International Review of Financial Analysis 93 (2024) 103138

Contents lists available at ScienceDirect

International Review of Financial Analysis

journal homepage: www.elsevier.com/locate/irfa

Analyst collaboration networks and earnings forecast performance

Shijiao Cao [a], Chao Liang [b,*]

[a] Business School, Chengdu University of Technology, Chengdu, China
[b] School of Economics and Management, Southwest Jiaotong University, Chengdu, China

ARTICLE INFO

Keywords:
Financial analysts
Collaboration
Collaboration network
Centrality
Earnings forecasts

ABSTRACT

This study establishes extensive analyst collaboration networks based on the coauthorships of analysts' research reports in Chinese financial markets. We focus on the centrality of analysts' positions in the networks, which represents the information access at their disposal, and explore the relationship between analysts' network positions and their earnings forecast performance. We obtain robust evidence showing that analysts occupying a more central position in the collaboration networks produce more accurate earnings forecasts and that the improvements in forecast accuracy are applicable to all analysts regardless of their experience and industry specialty. We further find that collaborating with analysts who have superior ability or specialize in a given industry is more beneficial for the focal analysts to improve forecast accuracy. Moreover, the documented effects are more pronounced when earnings are more difficult to forecast. Finally, our evidence demonstrates that analysts with higher collaboration network centrality generally take a longer period of time to issue forecasts. The findings help to further our understanding of the spread of information among analysts and highlight the value of collaborations to work performance in knowledge-based industries.

1. Introduction

Financial analysts, as important information intermediaries in financial markets, have drawn extensive research on their work performance, especially the accuracy of their earnings forecasts. Having established that analysts' forecasts are more accurate than those derived from time-series models (Brown, Hagerman, Griffin, & Zmijewski, 1987; Collins & Hopwood, 1980; Fried & Givoly, 1982), researchers seek to determine whether some analysts outperform others (Bradshaw, 2011; Cheng, Du, Wang, & Wang, 2016; Clement, 1999; Cohen, Frazzini, & Malloy, 2010; Li, Lin, & Lu, 2023) and highlight that analysts' information access is pivotal for forecast accuracy (Chen, Mayew, & Yan, 2022; Gibbons, Iliev, & Kalodimos, 2021; Hwang, Liberti, & Sturgess, 2019; Mayew, Sharp, & Venkatachalam, 2013). Most of these studies, however, assume that forecasts are made by individual analysts, ignoring the prevalent collaborations between analysts whereby information is exchanged and shared.

Very recently, He, Jackson, and Li (2020), Fang and Hope (2021), and Gao, Ji, and Rozenbaum (2022) studied the earnings forecast performance of analyst teams. They find that, owing to the division of labor and the exchange of knowledge and information between analysts, analyst teams generally produce forecasts better than those by individual analysts. While acknowledging the presence of analyst collaborations, these studies implicitly presume that information is shared only between analysts who sign the research report. As argued by Phua, Tham, and Wei (2023), information exchange among analysts is more common than we thought and not confined to analysts who authored the report. By assuming that two analysts in a brokerage house share an information link if they cover at least one common sector, they construct analyst information networks and find that analysts who occupy a central position in the network issue more accurate forecasts because of access to more information. Their findings shed valuable light on the information flows of analyst networks and on how these flows affect the performance of analysts at different positions in the network. However, the way in which they identify the information links between analysts is, to some degree, rough. Covering the same industry sector does not necessarily guarantee the information exchange between two analysts. In other words, a considerable number of information links they think exist may not exist. As such, seeking to identify real information links is an important mission in evaluating the effect of analyst networks on earnings forecast performance.

In this study, we establish novel collaboration networks based on the

* Corresponding author at: Southwest Jiaotong University, No.111, North 1st Section of Second Ring Road, Jinniu District, Chengdu 610031, Sichuan Province, China.

E-mail address: liangchaoswjt@163.com (C. Liang).

https://doi.org/10.1016/j.irfa.2024.103138
Received 3 November 2023; Received in revised form 26 January 2024; Accepted 28 February 2024
Available online 29 February 2024
1057-5219/

coauthorships of analysts' research reports and explore how analysts' network positions impact their forecast performance in Chinese financial markets. China is an ideal place to conduct such research for the indispensable role of financial analysts. Prior studies (e.g., Titman, Wei, & Zhao, 2022) show that nearly 90% of daily trading volume in Chinese stock markets is completed by retail investors who are relatively unsophisticated and heavily rely on the works of sell-side analysts (Bradshaw, 2011). The predominance of retail investors in Chinese stock markets yields a great demand for the information dissemination role of financial analysts, which magnifies the significance of this study.

As mentioned, in contrast to the *presumed* information links between analysts in Phua et al. (2023), we focus on the *real* collaboration relationships between analysts. By *real* collaboration, we mean that two analysts work together to issue at least one research report in a year. In the process of writing a research report of multiple signing analysts, each analyst must engage in substantial communications with other signing analysts whom he/she collaborates with. They typically need to get together to discuss key issues, have face-to-face conversations, and communicate by phone and email about the firm they are analyzing. Their communications may continue after the project is completed and involve firms that are not part of the project. Thus, the collaboration relationship between the two analysts is an effective conduit through which useful knowledge and information are transferred. Utilizing all research reports issued between 2007 and 2022, we determine whether there is a collaborative relationship between two analysts in a year by tracing if their names appear together on at least one research report. This enables us to map all the collaboration relationships between analysts and thereby construct collaboration networks that envelop all analysts. According to our data, analyst collaborations are becoming more and more prevalent, as the percentage of analysts who have collaborators has sharply increased from 46.53% in 2007 to 88.35% in 2022. We thus believe that this nonnegligible behavioral pattern of analysts should have a significant effect on their earnings forecast performance.

To understand how information spreads on analyst collaboration networks and how the network characteristics of analysts are associated with their forecast accuracy, we implement the methods of social network analysis (SNA) to compute the network centrality of each analyst, which indicates how close an analyst is to the center of the network and the amount of information access at his/her disposal. We then run regressions to estimate the effect of the network centrality of a research report's lead analyst on forecast accuracy, with many well-known determinants of forecast accuracy being controlled for. The results are summarized as follows. First, we demonstrate that collaboration network centrality has a positive effect on earnings forecast accuracy. This means that analysts who occupy a central position in the collaboration networks have superior access to useful information that helps to improve forecast accuracy. This inference survives a bunch of robustness tests, including adjustments to the collaboration boundaries and regressions with an instrumental variable that captures regional collaborative culture intensity. More importantly, we find that the documented effect applies to all analysts irrespective of their experience and industry specialty, which translates as that collaboration networks provide undifferentiated benefits to every analyst. Although the attributes of the focal analysts are irrelevant, we find that the characteristics of the collaborators do matter. Specifically, our results show that collaborating with analysts who specialize in the industry or have a greater ability to forecast would be particularly helpful in improving an analyst's earnings forecast accuracy. This implies that information and knowledge that are more relevant and advisable for the forecasting task would be more helpful to analysts.

Moreover, we extend our exploration to examine the effects of collaboration network centrality in settings where earnings are difficult to forecast. We use several measures to proxy for forecast difficulty and find that the effects of collaboration network centrality on forecast accuracy are more pronounced when the firm experiences greater market value volatility or engages in more accrual earnings management, or when the forecast is made at an earlier time. Finally, we investigate whether and how collaboration network centrality affects forecast timeliness. The results show that network centrality exerts a negative effect on forecast timeliness, consistent with the information overload theory which stresses the time needed for processing a large amount of information that comes in a short period of time from different sources.

This study makes several important contributions. First, it establishes extensive analyst collaboration networks and adds to the literature on social networks among financial analysts. Although the SNA methods have been widely applied to research in disciplines including, for example, sociology, management, and information systems, their application in accounting research lags far behind (de Villiers, La Torre, & Botes, 2022; Worrell, Wasko, & Johnston, 2013). In the scant accounting research, social networks of directors, managers, and brokerage houses have been studied (Bajo, Chemmanur, Simonyan, & Tehranian, 2016; Chuluun, 2015; Larcker, So, & Wang, 2013; Rumokoy, Neupane, Chung, & Vithanage, 2019; Schabus, 2022), but the social networks of financial analysts, whose primary duty is to discover and disseminate information, seem to be absent from the literature. The establishment of analyst collaboration networks would make it possible to recognize the characteristics and evolutionary trends of analyst collaborations and to understand analyst behaviors and performance in the presence of collaboration networks that envelop them.

Second, this study contributes to the literature dealing with the relationship between information links and earnings forecast accuracy from an emerging perspective of second-hand information links. Earlier studies concentrate on documenting the effects of analysts' links to first-hand information sources on their forecast performance (Bradley, Gokkaya, & Liu, 2020; Cohen et al., 2010; Fang & Huang, 2017), and Li, Wong, and Yu (2020) demonstrate that analysts with first-hand information play a role in disseminating it to those without, depicting how first-hand information is turned into second-hand information. Our study aligns with an emerging stream of literature that focuses on the exchanges of second-hand information between analysts (Chen et al., 2022; Huang, Lin, & Zang, 2022; Hugon, Lin, & Markov, 2023; Phua et al., 2023) and thereby helps to further our understanding of the spread of second-hand information among analysts and the attendant consequences.

Third, this study also provides valuable insights into the effects of collaboration on individual work performance, a classic topic that long appeals to scholars and practitioners in the fields of human resource and organization management. Collaborations between workers are common, but in most cases, we can only observe the team-level or firm-level outcomes but have no idea whether the performance of individuals is improved (Hwang et al., 2019). Analysts' earnings forecasts are ideal materials for detecting the work performance of individual analysts. Our findings support the anecdotal opinion that collaborations are beneficial for work performance, particularly in the knowledge-based industry.

The rest of this paper is organized as follows. Section 2 reviews relevant studies and develops the hypothesis. Section 3 introduces the research design, including the construction of analyst collaboration networks, sample selection, and variables. Section 4 examines the hypothesis, encompassing the baseline regression results and robustness tests. Section 5 presents the results of further analyses conducive to a comprehensive understanding of the effects of analyst collaboration networks. Finally, Section 6 concludes the paper.

