Connections over Competence:
The Impact of Political Ties on Sell-Side Research

Kaizhao Guo, University of Glasgow

2375890G@student.gla.ac.uk

Yanghua Shi, University of Mannheim

yanghua.shi@uni-mannheim.de

Jingshu Wen, University of Oxford

Jingshu.Wen.DPHIL@said.oxford.edu

Abstract

We document one form of political rent-seeking at Chinese state-owned (SOE) brokerages, where the managers and board members hire analysts connected with securities regulators to receive promotions. Using textual measures for analyst performance and kinship scores based on facial similarity, we find that politically connected analysts have lower research quality compared with merit-based hires, which affects investor returns negatively but brokerage officials' promotional prospects positively. China's anti-corruption campaign reduces the cronyism and improves analysts' research quality. Instead of mitigating rent extraction through internal party oversight, regulators in autocracies may share political rents with SOE officials.

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

JEL classification: G15; G24; D73; P26.

1. Introduction

In autocratic countries, bureaucrats are not fully accountable to citizens, corruption is rampant and economic efficiency is low (Chen and Kung, 2019; Djankov et al., 2010; Fisman et al., 2018; Fisman and Svensson, 2007; Fisman and Wang, 2015; Pei, 2016; Shleifer and Vishny, 1993; Treisman, 2000). However, the rulers in autocratic countries need to balance rent extraction with economic growth to avoid internal revolutions and external invasions. In stable autocracies, there is generally some democratic institutions at the grassroot level, such as the elections of village heads in China. Due to their information asymmetry, the top officials in China may loosen the vertical hierarchy and use local elections to hold village chairmen accountable when bureaucratic capacity is low (Martinez-Bravo et al., 2022). Besides villages, state owned enterprises (SOEs) are another type of institution with decentralized control and local activity. Unlike villages, SOEs can tolerate a higher level of rent extraction because their managers can select employees, they derive economic rents from state monopoly (Brødsgaard and Li, 2013; Unirule Institute of Economics, 2015, 2016), and they are less politically important.

To address the information asymmetry and keep vertical control, China's central government sets up various regulatory agencies, such as China Securities Regulatory Commission (CSRC) for monitoring firms in the securities industry, including SOE brokerages. The within-party monitoring for SOEs is unlikely to be as effective as the grassroot monitoring in villages, so SOE officials are likely to be less accountable to citizens. Where political accountability is lower, the rules for the political tournament are less based on performance, but more on privileges endowed from birth, such as connections and wealth, as the SOE officials with more resources can pay higher bribes to get promotions.

In this paper, we investigate whether Chinese SOE brokerage officials hire security analysts connected with financial regulators to benefit brokerages or themselves, and the social consequences of the connection-based hiring. We study the sell-side industry, because the performance of financial analysts has large variation and is publicly available. If SOE brokerage managers, who are Chinese Communist Party (CCP) officials, hire politically connected analysts to maximize brokerage profit, then they are likely to be accountable to the ultimate owners of the SOEs, or Chinese citizens. If the managers hire connected but unqualified analysts as bribes to CSRC officials, then it is a form of cronyism and political rent-seeking.

Opposite economic forces can explain the connection-based hiring. Connections may bring informational benefits that allow the SOE managers to select more competent employees and those in the same social circle share a higher level of trust, which may increase productivity. However, SOE brokerage managers may also hire analysts connected with financial regulators from the CSRC to increase their promotional prospects, and connections do not bring merit-based hires when cronyism dominates the information and trust effects. The CSRC is ranked one level up the SOE brokerages in CCP hierarchy, so CSRC officials can appoint, promote, or demote SOE brokerage officials. Because SOE officials' salaries are government controlled, they can derive personal benefits even if their political rent-seeking harms brokerage profitability.

To test the opposing hypotheses, we select Chinese A-share healthcare firms for our sample, as clinical trials and medical products are publicly available so that we can measure analysts' industry knowledge, which is consistently rank as the single most important quality of sell-side analysts by institutional investors (Bradley et al., 2017a; Brown et al., 2015 and 2016; Kadan et al., 2012). We measure industry knowledge as the frequency of industry-specific words in analyst reports, because analysts are unlikely to forecast firms' cash flows meaningfully without knowing their products and technology. We measure analysts' political connection as the probability of their kinship connection with financial regulators based on their facial features, as individuals with greater facial similarity have greater genetic similarity and are more likely to be kins (Torosin et al., 2020). 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.

There is large variability in our sample analysts' research quality, which is well captured by our industry knowledge measure. We sort our sample analysts into quintiles based on their average industry knowledge. Analysts in the top quintile also use more financial vocabulary, write longer reports and are less likely to plagiarize compared with analysts in the bottom quintile, all with a 1% level of significance. In addition, analysts in the top quintile also create more value for investors by generating higher abnormal returns with their recommendations.

We regress analyst performance on their political connection, controlling for other analyst characteristics, which shows that politically connected analysts have lower levels of industry knowledge than unconnected analysts. We divide the whole sample into pre-2015 and post-2015 subsamples, as China's anti-corruption campaign reached the financial industry and severed many analysts' political ties in 2015, when the State Council of China changed the chairman and most vice-chairmen of CSRC. We find that the average research quality of our sample analysts significantly improves after 2015, including their industry knowledge, report originality and recommendation profitability, and they are also less likely to piggyback, or recommend overvalued stocks with recent runups. Although they are not issuing reports randomly, the politically connected analysts are unlikely to be informed as they seem to follow salient news and plagiarize other analysts' reports.

To mitigate endogeneity and strengthen our causal interpretation, we use a difference-in-differences (DID) setting that exploits the exogenous shock of China's anti-corruption campaign. After 2015, the research quality of politically connected analysts significantly improves relative to that of unconnected analysts, in terms of the industry knowledge, length and originality of their reports as well as their recommendation profitability. Relative to the pre-2015 period, one percentage point increase in our kinship measure is associated with 3.54 and 6.45 percentage increase in one-month and two-month abnormal returns following analyst recommendations after 2015. Despite skepticism about the motives of the campaign (Griffin et al., 2022), it is effective in reducing corruption (Chen and Kung, 2019; Giannetti et al., 2021), and the campaign is likely to discourage the new CSRC officials from corrupt practices. After unqualified analysts lose their connections, they may face greater competitive pressure to conduct in-depth research and some of them are replaced by industry experts who can provide more valuable sell-side service.

As politically connected analysts are less competent, we further test whether SOE brokerage officials are promoted after hiring politically connected analysts as additional evidence for cronyism. To test the quid pro quo, we hand-collect SOE brokerage managers' and board members' employment history and regress their likelihood of being promoted on their brokerages' analysts' average political connection, controlling for official and brokerage characteristics. Officials at brokerages with more politically connected analysts are more likely to be promoted, which is statistically significant at the 5% level. However, the relationship between analyst political connection and brokerage official promotion is only statistically significant before 2015, during which one percentage increase in our kinship measure is associated with 2.5 percentage increase in the probability of brokerage official promotion. In our DID test, brokerage officials who hire politically connected analysts are significantly less likely to be promoted in the period after 2015 than before 2015. The anti-corruption campaign is likely to reduce the political rent-seeking so that new financial regulators and brokerage officials do not abuse their power as much as their predecessors.

Within our SOE brokerages, 23 officials of 19 brokerages are accused of misconduct after 2015, and most of them were promoted two to three times before 2015. On average, these brokerages' industry knowledge increases from 17.99 to 36.65, and their employee turnover increases from 9.92% to 11.06% after 2015. Among them, 13 officials are investigated and charged with corruption by the Central Commission for Discipline Inspection (CCDI), the highest anti-corruption government body in China.

These results support our cronyism hypothesis over the connection benefit hypothesis. The hiring of politically connected analysts is not merit-based, but is brokerage officials' bribe to financial regulators for promotions within the CCP hierarchy. The exchange of favors between the higher and lower ranked CCP officials is a form of corruption, which negatively affects sell-side research quality and investor returns following analyst recommendations. In effect, the cronyism transfers wealth from investors to CCP officials in charge of Chinese stock market.

Besides investor returns, we further test whether the corruption affects financial market efficiency. Our proxies for stock price informativeness include Amihud's stock illiquidity measure (2002) and Weller's price jump ratio (2018). Controlling for other factors that could affect firms' information asymmetry, we find that the crony hiring has insignificant effect on financial market efficiency, as the changes in price informativeness are indistinguishable for firms that are more or less intensely covered politically connected analysts. China's relatively underdeveloped stock market and weak legal infrastructure suggest larger roles for speculation, sentiment and even market manipulation than for fundamental research, which may explain the insignificant results. Despite the limited impact of the cronyism on financial market efficiency, A-share investors and Chinese citizens bear the costs of the opaque and corrupt hiring at the SOEs.

We document connection-based corrupt hiring in the financial market, which is a form of labor market discrimination and financial market friction, Exchange of favors between Chinese politicians for rent extraction has been documented in the real estate industry (Chen and Kung, 2019) and banking sector (Agarwal et al., 2020). Although corruption facilitates the exchange of resources and political competition in an autocracy, it occurs among the privileged few at the expense of citizens, which maintains the political hierarchy. Given the unbalanced power distribution in autocracies, the political rent-seeking at SOEs may also exist in other organizations whose leaders are eligible for promotions within the CCP. In contrast, village heads are not eligible for promotions within the CCP, so village elections do not threaten the rent extraction ability of the existing political elites, leading to grassroot democracy outside the elite circle.

We contribute to connection-based hiring. Most studies show the benefits of political connection to firms (Acemoglu et al., 2016; Amore and Bennedsen, 2013; Bertrand et al., 2014; Cooper et al., 2010; Faccio, 2006; Ferguson and Voth, 2008; Fisman, 2001; Goldman et al., 2009; Vidal et al., 2012), but connections can also facilitate rent extraction when the interests of politicians are misaligned with those of the organizations, as in patronage-based hiring (Colonnelli et al., 2020). Colonnelli, Li and Liu (2022) find that the net benefits of political connection to unconnected private equity investors are negative in China. Cruz et al. (2017) show that candidates in the center of social networks tend to win public sector jobs in Philippines due to their ability to practice clientelism. Individuals connected with current politicians obtain better paying jobs (Fafchamps and Labonne, 2017; Gagliarducci and Manacorda, 2020), potentially from exchange of favors between the politicians and firms.

We contribute to the academic debate on whether corruption is socially efficient.

Although corruption may allow economic resources to be allocated to individuals who value them the most (Beck and Maher, 1986; Leff, 1964; Lien, 1986; Lui, 1985), it benefits the privileged rather than the socially efficient recipients (Colonnelli et al., 2020; Esteban and Ray, 2006; Krueger, 1974; Shleifer and Vishny, 1993; Xu, 2018). We find that corrupt hires have worse performance than merit-based hires, while Weaver (2021) shows corrupt hiring does not reduce community health service quality, possibly as the health workers use cash bribes to compete for positions with limited autonomy and financial privilege, and performance dispersion is lower in the public than private sector (Borjas, 2002). Most studies show negative effects of corruption, including trade costs (Sequeira, 2016), regulatory non-compliance and worker mortality (Fisman and Wang, 2015), and distortion in investment efficiency (Duchin and Sosyura, 2012), license allocation (Bertrand et al., 2007) and knowledge production (Fisman et al., 2018), which may explain its negative correlation with economic growth in Mauro (1995).

We also use a textual measure of industry knowledge that is independent of analysts' professional connections, and our report-based measure captures analysts' industry expertise across portfolio firms and over time. Unlike our direct approach, Kadan et al. (2012) indirectly measure industry expertise as analysts' across-industry recommendation profitability, which is affected by many confounding factors and disconnected from the term's practical meaning of within-industry expertise (Bradshaw, 2012). Bradley et al. (2017a) proxy industry knowledge as pre-analyst work experience. The drawback is that previous work experience may lead to access to insiders and private information, which is still analysts' competitive advantage after Regulation Fair Disclosure (Green et al., 2014).

We also contribute to the literature on the cross-sectional variation of analyst performance (Asquith et al., 2005; Clement and Tse, 2003; Kadan et al., 2012; Stickel, 1992). Loh and Stulz (2011) show that only a small group of skilled analysts issue influential reports persistently. Our study suggests that institutional frictions may cause large variability in analyst skill as well as analysts' persistent relative performance as documented by Li (2005) and Mikhail et al. (2004).

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, and Section 4 presents the results and discussion for our main tests. Section 5 presents our additional tests, and Section 6 reports our robustness tests. Section 7 concludes the paper.

