109	0.014	0.000	-0.001	0.006	0.001
	(0.022)	(0.07)	(0.213)	(0.111)	(0.011)	(0.001)	(0.003)	(0.004)	(0.001)
Broker Fund Affiliation	0.543	-4.297*	4.398	-4.807	0.208	-0.009	0.022	0.039	0.079**
	(0.640)	(2.449)	(8.666)	(3.914)	(0.351)	(0.029)	(0.092)	(0.140)	(0.036)
R&D /Operating Cost	8.813	51.033*	-6.308	52.207	-3.916	-0.272	-0.534	0.590	0.070
	(7.660)	(30.632)	(136.799)	(48.952)	(4.225)	(0.364)	(1.157)	(1.748)	(0.447)

Continuation of Table 3

	(1) Star analyst	(2) ABR_All	(3) ABR_Aevision	(4) BHAR	(5) First Mover	(6) EARevisions	(7) Piggyback	(8) FScore	(9) Employment
$\ln Trading\,Volume$	0.493	2.523	15.823*	3.278	-0.742***	0.036*	0.145**	-0.008	-0.045*
	(0.539)	(1.847)	(8.737)	(2.952)	(0.260)	(0.022)	(0.070)	(0.105)	(0.027)
Firm Ownership Category	0.263	-4.751*	6.807	-4.472	0.135	0.006	-0.071	-0.555***	0.059
	(0.675)	(2.623)	(9.595)	(4.192)	(0.364)	(0.031)	(0.099)	(0.150)	(0.038)
Firm Size	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Momentum	0.012	-0.029	0.262	-0.025	-0.009	0.000	0.006***	0.001	0.000
	(0.011)	(0.049)	(0.223)	(0.079)	(0.007)	(0.001)	(0.002)	(0.003)	(0.001)
Intercept	-12.204	-83.570*	-501.190**	-178.699**	13.318**	-0.901*	-2.531	3.706	1.199*
	(13.249)	(45.624)	(239.98)	(72.911)	(6.277)	(0.543)	(1.723)	(2.604)	(0.665)
Observations	265	265	65	265	265	265	265	265	265
R^2		0.151	0.333	0.158		0.103	0.101	0.109	0.098

Table 4. T-tests on analyst qualification

T-tests on analyst performance for subsamples created based on the level of industry knowledge. We conduct t-tests using the whole sample, pre-2016 subsample, and post-2016 subsample. We sort sample analysts into quintiles based on their average industry knowledge, and conduct t tests on the performance of those in the top and bottom quintiles, respectively. We also use difference-in-mean t tests to compare their performance. Measures for analyst performance include the following: (1) Star Analyst, the percentage of years that New Fortune magazine ranks an analyst as a star in our sample period; (2) ABR_All, the buy-and-hold abnormal return following analyst recommendations with a holding period of 180 days based on all reports; (3) BHAR, the buy-and-hold abnormal return following analyst recommendations with a holding period of one year; (4) EARevisions, the percentage of reports issued within 5 days after earnings announcements; (5) Piggyback, the average daily return for seven days before analyst report issuance; (6) First Mover, the percentage of an analyst's reports that are the first among all reports issued after earnings announcements; (7) FScore, a composite index of changes in firms' accounting performance; (8) Employment, the percentage of years during which an analyst has a promotion. All variables are at analyst-year level. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Whole Sample			Pre-2016			Post-2016		
	Top	Bottom	Diff-in-Mean	Top	Bottom	Diff-in-Mean	Top	Bottom	Diff-in-Mean
Industry Knowledge	69.902***	3.314***	66.588***	60.510***	3.127***	57.384***	79.971***	3.562***	76.409***
	(2.843)	(0.155)	(2.847)	(3.979)	(0.208)	(3.985)	(4.008)	(0.230)	(4.015)
Star Analyst	0.048***	0.012	0.036*	0.022	0.021	0.000	0.071**	0.000	0.071**
	(0.017)	(0.008)	(0.019)	(0.015)	(0.015)	(0.021)	(0.031)	(0.000)	(0.031)
ABR_All	7.214***	0.425	6.790**	0.920	-3.588	4.508	9.742***	5.539*	3.694
	(1.606)	(1.838)	(2.441)	(2.279)	(2.368)	(3.286)	(2.050)	(2.802)	(3.473)
BHAR	11.223***	-2.792	14.016***	-3.634	-12.014***	8.381*	18.079***	9.290	8.002
	(2.756)	(3.566)	(4.506)	(3.137)	(3.421)	(4.642)	(4.201)	(6.754)	(7.954)
EARevisions	0.117***	0.149***	-0.032	0.135***	0.153***	-0.019	0.094***	0.139***	-0.045
	(0.016)	(0.025)	(0.029)	(0.020)	(0.032)	(0.039)	(0.021)	(0.037)	(0.043)
Piggyback	0.256***	0.213**	0.043	0.218***	0.187	0.031	0.338***	0.261***	0.089
	(0.052)	(0.083)	(0.098)	(0.062)	(0.128)	(0.141)	(0.065)	(0.094)	(0.113)
First Mover	0.263***	0.473***	-0.211***	0.224***	0.436***	-0.213***	0.256***	0.534***	-0.272***
	(0.028)	(0.036)	(0.046)	(0.033)	(0.047)	(0.057)	(0.046)	(0.054)	(0.071)
FScore	5.309***	5.434***	-0.125	4.991***	5.351***	-0.360**	5.562***	5.586***	-0.026
	(0.099)	(0.108)	(0.146)	(0.110)	(0.139)	(0.177)	(0.173)	(0.173)	(0.242)
Employment	0.116***	0.109***	0.007	0.086***	0.170***	-0.084*	0.086**	0.028	0.058
	(0.025)	(0.024)	(0.035)	(0.029)	(0.039)	(0.049)	(0.034)	(0.020)	(0.039)