2. Literature review and hypothesis development

2.1. Social network analysis

One feature that human societies share with biological systems is that they are both made up of interrelated components (Durkheim, 1951). Humans are constantly interacting with their surroundings, exchanging resources, information, and ideas with other persons. From this point of view, individuals are not isolated but embedded in networks of social interactions, and acknowledging this is one of the most salient ideas of social sciences (Borgatti, Mehra, Brass, & Labianca, 2009). Similar to the composition of traffic networks, individuals, coupled with the interaction relationships among them, constitute social networks. Due to the existence of social networks, individuals' behaviors, performances, and outcomes must be different from the predictions of models without considering the interaction patterns on social networks. More importantly, these micro-level interpersonal interactions ultimately have a decisive impact on the patterns of large-scale social movements (Granovetter, 1973). To better model human behaviors, a recent trend in economics research is the endeavor to take social ties and social networks into account (Jackson, 2014). With regard to social ties, based on the time of interaction, the emotional intensity, and other characteristics, the ties between individuals can be different in strength. Individuals linked by strong ties are similar in the information they each possess, while individuals to weak ties possess and can exchange information that is novel to each other (Granovetter, 1973).

In contrast to theories that explain individuals' outcomes by their attributes such as gender and age, social network perspectives tend to regard individuals' outcomes as a function of the social relations and surroundings in which they are embedded (Borgatti et al., 2009; Worrell et al., 2013). Studying the behaviors and outcomes of individuals on social networks is usually referred to as social network analysis (SNA). A social network, consisting of nodes and links between them, sometimes can be very complicated. To see through the complex structure and study the characteristics of social networks, researchers have developed many useful measures to characterize different aspects of a social network, for example, the position of a node, the density of relations, and clustering. Among these aspects, perhaps the most important one is a node's position in the network. The position, to some degree, determines the resources, influence, and constraints of a node, thereby playing an important role in its outcomes (Borgatti et al., 2009). The concept pertaining to nodes' positions is centrality, which measures the extent to which a node is close to the center of a network. The frequently used measures of centrality are degree, betweenness, closeness, and eigenvector centrality. The first three are proposed by Freeman (1978), capturing the number of direct connections a node has, the frequency with which a node lies on the shortest path between any other two nodes, and the farness from a node to other nodes, respectively. The last one, developed by Bonacich (1987), measures how a node is connected to other well-connected nodes. It is worth noting that the focus of network research in economics today has extended from individuals to organizations such as firms that are led by human beings.

2.2. *Social network analysis in accounting and finance research*

Links or connections, in the forms of, for example, lending, share-

Table 1
Selection of research reports and analyst distribution.

Panel A: Research report selection procedures

	No. of reports	No. of analysts
All research reports for Chinese A-share firms issued between 1/1/2007 and 12/31/2022 available on the CSMAR database	1,786,252	9376
Drop:		
Duplicated research reports	(1,169,346)	
Research reports without signing analysts	(5790)	
Total unique research reports	611,116	9376
Containing:		
Research reports authored by single analyst	273,441	3936
Research reports authored by multiple analysts	337,675	7629
Total unique analyst collaboration pairs matched up		56,846

Panel B: Analyst distribution by year

Year	No. of analysts	No. of collaborative analysts	% of collaborative analysts	Year	No. of analysts	No. of collaborative analysts	% of collaborative analysts
2007	1081	503	46.53	2015	1684	1228	72.92
2008	1601	915	57.15	2016	1637	1304	79.66
2009	1745	1092	62.58	2017	1948	1625	83.42
2010	2131	1499	70.34	2018	2224	1892	85.07
2011	1660	848	51.08	2019	2401	2021	84.17
2012	1912	1125	58.84	2020	2321	1985	85.52
2013	1908	1306	68.45	2021	2303	2024	87.89
2014	1799	1312	72.93	2022	2164	1912	88.35

Panel C: Distribution of analyst collaboration pairs by year

Year	No. of analyst pairs	%	Year	No. of analyst pairs	%
2007	788	1.39	2015	2818	4.96
2008	1574	2.77	2016	3760	6.61
2009	1824	3.21	2017	4436	7.80
2010	2844	5.00	2018	4898	8.62
2011	1374	2.42	2019	5502	9.68
2012	1980	3.48	2020	6402	11.26
2013	2572	4.52	2021	6784	11.93
2014	2850	5.01	2022	6440	11.33
			Total	56,846	100.00

Note: Panel A of this table records the procedures to select the research reports for constructing the analyst collaboration networks. A research report may make earnings forecasts for multiple firms and years, and the CSMAR database segments such a report into multiple observations, resulting in the duplication of research reports. That is the reason why the number of initial observations is 1,786,252 but the number of unique research reports is only 611,116. Because an analyst may issue research reports independently in one year but collaborate with others in another year, the sum of analysts of single-authored research reports and of multi-authored reports is more than the total number of unique analysts.

holding, underwriting, and even personal social relations between managers, are very common among financial participants including firms, brokerage houses, investors, and so on (Calluzzo, 2023; Dharwadkar, Harris, Shi, & Zhou, 2020; Engelberg, Gao, & Parsons, 2012; Luong, Qiu, & Wu, 2021). This means that financial participants are embedded in social networks as well. Apart from the flows of money and resources, information exchange is constantly taking place on these connections, and this information exchange has been found to have far-reaching impacts. For example, lenders that have relationships with a firm would be willing to lend more with a lower rate (Engelberg et al., 2012; Johan & Wu, 2014; Petersen & Rajan, 1994). A firm's socially connected underwriters and brokerage houses can help alleviate IPO underpricing and decrease the cost of equity because the connections ameliorate asymmetric information problems (Luong et al., 2021; Schenone, 2004). Mutual funds that are socially connected to a firm perform significantly better due to their information advantages (Calluzzo, 2023; Cohen, Frazzini, & Malloy, 2008). Moreover, information and knowledge are also exchanged between socially connected firms, resulting in better investment decisions and similar management choices (Cai, Dhaliwal, Kim, & Pan, 2014; Cheng, Rai, Tian, & Xu, 2021; Dharwadkar et al., 2020; Geng, Xiao, Yuan, & Zhang, 2021). All of these studies point to the important role of information transfer via social connections.

Leveraging the SNA methods, researchers of accounting and finance could provide more novel and interesting insights into the effects of social networks in financial markets. Given the decisive role of a node's position in determining its opportunities, power, and constraints (Borgatti et al., 2009; Freeman, 1978; Jackson, 2014), a large proportion of studies focus on the centrality of actors (e.g., individuals, firms, brokerage houses, etc.). The social networks among firms that are linked by director interlocks or managers' personal ties have been extensively studied (El-Khatib, Fogel, & Jandik, 2015; Feng, Song, & Tian, 2019; Larcker et al., 2013; Schabus, 2022). For example, with regard to director networks of firms, Larcker et al. (2013) find that firms occupying a central position in the director network will benefit from the information and resources exchanged. Schabus (2022) provides more direct evidence on the information flows of director networks by documenting that managers' forecasts are more accurate if the directors are better connected. Social networks also exist among financial intermediaries. For example, Chuluun (2015), Bajo et al. (2016), and Rumokoy et al. (2019) demonstrate that well-connected underwriters can make use of their connections to disseminate and extract information and thereby achieve better IPO outcomes. All of these studies highlight that a central position in the network endows the actor with more information access and thereby improves their outcomes. However, in contrast to the numerous studies on social networks at the firm/institution level, studies focusing on the social network of individual participants in financial markets are rarely seen. One exception is Shiue, Chiu, Kuo, and Yeh (2022), who find that a central position of signing auditor network benefits the auditor with incremental information and thus is positively related to audit quality. So far, virtually no attention has been paid to the social networks of financial analysts who serve as important information intermediaries in financial markets.

2.3. Hypothesis development

Information is one of the major ingredients for analysts to produce earnings forecasts, as evidenced by numerous studies showing that analysts with unique information access release more accurate forecasts and better recommendations (Bradley et al., 2020; Clement, 1999; Cohen et al., 2010; Fang & Huang, 2017; Huang et al., 2022; Li et al., 2020). Information, based on its origin, can be categorized as first-hand and second-hand information, and analysts' connections are useful in acquiring both of them. Earlier studies are more interested in the role of analysts' ties to management teams in creating the channel for extracting first-hand information. For example, Cohen et al. (2010) and Fang and Huang (2017) find that analysts who share educational ties with the managers of the firm they cover can access private information and thereby outperform analysts without such ties. Bradley et al. (2020) demonstrate that the overlapped employment history between an analyst and the management team members of a firm could help to create an information channel that enables the analyst to make more accurate forecasts for the firm. Li et al. (2020) document how analysts who have ties to managers disseminate the first-hand information they acquire to other analysts and thereby improve their forecast performance. These studies corroborate the story that financial analysts can extract information from managers and then disseminate it into financial markets through earnings forecasts.

As information intermediaries, financial analysts also communicate and exchange second-hand information with colleagues. Prior studies have shown that the information exchanged through analyst interaction includes, but is not limited to, firm-specific information, region-specific news, and cross-industry knowledge (Chen et al., 2022; Huang et al., 2022; Hwang et al., 2019). All of this information is helpful for forecasting, and analysts with better social skills would benefit more from social interactions (Li et al., 2023). Therefore, it is reasonable to predict that analysts with more connections to other analysts would outperform those without sufficient connections. Although the extent to which an analyst is engaged in the interactions with others is difficult to observe, some studies do show that analyst teams, wherein analyst interactions must take place, issue forecasts that are more accurate than those issued by individual analysts (Fang & Hope, 2021; Gao et al., 2022; He et al., 2020). Keeping in mind that analyst interactions and collaborations are not confined to those who author a single research report, we extend the scope of analyst collaboration and construct analyst collaboration networks based on the coauthorships of all research reports. Such a network makes it possible for analysts to disseminate and extract second-hand information. We expect that analysts located in different positions of the collaboration networks exhibit differentiated earnings forecast performance. A central position in the second-hand information network endows analysts with more access to information and experience from others. This would be helpful for earnings forecasts for the following reasons. First, it allows analysts who are excluded from the first-hand information channels to keep informed. Second, it brings complementary information and enables analysts to verify information from different sources. Third, the knowledge and ideas exchanged through the interactions make it possible for analysts to reconsider the materiality of different information and to generate new interpretations based on their existing information (Li et al., 2020; Li et al., 2023). Besides, the time saved from information searching can be reallotted to modeling and quantitative works, albeit not a direct information effect. Taken together, our research hypothesis is stated as follows:

Hypothesis: Analysts' centrality in collaboration networks is positively associated with earnings forecast accuracy.

