Pacific-Basin Finance Journal 61 (2020) 101320

Contents lists available at ScienceDirect

Pacific-Basin Finance Journal

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

Market information traveling on high-speed rails: The case of analyst forecasts☆

Dongmin Kong[a], Lihua Liu[a], Shasha Liu[b,*]

[a] School of Economics, Huazhong University of Science and Technology, Wuhan 430074, Hubei, China
[b] Department of Accounting, Jinan University, Wuhan 430074, China

ARTICLE INFO

Keywords:
High-speed rail
Sell-side analyst
Information dissemination
Forecast accuracy
China

JEL classifications:
G14
G28

ABSTRACT

This study examines the causal effects of high-speed rail (HSR) on the forecast accuracy of sell-side analysts. Although transportation infrastructure may reshape business activities, comparatively little attention has been paid to the impact of traveling cost on the information acquisition of analysts. In this study, we exploit a quasi-experiment of variation in China's HSR and use difference-in-differences estimation to show the following. (1) HSR substantially increase forecast accuracy for firms located in the cities connected to HSR. (2) To address endogeneity, we introduce a placebo test and instrumental variable based on the hypothetical least-cost HSR networks, and the results are highly consistent. (3) A plausible mechanism is that HSR increase the probability of analysts' visit of listed firms rather than via increased analyst coverage and an improved firm information environment. (4) Our findings are more pronounced for firms with high accrual quality, firms audited by a Big 4 auditor, and firms located nearer information centers. Overall, we provide the first empirical evaluation of the economic consequences of HSR in terms of analyst forecast and market information transmission.

1. Introduction

Sell-side analysts play important role in market information dissemination by gathering, analyzing, and delivering information based on publicly available data and private communication with managers (Soltes, 2014; Brown et al., 2015; Zhang and Wei, 2020). Given that sell-side analysts spend most time on primary research and communication with different market participants (Brown et al., 2015), the time constraints are vital to the quality of analysts' output.

In this paper, we use a large-scale transportation infrastructure (i.e., Chinese high-speed rail, HSR), which significantly shortens actual travel time and enhances the mobility of people, to investigate the causal effects of time-saving infrastructure on analysts' forecast accuracy and explore the underlying channels through which analysts might acquire information.

HSR was introduced in China in 2008 to facilitate the flow of information, capital, and labor among cities and to stimulate economic growth during the post-crisis recession (Ke et al., 2017; Lin, 2017; Qin, 2017). By 2017, the HSR network spanned 29 of China's 34 province-level administrative divisions, covering 25,000 km and providing 7 billion trips in total. By substantially

☆ We thank David Reeb (the editor), anonymous referee, Chen Lin, Bohui Zhang, Rui Shen, Hai Lu, Zhisheng Li, Wenfeng Wu, Jinqiang Yang, Bo Liu, Haijian Zeng, Kebin Deng, and seminar participants at Nanjing University, Wuhan University, andNanchang University, Zhongnan University of Economics and Law for helpful suggestions. This study was funded by the National Natural Science Foundation of China (grant no. 71772178; 71602074). All remaining errors are our own.

* Corresponding author.
E-mail addresses: kongdm@hust.edu.cn (D. Kong), lihualiu75@hust.edu.cn (L. Liu), sarahliu@jnu.edu.cn (S. Liu).

https://doi.org/10.1016/j.pacfin.2020.101320
Received 18 December 2019; Received in revised form 4 March 2020; Accepted 19 March 2020
Available online 21 March 2020
0927-538X/

shortening the travel time and enhancing the mobility of human capital, HSR fundamentally affects the business environment and information dissemination. As such, it is intriguing to explore whether and how HSR affects analysts' information production and dissemination.

Most of the previous studies focus on the impact of HSR on the local economy and provide the mixed results. Despite related studies on the impact of HSR on economic growth, limited focus has been given to the analyst forecast behavior of listed firms. Accordingly, to test whether large-scale transport infrastructure leads to the increase of analysts' visit of listed firms and improve forecast accuracy is an empirical issue that has yet to be addressed.