2. Background and hypothesis development

2.1. Background

Since China's economic reform in 1978, the political competition in China has increased relative to the pre-reform period, as the local politicians have gained de jure control over local resources such as land, and the development of their region or organization affects their promotional prospects within the CCP. To win the political tournament, local officials have reacted strategically to environmental monitoring (He et al., 2020) and sold land at steep discounts to princelings, or top CCP officials (Chen and Kung, 2019), the latter of which is an example of cronyism arising from the decentralized control over resources in an autocracy. Because the economic reform has barely changed the political power distribution between the elites and the average Chinese citizens, the de facto performance criteria for government officials, such as economic growth, may not be actually used, and the exchange of favors between the higher and lower ranked officials suggests that the political elites in China enjoy dynastic political rents (Chen and Kung, 2019).

Does the rampant corruption mean that the political tournament within CCP is ineffective and policies do not benefit citizens? Because China does not practice old school dictatorship like North Korea after the reform in 1978, its central government balances political elites' and citizens' interests to achieve political stability. Although the concentration of power is high within the CCP with relatively few checks and balances compared to democratic countries, there are grassroot elections of village chairmen in China, as the villagers' monitoring can reduce information asymmetry between top and local officials to reduce local officials' rent extraction, especially for remote villages and for periods with less bureaucratic capacity (Martinez-Bravo et al., 2022). Besides villages, state owned enterprises (SOEs) are another type of institution with decentralized control and local activity, so the information asymmetry between SOE officials and top officials is relatively high. To address the information asymmetry and keep vertical control, the central government sets up various regulatory agencies, such as China Securities Regulatory Commission (CSRC) for monitoring firms in the securities industry, including SOE brokerages.

The within-party monitoring is unlikely to be as effective as the grassroot monitoring in villages, as SOE officials can select the employees. In addition, SOEs enjoy economic rents due to state monopoly on resources and various benefits such as government subsidies, which cushion the bottom line for SOE managers' rent extraction. Although SOE officials' lack of performance incentives and potential rent-seeking do not benefit SOE performance, the discontent from minority shareholders is unlikely to cause widespread protests and uprising. Due to SOEs' lower political importance than villages, the political elites in China may allow higher level of rent extraction and lower political accountability in SOEs. Where political accountability is lower, the rules for the political tournament are less based on performance, but more on privileges endowed from birth, such as connections and wealth, as the SOE officials with more resources can pay higher bribes to get promotions.

2.2. Hypotheses

We examine the political accountability at SOEs by testing whether SOE brokerage officials pay bribes by hiring security analysts connected to financial regulators. Because bribery and cronyism are generally secretive, we select the sell-side sector as financial analysts' performance is publicly observable, which allows us to directly test the impact of crony hiring.

The tension in connection-based hiring is that connections can work in two ways, either reducing information asymmetry or facilitating cronyism. As financial analysts enjoy relatively high salaries and autonomy in report issuance, company visits and communication with clients, their skills are multi-dimensional and difficult to measure in interviews. To find suitable candidates with relevant skills and good work ethics, brokerage managers may use their informational advantage and hire competent candidates in their social circles, which include CSRC officials' friends and relatives. If connections reduce labor market search frictions but do not distort the merit-based hiring process, then the SOE brokerage directors have fulfilled their duty to citizens, which indicate a high level of political accountability at Chinese SOEs.

If politically connected analysts are less competent and brokerage directors hire them as bribes to CSRC officials for their own political career, then the political rent-seeking suggests a low level of accountability in SOEs.

Our two opposing hypotheses are as follows:

The Cronyism Hypothesis: SOE brokerage directors and managers hire politically connected but less qualified financial analysts to obtain personal benefits such as promotions.

The Connection Benefit Hypothesis: SOE brokerage directors and managers hire politically connected analysts due to connection benefits, including lower information asymmetry and more trust.

In the following sections, we test the hypotheses on connection-based hiring at SOE brokerages. If analysts' political connection is associated with better analyst and brokerage performance, then connection benefits are likely to explain the hiring decisions. If the opposite is true, we test potential exchange of favors between financial regulators and brokerage officials. If analysts' political connection is associated with worse analyst and brokerage performance and brokerage officials are likely to be promoted after hiring connected analysts, then the connection-based hiring is likely to be corruption in the form of exchange of favors between higher and lower ranked officials.

3. Data and Methodology

3.1. Data

We select A-share healthcare industry firms and analysts for our sample. Because the healthcare industry has high technological barriers to entry, rapid innovation and transparency, we study analysts who cover healthcare firms so that we can directly measure their industry knowledge based on their reports and public information. 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). Crane and Crotty (2020) show that the proportion of skilled analysts is increasing over time, and most Chinese brokerages started sell-side research service after 2000, so more recent samples may be more reliable.[1] 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 most analyst reports issued before 2012 are not publicly available. We collect company financial data from Wind Financial Terminal 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 download analyst reports from Hexun.com, Huibo and Wind.

Because all sell-side analysts are required to register their profiles on the Securities Association of China (SAC), we collect analysts' education level and 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 for pre-analyst and post-analyst work experience include financial websites like Hexun.com and Eastmoney.com, as well as the websites of asset management firms. SAC also provides brokerage ranking, revenue and profit data. We hand collect data on brokerage managers and directors (or brokerage officials for short) from brokerage disclosures as well as Shanghai and Shenzhen Stock Exchange, which include their age, gender, qualification, and employment history.

For analyst recommendation profitability, we use all analysts' ratings, including forecast revisions, initial coverages, and other events. Although analyst recommendation value is more accurate with time stamp data (Bradley et al., 2020), 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 B1 of the Appendix.

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

3.2. Analyst performance measures

3.2.1. Industry knowledge

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 et al., 2020), potentially due to similarity in technology and R&D as analysts need technological expertise to understand and forecast firm performance (Tan et al., 2019). We use industry knowledge as one of our analyst performance measures, as knowledge is the prerequisite of skills, and many papers show the importance of industry knowledge in investment. Hutton et al. (2012) find that analysts can forecast earnings as accurately as managers and attribute this to their industry expertise. Industry knowledge can help venture capital firms select and nurture innovative startups (Chemmanur et al., 2014), can benefit firms' innovation via knowledge spillovers (Martens and Sextroh, 2021), can improve corporate monitoring and reduce agency conflicts (Bradley et al., 2017b). A qualified analyst must possess an adequate amount of industry knowledge to understand firms' business models, competitiveness and growth potential to forecast future cash flows and estimate intrinsic values. However, long-term cash flows are uncertain and cannot be predicted based on knowledge of existing facts alone, so industry knowledge is a necessary, but insufficient condition for analyst skill.

Industry knowledge reflects analysts' value to investors, as 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). In addition, 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.

Using a bag-of-words approach, we measure industry knowledge as the number of occurrences of industry-specific terms in analyst reports, because knowledge of industry-specific jargons is a necessary condition for understanding business operations forecasting growth. Insightful analyst reports tend to focus on the key drivers of firm operations and growth, including the products, R&D, patents and services, rather than general policies or past financial performance. The key drivers of firm profits differ across sectors within the healthcare industry. 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 pharmaceutical and biotech firms, active pharmaceutical ingredient (API) manufacturers care less about innovation and more about the costs of ingredients. To reflect the focuses of different sectors, we aggregate the sector-specific terms into a healthcare industry knowledge dictionary.

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, and the websites contain useful information for analysts and investors. For example, Klein et al. (2020) show that healthcare analysts directly access information on US FDA (Food and Drug Administration) websites. A list of our word sources is in Section B2.1 of the Appendix.

Second, we add key terms from firm disclosures that are contextually similar to the jargons above. Based on previous studies, we use annual reports and IPO prospectus as our additional corpus, which include information that is both investment relevant and industry specific, such as firms' main products, R&D, and competitors. Hoberg and Phillips (2016) use 10-k business descriptions to classify firms' industries, because firms generally discuss their main products in annual fillings. Gibbons et al. (2021) find that analysts write more informative recommendation reports when they directly access corporate disclosures via EDGAR. Brown et al. (2016) show that financial reports like 10-k filings are more important for buy-side analysts than conference calls or management earnings guidance. 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 et al. (2021), to find terms in disclosures that are contextually similar as our precompiled words above. We provide the technical details in Section B2.2 of the Appendix. We give the same weight to all the words in our industry knowledge dictionary, as different word weighting schemes are unlikely to significantly change our results.

3.2.2. Other performance measures

This section reports the other proxies for analyst performance besides industry knowledge.

First, we calculate analysts' recommendation profitability based on the investment recommendations from their reports.[2] We use one-month to three-month recommendation profitability, rather than announcement day abnormal return, because we cannot differentiate among reports issued before or after trading hours and the large percentage of retail investors in China means that short-term price impact measures are highly noisy for most stocks. We study the investment profitability of analysts' ratings by trading on their recommendations at report issuance date T with a holding period of 30 to 90 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 i is as follows:

$$ABR_i(\mathrm{T}) = D_i(\prod_{t=T}^{T+n}(1 + r_{i,t}) - \prod_{t=T}^{T+n}(1 + r_{m,t})), n = 30, 60, 90$$

Where $r_{i,t}$ is the return on the target stock in report i, $r_{m,t}$ is the market return, and D_i is equal to 1, 0, and -1 for upgrades, neutral opinions and downgrades, respectively. We use all reports, including revisions, 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 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 each year.

[2] We first extract investment recommendations by algorithms and then manually check their accuracy.

Second, we use analysts' tendency to follow stock price trends or earnings announcements. Industry experts are more likely to provide new information to investors (Li, et al., 2015; Luo and Nagarajan, 2015), rather than to piggyback on financial news without providing new insight or investment value (Altinkilic and Hansen, 2009; Loh and Stulz, 2011). 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 (2011).

Third, we also proxy analyst performance by their employment outcome. 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.

Fourth, we also use plagiarism tendency to measure research quality. Due to the relatively weak protection on intellectual property rights in China (Fang et al., 2017), some financial analysts may directly copy the reports of other analysts. 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 B3 of the Appendix.

3.3. Analysts' political connection

We measure analysts' political connection as the probability of their kinship connection with financial regulators. Based on anecdotal evidence, some investment banking and sell-side analysts are Chinese officials' relatives, especially children, nephews or nieces.[3] There are also evidence of top Chinese officials' relatives profiting in the financial industry from political power.[4]In literature on social connections, ethnographic or genealogical data can be used to measure kinship tightness (Benzell and Cooke, 2021; Diao and Zhan, 2023; Enke, 2019; Giuliano and Nunn, 2018; Moscona et al., 2020). Because the family relationship of most analysts is private information, we estimate the probability of their kinship with financial regulators based on their facial features using analyst photos from SAC website and financial

[3] For example, see JP Morgan's Sons and Daughters Program: https://archive.nytimes.com/dealbook.nytimes.com/2013/08/30/morning-agenda-jpmorgans-sons-and-daughters-program/

[4] See the report of China's princelings in the finance industry: https://www.ft.com/content/e3e51a48-3b5d-11df-b622-00144feabdc0

regulator photos from CSRC annual reports. Individuals with greater facial similarity have greater genetic similarity and are more likely to be kins (Torosin et al., 2020).

We use the kinship prediction algorithm from Howard et al. (2019) and training dataset from Lu et al. (2012 and 2014).[5] We use the pictures of 310 Chinese parent-child pairs as our training set, and we do not use pictures of people from different ethnicities to avoid overestimating kinship.

3.4. Main tests

To test our hypotheses on connection-based hiring at SOE brokerages, we first investigate whether politically connected hires are more competent than nonconnected hires in Equation (1).

$$\begin{aligned}\text{Analyst performance}_{i,t} &= \alpha + \beta \cdot \text{Political connection}_{i,t} \\ &+ \gamma \cdot \text{Controls}_{i,t} + \delta_t + \varepsilon_{i,t}\end{aligned} \tag{1}$$

The subscript i denotes each analyst and t denotes each year. We aggregate industry knowledge, recommendation profitability and other report level variables to the analyst year or brokerage year level. We use heteroskedasticity-robust standard errors for regressions based on Equation (1). Because an analyst's industry knowledge and other measures of performance are likely to contain time-invariant components that are absorbed by analyst fixed effects, we do not include analyst fixed effects in Equation (1).

For the regressions on analyst performance, our control variables include analyst experience, education and portfolio complexity. Clement (1999) shows that analysts' experience and portfolio complexity affect their forecast accuracy. Mikhail et al. (1997, 2003) find that analysts tend to become more accurate as they become more experienced covering a firm. 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

[5] The training dataset is downloaded from https://www.kinfacew.com/download.html

experience or self-learning after becoming an analyst. For high-technology industries, education in relevant fields may contribute to an analyst's industry knowledge. To address analyst learning effects, we add analysts' work experience as a control variable. Besides experience, education level can also affect analysts' expertise and investment insight. Portfolio complexity may negatively affect analysts' accuracy, as busy analysts are likely to devote less time to each portfolio firm.