3. Research design

3.1. Construction of analyst collaboration network

To construct the collaboration network of analysts, we should first identify the between-analyst collaboration relationships. We consider two analysts as having a collaboration relationship in a year if they work together for at least one research report. This requires a dataset of research report authorships as complete as possible. We turn to the Analyst Forecast Indicators File of the China Stock Market and Accounting Research (CSMAR) database for such information. This dataset extracts the major information of a research report into a set of fields including, for example, a unique ID of the research report, the issuing date of the report, the ID of the firm being forecasted, the names (and corresponding IDs) of the analysts who author the report, and the forecasts of EPS, price-to-earnings ratio, earnings, and so on, of the firm.

To ensure the completeness of the analyst collaboration networks,

Table 2
Sample selection procedures.

	No. of observations
All forecasts for Chinese A-share firms issued between 1/1/2007 and 12/31/2022 available on the CSMAR database	1,786,252
Drop:	
Forecasts without analyst names	(14,880)
Forecasts for financial firms	(73,207)
Forecasts released outside the [−365, −30] window prior to the earnings announcement date	(1,191,659)
Observations with missing data	(55,244)
Observations of the final sample	451,262
Unique lead analysts covered	5853
Unique firms covered	4018

Note: This table records the procedures to construct the sample for the earnings forecast accuracy tests.

we download the data on all research reports for Chinese A-share firms issued between January 1, 2007, and December 31, 2022, which consist of 1,786,252 observations. Following the filtering procedures shown in Panel A, Table 1, we exclude duplicated reports and reports without analyst names. Our final dataset encompasses 611,116 research reports, covering 9376 unique analysts. Among these research reports, 273,441 are authored by single analysts, while 337,675 by more than one analyst. In Panel B, Table 1, we present the distribution of all analysts and collaborative analysts (i.e., analysts who have collaborators) by year. It shows that 503 out of 1081 analysts chose to collaborate with others in 2007, taking a proportion of 46.53%. By contrast, in 2022, 1912 out of 2164 analysts had collaborators, raising the proportion to 88.35%. This demonstrates a marked increasing trend of analyst collaborations.

Using each multi-authored research report, we trace the collaboration relationships between the analysts and form analyst pairs in the following way. If a research report is authored by analysts A and B, we determine that A and B share a collaboration relationship. By the same token, a research report signed by analysts B, C, and D indicates the collaboration relationships between analyst pairs of B—C, C—D, and B—D. By reviewing all the research reports of multiple authors, we map all the collaboration relationships between analysts. We end up with 56,846 unique analyst collaboration pairs and present their distribution by year in Panel C of Table 1. The number of identified analyst collaboration pairs was 788 in 2007, which markedly increased to 6440 in 2022.

Our analyst collaboration networks are constructed by year. For each year, we establish an n by n adjacency matrix A, where each row i and column j represent an active analyst (i.e., an analyst who issues at least one research report) in that year. We let cell $a_{i,j}$ take the value of one if analysts i and j have a collaboration relationship in that year and 0 otherwise. Since the collaboration relationships are undirected in this study, matrix A is symmetric, with the diagonal always equal to zero. The matrix is then input into UCINET to compute the centrality of each analyst in the analyst collaboration networks.

3.2. Models and variables

We use the following model to test our hypothesis:

$$\begin{aligned}AFE_{k,i,j,t} &= c + \alpha Centrality_{i,t} + \beta AnalystControls_{i,t} + \chi ReportControls_{i,k} \\ &+ \delta BrokerControls_{i,t} + \varphi FirmControls_{j,t-1} + \gamma_t + \eta_j + \lambda_i + \varepsilon_{k,i,j,t}\end{aligned} \quad (1)$$

In Eq. (1), subscript k denotes the research report, i the lead analyst of the report, j the firm that the report is about, and t the year in which the report is issued. Detailed definitions and measurements of all variables are shown in the Appendix. The dependent variable, $AFE_{k,i,j,t}$, is the absolute earnings forecast error of research report k issued by i as the lead analyst of the report for firm j in year t. Similar to He et al. (2020) and Chen et al. (2022), absolute earnings forecast error is calculated as the absolute difference between the actual earnings per share (EPS) and the EPS forecast, scaled by the stock price of the firm at the end of year t-

Table 3
Summary statistics and univariate analysis results.

Panel A: Summary statistics for main variables

Variables	N	Mean	SD	P1	P25	Median	P75	P99
AFE	451,262	1.394	2.456	0.006	0.194	0.567	1.495	16.727
Centrality	451,262	0.160	0.176	0.000	0.046	0.105	0.217	0.904
Closeness	451,262	0.046	0.026	0.000	0.043	0.052	0.061	0.086
Betweenness	451,262	0.092	0.363	0.000	0.000	0.000	0.006	2.688
Gen_Exp	451,262	271.749	298.176	3.000	64.000	167.000	376.000	1387.000
Firm_Exp	451,262	7.612	9.453	1.000	2.000	4.000	10.000	46.000
lnFirmNum	451,262	31.956	25.583	3.000	15.000	25.000	41.000	135.000
lnReportNum	451,262	88.630	73.533	4.000	37.000	70.000	117.000	347.000
Horizon	451,262	223.668	92.367	36.000	156.000	225.000	305.000	365.000
lnSignAnaNum	451,262	1.584	0.732	1.000	1.000	1.000	2.000	4.000
BrokerAge	451,262	20.033	7.192	4.000	15.000	21.000	25.000	34.000
ListedBroker	451,262	0.609	0.488	0.000	0.000	1.000	1.000	1.000
lnActAnalysts	451,262	53.075	27.746	6.000	33.000	51.000	69.000	142.000
lnFirmAnaNum	451,262	28.969	17.458	2.000	15.000	26.000	40.000	77.000
lnAsset	451,262	23.029	1.499	20.422	21.936	22.798	23.859	27.464
Leverage	451,262	0.436	0.191	0.069	0.285	0.437	0.582	0.844
Growth	451,262	0.250	0.334	−0.349	0.065	0.187	0.352	1.926
ROA	451,262	0.079	0.059	−0.063	0.039	0.070	0.110	0.277
BM	451,262	0.496	0.273	0.083	0.276	0.443	0.682	1.193
Big4	451,262	0.143	0.350	0.000	0.000	0.000	0.000	1.000
SOE	451,262	0.383	0.486	0.000	0.000	0.000	1.000	1.000
Loss	451,262	0.028	0.166	0.000	0.000	0.000	0.000	1.000

Panel B: T-test for *AFE* between low- and high-centrality analysts

	N	*AFE*	mean-diff.	*p*-value
AFE of low-centrality analysts	219,843	1.416	0.042	< 0.000
AFE of high-centrality analysts	231,419	1.374		

Note: To facilitate interpretation, variables defined as natural logarithms are not logarithmized in this table except for *lnAsset*.