Table 5. Analysts' plagiarism and trend following tendency

This table presents panel regressions of analysts' plagiarism and trend following tendency on their industry knowledge. The dependent variables are: (1) Plagiarism, the maximum cosine similarity of a report with all the reports issued within seven days before its issuance; (2) EARevisions, the percentage of reports issued within 5 days after target firms' earnings announcements; (3) Piggyback, the average daily return for seven days before analyst report issuance. Control variables include analyst experience and portfolio complexity, brokerage size and ranking, the beta of industry sectors under analyst coverage, firm size, momentum, R&D expense, log trading volume, firm ownership category, and brokerage's mutual fund affiliation. All control variables are defined in Table 12 of Appendix A. All variables are at the analyst-year level. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Plagiarism	(2) EARevisions	(3) Piggyback
Industry Knowledge	-0.000***	-0.001**	-0.002
	(0.000)	(0.001)	(0.001)
Sector Beta	0.038	-0.004	1.964***
	(0.063)	(0.189)	(0.731)
Portfolio Complexity	0.000	0.003	-0.009**
	(0.001)	(0.002)	(0.004)
Analyst Experience	0.009	-0.023	-0.148
	(0.014)	(0.044)	(0.136)
Brokerage Revenue	0.000	0.000**	0.000
	(0.000)	(0.000)	(0.000)
Brokerage Ranking	-0.000	0.000	0.008
	(0.001)	(0.001)	(0.005)
Broker Fund Affiliation	0.005	-0.112*	-0.117
	(0.017)	(0.058)	(0.184)
R&D Expense / Operating Cost	0.109	-0.366	-1.890
	(0.147)	(0.406)	(1.582)
$\ln Trading\,Volume$	0.005	-0.001	-0.190**
	(0.009)	(0.028)	(0.088)
Firm Ownership Category	-0.004	0.039	-0.076
	(0.008)	(0.027)	(0.103)
Firm Size	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Momentum	-0.000	-0.001	0.003
	(0.000)	(0.001)	(0.003)
Constant	-0.144	-0.033	2.801
	(0.216)	(0.701)	(2.100)
Year FE	YES	YES	YES
Observations	605	605	605
R^2	0.182	0.165	0.168

Table 6. Subsample t tests on analysts' political connection

T tests on analyst performance for subsamples created based on brokerage ownership category. SOEs are state-owned enterprises. Measures for analyst performance include the following: (1) Star Analyst, the percentage of years that New Fortune magazine ranks an analyst as a star in our sample period; (2) ABR_All, the buy-and-hold abnormal return following analyst recommendations with a holding period of 180 days based on all reports; (3) BHAR, the buy-and-hold abnormal return following analyst recommendations with a holding period of one year; (4) EARevisions, the percentage of reports issued within 5 days after earnings announcements; (5) Piggyback, the cumulative abnormal return for seven days before analyst report issuance; (6) First Mover, the percentage of an analyst's reports that are the first among all reports issued after earnings announcements; (7) FScore, a composite index of changes in firms' accounting performance; (8) Employment, dummy equal to 1 if the analyst is promoted or has moved to a higher-ranked brokerage firm and 0 otherwise. All variables are at the analyst-year level. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Non-SOE	SOE	Local SOE	Central SOE
Industry Knowledge	24.884***	25.777***	29.574***	20.052***
	(1.663)	(1.251)	(1.804)	(1.478)
Star Analyst	0.015**	0.049***	0.061***	0.031**
	(0.007)	(0.010)	(0.014)	(0.013)
ABR_All	1.770*	4.457***	3.881***	5.314***
	(1.034)	(0.949)	(1.266)	(1.424)
BHAR	0.581	4.290***	5.107**	3.075
	(1.936)	(1.575)	(2.030)	(2.498)
EARevisions	0.163***	0.148***	0.148***	0.147***
	(0.013)	(0.011)	(0.014)	(0.018)
Piggyback	0.200***	0.186***	0.207***	0.154***
	(0.043)	(0.034)	(0.044)	(0.050)
First Mover	0.294***	0.116***	0.101***	0.139***
	(0.020)	(.015)	(0.018)	(0.027)
FScore	5.159***	5.327***	5.350***	5.293***
	(0.062)	(0.051)	(0.069)	(0.074)
Employment	0.127***	0.279***	0.268***	0.294***
	(0.020)	(0.016)	(0.021)	(0.026)