Analyst forecast accuracy might improve for firms located in cities with HSR connections for several non-mutually exclusive reasons. First, connection to HSR could increase the probability of analysts visiting the listed firms, and communications between analysts and management might be more efficient (Soltes, 2014; Brown et al., 2015; Han et al., 2018),[1] because the time-space compression brought by the introduction of HSR could facilitate face-to-face contact in the knowledge-generating process (Chen and Hall, 2011). Second, connection to HSR could increase analyst coverage and competition, thus improving forecast accuracy. Geographical proximity could provide local analysts with an information advantage (Malloy, 2005; O'Brien and Tan, 2015). Given that the time-space compression with HSR may reduce coverage costs and enhance coverage from non-local analysts, we expect that the connection to HSR could improve forecast accuracy by increasing analyst coverage. Third, firms connected to HSR might improve information disclosure quality under monitoring pressure from stakeholders, who are more likely to visit the firm. If so, analysts could benefit from the improvement of the firm's information quality, which is an important input of analysts' forecasts.

To conduct the empirical examination, we construct a unique database that merges the listed firms and detailed opening dates of the HSR lines. The recent massive construction waves of HSR provide us with a quasi-natural experiment design. In China, the construction of HSR began in 2005, and the first line, between Beijing and Tianjin, opened in 2008. Since then, the HSR network has undergone unprecedented large-scale expansion. The construction boom of the HSR network is attributed to the Mid-to-Long Term Railway Development Plan, a national grid proposed by the Ministry of Railway (MOR) in 2004 and composed of four north–south corridors and four east–west corridors. One important policy objective of HSR is to provide rapid, reliable, and comfortable transportation for many passengers between densely populated metropolitan cities over long distances. Given that the peripheral cities are not explicitly targeted by this policy, the connections to HSR for these cities are plausibly exogenous.

We use difference-in-differences (DID) methodology to analyze the impact of HSR on the forecast accuracy with a sample period of 2006–2016. First, we find that after being connected to the HSR network, a firm experiences an increase in forecast accuracy compared with unaffected firms. To verify the quasi-random treatments on cities' connection to HSR, we use a dynamic DID approach to test whether the connected and unconnected cities have experienced parallel trends in analyst forecast accuracy prior to the initial connection to HSR, and the results support the parallel trend assumption. That is, the affected firms and unaffected firms would have experienced similar changes in the absence of HSR.

To some extent, the idea that HSR may have an impact on analyst forecast accuracy is quite natural. However, the attempts to identify a causal impact face significant challenges, just like the argument in Campante and Yanagizawa-Drott (2018). Therefore, to address the endogeneity, we conduct three further tests. 1). We conduct a placebo test using the randomly generated time of the HSR opening to show that our results are not driven by other confounding factors. We determine no effect of HSR using this pseudo opening year, thereby supporting the validity of our results. 2). We adopt a two-stage least square (2SLS) estimation with an instrumental variable to further rule out the endogeneity concerns that the HSR placements may be correlated with several district characteristics, such as economic importance and political favoritism. The main results remain valid. 3). We also try to exclude alternative explanations (the mobility of people) by splitting the sample into two subgroups according to mobility variables.

Second, we further investigate the possible mechanism. In particular, we present evidence that the HSR increases the probability of analysts' visit of listed firms rather than an increase in analyst coverage and the improvement of the firm information environment. We further find that the effect mainly exists in non-local firms rather than local firms.

Third, we cross-sectionally verify that the impact exists mainly in firms with lower information asymmetry by providing evidence on heterogeneity. The opening of HSR has its largest effect on firms with higher accrual quality, firms audited by a Big 4 auditor, and firms located nearer information centers.

This study makes two contributions to the literature. First, it adds to a growing body of literature on estimating the economic consequences of transport infrastructure. Prior studies have analyzed regional economic growth (Qin, 2017), the spillover effect of knowledge and the labor market (Lin, 2017), market integration and the housing price (Mayer and Trevien, 2017), trade costs (Faber, 2014) and cost of debt (Wang et al., 2019). However, surprisingly little attention has been paid to micro-level activities in capital markets.[2] This study differs from prior research because we focus on specific information transmission issues and use micro-level data to explore the impact of China's HSR on the forecast accuracy of analysts about the HSR-affected firms.