Table 4
Correlation coefficient matrix.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. *AFE*		**−0.030**	**−0.022**	**−0.027**	**−0.025**	**0.014**	**−0.058**	**−0.055**	**0.290**	−0.002	**−0.024**	**−0.006**	**0.023**	**−0.119**	**0.234**	**0.279**	**−0.154**	**−0.268**	**0.382**	**0.102**	**0.149**	**0.206**
2. *Centrality*	**−0.020**		**0.672**	**0.857**	**0.488**	**0.201**	**0.502**	**0.520**	**0.011**	**0.493**	**0.218**	**0.167**	**0.540**	**0.028**	**0.047**	**−0.016**	**0.019**	**−0.004**	**−0.063**	**−0.019**	**−0.102**	**0.016**
3. *Closeness*	−0.004	**0.541**		**0.550**	**0.244**	**0.127**	**0.269**	**0.291**	**0.006**	**0.340**	**0.107**	**0.097**	**0.375**	**−0.063**	0.003	0.002	**0.027**	**−0.032**	**−0.062**	**−0.020**	**−0.052**	**0.016**
4. *Betweenness*	**−0.010**	**0.393**	**0.281**		**0.465**	**0.169**	**0.486**	**0.487**	**0.011**	**0.331**	**0.242**	**0.160**	**0.464**	**0.057**	**0.058**	**−0.021**	**0.026**	**0.006**	**−0.066**	**−0.018**	**−0.116**	**0.018**
5. *Gen_Exp*	**0.011**	**0.434**	**0.327**	**0.184**		**0.487**	**0.625**	**0.714**	**0.014**	**0.251**	**0.274**	**0.233**	**0.337**	**0.080**	**0.115**	**−0.014**	**−0.027**	**0.008**	**−0.012**	**0.005**	**−0.108**	**0.023**
6. *Firm_Exp*	**0.040**	**0.149**	**0.179**	**0.058**	**0.506**		**0.117**	**0.304**	**−0.005**	**0.182**	**0.158**	**0.137**	**0.291**	**0.252**	**0.259**	**0.055**	**−0.051**	**0.051**	**0.041**	**0.112**	**0.013**	**0.004**
7. *lnFirmNum*	**−0.033**	**0.498**	**0.316**	**0.212**	**0.640**	**0.122**		**0.883**	**0.039**	**0.136**	**0.158**	**0.153**	**0.155**	**−0.020**	**−0.020**	**−0.056**	**0.028**	**0.007**	**−0.072**	**−0.057**	**−0.112**	**0.011**
8. *lnReportNum*	**−0.026**	**0.485**	**0.355**	**0.197**	**0.727**	**0.314**	**0.896**		**0.034**	**0.187**	**0.219**	**0.196**	**0.272**	**0.040**	**0.035**	**−0.042**	**0.005**	**0.015**	**−0.056**	**−0.024**	**−0.115**	**0.009**
9. *Horizon*	**0.142**	0.001	0.002	0.000	**0.015**	**−0.009**	**0.038**	**0.033**		**−0.017**	**0.021**	**0.014**	**0.023**	**0.007**	**0.039**	0.003	**−0.049**	**−0.015**	**0.044**	**0.011**	**0.010**	**0.018**
10. *lnSignAnaNum*	**0.015**	**0.394**	**0.383**	**0.088**	**0.238**	**0.175**	**0.135**	**0.185**	**−0.016**		**0.148**	**0.113**	**0.439**	**0.060**	**0.063**	**0.012**	**−0.022**	**−0.007**	−0.002	0.002	**−0.037**	**0.007**
11. *BrokerAge*	**0.014**	**0.182**	**0.165**	**0.101**	**0.273**	**0.157**	**0.160**	**0.223**	**0.009**	**0.137**		**0.435**	**0.287**	**0.122**	**0.125**	**−0.033**	**−0.027**	**0.019**	**−0.033**	**0.014**	**−0.115**	**0.027**
12. *ListedBroker*	**0.017**	**0.147**	**0.139**	**0.054**	**0.229**	**0.141**	**0.156**	**0.196**	**0.008**	**0.110**	**0.411**		**0.278**	**0.039**	**0.072**	**−0.015**	**−0.034**	**−0.013**	−0.000	0.002	**−0.070**	**0.025**
13. *lnActAnalysts*	**0.042**	**0.444**	**0.475**	**0.180**	**0.342**	**0.301**	**0.152**	**0.292**	**0.004**	**0.403**	**0.248**	**0.293**		**0.079**	**0.133**	**0.026**	**−0.042**	**−0.004**	**0.046**	**0.030**	**−0.033**	**0.013**
14. *lnFirmAnaNum*	**−0.066**	**0.015**	**0.013**	**0.033**	**0.074**	**0.249**	**−0.017**	**0.045**	**0.006**	**0.065**	**0.114**	**0.034**	**0.070**		**0.373**	**0.016**	**0.113**	**0.342**	**−0.191**	**0.191**	**0.018**	**−0.102**
15. *lnAsset*	**0.235**	**0.014**	**0.044**	**0.017**	**0.104**	**0.278**	**−0.032**	**0.028**	**0.037**	**0.056**	**0.111**	**0.067**	**0.121**	**0.348**		**0.538**	**−0.076**	**−0.204**	**0.509**	**0.413**	**0.376**	**0.019**
16. *Leverage*	**0.229**	**−0.021**	0.000	**−0.019**	**−0.014**	**0.064**	**−0.054**	**−0.040**	0.004	**0.017**	**−0.043**	**−0.015**	**0.033**	**0.009**	**0.540**		**0.022**	**−0.508**	**0.536**	**0.199**	**0.252**	**0.102**
17. *Growth*	**−0.060**	**0.034**	**0.035**	**0.034**	**−0.022**	**−0.063**	**0.028**	**0.005**	**−0.040**	**−0.011**	**−0.019**	**−0.024**	**−0.031**	**0.056**	**−0.049**	**0.049**		**0.270**	**−0.219**	**−0.082**	**−0.143**	**−0.163**
18. *ROA*	**−0.151**	**−0.014**	**−0.010**	**0.014**	**0.008**	**0.048**	**0.004**	**0.013**	**−0.011**	**−0.009**	**0.023**	**−0.012**	**−0.008**	**0.328**	**−0.168**	**−0.472**	**0.213**		**−0.578**	**−0.080**	**−0.165**	**−0.288**
19. *BM*	**0.280**	**−0.069**	**−0.048**	**−0.057**	0.004	**0.068**	**−0.059**	**−0.038**	**0.046**	**0.009**	**−0.022**	**0.008**	**0.061**	**−0.158**	**0.571**	**0.535**	**−0.159**	**−0.525**		**0.210**	**0.335**	**0.071**
20. *Big4*	**0.095**	**−0.028**	**−0.006**	**−0.010**	0.003	**0.125**	**−0.057**	**−0.024**	**0.015**	−0.000	**0.009**	0.002	**0.031**	**0.179**	**0.473**	**0.199**	**−0.070**	**−0.057**	**0.235**		**0.190**	**0.012**
21. *SOE*	**0.101**	**−0.105**	**−0.065**	**−0.061**	**−0.106**	**0.019**	**−0.108**	**−0.112**	**0.017**	**−0.035**	**−0.128**	**−0.070**	**−0.025**	**0.014**	**0.386**	**0.256**	**−0.122**	**−0.129**	**0.359**	**0.190**		**0.012**
22. *Loss*	**0.322**	**0.019**	**0.014**	0.004	**0.023**	**0.008**	**0.012**	**0.011**	**0.013**	**0.006**	**0.025**	**0.025**	**0.012**	**−0.117**	**0.017**	**0.108**	**−0.133**	**−0.345**	**0.071**	**0.012**	**0.012**	

Note: The lower triangle shows the Pearson correlation coefficients while the upper shows the Spearman's. The coefficients in bold are statistically significant at the 1% level.

1, multiplied by 100.

Centrality$_{i,t}$ is the centrality of analyst *i*' position in the analyst collaboration networks of year *t*. As mentioned, prior studies have developed several centrality measures, for example, degree, closeness, and betweenness, to characterize a node's position in different dimensions. In brief, degree centrality is based on the number of direct connections that a node has with others in the network. Closeness centrality reflects the farness on average from a node to other nodes, and it thus depends on not only the direct but also the indirect connections that the node has with others in the network. Betweenness centrality measures the frequency with which a node connects two nodes that are originally unconnected. In this study, we are more interested in degree centrality instead of closeness and betweenness centrality. This is because information spreads by replication rather than transference and thus can go through multiple possible paths concurrently, which is not in line with the implicit assumption of closeness and betweenness centrality that anything flowing through the network follows the shortest possible paths only (Borgatti, 2005; Larcker et al., 2013). Considering that the lead analyst of a research report is committed to discovering information and contributes most to the forecast (Fang & Hope, 2021; Gao et al., 2022), we use the degree centrality of a research report's lead analyst for analyses.

The control variables are largely overlapped with those used in He et al. (2020) and Phua et al. (2023), which can be classified into four categories. *AnalystControls*$_{i,t}$ denotes a group of variables that capture the attributes of the lead analyst of a research report, including his/her general experience (*Gen_Exp*), firm-specific experience (*Firm_Exp*), the number of firms he/she follows in that year (*lnFirmNum*), as a proxy for the portfolio complexity, and the number of reports he/she issues (*lnReportNum*) in that year, a proxy for the busyness of the analyst. The

Table 5
Analysts' collaboration network centrality and earnings forecast accuracy.

Variables	*AFE*					
	(1)	(2)	(3)	(4)	(5)	(6)
Centrality	−0.179***	−0.148**	−0.142***	−0.185***		
	(−2.663)	(−2.212)	(−3.628)	(−3.153)		
Closeness					0.409	
					(1.304)	
Betweenness						−0.026
						(−1.598)
Gen_Exp	0.033***	0.017	0.002	−0.071***	−0.070***	−0.070***
	(3.011)	(1.443)	(0.292)	(−5.856)	(−5.752)	(−5.775)
Firm_Exp	−0.010	−0.012	0.007	0.013*	0.013*	0.013*
	(−0.827)	(−0.925)	(0.950)	(1.716)	(1.722)	(1.738)
lnFirmNum	0.019	0.044	0.024	−0.004	−0.027	−0.021
	(0.562)	(1.248)	(1.065)	(−0.113)	(−0.692)	(−0.538)
lnReportNum	−0.130***	−0.102***	−0.027	0.011	0.013	0.012
	(−4.401)	(−3.446)	(−1.449)	(0.366)	(0.463)	(0.420)
Horizon	0.587***	0.591***	0.637***	0.640***	0.640***	0.640***
	(55.243)	(56.018)	(60.616)	(60.418)	(60.451)	(60.451)
lnSignAnaNum	0.036***	0.020	0.006	0.008	0.002	0.003
	(2.606)	(1.479)	(0.727)	(0.869)	(0.219)	(0.354)
BrokerAge	0.022	−0.011	−0.012	0.045	0.040	0.041
	(0.900)	(−0.417)	(−0.852)	(0.991)	(0.884)	(0.898)
ListedBroker	−0.002	0.001	−0.018	−0.049*	−0.046*	−0.048*
	(−0.084)	(0.053)	(−1.576)	(−1.874)	(−1.741)	(−1.821)
lnActAnalysts	0.086***	0.053***	0.037***	0.060***	0.031	0.040*
	(4.280)	(2.663)	(3.749)	(2.690)	(1.444)	(1.890)
lnFirmAnaNum	−0.326***	−0.298***	−0.227***	−0.214***	−0.215***	−0.215***
	(−23.988)	(−22.080)	(−14.741)	(−14.299)	(−14.335)	(−14.340)
lnAsset	0.200***	0.160***	0.023	0.060**	0.060**	0.060**
	(19.460)	(14.989)	(0.893)	(2.371)	(2.368)	(2.362)
Leverage	1.367***	1.310***	1.085***	1.058***	1.058***	1.058***
	(19.313)	(18.490)	(16.638)	(15.855)	(15.851)	(15.868)
Growth	−0.084***	0.004	0.002	0.002	0.002	0.002
	(−2.891)	(0.138)	(0.118)	(0.121)	(0.097)	(0.106)
ROA	7.454***	7.787***	4.781***	4.191***	4.197***	4.196***
	(23.605)	(24.420)	(17.481)	(15.667)	(15.666)	(15.665)
BM	1.905***	2.239***	1.077***	0.815***	0.817***	0.816***
	(35.557)	(35.646)	(17.371)	(13.459)	(13.510)	(13.496)
Big4	−0.062	−0.051	−0.428***	−0.383***	−0.383***	−0.383***
	(−1.635)	(−1.355)	(−12.296)	(−10.669)	(−10.678)	(−10.680)
SOE	−0.167***	−0.157***	0.191***	0.199***	0.198***	0.198***
	(−8.268)	(−7.810)	(3.971)	(3.912)	(3.903)	(3.903)
Loss	5.049***	5.098***	4.747***	4.726***	4.725***	4.725***
	(47.817)	(48.797)	(48.724)	(47.658)	(47.606)	(47.606)
_cons	−7.560***	−6.816***	−3.392***	−4.076***	−3.936***	−3.962***
	(−29.633)	(−26.011)	(−6.007)	(−6.800)	(−6.616)	(−6.656)
Year fixed effects		Yes	Yes	Yes	Yes	Yes
Firm fixed effects			Yes	Yes	Yes	Yes
Analyst fixed effects				Yes	Yes	Yes
N	451,262	451,262	451,059	450,412	450,412	450,412
Adjusted R^2	0.218	0.229	0.582	0.594	0.594	0.594

Note: This table reports the regression results of earnings forecast accuracy on analysts' collaboration network centrality. *AFE* is the absolute forecast error of an earnings forecast. *Centrality*, *Closeness*, and *Betweenness* are the degree, closeness, and betweenness centrality of a research report's lead analyst in the collaboration networks, respectively. Other variables are defined in the Appendix. T-statistics (in parentheses) are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

second category is to describe the nature of research reports, including two variables. *Horizon* is the time span between the issuing date of a research report and the actual earnings announcement date, and *lnSignAnaNum* is the number of analysts who author the report. Considering that the report-level analyst team size is positively related to forecast accuracy (Fang & Hope, 2021) and may overlap with analyst collaboration, this variable is indispensable and conducive to better estimation of the effect of analyst collaboration centrality.