To address endogeneity concerns and strengthen the causal inference, we use a differences-in-differences setting by exploiting the exogenous change in analysts' political connection due to the anti-corruption campaign. China's anti-corruption campaign launched by Xi Jinping touched the financial industry in 2015, starting from the banking sector.[6] In 2015, the campaign also reached the China Securities Regulatory Commission (CSRC), which is the highest oversight committee for the securities and asset management industry in China. In table A2, we list the turnover in CSRC top officials each year, who include CSRC chairman, vice chairmen, chairman assistants, and the leader of discipline inspection and supervision team. In 2015, 10 CSRC top officials left their positions, including the chairman and three out of the four vice-chairmen, and 4 were investigated for corruption. As CSRC top officials have the power to appoint SOE brokerage officials, to allocate brokerage business licenses and to approve IPOs, the drastic turnover in CSRC severs many analysts' political ties, so that the previously politically connected analysts become less valuable to SOE brokerages and may be fired or work harder to avoid being fired afterwards.

The DID test is specified by Equation (2) below, where Post is a dummy variable that equals one after China's anti-corruption campaign reached China's stock market in 2015 and zero otherwise. The coefficient on $\text{Political connection}_{i,t} \cdot \text{Post}_{t}$ is the DID coefficient and it captures the effect of losing political connection on analysts' performance.

$$\begin{aligned} \text{Analyst performance}_{i,t} = \alpha &+ \beta_1 \cdot \text{Political connection}_{i,t} \cdot \text{Post}_{t} \\ &+ \beta_2 \cdot \text{Political connection}_{i,t} + \beta_3 \cdot \text{Post}_{t} \\ &+ \gamma \cdot \text{Controls}_{i,t} + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{2}$$

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

As additional evidence, we test whether brokerage officials are likely to be promoted after hiring politically connected analysts. If politically connected analysts are more competent and their hiring benefits brokerage profitability, brokerage officials may be promoted as SOE officials have the dual objective of fulfilling political goals and increasing SOE profits. If politically connected analysts are less competent and brokerage officials are promoted for hiring them despite the negative effects on brokerage profitability, then the nonmeritocratic hiring is likely to be a form of bribe that brokerage officials give to financial regulators. Our baseline and DID test for brokerage officials' career outcome are specified in Equation (3) and (4), respectively, where the subscript k denotes each brokerage.

$$\begin{aligned} \text{Official promotion}_{k,t} = \alpha + \beta \cdot \text{Political connection}_{k,t} \\ +\gamma \cdot \text{Controls}_{k,t} + \delta_t + \eta_k + \varepsilon_{k,t} \end{aligned} \tag{3}$$

$$\begin{aligned} \text{Official promotion}_{k,t} = \alpha + \beta_1 \cdot \text{Political connection}_{k,t} \cdot \text{Post}_t \\ +\beta_2 \cdot \text{Political connection}_{k,t} + \beta_3 \cdot \text{Post}_t \\ +\gamma \cdot \text{Controls}_{k,t} + \delta_t + \eta_k + \varepsilon_{k,t} \end{aligned} \tag{4}$$

We use Cox regression models to estimate Equation (3) and (4). Official promotion is a dummy that equals one if any of the managers or directors at the brokerage that hires the analyst is promoted at least once after the hiring, and zero otherwise, where promotion includes moving to a higher ranked position at the current brokerage, other SOE brokerages or mutual funds, stock exchanges or the CSRC. Equation (3) is estimated first on the whole sample, and then on the pre-2015 and post-2015 subsamples.

For the regressions on brokerage official promotion, our control variables include brokerage revenue, brokerage official age, gender, education and certification. The definitions of all the variables are in Table A1 of Appendix A.

4. Empirical results and discussion

4.1. Summary statistics

We have downloaded 34,788 reports from Hexun.com, Huibo and Wind. The pharmaceuticals sector accounts for 24.9% of total coverage, the largest among all healthcare sectors. The second most popular sector is the traditional Chinese medicine sector, accounting for 21.30%, which slightly outnumbers the biotechnology sector (20.56%). We have 126 brokerages, 411 analysts, and 250 healthcare firms. After excluding analysts who do not issue reports with ratings and those with missing observations, we have 300 analysts with observations including photos. Each analyst has been issuing reports on average for 4.85 years and covers 11 firms on average over our sample period. Only one report has a "Sell" rating, while 86.31% of all reports give positive ratings, ranging from "Hold-outperform" to "Strong Buy". Most revision reports are upgrades.

Table 1 reports summary statistics of the variables used in analyst performance analysis. We collapse report-level observations to the analyst-brokerage-year level. The average value of kinship of the whole sample is 0.633, with a significant degree of variation across analysts, brokerages, and years. We classify the analysts with kinship bigger than 0.7 as connected and the others as unconnected, with their respective description statistics shown in Panel B and Panel C. The table shows that connected analysts have more work experience (5.26 years for connected analysts, and 4.56 years for unconnected analysts) and higher education levels than unconnected analysts, allowing them to accumulate more industry knowledge and financial knowledge. For other variables, there are also variations across analysts and years. The variation of recommendation profitability is larger for longer time horizons.

[Insert Table 1 here]

We report the pairwise correlation between the analyst characteristics in Table 2. The performance measures are highly correlated, including industry knowledge, financial knowledge, report length, and plagiarism tendency. The recommendation profitability for each analyst is quite persistent over different holding horizons. The correlation of our industry knowledge measure with analyst and brokerage characteristics is relatively low, suggesting that there is research quality heterogeneity even among analysts with similar background. If the low correlation is driven by time-series variation, analysts' research caliber changes over time.

There could be convergence or divergence in sell-side research quality across SOE and non-SOE brokerages over our sample period.

[Insert Table 2 here]

Table 3 displays the distribution of connected and unconnected analysts across brokerages of different ownership categories, as well as the analysts' performance. The average values of kinship are similar across non-SOEs, local SOEs, and central SOEs. While the analysts working in non-SOEs tend to be more experienced than those working in SOEs, they are slightly less educated and perform worse. There are 121, 488, and 347 observations for analysts at non-SOEs, local SOEs and central SOEs, respectively. The central SOE brokerages are larger than local SOE brokerages, potentially due to more government resources. Qin et al. (2018) find that local governments are more profit-oriented while the central government cares more about political goals in China.

[Insert Table 3 here]

Table 4 provides descriptive statistics for variables at the brokerage-year level, including brokerage official characteristics. The mean value of brokerage official promotion is 0.214, meaning that on average, 21.4% of the brokerage officials experience promotions in a given year during the sample period.

[Insert Table 4 here]

For Table 5, we investigate the performance variability of analysts with high and low levels of industry expertise. We first sort our sample analysts into quintiles based on their industry knowledge and then conduct t-tests on the other measures of their performance. Compared with analysts in the bottom quintile, those in the top quintile have significantly higher financial knowledge, tend to plagiarize less, and produce longer analyst reports. The patterns support the validity of our bag-of-words industry knowledge measure for distinguishing competent and incompetent analysts.

[Insert Table 5 here]

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 et al., 2006; Ivkovic and Jegadeesh, 2004;), while others cast doubt on the information discovery role of analysts (Kim and Song, 2015; Livnat and Zhang, 2012). As knowledge of vocabulary or jargons is a minimum requirement for fundamental analysis, unqualified analysts may hide their lack of industry knowledge for gathering and interpreting information by copying the content of news or other analysts' reports. This strategy takes little efforts and is not risky in a country of weak protection on intellectual property rights.[7]

In 2016, Bloomberg released an article criticizing Chinese sell-side analysts.[8] 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.[9] 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.

4.2. Connection benefits or cronyism

4.2.1. Analyst performance and political connection

This section reports the results for our hypothesis testing. Table 6 presents the results of our baseline regressions of analysts' performance on their political connection. Most of the coefficients on kinship are not significant, except that politically connected analysts use more financial vocabulary in their reports. Because most of the politically connected analysts are at

[7] 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

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

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

SOEs, which have more resources and benefits such as subsidies than non-SOEs, the performance of analysts may be driven by these confounding factors. However, the average research quality of our sample analysts improves after 2015, including their research quality, piggybacking tendency and recommendation profitability.

[Insert Table 6 here]

To mitigate endogeneity issues, we use the DID design in Equation (2) to isolate the effect of losing political connection on analysts' performance. Before showing our DID results, we first examine the strength and relevance of the exogenous shock. Table A2 in Appendix A shows CSRC official turnover in each year. In 2015, 10 CSRC officials were replaced and 4 of them were investigated for corruption and punished by the CCP, which exceeds the turnover in any year before or after 2015. The drastic turnover at the top of CSRC is likely to sever the political connection of many analysts, which is exogenous to performance confounders.

Table 7 reports our DID results for which the dependent variables are analysts' industry knowledge, financial knowledge, report length, plagiarism and piggybacking tendency. The coefficient of the DID terms Post × Kinship is significantly positive for industry knowledge and report length. After the exogenous shock in 2015 that severs many analysts' political ties, the average industry knowledge improves by 15.756 percentage points for each one percentage point increase in our kinship measure, which is significant at the 10% level.

[Insert Table 7 here]

Table 8 reports our DID results for which the dependent variables are different specifications of analysts' recommendation profitability. The coefficient for Post × Kinship is positive across all specifications, as well as statistically significant for most columns. Compared with the period before 2015, one percentage point increase in analysts' kinship likelihood with financial regulators increase the one-month abnormal return from following their recommendations by 3.54 after 2015.

[Insert Table 8 here]

The results support the cronyism hypothesis over the connection benefit hypothesis.

Although politically connected analysts enjoy more resources, their research quality is not better than unconnected analysts at lower ranked brokerages, and their research quality relatively improves after they lose their political connection due to the exogenous shock of the anti-corruption campaign. Both the loss of political ties and China's clamp-down on corruption likely reduce the rent seeking at SOEs, which may incentivize the crony analysts to work harder or leave the sell-side sector after 2015. Our results also suggest that top-down monitoring may be effective in reducing corruption (Olken, 2007), and that anti-corruption campaigns may reduce corruption and improve economic efficiency (Colonnelli and Prem, 2022).

Political connection is a barrier to entry in the sell-side market, which protects unqualified analysts from market competition. 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. While we show large variability in analyst performance in China, Crane and Crotty (2020) find that the majority of sell-side analysts in the US market are skilled. In more democratic countries with higher transparency and less political intervention, unqualified workers are less likely to obtain and stay in high-paying positions than more autocratic countries with more political rent-seeking.

One concern for the exogenous shock is that the new CSRC officials may continue to exchange favors with brokerage officials, so that new connections are formed after old ones are severed due to the anti-corruption campaign. According to the organization Transparency International, China's corruption perception score increases from 37 to 45 from 2015 to 2022, and a higher score indicates less corruption.[10] With the effectiveness of the anti-corruption campaign, cronyism is likely to decrease the new CSRC and SOE officials are unlikely to abuse their powers as much as their predecessors.

4.2.2. Brokerage director promotion and analyst political connection

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

---

[10] https://www.transparency.org/en/cpi/2022/index/chn

potentially commissions at the brokerages. As SOE directors' salaries are government-controlled, the directors may not lose much personally after they hire inept employees, and they can increase their promotion prospects after giving the favors to their superiors at CSRCs.

Table 9 reports our whole sample and subsample estimation results for Equation (3). Officials at brokerages with more politically connected analysts, proxied by their kinship with CSRC officials, are significantly more likely to be promoted, which is only statistically significant before 2015. One percentage point increase in the average analyst kinship measure increases the likelihood of brokerage official promotion by 2.459 percentage points, which is significant at the 5% level. The pattern is robust to alternative control variable specifications.

[Insert Table 9 here]

Table 10 reports our DID estimation results based on Equation (4). The coefficient of Post × Kinship is significantly negative, so brokerage officials who hire politically connected analysts are less likely to be promoted after 2015 than before. The loss of political ties also affects brokerage officials. In an autocracy with ill defined rights and weak rules, the proceeds of corruption hinges on the officials receiving the bribes remaining in power. The financial regulators personally benefit from the favor exchange, and political rent-seeking can explain politicians' high rates of asset growth, as documented by Fisman et al. (2014) for Italian politicians.

[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 (2019), 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. However, corruption-prone directors are likely to misuse their power in many ways, which can explain their corruption charges.

5. Additional tests

In this section, we test whether the connection-based hiring has real effects on market efficiency. If a country has weak legal institutions, it tends to have a smaller, less valuable and less efficient capital market (La Porta et al., 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 et al., 2005; Cheung et al., 2006; Jiang et al., 2010;). Does low-quality investment research from crony analysts also contribute to the inefficient capital market in China?