Second, this study enhances our understanding of the determinants of analysts' forecast accuracy. Previous studies indicate that geographical proximity can provide an information advantage for mutual funds, analysts, individual investors, and even regulators (Coval and Moskowitz, 1999; Malloy, 2005; Han et al., 2018). In this paper, we complement the literature by showing that the

[1] For example, Soltes (2014) finds that private communication is an important communication channel compared with public interaction with management. Brown et al. (2015) find that compared with 10-K and 10-Q reports, private communication with management is a more useful input. Han et al. (2018) find that firm site visits can improve the accuracy of the analysts' earnings forecasts.

[2] Giroud (2013) studies the impact of new airline routes on resource allocation within the company. Holl (2016) shows that highways raise firm-level productivity directly and beyond the effect of density.

business travel cost and time constraints are important factors affecting analysts' visiting decisions.

In addition, our analysis of the impact of China's HSR on analyst forecast accuracy and the heterogeneous effects for different firms should be of interest to management, investors, and regulators who are concerned with information quality and fair disclosures.

The remainder of this paper is organized as follows. Section 2 describes the data and descriptive statistics. Section 3 presents the empirical strategy. Section 4 provides the main results. Section 5 discusses the mechanism and heterogeneity. Section 6 concludes the paper.

2. Data and descriptive statistics

This section presents the data source and variables used in the regression. We collect the HSR information from the National MOR, including the line name, construction starting date, opening date, line length, operating speed, and stations along the HSR lines. We double-check the HSR opening date and stations of each line from the official website (http://news.gaotie.cn) and other online sources. The HSR sample includes 98 HSR lines in operation by the end of 2016. We measure whether a city is connected to the HSR network using the station information. We manually match each HSR station to its city based on the Baidu Map and other online resources. A city is defined as connected to the HSR network if at least one HSR station is present, while the opening year of the city is defined as the earliest year an HSR line goes through.

To estimate the impact of HSR on the affected firms, we combine the HSR information of cities with the financial data of publicly listed firms based on the location of each firm. We handle data at the analyst-firm-year level and further exclude observations without analyst name, those for which the analysts' broker location is Hong Kong,[3] and those for financial firms and firms with negative or zero net assets.

All of the continuous variables are winsorized at the 1% level at both tails of their distribution to control for extreme outliers. Our sample covers 2679 firms, 15,207 firm-year observations and 170,132 analyst-firm-year observations between 2006 and 2016. Firms in our sample are separately located in 31 provinces in mainland China and 263 cities in China. In total, 170 cities were connected to the HSR network between 2008 and 2016 (i.e., treatment group), while the other 93 cities remain unconnected (i.e., control group).

We define a dummy variable ***HSR*** as 1 for the treatment group after the HSR opening year and 0 before the opening year; ***HSR*** constantly takes a value of 0 for the control group.

To obtain the forecasts accuracy, we retrieve the analysts' earnings forecasts from the China Stock Market and Accounting Research Database (CSMAR). We calculate forecast accuracy using the company's annual earnings announcements and analysts' earnings-per-share forecasts (***FEPS***) for the current and next two fiscal years. Consistent with Merkley et al. (2017), we measure forecast accuracy (***FERROR***) as the absolute difference between the earnings per share forecast (***FEPS***) minus the actual ***EPS***, scaled by the absolute value of the consensus EPS forecast (***Consensus EPS***). We further calculate ***FERROR(1)*** and ***FERROR(2)*** for the firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year $t + 1$ and $t + 2$, respectively.

$$FERROR_{i,j,t} = \frac{|FEPS_{i,j,t} - EPS_{i,t}|}{|Consensus\ EPS_{i,t}|} \tag{1}$$

where t denotes year, j denotes analyst, and i denotes firm. $\boldsymbol{Consensus\ EPS_{i,t}}$ is the median earnings per share forecast of all the analysts.