The third category, $BrokerControls_{i,t}$ pertains to the attributes of the brokerage house that the lead analyst serves, including the number of years since the establishment of the brokerage house (*BrokerAge*), the number of active analysts of the brokerage house in that year (*lnActAnalysts*), usually used as a proxy for the resources of the brokerage house, and a dummy variable to indicate whether the brokerage house is a listed company (*ListedBroker*). The fourth category, $FirmControls_{j,t-1}$, includes some firm-level control variables such as *lnFirmAnaNum* (i.e., the number of analysts who cover the firm in that year), *lnAsset*, *Leverage*, *Growth*, *ROA*, *BM*, *Big4*, *SOE*, and *Loss*.

We also control for the year, firm, and analyst fixed effects. Controlling for analyst fixed effects is particularly important for this study because they absorb the impacts of analyst-specific factors that may be responsible for both the choice to join the collaboration network and the ability to issue accurate earnings forecasts. Since the dependent variable, *AFE*, inversely captures earnings forecast accuracy, we expect the coefficient on *Centrality* to be significantly negative.

3.3. Sample selection

We choose 2007 as the beginning year of our sample because all Chinese A-share firms switched to new accounting standards in that year. Since our interest is in the effects of analysts' collaboration network centrality on the accuracy of their earnings forecasts, we take each earnings forecast as the basic observation unit. We download all earnings forecasts for Chinese A-share firms released between January 1, 2007, and December 31, 2022, from the CSMAR database. As mentioned in Section 3.1, the initial sample is comprised of 1,786,252 forecasts from 616,906 unique research reports (this is because a research report may contain earnings forecasts for multiple firms and years). We drop forecasts without analyst names and forecasts for financial firms. Considering that forecasts made one year ago are stale while those released close to the earnings announcement date may be merely results of mimicking the forecasts of other analysts (Clement, 1999), we follow Phua et al. (2023) to exclude forecasts released earlier than 365 days or later than 30 days prior to the earnings announcement date. After deleting observations with missing data for calculating the variables used in the tests, our final sample consists of 451,262 observations. Table 2 traces the sample selection procedures. Besides the earnings forecast data, other data used in this study are also downloaded from the CSMAR database. All continuous variables are winsorized at the 1% and 99% levels to mitigate the impacts of outliers.

Table 7
Regression results using alternative measures for analyst collaboration network centrality.

Variables	(1)	(2)	(3)	(4)
	AFE			
	Coeff.	t-stat.	Coeff.	t-stat.
Centrality_MAX	−0.142***	−2.961		
Centrality_within			−0.004***	−3.392
_cons	−4.051***	−6.770	−3.971***	−6.681
Control variables	Yes		Yes	
Year FEs	Yes		Yes	
Firm FEs	Yes		Yes	
Analyst FEs	Yes		Yes	
N	450,412		450,354	
Adjusted R^2	0.594		0.594	

Note: This table reports the regression results using alternative measures for analyst collaboration network centrality as the main independent variables. Variables are defined in the Appendix. T-statistics are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

3.4. Descriptive statistics

We tabulate the summary statistics for the main variables in Panel A, Table 3. The mean value of *AFE* is 1.394, suggesting that the difference between a forecast EPS and the actual EPS is approximately 1.394% of the stock price. The standard deviation is 2.456. These two statistics indicate that there is a considerably large variance in the accuracy across

Table 8
Regression results of different sample compositions.

Variables	(1)	(2)	(3)	(4)
	AFE			
	Extended sample		Subsample of collaborative analysts	
	Coeff.	t-stat.	Coeff.	t-stat.
Centrality	−0.389***	−2.862	−0.210***	−3.188
_cons	−6.487***	−7.236	−3.397***	−4.941
Control variables	Yes		Yes	
Year FEs	Yes		Yes	
Firm FEs	Yes		Yes	
Analyst FEs	Yes		Yes	
N	1,506,801		360,645	
Adjusted R^2	0.532		0.604	

Note: This table reports the regression results of different sample compositions. Variables are defined in the Appendix. T-statistics are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Table 6
Regression results using alternative measures for analyst forecast performance.

Variables	(1)	(2)	(3)	(4)	(5)	(6)
	Norm_EPS_error		*Norm_Earnings_error*		*Norm_PE_error*	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
Centrality	−0.042**	−2.356	−0.034*	−1.897	−0.047***	−3.099
_cons	−0.615***	−5.913	−0.876***	−6.890	0.535***	5.314
Control variables	Yes		Yes		Yes	
Year FEs	Yes		Yes		Yes	
Firm FEs	Yes		Yes		Yes	
Analyst FEs	Yes		Yes		Yes	
N	450,412		409,810		415,646	
Adjusted R^2	0.137		0.119		0.036	

Note: This table reports the regression results using alternative measures for analyst forecast performance as the dependent variables. Variables are defined in the Appendix. T-statistics are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

forecasts. *Centrality* averages 0.160, meaning that the collaboration relationships that an average analyst has are only 0.16% of all possible relationships that he/she can establish in the whole network. The mean values of *Gen_Exp* and *Firm_Exp* mean that an analyst has issued, on average, 271.749 ordinary reports and 7.612 reports for a specific firm till the day when the forecast being studied is released. The mean values of *lnFirmNum* and *lnReportNum* indicate that an analyst averagely covers 31.956 firms and issues 88.630 reports per year. The descriptive statistics for *lnSignAnaNum* show that more than half of research reports are authored by single analysts. Besides, the mean values of *lnActAnalysts* and *lnFirmAnaNum* suggest that on average a brokerage house has 53.075 active analysts and that a firm is followed by 28.969 analysts. The summary statistics also indicate that the values of our variables are within reasonable ranges. Overall, the summary statistics suggest that the sample and variables are valid.

In Panel B of Table 3, we report the *t*-test results for earnings forecast accuracy between low- and high-centrality analysts. The results show that earnings forecasts made by high-centrality analysts are more accurate. Although the results do not account for other determinants of earnings forecast accuracy, they serve as a piece of preliminary evidence that analysts occupying a central position in the collaboration networks can produce better earnings forecasts.

4. Results

4.1. Correlation analysis

We present the pairwise correlation coefficients between variables in Table 4. The results show that all three centrality measures are negatively correlated with *AFE*, meaning that a central position in the collaboration networks is likely to enable analysts to issue earnings forecasts that are more accurate. We note again that this inference is drawn without taking other factors into account. However, the results corroborate the univariant analysis results above and support the idea that collaboration relationships are useful for forecasting earnings. Moreover, we do not find any of the control variables to be highly correlated with centrality. This would help to ease the multicollinearity concern. Another observation we wish to share is that the centrality variables are not highly correlated with *lnSignAnaNum* (i.e., the number of analysts of a research report). This is desirable because it clarifies that what the measures for collaboration network centrality capture is different from the within-report collaborations between analysts. We emphasize again that the collaboration network centrality provides novel information on analyst collaborations that break the boundaries of single research reports.

4.2. Baseline regression results

The results of estimating eq. (1) are reported in Table 5. Columns (1) to (4) use degree centrality (*Centrality*) as the main explanatory variable. The results show that the coefficients on *Centrality* are statistically negative across all models with fixed effects specified at different dimensions, suggesting that earnings forecasts made by analysts at central positions in the collaboration networks are more accurate than those by analysts at peripheral positions. Taking the results in column (4) to illustrate, in economic terms, a one-standard-deviation increase in analysts' degree centrality would cause the absolute forecast error to decrease by 0.03256, equivalent to 5.74% of the sample median. The results support our hypothesis that collaboration relationships are useful to extract valuable information and that analysts occupying positions more central in the analyst collaboration network can produce more accurate earnings forecasts.

As discussed in Section 3.2, we do not expect the closeness centrality and betweenness centrality of analysts to have a significant impact on their earnings forecast accuracy. Notwithstanding, we still present the regression results using them as the explanatory variables in columns (5) and (6) of Table 5, respectively. The coefficients on *Closeness* and *Betweenness* are not statistically significant, suggesting their subtle effects on earnings forecast accuracy. This is not surprising because the spread of information and knowledge is unlike the transference of packages that are optimized to move along the shortest possible paths. Information can flow through any possible paths, which does not agree with the shortest-path assumption of closeness and betweenness centrality. Collectively, the results imply that the direct collaboration relationships an analyst has with other analysts are of paramount importance. Admittedly, indirect collaboration relationships are also useful, but the shortest paths to information sources are relatively less important for analysts because longer paths potentially serve a similar role. As Borgatti (2005) argues, the probability of receiving something that can be parallelly duplicated is entirely a function of the number of ties that a given node has, and thus degree centrality is appropriate for all parallel duplication flow processes.

Table 9
2SLS regression results with an instrumental variable.

Variables	(1)	(2)	(3)	(4)
	First stage		Second stage	
	Centrality		*AFE*	
	Coeff.	t-stat.	Coeff.	t-stat.
Rice_Areas	0.036***	4.370		
Centrality			−1.242**	−2.034
_cons	−0.723***	−11.535	−4.196***	−5.845
Control variables	Yes		Yes	
Year FEs	Yes		Yes	
Firm FEs	Yes		Yes	
Analyst FEs	Yes		Yes	
N	450,412		450,412	
Kleibergen-Paap rk LM statistic	18.247			
Kleibergen-Paap rk Wald F statistic	19.093			

Note: This table reports the 2SLS regression results using the percentage of rice cultivation areas in areas for all crops of the province where the brokerage house is headquartered as the instrumental variable (*Rice_Areas*) for *Centrality*. Variables are defined in the Appendix. T-statistics are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Table 10
Heterogenous effects across analysts.