Studies on the US market show that analysts are important information intermediaries who can affect firm policies (Derrien and Kecskes, 2013; Guo et al., 2019). We test how crony analysts affect Chinese financial market efficiency in the baseline and DID tests specified by Equation (5) to (7).

$$\text{Price informativeness}_{j,t} = \alpha + \beta \cdot \text{Post}_{j,t} + \gamma \cdot \text{Controls}_{j,t} + \delta_t + \eta_j + \varepsilon_{j,t} \tag{5}$$

$$\text{Price informativeness}_{j,t} = \alpha + \beta \cdot \text{Political connection}_{j,t} + \gamma \cdot \text{Controls}_{j,t} + \delta_t + \eta_j + \varepsilon_{j,t} \tag{6}$$

$$\text{Price informativeness}_{j,t} = \alpha + \beta_1 \cdot \text{Political connection}_{j,t} \cdot \text{Post}_t$$

$$+\beta_2 \cdot \text{Political connection}_{j,t} + \beta_3 \cdot \text{ Post}_t \qquad (7)$$
$$+\gamma \cdot \text{ Controls}_{j,t} + \delta_t + \eta_j + \varepsilon_{j,t}$$

Where the subscript j denotes each healthcare firm. The firm year level $\text{Political connection}_{j,t}$ measures the intensity of firms receiving coverage from politically connected analysts. Equation (6) is estimated first on the whole sample and then on the pre-2015 and post-2015 subsamples.

Our proxies for stock price informativeness include Amihud's stock illiquidity measure (2002) and Weller's price jump ratio (2018). We control for factors that could affect firms' information asymmetry, including firms' market capitalization, stock price and return volatility, following Harford et al. (2019) and Weller's (2018). We also use firm and year fixed effects to control for time invariant or macroeconomic confounders.

Relative to private brokerages, large SOE brokerages have comparative advantage in broker services, which are often bundled with security research services, and more than 70% of money in the Chinese stock market is from retail investors, institutional investors are likely to continue using the service of a brokerage even if its research in certain industries is not useful due to crony hiring. In addition, the Chinese government has administrative monopoly on the financial market, so the unqualified financial analysts crowd out more competent competitors due to the limited brokerage licenses and jobs available. The lower research quality means less informative analysis on companies and less timely responses to changes in fundamentals, but the connected analysts may have access to more insider information that increases the price informativeness of the stocks they cover, so crony hiring has overall ambiguous effects on financial market efficiency.

The results are in Table 11 to Table 14. Using either measure, we find that informational efficiency does not change significantly for A-share healthcare firms on average after 2015, and there is a small difference between stocks most intensely covered by politically connected analysts and those by unconnected analysts. The DID term is not statistically significant either.

[Insert Table 11 here]

[Insert Table 12 here]

[Insert Table 13 here]

[Insert Table 14 here]

A larger number of analysts covering an industry can improve information efficiency in the U.S. market (Merkley et al., 2017). However, we show that politically connected analysts do not contribute to financial market efficiency. Besides potential access to insider information, the crony analysts have relatively little impact on the financial market, possibly because investors can select the research service from competent analysts or conduct their own analysis. In addition, the Chinese market has capital control, a high percentage of retail investors and stringent short-sale constraints (Mei et al., 2009), where the roles of institutional investors are relatively limited.[11] These characteristics contribute to speculations, leading to drastic bubbles and busts in the A-share market (Xiong and Yu (2011), which is another reason for the negligible effects of crony hiring on average.

6. Robustness tests

Our results are robust to various variable and estimation method specifications. Table 15 presents the result for Equation (2) using winsorized kinship and analyst performance. In Table 16 and Table 17, we adopt alternative measures of kinship to conduct analysis for Equation (2), which is the dummy variable Kinshipcat, that equals 1 when the continuous kinship measure is bigger or equals 0.7. These results are consistent with the main result.

In Table 18, we test whether the results for Equation (3) would persist when the dummy measure of kinship is involved instead of the continuous measure, as done in Table 16 and Table 17. Besides treating age as a continuous variable, we also account for age in an alternative way by creating a dummy variable for it, Agecat, which equals 1 if the age is between 50 and

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

60. The results do not change qualitatively.

Table 19 presents the results of linear probability model regression for Equation (4), with different specifications for kinship and age. Table 20 presents the results of logit regression for Equation (4), with respectively various specifications for kinship and age. These results are consistent with the results obtained in the main analysis.

In untabulated results for Equation (7), we additionally include the proportion of institutional holding and the number of analysts that covered the respective firms in a year as control variables. We also regress brokerage official promotion on lagged brokerage-level kinship. The results are qualitatively similar to the main results.

7. Conclusions

Using a novel measure for analysts' industry knowledge, we document a form of connection-based corrupt hiring at state-owned brokerages and show that the nonmeritocratic hires tend to lack industry knowledge and investment insight, who negatively affect market efficiency and impose real costs on Chinese A-share investors. As exchange of favors, securities regulators are likely to promote brokerage directors after they hire unqualified analysts. However, political connection is only one of the possible explanations for the existence of unqualified professionals in a state-regulated industry, where political barriers to entry and ill-defined property rights result in imperfect market competition. In addition, we only investigate one form of corrupt hiring at Chinese SOE brokerages. Nepotism is also possible. Besides hiring analysts connected with financial regulators, SOE brokerage directors may hire analysts connected with themselves to transfer resources to their own networks. In addition, not all CSRC officials have the power to appoint personnel, so some SOE directors may receive other forms of benefits in return, such as business advantages or nonpublic information.

Unlike Weaver (2021), we find that corrupt hires are less competent and corrupt hiring negatively affects market efficiency. Compared with community health workers in Weaver's sample, financial analysts require more specialized knowledge and higher skills to deliver useful services to investors and their performance variation is much larger. In addition, financial analysts are likely to have more rent extraction opportunities on the job than community health workers. For example, security analysts could issue biased reports for brokerage commissions and investment banking fees (Groysberg et al., 2011). Besides the difference in worker skill and task complexity and autonomy, our results are different from Weaver (2021) due to the different institutional context. The inter-generational correlations of wealth are relatively high in China, so wealth may not be indicative of individuals' performance on the job, especially where abstract thinking and initiatives are required. In autocratic countries with low social mobility, corruption based on political connection creates large distortions in resource allocation and negatively impacts social welfare.

References

Acemoglu, D., Johnson, S., Kermani, A., Kwak, J., Mitton, T., 2016. The Value of Connections in Turbulent Times: Evidence from the United States. Journal of Financial Economics 121, 368–391.

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

Amihud, Y., 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. Journal of Financial Markets 5(1), 31–56.

Amore, M. D., Bennedsen, M., 2013. The Value of Local Political Connections in a Low-Corruption Environment. Journal of Financial Economics 110, 387–402.

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

Beck, P. J., Maher, M. W., 1986. A Comparison of Bribery and Bidding in Thin Markets. Economics Letters 20, 1–5.

Benzell, S. G., Cooke, K., 2021. A Network of Thrones: Kinship and Conflict in Europe, 1495–1918. American Economic Journal: Applied Economics 13, 102–33.

Bertrand, M., Bombardini, M., Trebbi, F., 2014. Is It Whom You Know or What You Know? An Empirical Assessment of the Lobbying Process. American Economic Review 104, 3885–3920.

Bertrand, M., Djankov, S., Hanna, R., Mullainathan, S., 2007. Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption. Quarterly Journal of Economics 122, 1639–1676.

Borjas, G. J., 2002. The Wage Structure and the Sorting of Workers into the Public Sector. Working Paper 9313, National Bureau of Economic Research.

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.

Brodsgaard, K. E., Li, X., 2013. SOE Reform in China: Past, Present and Future. Copenhagen Journal of Asian Studies 31, 54–78.

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.

Chemmanur, T. J., Loutskina, E., Tian, X., 2014. Corporate venture capital, value creation, and innovation. The Review of Financial Studies 27, 2434-2473.

Chen, T., Kung, J., 2019. Busting the "Princelings": The Campaign Against Corruption in China's Primary Land Market. Quarterly Journal of Economics 134(1), 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.

Clement, M. B., 1999. Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter? Journal of Accounting and Economics 27(3), 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.

Colonnelli, E., Li, B., Liu, E., 2022. Investing with the Government: A Field Experiment in China. SSRN Working Paper.

Colonnelli, E., Prem, M., 2021. Corruption and Firms. Review of Economic Studies 89, 695–732.

Colonnelli, E., Prem, M., Teso, E., 2020. Patronage and Selection in Public Sector Organizations. American Economic Review 110, 3071–99.

Cooper, M. J., Gulen, H., Ovtchinnikov, A., 2010. Corporate Political Contributions and Stock Returns. Journal of Finance 65, 687–724.

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

Cruz, C., Labonne, J., Querubin, P., 2017. Politician Family Networks and Electoral Outcomes: Evidence from the Philippines. American Economic Review 107, 3006–37.

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.

Diao, W., Zhan, C., 2023. Rooted in the Land: Clanship and Land Market in China. Available at SSRN 4609646.

Djankov, S., La Porta, R., Lopez-de Silanes, F., Shleifer, A., 2010. Disclosure by Politicians. American Economic Journal: Applied Economics 2, 179–209.

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

Enke, B., 2019. Kinship, Cooperation, and the Evolution of Moral Systems. Quarterly Journal of Economics 134, 953–1019.

Esteban, J., Ray, D., 2006. Inequality, Lobbying, and Resource Allocation. American Economic Review 96, 257–279.

Faccio, M., 2006. Politically Connected Firms. American Economic Review 96, 369–386.

Fafchamps, M., Labonne, J., 2017. Do Politicians' Relatives Get Better Jobs? Evidence from Municipal Elections. Journal of Law, Economics, and Organization 33, 268–300.

Fang, L. H., Lerner, J., Wu, C., 2017. Intellectual Property Rights Protection, Ownership, and Innovation: Evidence from China. Review of Financial Studies 30(7), 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., 2001. Estimating the Value of Political Connections. American Economic Review 91, 1095– 1102.

Fisman, R., Schulz, F., Vig, V., 2014. The Private Returns to Public Office. Journal of Political Economy 122, 806–862.

Fisman, R., Shi, J., Wang, Y., Xu, R., 2018. Social Ties and Favoritism in Chinese Science. Journal of Political Economy 126, 1134–1171.

Fisman, R., Svensson, J., 2007. Are Corruption and Taxation Really Harmful to Growth? Firm Level Evidence. Journal of Development Economics 83, 63–75.

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.

Gagliarducci, S., Manacorda, M., 2020. Politics in the Family: Nepotism and the Hiring Decisions of Italian Firms. American Economic Journal: Applied Economics 12, 67–95.

Giannetti, M., Liao, G., You, J., Yu, X., 2020. The Externalities of Corruption: Evidence from Entrepreneurial Firms in China. Review of Finance 25, 629–667.

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

Giuliano, P., Nunn, N., 2018. Ancestral Characteristics of Modern Populations. Economic History of Developing Regions 33, 1–17.

Goldman, E., Rocholl, J., So, J., 2009. Do Politically Connected Boards Affect Firm Value? Review of Financial Studies 22, 2331–2360.

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.

Griffin, J. M., Liu, C., Shu, T., 2022. Is the Chinese Anticorruption Campaign Authentic? Evidence from Corporate Investigations. Management Science 68, 7248–7273.

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., 2019. Analyst Career Concerns, Effort Allocation, and Firms' Information Environment. Review of Financial Studies 32, 2179–2224.

He, G., Wang, S., Zhang, B., 2020. Watering Down Environmental Regulation in China. Quarterly Journal of Economics 135, 2135–2185.

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

Howard, A., Robinson, J., Dane, S., Kan, W., Kong, Z. (2019). Northeastern SMILE Lab - Recognizing Faces in the Wild. Retrieved from https://kaggle.com/competitions/recognizing-faces-in-the-wild.

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.

Jegadeesh, N., Kim, W., 2010. Do Analysts Herd? An Analysis of Recommendations and Market Reactions. Review of Financial Studies 23(2), 901–937.

Jiang, G., Lee, C., 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.

Krueger, A. O., 1974. The Political Economy of the Rent-Seeking Society. American Economic Review 64, 291–303.

Kwon, O. W., Lee, J. H., 2003. Text Categorization Based on K-Nearest Neighbor Approach for Web Site Classification. Information Processing & Management 39(1), 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.

Leff, N. H., 1964. Economic Development Through Bureaucratic Corruption. American

Behavioral Scientist 8, 8–14.

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(7), 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.

Lien, D.-H. D., 1986. A Note on Competitive Bribery Games. Economics Letters 22, 337–341.

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

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:1907.1169.

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., 2011. When Are Analyst Recommendation Changes Influential? Review of Financial Studies 24, 593–627.

Lu, J., Zhou, X., Tan, Y.-P., Shang, Y., Zhou, J., 2013. Neighborhood Repulsed Metric Learning for Kinship Verification. IEEE Transactions on Pattern Analysis and Machine Intelligence 36, 331–345.

Lui, F. T., 1985. An Equilibrium Queuing Model of Bribery. Journal of Political Economy 93, 760–781.

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., Sextroh, C. J., 2021. Analyst Coverage Overlaps and Interfirm Information Spillovers. Journal of Accounting Research 59(4), 1425-1480.