We also control for several firm-level characteristics that have been documented important for ***FERROR***. The financial and corporate governance data of publicly listed firms are from the CSMAR database. Following the literature, we control for a company's operating cash flow divided by total debt (***OP_DEBT***) (Demers, 2002), firm size measured by the natural logarithm of the total assets (***Size***) and firms' profitability (***LOSS***) (Dhaliwal et al., 2012), financial leverage measured by the debt-to-assets ratio of a company (***LEV***) (Lang and Lundholm, 1996), the standard deviation of actual earnings per share for the last 3 years (***EV***) (Kross and Ro, 1990), the standard deviation of stock returns over the past 36 months (***VAR_RET***) (Platikanova and Mattei, 2016), the natural logarithm of the number of years since the firm was established (***AGE***) (Coval and Moskowitz, 1999), ownership concentration measured by the shareholding ratio of the top ten shareholders (***Top10***) (Sabherwal and Smith, 2008) and intangible assets measured by the ratio of intangible assets to total assets (***Intan***). To control for the impact of airlines, we also control for air transport using a dummy variable ***Air*** that takes a value of 1 if the city in which a firm is located has at least one airport in use, and 0 otherwise. Definitions of the variables (including the accuracy of earnings forecasts and the HSR connections) used in our empirical tests are summarized in Appendix 1.

Table 1 reports the summary statistics of the aforementioned variables. The first three rows report the statistics of the variables that we are interested in: the absolute value of the firm's analyst forecast errors for forecasts made in year t for the earnings of year t, $t + 1$, and $t + 2$ (***FERROR, FERROR(1), FERROR(2)***, respectively). Intuitively, ***FERROR*** is the smallest and ***FERROR (2)*** is the largest.[4] In our sample, 63.3% of the observations are connected to the HSR network, while a high proportion of observations (77.7%) are affected by the airlines. For an average firm, the financial leverage is 44.1%, and the loss is 3.2%. Overall, the sample is comparable to the data in the related studies.

[3] We exclude observations for which the analysts' broker location is Hong Kong because the coverage decisions of those analysts does not only depend on the linear distance between analysts and firm based on latitudes and longitudes of their city locations but also on other factors. Usually, they need apply for a Mainland Travel Permit for Hong Kong and Macao Residents.

[4] This is consistent with Dhaliwal et al. (2012), in which *FERROR* had the highest accuracy, followed by *FERROR* (1) and then *FERROR* (2).

Table 1
Summary statistics.

Variable	Obs	Mean	SD	Min	P25	Median	P75	Max
Main variables								
FERROR	170,132	0.268	0.406	0.002	0.050	0.134	0.320	2.801
FERROR(1)	144,222	0.589	0.648	0.007	0.199	0.430	0.735	4.305
FERROR(2)	102,071	0.998	1.056	0.014	0.364	0.745	1.223	6.920
HSR	170,132	0.633	0.482	0.000	0.000	1.000	1.000	1.000
TREAT	170,132	0.909	0. 288	0.000	1.000	1.000	1.000	1.000
HSR_D	170,132	0.283	0.451	0.000	0.000	0.000	1.000	1.000
Dis_Broker	170,132	1071.305	723.164	0.000	568.274	1066.617	1463.217	3834.707
Visit	69,930	0.391	0.488	0.000	0.000	0.000	1.000	1.000
Coverage	20,712	1.572	1.216	0.000	0.000	1.609	2.639	4.263
DA	170,132	0.076	0.072	0.001	0.025	0.055	0.102	0.377
Big4	170,132	0.112	0.316	0.000	0.000	0.000	0.000	1.000
Dis_Center	170,132	377.383	435.618	0.412	21.315	270.576	566.939	3046.423
Mob_tt/10^3	57,129	200.889	73.674	19.932	169.300	215.700	259.585	346.819
Mob_in/10^3	57,129	146.642	56.259	4.982	116.100	161.638	190.510	223.715
Mob_out/10^3	57,129	54.247	31.339	4.474	21.056	57.800	83.143	204.043
Control variables								
OP_DEBT	170,132	0.223	0.344	−0.379	0.035	0.138	0.306	1.860
SIZE	170,132	22.572	1.376	20.190	21.567	22.342	23.358	26.660
LEV	170,132	0.441	0.201	0.053	0.281	0.442	0.602	0.848
EV	170,132	0.232	0.232	0.010	0.083	0.158	0.295	1.320
LOSS	170,132	0.032	0.177	0.000	0.000	0.000	0.000	1.000
VAR_RET	170,132	0.146	0.048	0.068	0.112	0.137	0.171	0.327
AGE	170,132	2.623	0.406	0.693	2.398	2.708	2.890	7.609
Top10	170,132	60.915	14.433	25.090	51.470	61.680	71.230	91.870
Intan	170,132	0.046	0.051	0.000	0.016	0.033	0.057	0.313
Air	170,132	0.777	0.416	0.000	1.000	1.000	1.000	1.000