Table 7. Panel regression on analysts' political connection

Random effects regressions of industry knowledge on brokerage ownership category at the analyst-year level. The outcome variable is the average industry knowledge of each analyst for each year. Ownership category is a categorical variable that is equal to 1 for local State-Owned Enterprises (SOEs), 2 for central SOEs, and 3 for non-SOEs. Non-SOEs are the reference group. Control variables include plagiarism likelihood, analyst education, experience and portfolio complexity, brokerage size and ranking, the beta of industry sectors under analyst coverage, firm size, momentum, R&D expense, log trading volume, firm ownership category, and brokerage's mutual fund affiliation. Two sets of control variables are used. Year fixed effects are included. All control variables are defined in Table 12 of the Appendix A. Standard errors are in parenthesis.
*, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Whole Sample		Pre-2016		Post-2016	
Central SOE	0.633 (3.917)	1.353 (3.873)	-1.143 (6.516)	-1.206 (6.330)	6.801 (6.783)	8.382 (6.881)
Local SOE	12.347*** (3.805)	12.289*** (3.694)	1.719 (5.566)	1.325 (5.686)	25.967*** (5.862)	23.776*** (5.862)
Plagiarism	-73.379*** (21.250)	-72.345*** (21.249)	-23.411 (35.202)	-23.181 (31.405)	-92.810*** (25.102)	-88.891*** (24.139)
Portfolio Complexity	-0.196 (0.123)	-0.193 (0.126)	-0.107 (0.201)	-0.072 (0.197)	-0.236 (0.169)	-0.247 (0.169)
Analyst Education	-1.952 (3.201)	-2.140 (3.047)	-8.421** (3.706)	-8.635** (3.704)	4.794 (4.872)	4.002 (4.776)
Analyst Experience	-0.123 (0.556)	-0.029 (0.550)	0.410 (1.144)	0.421 (1.114)	-1.074* (0.647)	-0.858 (0.665)
Brokerage Ranking	-0.069 (0.101)	-0.069 (0.103)	-0.179 (0.116)	-0.141 (0.112)	0.271* (0.159)	0.273 (0.181)
Broker Fund Affiliation	7.330** (3.593)	7.128** (3.372)	0.476 (3.336)	-0.657 (3.230)	14.181** (6.243)	13.428** (5.485)
Firm Ownership Category	2.662 (2.892)	1.790 (2.780)	1.781 (4.693)	1.084 (4.340)	6.199* (3.407)	3.122 (2.726)
$\ln Trading\,Volume$	3.427 (2.420)	3.206 (2.665)	5.620 (4.727)	5.590 (4.696)	2.533 (3.046)	1.103 (3.279)
Sector Beta (VW)	20.628 (24.833)		27.294 (39.971)		15.186 (33.686)	
Sector Beta (EW)		45.035* (27.053)		10.976 (52.111)		59.391* (31.236)
Brokerage Size	0.000 (0.000)		-0.001* (0.000)		0.001*** (0.000)	
Brokerage Profit		-0.000 (0.001)		-0.001* (0.001)		0.003*** (0.001)

Continuation of Table 7

	Whole Sample		Pre-2016		Post-2016	
Momentum (6-Month)	-0.063 (0.064)		-0.086 (0.095)		-0.095 (0.108)	
Momentum (3-Month)		-0.056 (0.077)		-0.137 (0.113)		-0.013 (0.152)
R&D Expense /Operating Cost	114.046* (58.634)		103.222 (137.003)		52.234 (65.543)	
R&D Expense / Total Assets		264.479* (146.878)		146.244 (365.246)		173.076 (118.875)
Firm Size	0.000 (0.000)		0.001 (0.001)		0.000 (0.000)	
Firm Total Assets		0.000 (0.000)		0.000 (0.001)		0.000 (0.000)
Intercept	-106.896* (64.022)	-124.642* (72.073)	-129.301 (117.094)	-114.573 (132.572)	-74.782 (79.449)	-89.188 (87.706)
Observations	462	466	251	253	211	213
R^2	0.127	0.142	0.093	0.090	0.266	0.280