Martinez-Bravo, M., Padro i Miquel, G., Qian, N., Yao, Y., 2022. The Rise and Fall of Local Elections in China. American Economic Review 112, 2921–58.

Mauro, P., 1995. Corruption and Growth. Quarterly Journal of Economics 110, 681–712.

Mei, J., Scheinkman, J. A., Xiong, W., 2009. Speculative Trading and Stock Prices: Evidence from Chinese A-B Share Premia. Annals of Economics and Finance 10, 225–255.

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.

Moscona, J., Nunn, N., Robinson, J. A., 2020. Segmentary Lineage Organization and Conflict in Sub-Saharan Africa. Econometrica 88, 1999–2036.

Olken, B. A., 2007. Monitoring Corruption: Evidence from a Field Experiment in Indonesia. Journal of Political Economy 115, 200–249.

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

Pei, M., 2016. China's crony capitalism: The dynamics of regime decay. Harvard University Press.

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

Sequeira, S., 2016. Corruption, Trade Costs, and Gains from Tariff Liberalization: Evidence from Southern Africa. American Economic Review 106, 3029–63.

Shleifer, A., Vishny, R. W., 1993. Corruption. Quarterly Journal of Economics 108, 599–617.

Shleifer, A., Wolfenzon, D., 2002. Investor Protection and Equity Markets. Journal of Financial Economics 66(1), 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.

Torosin, N. S., Ward, J., Bell, A. V., Knapp, L. A., 2020. Genetic Similarity and Facial Cues for Kin Recognition in Humans. bioRxiv pp. 2020–04.

Treisman, D., 2000. The Causes of Corruption: A Cross-National Study. Journal of Public Economics 76, 399–457.

Unirule Institute of Economics, 2015. The Nature, Performance, and Reform of the State-owned Enterprises (Second Edition). World Scientific.

Unirule Institute of Economics, 2016. The Causes, Behaviors, and Termination of Administrative Monopoly in China (Second Edition). World Scientific.

Vidal, J. B., Draca, M., Fons-Rosen, C., 2012. Revolving Door Lobbyists. American Economic Review 102, 3731–3748.

Weaver, J., 2021. Jobs for Sale: Corruption and Misallocation in Hiring. American Economic Review 111, 3093–3122.

Weller, B. M., 2018. Does Algorithmic Trading Reduce Information Acquisition? Review of Financial Studies 31, 2184–2226.

Xiong, W., Yu, J., 2011. The Chinese Warrants Bubble. American Economic Review 101(6), 2723-2753.

Xu, G., 2018. The Costs of Patronage: Evidence from the British Empire. American Economic Review 108, 3170–98.

Table 1

Descriptive statistics for analyst performance.

This table reports the summary statistics of the variables related to analyst performance analysis in this paper. We report the number of observations (Obs.), mean (Mean), standard deviation (Std. Dev.), minimum value (Min), 1st percentile (P1), 50th percentile (P50), and 99th percentile (P99), maximum value (Max.), skewness (Skew.), kurtosis (Kurt.). All variables are defined in Table A1 of Appendix A. All variables in this table are at the analyst-brokerage-year level.

	Obs.	Mean	Std. Dev.	Min	P1	P50	P99	Max	Skew.	Kurt.
Panel A: Full sample										
Kinship	956	0.633	0.180	0.134	0.141	0.675	0.922	0.944	-0.732	2.814
Employment at Central SOE	956	0.363	0.481	0	0	0	1	1	0.570	1.325
Employment at Local SOE	956	0.510	0.500	0	0	1	1	1	-0.042	1.002
Employment at SOE	956	0.873	0.333	0	0	1	1	1	-20.250	6.046
Analyst Experience	956	4.856	3.292	0	0	4	14	16	0.712	2.951
Analyst Education	956	0.934	0.248	0	0	1	1	1	-3.499	13.245
Portfolio Complexity	956	11.017	18.098	0	0	3	79	190	3.625	23.271
Industry Knowledge	956	24.835	26.109	0.500	2	18	146.500	320	3.979	29.695
Financial Knowledge	956	43.860	11.455	6	17.667	44	71	100	0.118	3.896
Report Length	956	9.906	8.080	2	2	7.873	44	81	3.240	20.004
Plagiarism	956	0.944	0.037	0.492	0.793	0.952	0.990	1	-3.836	33.343
Piggyback	536	0.064	0.159	-0.307	-0.216	0.055	0.464	1.327	1.896	15.508
AR_{1m} (Neutral)	536	0.969	2.742	-5.883	-5.170	0.814	8.609	21.509	1.233	9.750
AR_{1m} (Recommend)	536	0.797	2.625	-16.069	-5.788	0.680	7.993	13.415	-0.015	7.374
AR_{2m} (Neutral)	536	1.643	5.193	-11.121	-10.498	1.398	15.449	26.342	0.710	5.251
AR_{2m} (Recommend)	536	1.427	4.951	-12.425	-10.498	1.168	14.673	25.333	0.529	4.786
AR_{3m} (Neutral)	536	2.287	7.607	-16.242	-14.388	2.053	23.669	40.141	0.658	5.072
AR_{3m} (Recommend)	536	1.970	7.272	-19.029	-14.388	1.663	22.061	39.662	0.540	4.961
Panel B: Connected Analysts										
Kinship	400	0.797	0.058	0.701	0.701	0.793	0.944	0.944	0.381	2.536
Employment at Central SOE	400	0.393	0.489	0	0	0	1	1	0.440	1.194
Employment at Local SOE	400	0.475	0.500	0	0	0	1	1	0.100	1.010
Employment at SOE	400	0.868	0.339	0	0	1	1	1	-2.168	5.700

Table 1
(*continued*)

	Obs.	Mean	Std. Dev.	Min	P1	P50	P99	Max	Skew.	Kurt.
Panel B: Connected Analysts										
Analyst Experience	400	5.263	3.354	0	0	5	14.500	16	0.680	3.010
Analyst Education	400	0.920	0.272	0	0	1	1	1	-3.096	10.587
Portfolio Complexity	400	10.912	20.796	0	0	2	106	190	4.179	26.999
Industry Knowledge	400	26.617	32.244	.5	2.250	17.197	168.700	320	4.018	26.529
Financial Knowledge	400	44.135	11.563	6	17.500	44.659	71.200	79	-0.160	3.130
Report Length	400	10.460	9.653	2	2	8	56	81	3.435	19.698
Plagiarism	400	0.942	0.035	0.740	0.799	0.951	0.991	0.993	-2.187	10.401
Piggyback	214	0.073	0.153	-0.216	-0.187	0.062	0.456	0.466	0.395	2.577
AR_{1m} (Neutral)	214	1.165	2.953	-5.883	-5.883	0.730	10.624	13.745	0.722	4.763
AR_{1m} (Recommend)	214	1.079	2.854	-5.883	-5.883	0.726	8.609	13.415	0.633	4.819
AR_{2m} (Neutral)	214	2.071	5.847	-11.121	-11.121	1.407	21.265	25.426	0.764	4.654
AR_{2m} (Recommend)	214	1.959	5.608	-11.121	-11.121	1.407	20.066	25.333	0.738	4.796
AR_{3m} (Neutral)	214	2.858	8.644	-15.596	-14.388	1.670	31.103	40.141	0.833	4.951
AR_{3m} (Recommend)	214	2.636	8.325	-15.596	-14.388	1.587	29.291	39.662	0.795	5.053
Panel C: Unconnected Analysts										
Kinship	556	0.515	0.152	0.134	0.141	0.562	0.695	0.699	-0.743	2.543
Employment at Central SOE	556	0.342	0.475	0	0	0	1	1	0.667	1.445
Employment at Local SOE	556	0.536	0.499	0	0	1	1	1	-0.144	1.021
Employment at SOE	556	0.878	0.328	0	0	1	1	1	-2.306	6.316
Analyst Experience	556	4.563	3.218	0	0	4	13	14	0.736	2.877
Analyst Education	556	0.944	0.230	0	0	1	1	1	-3.872	15.995
Portfolio Complexity	556	11.092	15.897	0	0	4	70	115	2.494	10.649
Industry Knowledge	556	23.553	20.533	1	2	18.829	90.867	186	2.627	14.451
Financial Knowledge	556	43.662	11.383	11.333	17.667	43.731	70	100	0.327	4.516
Report Length	556	9.508	6.707	2	2	7.822	36	51	2.203	9.867
Plagiarism	556	0.945	0.038	0.492	0.793	0.953	0.990	1	-4.745	44.859
Piggyback	322	0.058	0.164	-0.307	-0.237	0.052	0.477	1.327	2.732	22.442
AR_{1m} (Neutral)	322	0.838	2.588	-5.883	-4.480	0.818	7.993	21.509	1.688	15.217
AR_{1m} (Recommend)	322	0.609	2.448	-16.069	-5.170	0.667	7.053	9.495	-0.780	9.722
AR_{2m} (Neutral)	322	1.358	4.697	-11.121	-9.553	1.398	14.349	26.342	0.525	5.389
AR_{2m} (Recommend)	322	1.074	4.435	-12.425	-9.924	1.075	14.278	15.449	0.095	3.705
AR_{3m} (Neutral)	322	1.908	6.820	-16.242	-13.938	2.112	18.603	33.463	0.296	4.264
AR_{3m} (Recommend)	322	1.528	6.455	-19.029	-14.093	1.736	17.692	22.061	0.021	3.515

Table 2
Analyst performance correlations.

This table shows correlations across the full sample for the key variables about analysts as defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) Kinship	1.00													
(2) Central SOE	0.00	1.00												
(3) Local SOE	-0.01	-0.79***	1.00											
(4) Industry Knowledge	-0.01	0.03	0.00	1.00										
(5) Financial Knowledge	0.06	0.20***	-0.17***	0.25***	1.00									
(6) Report Length	0.04	0.08*	-0.09**	0.74***	0.37***	1.00								
(7) Plagiarism	-0.04	-0.10***	0.11***	-0.50***	-0.37***	-0.71***	1.00							
(8) Piggyback	-0.03	-0.06	0.05	-0.03	-0.01	-0.01	-0.02	1.00						
(9) AR_{1m} (Neutral)	0.00	-0.05	0.05	-0.05	0.03	-0.01	-0.05	0.69***	1.00					
(10) AR_{1m} (Recommend)	0.06	-0.03	0.02	-0.08*	0.07*	-0.01	-0.04	0.43***	0.86***	1.00				
(11) AR_{2m} (Neutral)	0.02	-0.04	0.05	-0.04	0.07	0.01	-0.07	0.66***	0.98***	0.88***	1.00			
(12) AR_{2m} (Recommend)	0.05	-0.03	0.03	-0.05	0.10**	0.01	-0.07	0.51***	0.88***	0.96***	0.93***	1.00		
(13) AR_{3m} (Neutral)	0.02	-0.03	0.05	-0.02	0.08*	0.02	-0.07*	0.65***	0.96***	0.87***	0.99***	0.92***	1.00	
(14) AR_{3m} (Recommend)	0.04	-0.03	0.04	-0.04	0.12***	0.02	-0.08*	0.53***	0.86***	0.93***	0.92***	0.99***	0.93***	1.00

Table 3

Descriptive statistics for analyst characteristics by brokerage ownership categories.

This table reports the distribution of connected and unconnected analysts in central, local SOEs and non-SOE brokerages in this paper, as well as the summary statistics of their respective characteristics and performance. We report the number of observations (Obs.), mean (Mean), standard deviation (Std. Dev.), minimum value (Min), 1st percentile (P1), 50th percentile (P50), and 99th percentile (P99), maximum value (Max.), skewness (Skew.), kurtosis (Kurt.). All variables are defined in Table A1 of Appendix A. All variables in this table are at the analyst-brokerage-year level.