Variable	(1)	(2)	(3)	(4)
	AFE			
	Analyst's experience		Is the analyst specialized in this industry	
	High	Low	Yes	No
Centrality	−0.212***	−0.182***	−0.167**	−0.201***
	(−2.821)	(−2.765)	(−2.176)	(−2.667)
_cons	−3.969***	−4.523***	−3.830***	−4.440***
	(−5.250)	(−6.638)	(−4.708)	(−6.026)
Control variables	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes	Yes
N	257,947	191,853	269,940	179,594
Adjusted R^2	0.622	0.569	0.556	0.668

Note: This table reports the regression results of examining the heterogenous effects of analyst collaboration network centrality on earnings forecast accuracy across analysts. Variables are defined in the Appendix. T-statistics (in parentheses) are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

4.3. Robustness checks

4.3.1. Using alternative measures for earnings forecast accuracy

We use three alternative measures for earnings forecast accuracy to conduct a set of robustness checks. The first one is borrowed from Clement (1999) and Phua et al. (2023), which is the absolute value of a forecast EPS minus the actual EPS, deflated by the average firm-year absolute forecast error, labeled *Norm_EPS_error*. The other two measures are calculated on forecast earnings, instead of forecast EPS, and forecast PE ratios, with an analogous algorithm as computing *Norm_EPS_error*, denoted *Norm_Earnings_error* and *Norm_PE_error*, respectively. According to the regression results reported in Table 6, the coefficients on *Centrality* are all negatively significant, indicating that our findings are robust to different measures of analyst forecast accuracy.

4.3.2. Using alternative explanatory variables

Next, we compute two new measures for collaboration network centrality to replace the original one. In the baseline regressions, we use the collaboration network centrality of the lead analyst of a research report as the explanatory variable. Considering that the associate analysts of a research report can also play a role in collecting information, we hence use the maximum value of the centrality of all analysts who author a research report as the explanatory variable, which is labeled *Centrality_MAX*. As shown in columns (1) and (2) of Table 7, the coefficient on *Centrality_MAX* is also significantly negative.

Moreover, we consider adjusting the boundary of analyst collaborations. When constructing the original collaboration network, we allow an analyst to establish collaboration relationships with analysts outside the same brokerage house. However, the fact is that, if an analyst decides to collaborate with others, more likely he/she will choose collaborators from the same brokerage house. In other words, brokerage houses are generally the tacit boundaries of analyst collaborations. Given that, we re-construct analyst collaboration networks within each brokerage house and re-compute the collaboration network centrality of each analyst, which is denoted *Centrality_within*. The results of using this explanatory variable are tabulated in columns (3) and (4) of Table 7, and the interpretation stays unchanged.

4.3.3. Changing the sample compositions

We also check the robustness of our results by restructuring the sample. First, we extend our sample by retrieving the forecasts released outside the [−365, −30] window prior to the earnings announcement date. The results shown in columns (1) and (2) of Table 8 continue to uphold our hypothesis. While the added forecasts might be, as mentioned, stale or merely results of mimicking other forecasts, we believe that the regression results provide further support for our claim.

Second, we re-estimate Eq. (1) by excluding forecasts issued by analysts who do not collaborate with anyone. We do this because the decision to join the collaboration network may be a problem of self-selection, which if true, implies that there may be some inherent differences between analysts who collaborate with others and those who do not. Therefore, if we concentrate on a subsample of forecasts by analysts who join the collaboration networks only, we can estimate the effects of collaboration network centrality among analysts who collaborate. The results reported in columns (3) and (4) of Table 8 are qualitatively similar to the baseline regression results, thereby verifying the benefits of analyst collaborations.

4.3.4. Discussion on the endogenous concern and the IV regression results

One endogenous concern this study suffers is that the decision to join the collaboration networks and the forecast performance of analysts may be a reverse-causality problem or both related to the attributes of analysts that are omitted in our model. In the baseline regressions, we have included analyst fixed effects and analysts' general and firm-specific experience to control for the analyst-level time-constant and time-variant factors. Besides, in the above robustness tests, we focus only on a subsample that excludes forecasts made by analysts who are not in the collaboration networks. The results show that collaboration network centrality can still improve the forecast accuracy of analysts who choose to collaborate with others. These treatments, to some degree, help to mitigate the endogenous concern.

Here, we take a further approach to addressing the endogenous concern with an instrumental variable. We hold that analysts' willingness to collaborate with others would be shaped by the circumstances, for example, the culture, in which they are working. For example, Hugon et al. (2023) show that brokerage houses with a collaborative culture do exhibit more intrafirm collaborations, highlighting the role of culture in shaping analyst behaviors. As such, we argue that brokerage houses headquartered in regions of intensive collaborative culture would be more likely to advocate analyst collaborations and create convenient conditions. Accordingly, the analysts employed by such brokerage houses would have more willingness and opportunities to collaborate with colleagues. More importantly, if collaborative culture has an impact on analyst forecast accuracy, the only channel through which the effect is exerted is surely to facilitate collaboration. We thus believe that a variable that captures the regional collaborative culture intensity is an ideal instrument. Prior studies suggest that rice cultivation areas are a good proxy for collaborative culture. As Talhelm et al. (2014) document, paddy rice, compared with other crops like wheat, requires intensive cooperation between farmers to, for example, build irrigation systems, coordinate water use, and harvest, thereby making the society more interdependent and collectivistic. They also argue that these agricultural legacies have a persistent influence in modern societies. Ensuing studies, like Fan, Gu, and Yu (2022), demonstrate that rice paddles are a valid proxy for collectivist culture in microeconomic research. Therefore, we use the percentage of land for rice cultivation in all cropland of a province (autonomous region, or municipality, hereafter) where the brokerage house is headquartered as the instrumental variable. To exclude the interferences of short-term non-agricultural factors, it is advisable to use historical data. We thus take the data from the China Statistical Yearbook of 1998, the year after China established the current administrative divisions. The instrumental variable is denoted *Rice_Areas*. We present the 2SLS regression results in Table 9. The first-stage results show that the coefficient on *Rice_Areas* is significantly positive, which indicates that the regional collaborative culture intensity does have a positive impact on analyst collaborations. This is consistent with

Table 11
Impacts of the characteristics of collaborators.

Variables	(1)	(2)	(3)	(4)
	AFE			
Network_Ind_Specialists	−0.039***			
	(−2.637)			
L_Network_AFE_mean		0.557**		
		(2.100)		
L_Network_AFE_median			2.655***	
			(4.095)	
L_Network_AFE_min				3.008***
				(3.345)
_cons	−3.299***	−3.440***	−3.562***	−3.607***
	(−4.813)	(−4.770)	(−4.979)	(−5.063)
Control variables	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes	Yes
N	360,645	319,499	319,499	319,499
Adjusted R^2	0.604	0.610	0.611	0.611

Note: This table reports the regression results of examining the impacts of the characteristics of an analyst's collaborators on his/her forecast accuracy. Variables are defined in the Appendix. T-statistics (in parentheses) are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

Table 12
Moderating effects of forecast difficulty.

Variables	(1)	(2)	(3)
	AFE		
Centrality	−0.063	−0.131*	0.309
	(−0.679)	(−1.919)	(1.315)
Centrality×Volatility	−0.850*		
	(−1.690)		
Centrality×Discre_Accruals		−1.258***	
		(−3.177)	
Centrality×Horizon			−0.093**
			(−2.036)
Volatility	1.539***		
	(11.180)		
Discre_Accruals		0.309**	
		(2.402)	
Horizon	0.637***	0.644***	1.310
	(60.342)	(57.926)	(102.996)
_cons	−4.465***	−4.203***	−4.158***
	(−7.465)	(−6.512)	(−6.925)
Control variables	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes
Firm FEs	Yes	Yes	Yes
Analyst FEs	Yes	Yes	Yes
N	450,412	420,015	450,412
Adjusted R^2	0.594	0.595	0.594

Note: This table reports the regression results of examining the moderating effects of forecast difficulty on the relationship between analyst centrality and earnings forecast accuracy. Variables are defined in the Appendix. T-statistics (in parentheses) are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

our expectations as well as prior studies (Fan et al., 2022; Talhelm et al., 2014). It means that analysts employed by brokerage houses headquartered in provinces where rice cultivation intensified the collaborative culture have more willingness and opportunities to collaborate with others. The coefficient on *Centrality* in the second-stage regression is still negatively significant, continuing to support that a central position in the analyst collaboration networks is helpful for an analyst to improve the forecast accuracy. Taken together, the IV 2SLS regression results suggest a robust causal relationship between collaboration network centrality and analyst forecast accuracy when the endogenous concern is appropriately addressed.

5. Further analyses

5.1. Heterogenous effects across analysts

5.1.1. Analyst experience

Having verified the effect of collaboration network centrality on analyst earnings forecast accuracy, we proceed to explore whether this effect differs across analysts. First, we attempt to examine the heterogeneity by analyst experience. We consider an analyst as an experienced analyst in forecasting a firm if his/her firm-specific experience at the issuance of a given forecast exceeds the median value of his/her firm-specific experience. Then the sample is correspondingly partitioned into two groups, encompassing forecasts made by experienced analysts and by inexperienced analysts, respectively. We re-estimate Eq. (1) using each subsample and present the results in columns (1) and (2) of Table 10. The results show that the coefficients on *Centrality* are significantly negative in both subsamples, indicating that the effects of collaboration network centrality on analyst earnings forecast accuracy are almost tantamount. We interpret the findings as that collaboration relationships are valuable for analysts to acquire useful information regardless of their experience.

Table 13
Regression results of forecast timeliness on analyst collaboration network centrality.

Variables	(1)		(2)	
	Timeliness_First		*Timeliness_Last*	
	Coeff.	t-stat.	Coeff.	t-stat.
Centrality	−0.036***	−3.934	−0.046***	−5.442
_cons	−2.319***	−47.095	−2.284***	−51.025
Control variables	Yes		Yes	
Year FEs	Yes		Yes	
Firm FEs	Yes		Yes	
Analyst FEs	Yes		Yes	
N	205,360		205,360	
Adjusted R^2	0.543		0.623	

Note: This table reports the regression results of forecast timeliness on analyst collaboration network centrality. Variables are defined in the Appendix. T-statistics are based on standard errors clustered at the analyst level. *, **, and *** denote significance levels of 10%, 5%, and 1%, respectively.

5.1.2. Analyst specialty

Next, we turn to examine whether the effects of collaboration network centrality are heterogeneous between analysts with and without specialty in a given industry. Like Hwang et al. (2019), we determine each analyst's industry specialty by identifying the most frequent industry that his/her earnings forecasts belong in. In cases where an analyst's most frequent industries are more than one, we take all of them into account. We then split the sample into two categories. One consists of forecasts made by analysts who specialize in that industry, while the other is comprised of forecasts by analysts who are not specialized in that industry. The regression results tabulated in columns (3) and (4) of Table 10 show that the coefficients on *Centrality* in two subsamples are both significantly negative, albeit a subtle difference in magnitude. This means that regardless of whether an analyst has a specialty in an industry or not, collaboration networks would offer him/her incremental information that is useful to improve the accuracy of earnings forecasts.