Table 8. DID test on analysts' political connection

This table shows the difference-in-difference test results for brokerages with different ownership categories. The dependent variable is the average industry knowledge of all sample analysts at each brokerage for each year. Our control group is non-SOE brokerages. Our treatment group is state-owned enterprises (SOEs) for Regression (1) and (2), which is further divided into local and central SOEs for Regression (3) and (4). Post is equal to 1 for observations after 2016, and 0 otherwise. Control variables include plagiarism likelihood, analyst education, experience and portfolio complexity, brokerage size and ranking, the beta of industry sectors under analyst coverage, firm size, momentum, R&D expense, trading volume, firm ownership category, and brokerage's mutual fund affiliation. Two sets of control variables are used. All control variables are defined in Table 12 of the Appendix A. Year fixed effects are included. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	2-Group DID		3-Group DID	
	(1)	(2)	(3)	(4)
Post	12.520 (29.998)	17.213 (29.710)	12.374 (29.911)	17.340 (29.678)
SOE	3.928 (4.963)	4.645 (4.895)		
Central SOE			-0.219 (5.034)	0.822 (4.956)
Local SOE			5.582 (4.696)	6.447 (4.639)
SOE × Post	10.041 (6.165)	8.171 (6.068)		
Central SOE × Post			-1.163 (7.662)	-2.542 (7.533)
Local SOE × Post			13.249** (6.514)	10.956* (6.433)
Plagiarism	-70.841*** (20.743)	-71.501*** (20.446)	-75.944*** (20.314)	-76.327*** (20.035)
Portfolio Complexity	-0.147 (0.149)	-0.136 (0.147)	-0.149 (0.147)	-0.130 (0.146)
Analyst Education	-1.410 (3.006)	-1.612 (2.955)	-2.234 (2.995)	-2.530 (2.950)
Analyst Experience	-0.274 (0.449)	-0.197 (0.442)	-0.151 (0.443)	-0.076 (0.437)
Brokerage Ranking	-0.039 (0.108)	-0.050 (0.108)	-0.120 (0.100)	-0.133 (0.100)
Broker Fund Affiliation	3.955 (3.782)	3.523 (3.729)	6.300* (3.250)	5.763* (3.188)
Firm Ownership Category	2.659 (2.720)	2.092 (2.573)	3.645 (2.721)	2.764 (2.576)

Continuation of Table 8

	2-Group DID		3-Group DID	
	(1)	(2)	(3)	(4)
$\ln Trading\,Volume$	2.437	2.470	3.365	3.557
	(2.541)	(2.671)	(2.528)	(2.661)
Sector Beta (VW)	11.393		15.369	
	(20.878)		(20.846)	
Sector Beta (EW)		33.103		35.049
		(21.781)		(21.779)
Brokerage Size	0.000		0.000	
	(0.000)		(0.000)	
Brokerage Profit		0.000		0.000
		(0.001)		(0.001)
Momentum (6-month)	-0.060		-0.067	
	(0.063)		(0.063)	
Momentum (3-month)		-0.066		-0.071
		(0.085)		(0.085)
R&D Expense/Operating Cost	112.678***		103.291***	
	(36.052)		(35.998)	
R&D Expense/Total Assets		263.852***		239.266***
		(75.275)		(75.432)
Firm Size	0.000		0.000	
	(0.000)		(0.000)	
Firm Total Assets		-0.000		-0.000
		(0.000)		(0.000)
Intercept	-69.107	-90.509	-95.452	-117.500
	(69.913)	(74.170)	(69.464)	(73.760)
Observations	462	466	462	466
R^2	0.108	0.124	0.141	0.154

Table 9. Subsample t tests on SOE brokerage characteristics

T tests on brokerage characteristics for subsamples created based on ownership category and observation period (before or after 2016). SOEs are state-owned enterprises. We conduct difference-in-mean tests to compare the difference across subsamples. Brokerage characteristics include the following: (1) Brokerage revenue; (2) Brokerage book value of equity; (3) Brokerage net profit; (4) Brokerage ranking; (5) Brokerage ROE; (6) Brokerage net profit margin. All variables are at the brokerage-year level. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	SOE			Local SOE			Central SOE		
	Pre-2016	Post-2016	Diff-in-Mean	Pre-2016	Post-2016	Diff-in-Mean	Pre-2016	Post-2016	Diff-in-Mean
Revenue	3264.066***	5306.814***	2042.748*	4739.319***	8886.428***	4147.109***	7276.185***	8531.979***	1255.794
	(556.629)	(1108.315)	(1110.117)	(485.309)	(1042.791)	(1022.252)	(1195.144)	(1191.449)	(1782.082)
Equity	28750.095***	47674.139***	18924.044	17763.822***	40712.900***	22949.078***	25163.872***	37560.107***	12396.235**
	(6751.927)	(12165.021)	(12842.513)	(1851.757)	(4511.985)	(4272.147)	(3481.256)	(4548.479)	(5700.163)
Net Profit	2789.751***	4602.855***	1813.104	1735.011***	2519.004***	783.993**	2759.356***	2353.145***	-406.211
	(656.147)	(1458.545)	(1383.84)	(201.865)	(334.394)	(367.354)	(488.297)	(392.884)	(691.964)
Brokerage Ranking	78.473***	79.694***	1.221	54.196***	53.613***	-0.583	52.875***	53.364***	0.489
	(1.779)	(2.476)	(3.017)	(2.077)	(2.587)	(3.355)	(3.021)	(3.820)	(4.923)
ROE	0.086***	0.066***	-0.020*	0.098***	0.054***	-0.044***	0.097***	0.059***	-0.038***
	(0.007)	(0.007)	(0.012)	(0.005)	(0.003)	(0.008)	(0.007)	(0.005)	(0.013)
Net Margin	0.269***	0.320***	0.051	0.305***	0.250***	-0.055***	0.333***	0.265***	-0.068***
	(0.018)	(0.054)	(0.046)	(0.008)	(0.010)	(0.011)	(0.013)	(0.017)	(0.020)