	Obs.	Mean	Std. Dev.	Min	P1	P50	P99	Max	Skew.	Kurt.
Panel A: Non-SOE										
Kinship	121	0.634	0.182	0.141	0.141	0.676	0.854	0.896	-0.824	2.820
Analyst Experience	121	5.223	3.646	0	0	5	15	16	0.605	2.757
Analyst Education	121	0.917	0.276	0	0	1	1	1	-3.032	10.190
Portfolio Complexity	121	7.405	10.978	0	0	3	55	63	2.769	11.843
Industry Knowledge	121	21.718	23.560	1	3	15.571	149	163.400	3.546	19.354
Financial Knowledge	121	43.302	10.836	15	18	43	70	74	0.115	3.638
Report Length	121	10.734	8.921	2	3	8	44	44.500	1.983	6.938
Plagiarism	121	0.942	0.033	0.758	0.847	0.949	0.988	0.990	-2.181	10.636
Piggyback	60	0.065	0.128	-0.187	-0.187	0.057	0.348	0.348	0.034	2.237
AR_{1m} (Neutral)	60	0.953	2.181	-4.039	-4.039	1.097	6.286	6.286	-0.239	2.969
AR_{1m} (Recommend)	60	0.947	2.183	-4.039	-4.039	1.097	6.286	6.286	-0.232	2.958
AR_{2m} (Neutral)	60	1.384	4.274	-8.589	-8.589	1.817	12.822	12.822	-0.223	3.189
AR_{2m} (Recommend)	60	1.385	4.273	-8.589	-8.589	1.817	12.822	12.822	-0.224	3.191
AR_{3m} (Neutral)	60	1.537	6.433	-13.481	-13.481	1.887	19.105	19.105	-0.257	3.357
AR_{3m} (Recommend)	60	1.548	6.430	-13.481	-13.481	1.887	19.105	19.105	-0.262	3.366
Panel B: Local SOEs										
Kinship	488	0.633	0.187	0.134	0.167	0.675	0.891	0.913	-0.671	2.781
Analyst Experience	488	4.625	3.234	0	0	4	14	15	0.794	3.037
Analyst Education	488	0.932	0.251	0	0	1	1	1	-3.444	12.860
Portfolio Complexity	488	13.025	21.826	0	0	4	115	190	3.443	19.518
Industry Knowledge	488	24.575	22.236	0.500	1.500	19.100	114.600	208	3.057	18.749
Financial Knowledge	488	42.587	11.311	11.333	17.333	43.133	69	100	0.351	4.817
Report Length	488	9.313	7.401	2	2	7.702	41.500	73	3.655	24.949
Plagiarism	488	0.946	0.032	0.707	0.793	0.955	0.990	1	-2.810	16.169
Piggyback	286	0.072	0.175	-0.216	-0.204	0.074	0.491	1.327	2.534	18.507

Table 3
(*continued*).

	Obs.	Mean	Std. Dev.	Min	P1	P50	P99	Max	Skew.	Kurt.
Panel B: Local SOEs										
AR_{1m} (Neutral)	286	1.105	2.802	-5.883	-4.480	0.897	9.495	21.509	1.921	13.546
AR_{1m} (Recommend)	286	0.844	2.699	-16.069	-5.788	0.684	8.609	13.415	-0.188	9.785
AR_{2m} (Neutral)	286	1.864	5.162	-11.121	-9.553	1.582	20.066	26.342	0.945	6.048
AR_{2m} (Recommend)	286	1.576	4.925	-12.425	-10.118	1.221	15.449	25.333	0.690	5.410
AR_{3m} (Neutral)	286	2.641	7.585	-15.596	-13.938	2.279	29.291	40.141	0.806	5.630
AR_{3m} (Recommend)	286	2.246	7.246	-19.029	-14.388	1.967	23.669	39.662	0.683	5.694
Panel C: Central SOE										
Kinship	347	0.633	0.183	0.141	0.177	0.675	0.907	0.944	-0.793	2.858
Analyst Experience	347	5.052	3.228	0	0	5	13	15	0.631	2.900
Analyst Education	347	0.942	0.233	0	0	1	1	1	-3.796	15.411
Portfolio Complexity	347	9.452	13.383	0	0	3	61	66	2.063	7.091
Industry Knowledge	347	26.287	31.424	2	2.500	18	174	320	4.200	29.854
Financial Knowledge	347	45.845	11.622	6	17.667	45.308	71.400	79	-0.208	3.205
Report Length	347	10.452	8.631	2	3	8.015	48	81	3.260	20.468
Plagiarism	347	0.940	0.043	0.492	0.784	0.951	0.993	1	-4.495	39.754
Piggyback	190	0.051	0.142	-0.307	-0.307	0.045	0.441	0.441	0.273	2.978
AR_{1m} (Neutral)	190	0.769	2.808	-5.883	-5.883	0.624	10.624	10.624	0.408	4.047
AR_{1m} (Recommend)	190	0.679	2.647	-5.883	-5.883	0.521	7.053	10.624	0.313	3.919
AR_{2m} (Neutral)	190	1.391	5.506	-11.121	-11.121	1.125	21.265	21.265	0.551	4.270
AR_{2m} (Recommend)	190	1.218	5.199	-11.121	-11.121	0.808	14.673	21.265	0.460	4.105
AR_{3m} (Neutral)	190	1.991	7.980	-16.242	-15.517	1.661	31.103	31.103	0.600	4.370
AR_{3m} (Recommend)	190	1.688	7.575	-16.242	-14.388	1.040	22.061	31.103	0.501	4.137

Table 4

Descriptive statistics at the brokerage level.

This table reports the summary statistics of the variables involved in official promotion analysis in this paper. We report the number of observations (Obs.), mean (Mean), standard deviation (Std. Dev), minimum value (Min.), 1st percentile (P1), 50th percentile (P50), and 99th percentile (P99), maximum value (Max.), skewness (Skew.), kurtosis (Kurt.). All variables are defined in Table A1 of Appendix A. All variables in this table are at the official/brokerage-year level.

Variables	Obs.	Mean	Std. Dev.	Min	P1	P50	P99	Max	Skew.	Kurt.
Official Promotion	2583	0.214	0.410	0	0	0	1	1	1.396	2.950
Kinship	3727	0.612	0.135	0.141	0.144	0.632	0.837	0.887	-0.877	3.980
Age	3727	49.831	6.936	29	35	50	66	75	0.169	2.918
Gender	3727	0.858	0.350	0	0	1	1	1	-2.046	5.185
Official Education	3727	0.798	0.401	0	0	1	1	1	-1.486	3.209
Certified	3727	0.464	0.499	0	0	0	1	1	0.144	1.021
Industry Knowledge	3727	24.329	18.701	1	3	19.263	114.600	116.167	2.715	12.149
Financial Knowledge	3727	44.716	8.865	18	21.25	44.138	64.792	64.987	-0.100	3.102
Report Length	3727	9.053	4.924	3	3	8.049	33	44.458	2.907	17.560
Plagiarism	3727	0.946	0.023	0.851	0.877	0.951	0.992	1	-1.152	4.959
Brokerage Revenue	3727	10363.700	10259.900	764.900	941.340	6086.570	43139.700	56013.400	1.660	5.475

Table 5
T-tests on analyst performance.

This table shows the variability in analyst performance. We first sort our sample analysts into quintiles based on their industry knowledge, and then conduct t tests on the other measures of their performance. Column "Top" contains the mean of the performance measure of the analysts who are in the top quintile in terms of industry knowledge; column "Bottom" contains the mean of the performance measure of the analysts who are in the bottom quintile in terms of industry knowledge. All variables are defined in Table A1 of Appendix A. The granularity of the t-tests is analyst-brokerage-year. Standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	Top	Bottom	Diff-in-Mean
Financial Knowledge	45.196	37.595	7.601***
			(1.335)
Plagiarism	0.918	0.964	-0.046***
			(0.004)
Report Length	16.676	4.421	12.255***
			(0.743)
Piggyback	0.054	0.018	0.036
			(0.025)
AR_{1m} (Neutral)	0.526	0.306	0.221
			(0.380)
AR_{1m} (Recommend)	0.199	0.437	-0.238
			(0.408)
AR_{2m} (Neutral)	0.909	0.190	0.720
			(0.716)
AR_{2m} (Recommend)	0.598	0.487	0.111
			(0.707)
AR_{3m} (Neutral)	1.432	0.135	1.298
			(1.046)
AR_{3m} (Recommend)	0.910	0.579	0.331
			(1.026)

Table 6

Analyst performance and political connection: baseline regressions.

This table shows regression results of analyst performance on Kinship in Panel A and Post in Panel B. The dependent variable is analyst performance, measured by: (1) Industry Knowledge: the average number of occurrences of medical words by an analyst in a report written by the analyst in a particular year; (2) Financial Knowledge: the average number of occurrences of financial technical words in a report written by the analyst in a particular year; (3) Report Length: the average number of pages in a report written by the analyst in a particular year; (4) Plagiarism: the average of the highest cosine similarity between the reports, written by an analyst within a specific year, and other reports within the seven days preceding its publication (excluding the reports written by the same authors and those written by some of the same authors in the same brokerages); (5) Piggyback is cumulative abnormal return for seven days before analyst report issuance; (6) AR is calculated by multiplying the corresponding abnormal return of following analyst recommendation and holding for the respective periods as indicated in subscript (e.g. "1m" means one month), with a ternary variable, which takes the value 1 if the rating is better than the threshold, as indicated in the bracket in the header of the corresponding column; 0 if the rating equals the threshold; -1 if the rating is worse than the threshold. Kinship is a proxy of maximum kinship between an analyst and the members of CSRC management. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Industry Knowledge	(2) Financial Knowledge	(3) Report Length	(4) Plagiarism	(5) Piggyback	(6) AR_{1m} (Neutral)	(7) AR_{2m} (Neutral)	(8) AR_{3m} (Neutral)	(9) AR_{1m} (Recommend)	(10) AR_{2m} (Recommend)	(11) AR_{3m} (Recommend)
Panel A											
Kinship	-2.288	3.814*	0.975	-0.006	-0.024	0.112	0.794	1.011	0.788	1.410	1.589
	(4.357)	(2.187)	(1.239)	(0.006)	(0.050)	(0.793)	(1.342)	(1.935)	(0.664)	(1.166)	(1.707)
Analyst Experience	0.400	-0.002	0.121	0	-0.004**	-0.015	-0.007	-0.009	0.027	0.051	0.057
	(0.284)	(0.124)	(0.086)	(0)	(0.002)	(0.035)	(0.063)	(0.093)	(0.030)	(0.058)	(0.087)
Analyst Education	-7.547	-1.998	-2.428	0.008	-0.010	-0.547	-0.920	-0.962	-0.639	-1.052	-1.204
	(4.950)	(2.133)	(1.596)	(0.007)	(0.029)	(0.492)	(1.051)	(1.605)	(0.485)	(1.039)	(1.591)
Portfolio Complexity	0.039	0.035**	-0.030***	0***	0	0.005	0.013	0.022*	0	0.007	0.014
	(0.030)	0(.016)	(0.009)	(0)	(0)	(0.004)	(0.008)	(0.012)	(0.005)	(0.008)	(0.011)
Observations	678	678	678	678	536	536	536	536	536	536	536
R-squared	0.013	0.009	0.023	0.015	0.009	0.003	0.005	0.004	0.008	0.007	0.005
Panel B											
Post	2.015	3.107***	0.108	0.003	-0.132***	0.035	0.687	1.552*	0.307	1.177**	2.197***
	(1.891)	(0.911)	(0.581)	(0.003)	(0.014)	(0.312)	(0.570)	(0.828)	(0.282)	(0.527)	(0.770)
Analyst Experience	0.355	-0.037	0.123	0	-0.002	-0.015	-0.015	-0.030	0.025	0.037	0.027
	(0.290)	(0.121)	(0.086)	(0)	(0.002)	(0.034)	(0.063)	(0.092)	(0.030)	(0.058)	(0.087)
Analyst Education	-7.445	-1.951	-2.439	0.008	-0.008	-0.549	-0.939	-0.996	-0.653	-1.086	-1.254
	(4.993)	(2.206)	(1.604)	(0.007)	(0.025)	(0.494)	(1.053)	(1.599)	(0.490)	(1.043)	(1.584)
Portfolio Complexity	0.038	0.027*	-.031***	0***	0	0.004	0.012	0.019	-0.001	0.005	0.011
	(0.029)	(0.015)	(0.009)	(0)	(0)	(0.004)	(0.008)	(0.012)	(0.005)	(0.008)	(0.011)
Observations	678	678	678	678	536	536	536	536	536	536	536
R-squared	0.014	0.023	0.023	0.016	0.152	0.003	0.007	0.013	0.008	0.016	0.023

Table 7

Analyst performance and political connection: DID tests.

This table shows the difference-in-difference test results. The dependent variable is analyst performance, measured by: (1) Industry Knowledge: the average number of occurrences of medical words by an analyst in a report written by the analyst in a particular year; (2) Financial Knowledge: the average number of occurrences of financial technical words in a report written by the analyst in a particular year; (3) Report Length: the average number of pages in a report written by the analyst in a particular year; (4) Plagiarism: the average of the highest cosine similarity between the reports, written by an analyst within a specific year, and other reports within the seven days preceding its publication (excluding the reports written by the same authors and those written by some of the same authors in the same brokerages); (5) Piggyback is cumulative abnormal return for seven days before analyst report issuance. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)
	Industry Knowledge	Financial Knowledge	Report Length	Plagiarism	Piggyback
Post	-6.198	2.167	-2.867*	0.013*	-0.185**
	(5.320)	(2.655)	(1.567)	(0.008)	(0.076)
Kinship	-6.587	1.371	-1.543	0.004	-0.079
	(6.353)	(3.136)	(1.878)	(0.010)	(0.099)
Post × Kinship	**15.756***	**2.873**	**5.702****	**-0.017**	**0.084**
	(9.244)	**(4.003)**	**(2.656)**	**(0.012)**	**(0.113)**
Analyst Experience	0.157	-0.202*	0.079	0	-0.002
	(0.307)	(0.111)	(0.088)	(0)	(0.002)
Analyst Education	-6.372	-1.924	-1.940	0.005	-0.006
	(4.222)	(1.810)	(1.403)	(0.006)	(0.026)
Portfolio Complexity	0.048*	0.019	-0.032***	0***	0
	(0.028)	(0.015)	(0.009)	(0)	(0)
Observations	956	956	956	956	536
R-squared	0.015	0.031	0.021	0.011	0.154

Table 8

Analyst recommendation profitability and political connection: DID tests.