5.2. Impacts of the characteristics of collaborators

5.2.1. Collaborators' industry specialty

We move on to discuss the impacts of the characteristics of collaborators on the forecast accuracy of focal analysts. First, we investigate how collaborators' industry specialty affects the accuracy of forecasts made by an analyst. We expect that collaborating with industry-specialized analysts would be particularly helpful in improving the accuracy of earnings forecasts issued for firms in that industry. We construct a variable, *Network_Ind_Specialists*, to indicate the number of an analyst's collaborators who are specialists in the industry of the forecast being examined. Similar to Hwang et al. (2019), we consider the most frequent industry that a collaborator's earnings forecasts belong in as his/her specialized industry. Some collaborators may specialize in multiple industries, we take each of them into account. According to the regression results reported in column (1), Table 11, the coefficient on *Network_Ind_Specialists* is significantly negative. This means that if an analyst has more collaborators who are specialized in the industry in which the firm being forecasted belongs, the analyst can produce more accurate earnings forecasts. This also implies that communicating with specialized collaborators is particularly efficient for discovering valuable information to make accurate forecasts.

5.2.2. Collaborators' ability

Second, we explore how collaborators' abilities matter. We hold an intuitive conjecture that collaborating with analysts who exhibit superior forecast performance would improve the performance of the focal analyst to a larger degree. We construct a group of variables, labeled *L_Network_AFE_mean*, *L_Network_AFE_median*, and *L_Network_AFE_min*, which are the mean, median, and minimum values of the previous-year absolute forecast errors of an analyst's current-year collaborators, respectively. A higher value of these variables indicates that the collaborators of an analyst exhibited poorer forecast performance in the previous year. The regression results presented in columns (2) to (4) of Table 11 show that, the dependent variable, *AFE*, is positively associated with the absolute forecast errors of the collaborators in the preceding year. This translates as more able collaborators are more helpful for accurate forecasts, which is consistent with our conjecture.

5.3. Moderating effects of task difficulty

Since collaboration networks benefit analysts with valuable information and knowledge and thereby improve earnings forecast accuracy, we predict that the effects of collaboration network centrality would be more prominent when earnings are difficult to forecast. To test this prediction, following prior studies like Schabus (2022), we use several variables to proxy for the complexity or uncertainty of forecasting, namely *Volatility*, *Horizon*, and *Discre_Accruals*. *Volatility* is the annual stock return volatility of the firm being forecasted, proxying for the volatility of the firm's economic conditions. *Horizon*, as defined before, indicates the time span between the issuing date of a forecast and the earnings announcement date, and long-horizon forecasts entail dealing with great uncertainty. *Discre_Accruals* is the firm's absolute abnormal discretionary accruals in the previous year, derived from the modified Jones model (Dechow, Sloan, & Sweeney, 1995), usually used as a proxy for opaque information. We interact these variables with *Centrality* and present the regression results in Table 12. According to the results, all the coefficients on these interaction terms are estimated to be significantly negative, meaning that collaboration network centrality is more effective in improving forecast accuracy when earnings are more difficult to forecast. The findings highlight the desirable information extraction role of analysts' collaboration networks when dependable information is publicly sparse.

5.4. Effects of collaboration networks on forecast timeliness

As Bradshaw (2011) and Gao et al. (2022) note, timeliness is another attribute of earnings forecasts valued by investors. Here, we are going to examine how collaboration network centrality affects the timeliness of earnings forecasts, which can be predicted with different theories. On the one hand, since a central position in the collaboration networks equips analysts with more information access, they could discover and internalize useful information in a timely manner and then release timely earnings forecasts. If so, we expect forecast timeliness to be positively related to collaboration network centrality. On the other hand, however, a higher collaboration network centrality means repetitive communications with different collaborators and may cause an information overload problem (Impink, Paananen, & Renders, 2022). When a large amount of information from different collaborators is received, analysts must perform a series of procedures to filter, verify, compare, integrate, and interpret the information, and these procedures cannot be completed in a short time. If this is the case, analysts with higher collaboration network centrality would take a longer period of time to issue forecasts before addressing the information overload problem.

To determine which theory is true, we design two variables to measure the forecast timeliness of analysts. Specifically, we first focus on each analyst's first forecast for a given firm and year. We sort these forecasts by their issuing dates and then compute a variable, *Timeliness_First*, as a forecast's temporal ordinal among all forecasts for that firm and year, divided by the total number of forecasts and multiplied by negative one. A higher value of *Timeliness_First* means that the forecast emerges among the earliest ones. Analogously, we compute the second variable, *Timeliness_Last*, which is analogously calculated based on each analyst's last forecast for a given firm and year. The regression results are reported in Table 13. The coefficients on *Centrality* are negatively significant no matter which variable is used as the dependent variable, suggesting that collaboration network centrality exerts a negative impact on earnings forecast timeliness. This coincides with the second theory, implying that analysts with more collaborators tend to await the arrival of information from all possible sources and have to spend more time processing and analyzing the information collected before incorporating it into earnings forecasts.

6. Conclusion

As important information intermediaries, financial analysts generally involve communications with others and exchange valuable information and knowledge. Mapping the information networks among analysts would be conducive to understanding the spread of information and its attendant consequences on analyst forecast performance. This study constructs extensive analyst collaboration networks based on their coauthorships of research reports. We argue that the collaboration relationships serve as conduits for information and knowledge flows and that analysts closer to the center of the collaboration networks have more opportunities to extract useful information to forecast earnings. Our empirical results show that analysts who occupy a central position in the collaboration networks issue more accurate earnings forecasts and that the effects of collaboration network centrality are applicable to all analysts regardless of their experience and industry specialty. Our further analyses indicate that collaborating with analysts who are more able or specialized in a given industry is more beneficial for the focal analysts to improve forecast accuracy. Moreover, the documented effects are more pronounced when earnings are more difficult to forecast, highlighting the information extraction role of analyst collaboration networks. In addition, our evidence demonstrates that analysts with higher collaboration network centrality generally take a longer period of time to issue forecasts, which is consistent with the prediction of information overload theory. Overall, this study presents evidence on whether and how a central position in the analyst collaboration networks benefits analyst forecast performance with incremental information. Our results are also in line with the idea of behavioral finance. Chen and Jiang (2006) document that, on average, analysts tend to overweight their private information to signal their ability and thus lose efficiency. In such cases, analysts in the central position of collaboration networks have an advantage in extracting and integrating private information. Sufficient and high-quality private information would compensate for the side effects of analysts overweighting private information.

The implications of this study are straightforward. First, analysts, no matter how experienced or specialized they are, are encouraged to collaborate with others. This enables them to effectively discover valuable information and knowledge that spread among analysts. Besides, the findings also suggest that analysts, particularly those who are unsophisticated and eager to achieve a better performance, find more able and specialized collaborators. Of course, the reason why analysts need to collaborate is to make use of the valuable information and knowledge spreading among analyst networks. If collaboration is difficult, communicating with more analysts would also be helpful in producing research reports. Second, as to brokerage houses, this study suggests that advocating an intrafirm collaborative culture and creating convenience for information sharing are of much help to improve the average forecast performance of their employees. Third, since earnings forecasts made by analysts with more collaboration relationships are more informative, investors are advised to pay attention to the network position of analysts and rely more on central analysts. For academics, a constructive suggestion is to consider analysts' network positions or collaboration relationships when understanding their behavior and performance.

Funding

This work was supported by the National Natural Science Foundation of China [72301224] and Sichuan Provincial Science and Technology Planning Project [2023NSFSC1030].

Declaration of competing interest

None.

Data availability

Data will be made available on request.

Appendix A. Variable definitions

Variables	Descriptions
AFE	The absolute value of an EPS forecast minus the actual EPS, scaled by the stock price of the firm at the end of the previous year, multiplied by 100.
Betweenness	The normalized betweenness centrality of analyst i, computed as $\frac{\sum_{j<k}\frac{g_{jk}(i)}{g_{jk}}}{(n-1)(n-2)/2}$, where n is the number of all analysts in the network, g_{jk} is the number of the shortest paths between analysts j and k, and $g_{jk}(i)$ expresses the number of the shortest paths passing through analyst i. The measure is calculated with UCINET and expressed in percentage.
Big4	A dummy variable that takes the value of 1 if the annual report of a firm is reviewed by a Big-Four accounting firm and 0 otherwise.
BM	The book-to-market ratio of a firm.
BrokerAge	The natural logarithm of 1 plus the number of years since the establishment of a brokerage house.
Centrality	The normalized degree centrality of analyst i, computed as $\frac{\sum_{j=1}^{n}x_{i,j}}{n-1}$, where n is the number of analysts in the network, and $x_{i,j}$ equals 1 if analyst i has a collaboration relationship with analyst j. The measure is calculated with UCINET and expressed in percentage.
Centrality_MAX	The maximum value of normalized degree centrality of all analysts who author a research report.
Centrality_within	The normalized degree centrality of analyst i in the within-brokerage analyst collaboration network; algorithms are analogous to computing *Centrality*.
Closeness	The normalized closeness centrality of analyst i, computed as $\frac{n-1}{\sum_{j=1}^{n}d_{i,j}}$, where n is the number of analysts in the network, and $d_{i,j}$ is the shortest distance from analyst i to analyst j. The measure is calculated with UCINET and expressed in percentage.
Discre_Accruals	The absolute abnormal discretionary accruals, derived from the modified Jones model (Dechow et al., 1995).
Firm_Exp	The natural logarithm of 1 plus the cumulative number of research reports issued by an analyst for a specific firm since the issuance of his/her first research report for the same firm up to the issuing date of the report being tested.
Gen_Exp	The natural logarithm of 1 plus the cumulative number of research reports issued by an analyst since the issuance of his/her first research report up to the issuing date of the report being tested.
Growth	The increase in sales of a firm divided by the sales from the previous year.
Horizon	The natural logarithm of 1 plus the number of days between the issuing day of a research report and the earnings announcement day.
lnActAnalysts	The natural logarithm of 1 plus the number of analysts of a brokerage who issue at least one research report in a year.
lnAsset	The natural logarithm of the total assets of a firm.
lnFirmAnaNum	The natural logarithm of 1 plus the number of analysts who follow a specific firm in a year.
lnFirmNum	The natural logarithm of 1 plus the number of firms that an analyst covers in a year.
lnReportNum	The natural logarithm of 1 plus the number of research reports authored by an analyst in a year.
lnSignAnaNum	The natural logarithm of 1 plus the number of analysts who author a research report.
L_Network_AFE_mean	The mean value of an analyst's current-year collaborators' previous-year absolute forecast errors.
L_Network_AFE_median	The median value of an analyst's current-year collaborators' previous-year absolute forecast errors.
L_Network_AFE_min	The minimum value of an analyst's current-year collaborators' previous-year absolute forecast errors.
Leverage	The ratio of debts to assets of a firm.
ListedBroker	A dummy variable that takes the value of 1 if a brokerage house is a listed company and 0 otherwise.
Loss	A dummy variable that takes the value of 1 if a firm reports negative earnings and 0 otherwise.
Network_Ind_Specialists	The natural logarithm of 1 plus the number of an analyst's collaborators who are specialized in the industry of the forecast being examined.
Norm_Earnings_error	The absolute value of an earnings forecast minus the actual earnings, scaled by the average firm-year absolute forecast error of earnings.
Norm_EPS_error	The absolute value of an EPS forecast minus the actual EPS, scaled by the average firm-year absolute forecast error of EPS.
Norm_PE_error	The absolute value of a price-to-earnings (PE) ratio forecast minus the actual PE ratio, scaled by the average firm-year absolute forecast error of PE ratio.
Rice_Areas	The proportion of areas cultivated for paddy rice in areas for all crops of the province where a brokerage house is headquartered.
ROA	The earnings scaled by the average total assets of a firm.
SOE	A dummy variable that takes the value of 1 if a firm is a state-owned enterprise and 0 otherwise.
Timeliness_First	The timeliness of each analyst's first forecast for a given firm and year, computed as $-\frac{n_{ijt}}{N_{jt}}$, where n_{ijt} is the temporal ordinal of analyst i's first forecast for firm j and year t among all the first forecasts of each analyst, and N_{jt} is the total number of first forecasts for firm j and year t.