Table 10. Panel regressions on brokerage director promotion

Panel regressions of director promotion on brokerage industry knowledge. We divide the whole sample into subsamples based on the brokerage's ownership category and the observation period (before or after 2016). The outcome variable is a dummy that is equal to 1 if any of the directors at a brokerage is promoted to a higher ranked position within the financial industry or Chinese political system in each year. Industry knowledge is the average industry knowledge of all the sample analysts at a brokerage in each year. Control variables include brokerage size, brokerage fund affiliation, brokerage ranking, director's gender, age, professional certification, education, and political status. All control variables are defined in Table 12 of the Appendix A. Year fixed effects are included. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	SOE			Local SOE			Central SOE		
	Whole	Pre-2016	Post-2016	Whole	Pre-2016	Post-2016	Whole	Pre-2016	Post-2016
Industry Knowledge	-0.009***	-0.016***	-0.012***	-0.010***	-0.013**	-0.007	-0.004	-0.000	-0.008
	(0.003)	(0.005)	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)	(0.016)	(0.010)
Brokerage Size	0.000**	-0.000	0.000***	0.000***	0.000***	0.000***	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Broker Fund Affiliation	0.035	-0.090	0.190	0.189	0.184	0.247	-0.592**	-0.575	-0.690
	(0.119)	(0.165)	(0.180)	(0.167)	(0.241)	(0.242)	(0.261)	(0.363)	(0.422)
Brokerage Ranking	0.005	0.002	0.011**	0.014**	0.011	0.013	-0.014	-0.006	-0.006
	(0.004)	(0.007)	(0.005)	(0.006)	(0.012)	(0.008)	(0.010)	(0.011)	(0.030)
Gender	0.058	0.138	0.002	0.043	0.086	0.003	0.019	-0.022	0.026
	(0.148)	(0.211)	(0.212)	(0.222)	(0.336)	(0.300)	(0.238)	(0.339)	(0.342)
Age	-0.032***	-0.007	-0.063***	-0.028**	-0.002	-0.056***	-0.043***	-0.023	-0.072***
	(0.009)	(0.012)	(0.014)	(0.013)	(0.017)	(0.018)	(0.015)	(0.020)	(0.023)
Certification	-0.146	-0.214	0.005	0.063	0.148	-0.018	-0.029	-0.484	0.467
	(0.115)	(0.159)	(0.170)	(0.160)	(0.230)	(0.225)	(0.232)	(0.338)	(0.333)
Education	0.041	0.149	-0.087	0.002	0.127	-0.116	0.097	0.091	0.074
	(0.090)	(0.126)	(0.131)	(0.121)	(0.173)	(0.172)	(0.169)	(0.246)	(0.243)
Political status	-0.019	-0.171	0.083	-0.035	-0.222	0.102	0.244	0.179	0.327
	(0.116)	(0.167)	(0.167)	(0.162)	(0.245)	(0.224)	(0.209)	(0.290)	(0.322)
Constant	0.253	-0.648	1.836**	-0.898	-2.546**	1.047	1.617*	0.533	2.972*
	(0.584)	(0.798)	(0.845)	(0.847)	(1.195)	(1.185)	(0.969)	(1.309)	(1.634)
Observations	1993	1040	953	1239	617	622	633	344	289