This table shows the difference-in-difference test results. The dependent variable is calculated by multiplying the corresponding abnormal return of following analyst recommendation and holding for the respective periods as indicated in subscript (e.g. "1m" means one month), with a ternary variable, which takes the value 1 if the rating is better than the threshold, as indicated in the bracket in the header of the corresponding column; 0 if the rating equals the threshold; -1 if the rating is worse than the threshold. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1) AR_{1m} (Neutral)	(2) AR_{2m} (Neutral)	(3) AR_{3m} (Neutral)	(4) AR_{1m} (Recommend)	(5) AR_{2m} (Recommend)	(6) AR_{3m} (Recommend)
Post	-3.514**	-5.691**	-6.571	-2.817*	-4.624*	-5.548
	(1.505)	(2.800)	(4.080)	(1.480)	(2.793)	(4.093)
Kinship	-3.654**	-6.096*	-7.645	-2.916	-4.803	-5.821
	(1.822)	(3.372)	(4.908)	(1.810)	(3.372)	(4.905)
Post × Kinship	**4.208****	**7.509****	**9.606***	**3.537***	**6.447***	**8.680**
	(1.930)	**(3.623)**	**(5.290)**	**(1.917)**	**(3.623)**	**(5.318)**
Analyst Experience	-0.007	0.007	0.013	0.008	0.025	0.027
	(0.027)	(0.050)	(0.074)	(0.025)	(0.048)	(0.073)
Analyst Education	-0.447	-0.702	-0.701	-0.569	-0.887	-1.018
	(0.496)	(0.987)	(1.463)	(0.494)	(0.980)	(1.452)
Portfolio Complexity	-0.004	-0.004	-0.005	-0.003	-0.004	-0.003
	(0.003)	(0.006)	(0.009)	(0.003)	(0.006)	(0.009)
Observations	658	658	658	658	658	658
R-squared	0.022	0.013	0.012	0.015	0.012	0.013

Table 9

Official promotion and analyst political connection: baseline tests.

This table shows the results of Cox proportional-hazards model for the full sample and the subsamples, respectively before 2015 and after 2015. The dependent variable is a dummy variable, which equals 1 if the brokerage official got promoted in the year. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management. Age is brokerage officials' age, and Age_{cat} is brokerage officials' age, which equals 1 if the Age of the official is between 50 and 60. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is official/brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	Full Sample		Before 2015		After 2015	
	(1)	(2)	(3)	(4)	(5)	(6)
Kinship	1.392**	1.398**	2.459**	2.336**	0.723	0.662
	(0.710)	(0.710)	(0.984)	(0.981)	(1.054)	(1.058)
Age	-0.005		0.021		-0.016	
	(0.009)		(0.015)		(0.011)	
Age_{cat}		0.225*		0.285		0.144
		(0.116)		(0.186)		(0.152)
Gender	-0.227	-0.274*	-0.34	-0.324	-0.247	-0.321
	(0.161)	(0.159)	(0.255)	(0.254)	(0.211)	(0.209)
Official Education	0.134	0.159	0.255	0.260	0.044	0.075
	(0.145)	(0.145)	(0.242)	(0.242)	(0.184)	(0.184)
Certified	0.080	0.038	0.222	0.241	-0.022	-0.084
	(0.117)	(0.115)	(0.188)	(0.185)	(0.155)	(0.152)
Brokerage Revenue	0***	0***	0	0	0***	0***
	(0)	(0)	(0)	(0)	(0)	(0)
Observations	2583	2583	1158	1158	1425	1425
Pseudo R^2	0.015	0.016	0.010	0.010	0.019	0.019

Table 10

Official promotion and analyst political connection: DID tests.

This table shows the difference-in-difference test results. The dependent variable is a dummy variable, which equals 1 if the brokerage official got promoted in the year. The explanatory variables are Kinship (a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management). Post equals 1 for observations after 2015, and 0 otherwise. Age is brokerage officials' current age. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is official/brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
Post	0.040**		0.371**
	(0.018)		(0.153)
Kinship		-0.056	0.221
		(0.104)	(0.141)
Post × Kinship			**-0.432****
			(0.201)
Age	-0.007***	-0.006***	-0.007***
	(0.001)	(0.001)	(0.001)
Gender	-0.002	-0.004	-0.011
	(0.023)	(0.023)	(0.023)
Official Education	0.012	0.012	0.008
	(0.020)	(0.020)	(0.02)
Certified	-0.009	-0.012	-0.011
	(0.016)	(0.016)	(0.016)
Brokerage Revenue	0**	0**	0**
	(0)	(0)	(0)
Observations	2583	2583	2583
R-squared	0.014	0.012	0.019

Table 11

Analyst political connection and Weller's price jump ratio: whole sample regressions.

This table shows the difference-in-difference test results. The dependent variable is Weller's price jump ratio. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management; and $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is on the report level. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)
Post	-0.006			0.031	0.031
	(0.009)			(0.044)	(0.044)
Kinship		0.046		0.075***	
		(0.030)		(0.022)	
Kinship × Post				**-0.049**	
				(0.055)	
$Kinship_{cat}$			-0.002		-0.049
			(0.015)		(0.055)
$Kinship_{cat}$ × Post					**-0.034**
					(0.021)
Portfolio Complexity	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)
Analyst Education	0.012	0.011	0.013	0.012	0.014
	(0.015)	(0.015)	(0.016)	(0.015)	(0.016)
Analyst Experience	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Market Cap.	0.016***	0.016***	0.016***	0.016***	0.016***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Price	0.043***	0.042***	0.042***	0.043***	0.043***
	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)
Volatility	-0.029***	-0.028***	-0.028***	-0.028***	-0.028***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Observations	13745	13745	13745	13745	13745
R-squared	0.005	0.005	0.005	0.005	0.005

Table 12

Analyst political connection and Weller's price jump ratio: subsample regressions.

This table shows the regression results for the subsamples, respectively before 2015 and after 2015. The dependent variable is Weller's price jump ratio. The explanatory variables are Kinship (a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management); and $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is on the report level. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	Before 2015		After 2015	
	(1)	(2)	(3)	(4)
Kinship	0.034		0.033	
	(0.024)		(0.057)	
$Kinship_{cat}$		0.010		-0.010
		(0.010)		(0.025)
Portfolio Complexity	0**	0**	0	0
	(0)	(0)	(0)	(0)
Analyst Education	0.022	0.019	0.010	0.016
	(0.018)	(0.018)	(0.027)	(0.030)
Analyst Experience	0.001	0.001	0	0
	(0.002)	(0.002)	(0.002)	(0.003)
Market Cap.	0.033***	0.033***	0.009	0.009
	(0.005)	(0.005)	(0.009)	(0.009)
Price	0.004	0.004	0.070***	0.070***
	(0.009)	(0.009)	(0.024)	(0.024)
Volatility	-0.013***	-0.014***	-0.040***	-0.040***
	(0.005)	(0.005)	(0.008)	(0.008)
Observations	3434	3434	7774	7774
R-squared	0.020	0.020	0.005	0.005

Table 13

Analyst political connection and Amihud illiquidity measure: whole sample regressions.

This table presents difference-in-difference (DID) tests on market informational efficiency, which is measured as the change in Amihud illiquidity from the previous year. The explanatory variables are Kinship (a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management); and $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. Post equals 1 for observations after 2015, and 0 otherwise. All control variables except volatility are lagged by 1 year. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is healthcare firm-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)	(5)
Post	-0.343***			-1.065	-0.487**
	(0.109)			(0.816)	(0.207)
Kinship		-0.004		-0.535	
		(0.544)		(0.725)	
Kinship × Post				**0.943**	
				(1.054)	
$Kinship_{cat}$			0.019		-0.080
			(0.118)		(0.130)
Post × $Kinship_{cat}$					**0.166**
					(0.208)
Portfolio Complexity	0	-0.002	-0.002	0	0
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Analyst Education	0.067	-0.065	-0.065	0.048	0.068
	(0.202)	(0.186)	(0.186)	(0.209)	(0.204)
Analyst Experience	0.069**	0.053**	0.053**	0.071**	0.070**
	(0.028)	(0.026)	(0.026)	(0.028)	(0.028)
Market Cap.	0	0	0	0	0
	(0)	(0)	(0)	(0)	(0)
Price	-0.015***	-0.017***	-0.017***	-0.015***	-0.015***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Volatility	-33.258***	-26.325***	-26.285***	-33.560***	-33.383***
	(5.770)	(4.914)	(4.897)	(5.877)	(5.832)
Observations	739	739	739	739	739
R-squared	0.093	0.081	0.081	0.094	0.093

Table 14

Analyst political connection and Amihud illiquidity measure: subsample regressions.

This table shows the regression results for the subsamples, respectively before 2015 and after 2015, on market informational efficiency, which is measured as the change in Amihud illiquidity measure from the previous year. The explanatory variables are Kinship (a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management); $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. Post equals 1 for observations after 2015, and 0 otherwise. All control variables except volatility are lagged by 1 year. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is healthcare firm-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	Before 2015		After 2015	
	(1)	(2)	(3)	(4)
Kinship	0.250		0.699	
	(0.647)		(0.729)	
$Kinship_{cat}$		0.069		0.082
		(0.113)		(0.157)
Portfolio Complexity	-0.009*	-0.009*	0.001	0.001
	(0.005)	(0.005)	(0.003)	(0.003)
Analyst Education	0.425**	0.422**	-1.211***	-1.170***
	(0.213)	(0.213)	(0.261)	(0.256)
Analyst Experience	0.009	0.010	0.060*	0.059*
	(0.037)	(0.036)	(0.034)	(0.034)
Market Cap.	0	0	0	0
	(0)	(0)	(0)	(0)
Price	0.002	0.002	-0.014***	-0.015***
	(0.005)	(0.005)	(0.005)	(0.005)
Volatility	5.826	5.782	-82.826***	-82.510***
	(4.624)	(4.490)	(10.811)	(10.796)
Observations	207	207	532	532
R-squared	0.054	0.054	0.197	0.196

Table 15

Winsorized kinship and analyst performance.

This table shows the difference-in-difference test results. The dependent variable is analyst performance: (1) Industry Knowledge: the average number of occurrences of medical words by an analyst in a report written by the analyst in a particular year; (2) Financial Knowledge: the average number of occurrences of financial technical words in a report written by the analyst in a particular year; (3) Report Length: the average number of pages in a report written by the analyst in a particular year; (4) Plagiarism: the average of the highest cosine similarity between the reports, written by an analyst within a specific year, and other reports within the seven days preceding its publication (excluding the reports written by the same authors and those written by some of the same authors in the same brokerages; (5) Piggyback is cumulative abnormal return for seven days before analyst report issuance. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management, and the variable is winsorized. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Industry Knowledge (w)	(2) Financial Knowledge (w)	(3) Report Length (w)	(4) Plagiarism (w)	(5) Piggyback (w)
Post	-5.771	2.564	-1.917	0.013**	-0.177**
	(3.891)	(2.619)	(1.293)	(0.006)	(0.077)
Kinship (w)	-9.738**	1.887	-2.215	0.006	-0.059
	(4.676)	(3.210)	(1.570)	(0.008)	(0.094)
Post × Kinship (w)	**12.085****	**0.952**	**3.725***	**-0.019****	**0.072**
	(5.959)	**(3.952)**	**(1.992)**	**(0.009)**	**(0.114)**
Analyst Experience	0.152	-0.156*	0.077	0	-0.002
	(0.154)	(0.089)	(0.051)	(0)	(0.002)
Analyst Education	-2.816	-1.322	-0.813	0.003	-0.006
	(2.205)	(1.234)	(0.786)	(0.003)	(0.026)
Portfolio Complexity	0.101***	0.016	-0.013**	0	0
	(0.018)	(0.013)	(0.006)	(0)	(0)
Observations	956	956	956	956	536
R-squared	0.030	0.030	0.014	0.016	0.153

Table 16

Kinship$_{cat}$ and analyst performance.