(*continued on next page*)

(*continued*)

Variables	Descriptions
Timeliness_Last	The timeliness of each analyst's last forecast for a given firm and year, computed as $-\frac{n_{ijt}}{N_{jt}}$, where n_{ijt} is the temporal ordinal of analyst i's last forecast for firm j and year t among all the last forecasts of each analyst, and N_{jt} is the total number of last forecasts for firm j and year t.
Volatility	The volatility of a firm's stock returns, measured as the standard deviation of monthly stock returns.

References

Bajo, E., Chemmanur, T. J., Simonyan, K., & Tehranian, H. (2016). Underwriter networks, investor attention, and initial public offerings. *Journal of Financial Economics, 122*(2), 376–408.

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology, 92*(5), 1170–1182.

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science, 323*(5916), 892–895.

Bradley, D., Gokkaya, S., & Liu, X. (2020). Ties that bind: The value of professional connections to sell-side analysts. *Management Science, 66*(9), 4118–4151.

Bradshaw, M. T. (2011). *Analysts' forecasts: What do we know after decades of work?*. Working paper, Boston College.

Brown, L. D., Hagerman, R. L., Griffin, P. A., & Zmijewski, M. E. (1987). Security analyst superiority relative to univariate time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics, 9*(1), 61–87.

Cai, Y., Dhaliwal, D. S., Kim, Y., & Pan, C. (2014). Board interlocks and the diffusion of disclosure policy. *Review of Accounting Studies, 19*(3), 1086–1119.

Calluzzo, P. (2023). Experts in the boardroom: Director connections in the mutual fund industry. *Contemporary Accounting Research, 40*(2), 1210–1249.

Chen, Q., & Jiang, W. (2006). Analysts' weighting of private and public information. *The Review of Financial Studies, 19*(1), 319–355.

Chen, Q., Mayew, W. J., & Yan, H. (2022). Equity analyst social interactions and geographic information transmission. *Review of Accounting Studies*. https://doi.org/10.1007/s11142-022-09714-z

Cheng, Q., Du, F., Wang, X., & Wang, Y. T. (2016). Seeing is believing: Analysts' corporate site visits. *Review of Accounting Studies, 21*(4), 1245–1286.

Cheng, Z., Rai, A., Tian, F., & Xu, S. X. (2021). Social learning in information technology investment: The role of board interlocks. *Management Science, 67*(1), 547–576.

Chuluun, T. (2015). The role of underwriter peer networks in IPOs. *Journal of Banking & Finance, 51*, 62–78.

Clement, M. B. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics, 27*(3), 285–303.

Cohen, L., Frazzini, A., & Malloy, C. (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy, 116*(5), 951–979.

Cohen, L., Frazzini, A., & Malloy, C. (2010). Sell-side school ties. *The Journal of Finance, 65*(4), 1409–1437.

Collins, W. A., & Hopwood, W. S. (1980). A multivariate analysis of annual earnings forecasts generated from quarterly forecasts of financial analysts and univariate time-series models. *Journal of Accounting Research, 18*(2), 390–406.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review, 70*(2), 193–225.

Dharwadkar, R., Harris, D., Shi, L. N., & Zhou, N. (2020). The initiation of audit committee interlocks and the contagion of accounting policy choices: Evidence from special items. *Review of Accounting Studies, 25*(1), 120–158.

Durkheim, E. (1951). *Suicide: A study in sociology* (2nd ed.). London: Routledge.

El-Khatib, R., Fogel, K., & Jandik, T. (2015). CEO network centrality and merger performance. *Journal of Financial Economics, 116*(2), 349–382.

Engelberg, J., Gao, P., & Parsons, C. A. (2012). Friends with money. *Journal of Financial Economics, 103*(1), 169–188.

Fan, J. P. H., Gu, Q., & Yu, X. (2022). Collectivist cultures and the emergence of family firms. *The Journal of Law and Economics, 65*(S1), S293–S325.

Fang, B., & Hope, O.-K. (2021). Analyst teams. *Review of Accounting Studies, 26*(2), 425–467.

Fang, L. H., & Huang, S. (2017). Gender and connections among wall street analysts. *The Review of Financial Studies, 30*(9), 3305–3335.

Feng, Y., Song, K., & Tian, Y. S. (2019). Director networks and initial public offerings. *Journal of Banking & Finance, 106*, 246–264.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks, 1*(3), 215–239.

Fried, D., & Givoly, D. (1982). Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics, 4*(2), 85–107.

Gao, M., Ji, Y., & Rozenbaum, O. (2022). When do associate analysts matter? *Management Science, 68*(5), 3925–3948.

Geng, C. X., Xiao, T. S., Yuan, C., & Zhang, Y. A. (2021). Follow friends to invest: Evidence from the information role of weak ties. *Accounting and Finance, 61*(5), 5999–6035.

Gibbons, B., Iliev, P., & Kalodimos, J. (2021). Analyst information acquisition via EDGAR. *Management Science, 67*(2), 769–793.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360–1380.

He, W., Jackson, A. B., & Li, C. K. (2020). Hierarchy and performance of analyst teams. *The European Accounting Review, 29*(5), 877–900.

Huang, A. H., Lin, A.-P., & Zang, A. Y. (2022). Cross-industry information sharing among colleagues and analyst research. *Journal of Accounting and Economics, 74*(1), Article 101496.

Hugon, A., Lin, A.-P., & Markov, S. (2023). Intrafirm knowledge sharing in the investment research industry. *Management Science*. https://doi.org/10.1287/mnsc.2023.4809

Hwang, B.-H., Liberti, J. M., & Sturgess, J. (2019). Information sharing and spillovers: Evidence from financial analysts. *Management Science, 65*(8), 3624–3636.

Impink, J., Paananen, M., & Renders, A. (2022). Regulation-induced disclosures: Evidence of information overload? *Abacus, 58*(3), 432–478.

Jackson, M. O. (2014). Networks in the understanding of economic behaviors. *Journal of Economic Perspectives, 28*(4), 3–22.

Johan, S. A., & Wu, Z. (2014). Does the quality of lender-borrower relationships affect small business access to debt? Evidence from Canada and implications in China. *International Review of Financial Analysis, 36*, 206–211.

Larcker, D. F., So, E. C., & Wang, C. C. Y. (2013). Boardroom centrality and firm performance. *Journal of Accounting and Economics, 55*(2), 225–250.

Li, C., Lin, A.-P., & Lu, H. (2023). The effect of social skills on analyst performance. *Contemporary Accounting Research, 40*(2), 1418–1447.

Li, Z., Wong, T. J., & Yu, G. (2020). Information dissemination through embedded financial analysts: Evidence from China. *The Accounting Review, 95*(2), 257–281.

Luong, T. S., Qiu, B., & Wu, Y. (2021). Does it pay to be socially connected with wall street brokerages? Evidence from cost of equity. *Journal of Corporate Finance, 68*, Article 101939.

Mayew, W. J., Sharp, N. Y., & Venkatachalam, M. (2013). Using earnings conference calls to identify analysts with superior private information. *Review of Accounting Studies, 18*(2), 386–413.

Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The Journal of Finance, 49*(1), 3–37.

Phua, K., Tham, M., & Wei, C. (2023). Peer effects in equity research. *Journal of Financial and Quantitative Analysis, 58*(2), 647–676.

Rumokoy, L. J., Neupane, S., Chung, R. Y., & Vithanage, K. (2019). Underwriter network structure and political connections in the Chinese IPO market. *Pacific-Basin Finance Journal, 54*, 199–214.

Schabus, M. (2022). Do director networks help managers forecast better? *The Accounting Review, 97*(2), 397–426.

Schenone, C. (2004). The effect of banking relationships on the firm's IPO underpricing. *The Journal of Finance, 59*(6), 2903–2958.

Shiue, M.-J., Chiu, P.-C., Kuo, L.-C., & Yeh, S.-L. (2022). Signing networks of audit partners and accruals quality: Evidence from Taiwan. *Managerial Auditing Journal, 37*(4), 464–485.

Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science, 344*(6184), 603–608.

Titman, S., Wei, C., & Zhao, B. (2022). Corporate actions and the manipulation of retail investors in China: An analysis of stock splits. *Journal of Financial Economics, 145*(3), 762–787.

de Villiers, C., La Torre, M., & Botes, V. (2022). Accounting and social capital: A review and reflections on future research opportunities. *Accounting and Finance, 62*(4), 4485–4521.

Worrell, J., Wasko, M., & Johnston, A. (2013). Social network analysis in accounting information systems research. *International Journal of Accounting Information Systems, 14*(2), 127–137.