Table 11. The real effects on market informational efficiency

This table presents panel regressions and difference-in-difference (DID) tests on market informational efficiency, which is measured as the change in Amihud illiquidity from the previous year. Healthcare firms are sorted into groups based on the ownership category of the brokerages that issue the largest number of reposts for each firm. Regressions 1-4 are subsample panel regressions. Regression 5 and 6 are DID tests where the control group are those firms that receive the most coverage from non-SOE brokerages. Post is a dummy that is equal to 1 for years after 2016, and zero otherwise. All control variables except volatility are lagged by 1 year. All control variables are defined in Table 12 of the Appendix A. All regressions are at the firm-year level. Year and firm fixed effects are included for panel regressions. Standard errors are in parenthesis. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	(1) SOE	(2) Non-SOE	(3) Local SOE	(4) Central SOE	(5) 2-Group DID	(6) 3-Group DID
SOE					0.028	
					(0.029)	
Post	-0.538***	-0.365	-0.525***	-0.378**	-0.459***	-0.460***
	(0.171)	(0.276)	(0.189)	(0.179)	(0.095)	(0.095)
SOE × Post					-0.058*	
					(0.034)	
Central SOE						0.045
						(0.042)
Local SOE						0.022
						(0.030)
Central SOE × Post						-0.067
						(0.052)
Local SOE × Post						-0.055*
						(0.033)
Firm Size	-0.402	-0.123	-0.517	-0.377***	-0.169***	-0.168***
	(0.265)	(0.320)	(0.432)	(0.099)	(0.042)	(0.042)
Trading Volume	0.302***	0.441*	0.362***	0.158	0.244***	0.244***
	(0.086)	(0.239)	(0.125)	(0.099)	(0.044)	(0.044)

Continuation of Table 11

	(1) SOE	(2) Non-SOE	(3) Local SOE	(4) Central SOE	(5) 2-Group DID	(6) 3-Group DID
Institutional Ownership	0.009	0.007	0.010	0.002	0.003**	0.003**
	(0.006)	(0.006)	(0.008)	(0.004)	(0.001)	(0.001)
Book-to-Market	0.842*	-0.533	1.105	0.110	-0.078	-0.078
	(0.501)	(0.466)	(0.717)	(0.326)	(0.162)	(0.163)
Leverage	0.004***	0.020	0.004***	-0.005	0.004***	0.004***
	(0.000)	(0.140)	(0.001)	(0.054)	(0.000)	(0.000)
Price	0.010	-0.032	0.010	0.008	0.002	0.002
	(0.007)	(0.040)	(0.009)	(0.006)	(0.001)	(0.001)
Return	-0.012	-0.530	0.119	-0.200	-0.564*	-0.564*
	(0.090)	(0.368)	(0.149)	(0.209)	(0.333)	(0.333)
ROA	-0.140	0.299	-0.447	1.521*	0.471	0.471
	(0.634)	(1.424)	(0.795)	(0.762)	(0.492)	(0.492)
Volatility	25.328	-1.251	33.956	-3.045	1.092	1.086
	(34.547)	(12.764)	(47.630)	(7.740)	(10.622)	(10.632)
Constant	-4.760***	-8.843	-5.543***	-0.399	-4.148***	-4.160***
	(1.296)	(6.341)	(1.371)	(2.286)	(1.033)	(1.035)
Firm FE	YES	YES	YES	YES	NO	NO
Year FE	YES	YES	YES	YES	YES	YES
Observations	868	321	675	193	2001	2001
R^2	0.035	0.096	0.038	0.247	0.065	0.065

Table A1: Definitions of variables

Balance sheet items and market cap are measured as the average of beginning-of and end-of-year values, while income statement items are the reported annual values.

Variable	Definition
ABR	Buy and hold abnormal return following analyst recommendations with holding period of six months.
Age	Brokerage directors' age.
Analyst Education	Categorical variable for analyst's highest level of education: 1 for undergraduate, 2 for master's, 3 for doctoral.
Analyst Experience	The number of years of work experience as a stock analyst.
BHAR	Buy and hold abnormal return following analyst recommendations with holding period of one year.
Book-to-Market	Book value of the firm's stockholder equity divided by market cap.
Brokerage Fund Affiliation	Dummy variable equal to 1 if the brokerage is affiliated with a mutual fund, 0 otherwise.
Brokerage Ranking	Ranking of brokerages based on annual net profit.
Brokerage ROE	Brokerage's net profit divided by its book value of equity.
Brokerage Salary	Average salary at a brokerage in million CNY.
Brokerage Size	Brokerage revenue in million CNY.
Certification	Dummy variable equal to 1 if the brokerage director holds one or more professional certificates, 0 otherwise.
Director Education	Brokerage director's highest level of education that is equal to 1 for undergraduate, 2 for master's and 3 for doctoral degree.
EARevisions	Percentage of analyst's revision reports that are issued within 5 days of target firm's earnings announcements.
Employment	Dummy variable equal to 1 if analyst moves to the buyside or a higher-ranked brokerage, 0 otherwise.
Firm Size	Revenue of healthcare firm in million CNY.
First Mover	A report is a first mover if it is the first among all reports issued within 7 days after target firm's earnings announcement. Averaged at the analyst level.
FScore	A composite index that measures changes in firms' accounting performance.
Institutional Ownership	Institutional investors' ownership percentage.
Leverage	Book value of total liabilities divided by book value of equity.
Momentum	Buy and hold abnormal return from six months earlier to analyst report issuance date.
Ownership Category	Dummy variable equal to 1 for State-Owned Enterprises (SOEs), 0 for others.
Piggyback	Cumulative abnormal return for seven days before analyst report issuance.
Plagiarism	Maximum cosine similarity of a report with all the reports issued within seven days before its issuance.
Political Status	Dummy variable equal to 1 if the brokerage director is a member of the Chinese Communist Party, 0 otherwise.
Portfolio Complexity	Number of firms covered by an analyst in a year
Price	Firm's average daily stock price in each year.
Return	Arithmetic average of daily stock return in each year.
ROA	Firm's net income divided by its book value of assets.
Sector Beta	Regression coefficient of the daily returns of healthcare industry sectors on that of A-share market index.
Volatility	Standard deviation of a firm's daily return in each year.