This table shows the difference-in-difference test results. The dependent variable is analyst performance: (1) Industry Knowledge: the average number of occurrences of medical words by an analyst in a report written by the analyst in a particular year; (2) Financial Knowledge: the average number of occurrences of financial technical words in a report written by the analyst in a particular year; (3) Report Length: the average number of pages in a report written by the analyst in a particular year; (4) Plagiarism: the average of the highest cosine similarity between the reports, written by an analyst within a specific year, and other reports within the seven days preceding its publication (excluding the reports written by the same authors and those written by some of the same authors in the same brokerages; (5) Piggyback is cumulative abnormal return for seven days before analyst report issuance. Kinship$_{cat}$ is a dummy variable, which equals 1 when the corresponding continuous kinship measure is bigger than or equals 0.7. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1) Industry Knowledge	(2) Financial Knowledge	(3) Report Length	(4) Plagiarism	(5) Piggyback
Post	2.265	3.550***	-0.430	0.008*	-0.148***
	(1.675)	(1.107)	(0.639)	(0.004)	(0.037)
Kinship$_{cat}$	0.340	0.093	-1.020	0.006	-0.039
	(2.693)	(1.319)	(0.837)	(0.005)	(0.043)
Post × Kinship$_{cat}$	**3.961**	**1.028**	**2.797****	**-0.013****	**0.063**
	(3.808)	**(1.621)**	**(1.158)**	**(0.006)**	**(0.058)**
Analyst Experience	0.135	-0.199*	0.068	0	-0.001
	(0.309)	(0.112)	(0.089)	(0)	(0.001)
Analyst Education	-6.371	-2.003	-1.941	0.005	-0.006
	(4.203)	(1.813)	(1.412)	(0.006)	(0.021)
Portfolio Complexity	0.045	0.018	-0.034***	0***	0
	(0.028)	(0.015)	(0.009)	(0)	(0)
Observations	956	956	956	956	536
R-squared	0.015	0.030	0.024	0.016	0.185

Table 17
$Kinship_{cat}$ and recommendation profitability.

This table shows the difference-in-difference test results. The dependent variable is calculated by multiplying the corresponding abnormal return of following analyst recommendation and holding for the respective periods as indicated in subscript (e.g. "1m" means one month), with a ternary variable, which takes the value 1 if the rating is better than the "neutral"; 0 if the rating is "neutral"; -1 if the rating is worse than "neutral". $Kinship_{cat}$ is a dummy variable, which equals 1 when the corresponding continuous kinship measure is bigger than or equals 0.7. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is analyst-brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1) AR_{1m}	(2) AR_{2m}	(3) AR_{3m}
Post	-1.474**	-1.842*	-1.506
	(0.618)	(1.026)	(1.457)
$Kinship_{cat}$	-1.240*	-1.836*	-2.134
	(0.661)	(1.097)	(1.555)
Post × $Kinship_{cat}$	**1.556****	**2.493****	**2.994***
	(.687)	**(1.162)**	**(1.655)**
Analyst Experience	-0.007	0.006	0.012
	(0.027)	(0.050)	(0.074)
Analyst Education	-0.402	-0.665	-0.677
	(0.509)	(1.003)	(1.487)
Portfolio Complexity	-0.004	-0.005	-0.005
	(0.003)	(0.006)	(0.009)
Observations	658	658	658
R-squared	0.024	0.014	0.011

Table 18

Cox hazard model for official promotion (with $Kinship_{cat}$).

This table shows the regression results. The dependent variable is a dummy variable, which equals 1 if the brokerage official got promoted in the year. $Kinship_{cat}$ is a dummy variable, which equals 1 if the continuous kinship measure is bigger than or equals 0.7. Age is brokerage officials' age, and Age_{cat} is brokerage officials' age, which equals 1 if the corresponding Age of the official is between 50 and 60. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is official/brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)
$Kinship_{cat}$	0.448***	0.452***
	(0.155)	(0.155)
Age	-0.005	
	(0.009)	
Age_{cat}		0.229**
		(0.116)
Gender	-0.226	-0.272*
	(0.161)	(0.159)
Official Education	0.128	0.152
	(0.145)	(0.145)
Certified	0.095	0.053
	(0.117)	(0.115)
Brokerage Revenue	0***	0***
	(0)	(0)
Observations	2583	2583
Pseudo R^2	0.016	0.017

Table 19

Official promotion and analyst political kinship.

This table shows the difference-in-difference test results. The dependent variable is a dummy variable, which equals 1 if the brokerage official got promoted in the year. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management; and $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. Age is brokerage officials' age, and Age_{cat} is brokerage officials' age, which equals 1 if the corresponding Age of the official is between 50 and 60. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is official/brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)
Post	0.376**	0.107***	0.094***
	(0.153)	(0.036)	(0.035)
Kinship	0.250*		
	(0.140)		
Post × Kinship	**-0.462****		
	(0.202)		
$Kinship_{cat}$		0.046	0.056*
		(0.031)	(0.031)
Post × $Kinship_{cat}$		**-0.080****	**-0.087****
		(0.040)	**(0.040)**
Age		-0.007***	
		(0.001)	
Age_{cat}	-0.022		-0.022
	(0.016)		(0.016)
Gender	-0.025	-0.009	-0.023
	(0.023)	(0.023)	(0.023)
Official Education	0.008	0.009	0.010
	(0.020)	(0.020)	(0.020)
Certified	-0.026	-0.012	-0.026
	(0.016)	(0.016)	(0.016)
Brokerage Revenue	0*	0**	0*
	(0)	(0)	(0)
Observations	2583	2583	2583
R-squared	0.006	0.015	0.006

Table 20

Official Promotion and Analyst Political Kinship (with Kinship) – Logit.

This table shows the difference-in-difference test results based on logit regression. The dependent variable is a dummy variable, which equals 1 if the brokerage official got promoted in the year. Kinship is a continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management; and $Kinship_{cat}$ is a dummy variable which equals 1 when the continuous kinship measure is bigger than or equals 0.7. Age is brokerage officials' age, and Age_{cat} is brokerage officials' age, which equals 1 if the Age of the official is between 50 and 60. Post equals 1 for observations after 2015, and 0 otherwise. All control variables are defined in Table A1 of Appendix A. The granularity of the regression is official/brokerage-year. Robust standard errors are in parentheses. *, **, and ***denote significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
Post	2.391**	2.372**	0.693***	0.611**
	(0.999)	(0.983)	(0.251)	(0.247)
Kinship	1.517	1.661*		
	(0.976)	(0.954)		
Post × Kinship	**-2.792****	**-2.913****		
	(1.302)	**(1.282)**		
$Kinship_{cat}$			0.335	0.392*
			(0.237)	(0.234)
Post × $Kinship_{cat}$			**-0.531***	**-0.569****
			(0.274)	**(0.272)**
Age	-0.040***		-0.040***	
	(0.008)		(0.008)	
Age_{cat}		-0.132		-0.131
		(0.098)		(0.098)
Gender	-0.060	-0.146	-0.047	-0.133
	(0.133)	(0.131)	(0.134)	(0.131)
Official Education	0.047	0.050	0.054	0.057
	(0.124)	(0.123)	(0.124)	(0.123)
Certified	-0.064	-0.157	-0.066	-0.158
	(0.100)	(0.097)	(0.100)	(0.098)
Brokerage Revenue	0**	0*	0**	0*
	(0)	(0)	(0)	(0)
Observations	2583	2583	2583	2583
Pseudo R^2	0.015	0.006	0.015	0.006

Appendix A

Table A1
Definitions of variables.

Variable	Definition
AR_{xm}	A variable constructed from the abnormal return of following analyst recommendation and hold for x months (x is a number). It is calculated by multiplying the abnormal return with a ternary variable, which takes the value 1 if the rating is better than a threshold; 0 if the rating equals a threshold; -1 if the rating is worse than a threshold. Threshold is indicated in the bracket following AR_{xm}, which can be e.g. "neutral".
Age	Brokerage officials' age.
Age_{cat}	Brokerage officials' age, which equals 1 if the Age of the official is between 50 and 60.
Analyst Education	Dummy variable, which equals 1 if the analyst has a degree higher than bachelor and 0 otherwise.
Analyst Experience	The number of years of analysts working in security sector (not counting internship).
Book to Market	Book value of the firm's stockholder equity divided by market cap.
Brokerage Revenue	Brokerage revenue in million CNY.
Certified	Dummy variable, which equals 1 if the brokerage official holds one or more professional certificates, 0 otherwise.
Official Education	Dummy variable, which equals 1 if the brokerage official has a degree higher than bachelor and 0 otherwise.
Official Promotion	Dummy variable, which equals 1 if the brokerage official got promoted in the year.
Employment at Central SOE	Dummy variable, which equals 1 if the person is employed at central SOE.
Employment at Local SOE	Dummy variable, which equals 1 if the person is employed at local SOE.
Employment at SOE	Dummy variable, which equals 1 if the person is employed at SOE.
Firm Size	Revenue of healthcare firm in million CNY.
Gender	Brokerage official's gender.
Industry Knowledge	The average number of occurrences of medical words by an analyst in a report written by the analyst in a particular year.
Institutional Ownership	The percentage of a firm's shares owned by institutional investors.

Table A1
(*continued*)

Variable	Definition
Kinship	Continuous measure that serves as a proxy of maximum kinship between an analyst and the members of China Securities Regulatory Commission (CSRC) management.
Kinship (w)	Winsorized measurement of Kinship.
Kinship_{cat}	Dummy variable, which equals 1 when Kinship is bigger than or equals 0.7 and equals 0 if otherwise
Leverage	Book value of total liabilities divided by book value of equity.
Plagiarism	The average of maximum cosine similarity of a report with all the reports issued within seven days before its issuance written by the analyst in a particular year.
Post	Dummy variable, and it equals 1 for observations after 2015, and 0 otherwise.
Piggyback	Cumulative abnormal return for seven days before analyst report issuance.
Plagiarism	The average of the highest cosine similarity between the reports, written by an analyst within a specific year, and other reports within the seven days preceding its publication (excluding the reports written by the same authors and those written by some of the same authors in the same brokerages).
Portfolio Complexity	The number of companies the analyst covered in a specific year.
Report Length	The average number of pages in a report written by the analyst in a particular year.
Financial Knowledge	The average number of occurrences of financial technical words in a report written by the analyst in a particular year.
Market Cap.	Daily market capitalization.
Price	Firm's average daily stock price in each year.
Volatility	Standard deviation of a firm's daily return in each year.

Table A2

CSRC official turnover

This table describes the CSRC management level replacement per year. Column (1) is the total number of people in CSRC management, including president, vice president, assistant of president, and leader of discipline inspection and supervision team. Information about assistants of president are not available after 2019. Column (2) is the number of people changed from the previous year. Those who left and those who entered CSRC management are both treated as changes. (For example, in year 2021, Jianjun Wang became vice president of CSRC for the first time, and Qingmin Yan and Zhengping Zhao were no longer vice president, so the number of people who changed in the previous year is 3). Column (3) is the number of personnel who left the position due to normal reasons, e.g. retirement. Column (4) is the number of personnel who left the position in that particular year due to abnormal reasons such as criminal investigation. Column (5) is the publication time of publication of the reports. Column (6) is the total number of position appointments at the management level indicated in the footnote. Column (7) is the number of personnel at the management level whose starting dates of the position were after the publication of the report.

Year	(1) Mng.	(2) Change	(3) Normal	(4) Abnormal	(5) Pub Time	(6) Footnotes	(7) After Pub.
2007	9	NA	NA	NA	04/2008	0	0
2008	9	3	3	0	05/2009	0	0
2009	9	0	0	0	05/2010	0	0
2010	9	0	0	0	07/2011	0	0
2011	9	2	2	0	05/2012	0	0
2012	8	5	4	1	06/2013	0	0
2013	8	1	1	0	06/2014	2	0
2014	7	1	1	0	04/2015	1	0
2015	8	10	6	4	08/2016	6	1
2016	8	0	0	0	06/2017	3	0
2017	8	0	0	0	05/2018	2	0
2018	10	2	2	0	05/2019	2	0
2019	6	3	2	1	05/2020	2	0
2020	6	0	0	0	05/2021	0	0
2021	5	3	3	0	06/2022	1	0
2022	5	0	0	0	08/2023	0	0
2023	5	0	0	0	NA	NA	NA

Appendix B

B1. 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 et al. (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 et al., 2018; Liu et al., 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

---

[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 et al. (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.

[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.

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.

B2. Construction of industry knowledge dictionary

B2.1. 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 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: a-hospital.com

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 73651 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).

B2.2. 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 et al. (2013a and 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.

B3. Plagiarism measure

We define the likelihood of plagiarism as the similarity between a report and all reports

[16] The training algorithm for the neural network is stochastic gradient descent with backpropagation.
[17] See the definition of cosine similarity in Section B3 of the Appendix.

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.

$$Report\ similarity = \frac{\mathrm{Vector}_i \cdot \mathrm{Vector}_j}{|\mathrm{Vector}_i||\mathrm{Vector}_j|} \quad (5)$$