Appendix B. Data cleaning and variables

B.1. Textual data cleaning

All the documents are in PDF format, which we use $PDFMiner$ to parse. We remove tables, graphics, exhibits and other non-text items. We also remove the appendix section of analyst reports, as the standardized expressions in this section can affect the calculation of our plagiarism measure. Because most words in our corpora are in Chinese, which is not an inflected language, we do not lemmatize (remove the inflectional endings of words)[12]. Because there are no white spaces between words in Chinese texts, we first segment our corpora (analyst reports and firm disclosures) into words using the $PKUSEG$ toolkit developed by Luo, Xu, Zhang, Ren, and Sun (2019)[13]. By training the domain-specific model, we have identified the general and corpus-specific phrases in our corpora. After segmenting the texts, white spaces delimit all the words and phrases so that our algorithm treats them as single words. Then we process the words and phrases in our corpora into tokens[14]. After tokenization, we remove fillers, punctuations and other stop words (generally articles, auxiliary verbs, conjunctions, prepositions and pronouns).

We parse analyst reports to identify the issue dates, recommendations, brokerages, and the number of pages of the reports. We use Named Entity Recognition (NER) to identify and tag named entities, such as places, companies, persons and dates in the downloaded analyst reports. The NER we use follows an optimized BERT pretraining approach (Devlin, Chang, Lee, and Toutanova, 2018; Liu, Ott, Goyal, Du,

[12]The non-Chinese words in analyst reports and firm disclosures are financial or medical jargons, such as EPS or English acronyms for cancer drug targets, so we do not lemmatize them.

[13]This segmentation method is based on Sun, Wang, and Li (2012). $PKUSEG$ provides domain-specific pretrained models, and allows users to add additional training data. We use the pretrained model for medicine and use the products in the healthcare industry as our user-defined dictionary. We keep punctuations to identify sentence boundaries before segmenting phrases.

[14]After cleaning, we use "words" to refer to both words and phrases in our corpora, which our algorithms regard as single words.

Joshi, Chen, Levy, Lewis, Zettlemoyer, and Stoyanov, 2019)[15]. Because $Hexun.com$ lists the issue date, brokerage, rating, target firm, and analyst, we directly scrape these variables for reports downloaded from $Hexun$. For reports downloaded from $Wind$ and $Huibo$, we manually extract the issue date. We use the NER algorithm to extract target firms, brokerages and analysts. To increase the accuracy, we set a condition that analyst names appear next to their registration number at SAC before extracting analysts. We extract the ratings by searching keywords related to investment recommendations on the first page of analyst reports. Then we manually check the company names, brokerages, recommendations, and analysts to minimize the possibility for errors.

B.2. Construction of industry knowledge dictionary

B.2.1. Precompiled word list

We first compile a word list for the healthcare industry using a top-down approach. The Global Industry Classification Standard (GICS) divides the healthcare industry into two industry groups- healthcare equipment & services, pharmaceuticals & biotechnology & life sciences, which are further divided into 10 sub-industries. Out of the 10 segments, managed healthcare and healthcare technology are more closely related to the insurance and IT industry, respectively, so we remove them from our segment list. For each of the eight remaining segments, we search for relevant jargons and terms.

Most healthcare products require regulatory approval and are registered online, so from China's National Medical Products Administration (NMPA), we download the product names for four industry segments- healthcare equipment, healthcare supplies, biotechnology and pharmaceuticals. For the healthcare distributors segment, we obtain words related to the wholesale and retail of healthcare products. For the healthcare facilities segment, we obtain words related to hospitals and clinical centers. Finally, we

[15]The package we use comes from this website: $https://huggingface.co/uer/roberta{-}base{-}finetuned{-}cluener2020 - chinese$. Our loss function is binary cross entropy.

gather words related to clinical, manufacturing, or other outsourcing for the last two segments: healthcare services, life sciences tools & services.

The words related to each category are from the following websites:

- *Drugs, Medical Equipment and Supplies*: National Medical Products Administration
  - $https://www.nmpa.gov.cn/datasearch/home-index.html\#category=hzp$
- *Drug and Treatment Categories*: $Drug.com$ and $DXY.cn$
  - $https://www.drugs.com/$
  - $https://portal.dxy.cn/$
- *CRO, CDMO and Other Outsourcing*: Websites of large Contract Research Organizations (CROs), Contract Development and Manufacturing Organization (CDMO), and other healthcare service firms
  - IQVIA: $https://www.iqvia.com/$
  - Labcorp: $https://drugdevelopment.labcorp.com/$
  - PPD: $https://www.ppd.com/$
  - Parexel: $https://www.parexel.com/$
  - WuXi AppTec: $https://www.wuxiapptec.com/$
  - Hangzhou Tigermed: $https://www.tigermed.net/$
  - AmerisourceBergen Corp. (ABC): $https://www.amerisourcebergen.com/$
  - Cardinal Health Inc. (CAH): $https://www.cardinalhealth.com/en.html$
  - KingMed: $http://www.kingmed.com.cn/$

– Dian Diagnostics Group: $http://www.dazd.cn/$

- *Hospitals and Clinical Centers*: $http://www.a-hospital.com/$

- *Wholesale and Retail of Healthcare Products*: The Ministry of Commerce of China

  – $https://yplt.mofcom.gov.cn/stat/page/auth/DrugWall.html$

After collecting all the words from the sources above, we manually inspect and remove ambiguous words that have meanings in other fields. For example, EPS may stand for both Epstein–Barr virus in medical context or earnings per share in financial contexts, which may bias our industry knowledge measure, so we remove it from the dictionary. As many companies manufacture the same products, we only keep unique product names. For example, there are 149,402 domestic drugs listed on the NMPA by the end of 2021, but there are only 17,856 unique domestic drug names. After removing duplicates, we have 19,185 drugs and 42,333 medical devices/equipments from the NMPA. The precompiled list contains a total number of 73,651 unique specialized terms in the healthcare industry words. Most of these words are in Chinese, exceptions include imported products and cancer drug targets such as $PD-1$ (Programmed cell death protein 1).

B.2.2. Word embedding for identifying additional words

To supplement our precompiled words above, we extract additional words from firm disclosures, as managers are likely to list the relatively important products, services, ingredients, and innovations in disclosures such as the annual reports. From firm disclosures, we look for words that are contextually similar to those in our precompiled word list through word embedding, a method that maps words and phrases into vectors of real numbers through their likelihood of cooccurrence with neighboring words. Vector values capture the semantic similarity of words in the corpus. We do not set a minimum requirement for word frequency so that we can capture medical and pharmaceutical jargons that are relatively rare. We use the $word2vec$ method developed by (Mikolov, Chen, Corrado, and Dean, 2013a; Mikolov, Sutskever, Chen, Corrado, and Dean, 2013b), and we use the $Gensim$ library to train the model. We use continuous bag-of-words (CBOW) approach with 2 layers of neural network[16] to learn the embeddings, and our training algorithm is hierarchical softmax. Our context window size is seven, meaning that we use the three neighboring words before and after each target word for prediction. The size of the word vectors is 100.

After we obtain the word vector for each word in our corpus, we compare the vector values of our seed words (those that appear in both our precompiled word list and firm disclosures) with those of all the other words. We calculate the cosine similarity between the vector of each seed word and that of each word in our corpus, and extract words whose vectors have cosine similarities[17] of at least 0.7 with that of one or more seed words. We have culled 6523 words from firm disclosures in this way.

Then, two coauthors manually sift out irrelevant words and phrases, and we Google their definitions for cross reference. We make sure that each word extracted by our algorithm belongs to a category in our precompiled word list, and we also remove ambiguous acronyms with multiple meanings. We compare our chosen words to ensure that our interpretations of web definitions are consistent. Human inspection removed 58% of the words identified by the algorithm above, so we add 2744 words to our healthcare word list. After removing duplicate words, we have a total of 75848 words in our healthcare industry dictionary.

[16] The training algorithm for the neural network is stochastic gradient descent with backpropagation.
[17] See the definition of cosine similarity in Section B.2.3 of the Appendix.

B.2.3. Plagiarism measure

We define the likelihood of plagiarism as the similarity between a report and all reports issued in the previous 7 days. We measure the similarity between two reports as the cosine similarity between their word vectors, or the dot product of the word vectors normalized by their vector lengths (Kwon and Lee, 2003). The angle between the two vectors is inversely related to their closeness, as shown in the formula below. This measure is in the interval of $[0, 1]$ and the closer to one, the more similar two reports are. We define the variable Plagiarism as the maximum cosine similarity between a report and all the reports issued in the previous seven days.

$$Report\,Similarity = \frac{Vector_i \times Vector_j}{|Vector_i| \times |Vector_j|} \tag{5}$$