3. Empirical strategy

To estimate the effects of the HSR connections on forecast accuracy among the listed firms, we use the DID approach of the following form as the baseline estimation strategy:

$$FERROR_{i,j,t} = \alpha_0 + \alpha_1 HSR_{i,t} + \alpha_n Controls_{i,t} + Year/Industry/Stock/Analyst\ Fixed\ Effects + \epsilon_{it} \tag{2}$$

where the dependent variable $\boldsymbol{FERROR_{i,j,t}}$ is the forecast error of analyst *j* for firm *i* at year *t* and a dummy variable *HSR* takes 1 for the treatment group after the HSR opening year and 0 before the opening year; *HSR* constantly takes 0 for the control group. Our primary interest is the coefficient α_1 because it determines the marginal effect of the HSR connection on analysts' forecast accuracy. The vector $\boldsymbol{Controls_{i,t}}$ includes a series of control variables described in the previous section that capture other relevant firm-level factors. We also include the year fixed effect to eliminate common time trends, the industry fixed effect to control for time-invariant industry-specific factors, the stock fixed effect to control for time-invariant stock-specific factors, and the analyst fixed effect to control for the time-invariant analyst-specific factors.[5] ϵ_{it} is the error term.

4. Main results

4.1. DID estimation

We use the baseline estimation specification (2) to explore the relationship between the HSR connection and forecast error. Table 2 presents the DID estimation results. Column (1) reports the estimation results when we include the year, industry and stock fixed effects. Column (2) shows the estimation results when we include the control variables of the aforementioned variables. Column (3) reports the results including the analyst fixed effect and all of the control variables, and the coefficient of ***HSR*** is around 1.951, which is statistically significant at the 1% level, suggesting that the HSR connection has a negative and significant effect on the forecast errors for firms.

As for the control variables, the coefficients of ***OP_DEBT*** and ***SIZE*** are −2.715 and −6.558, respectively, both of which are statistically significant at the 1% level, which is consistent with our prediction. The coefficient of ***SIZE*** is consistent with the findings in Lang and Lundholm (1996), who find a positive relationship with forecast accuracy. In contrast, the coefficients of ***LEV*** and ***AGE*** are 22.637 and 5.375, respectively, and they are statistically significant at the 1% and 10% level which mean that firms with higher

[5] We control the industry fixed effects and stock fixed effects simultaneously for some firms' industry changed in our sample period. Results are qualitatively similar when we don't control the industry fixed effects.

Table 2
Baseline results.

	Dep. Var.: *FERROR*		
	(1)	(2)	(3)
HSR	−2.272***	−1.983***	−1.951***
	(−6.05)	(−5.93)	(−3.64)
OP_DEBT		−2.531***	−2.715***
		(−6.25)	(−4.69)
SIZE		−6.919***	−6.558***
		(−15.40)	(−8.39)
LEV		25.381***	22.637***
		(17.62)	(9.33)
EV		26.923***	24.722***
		(35.68)	(21.64)
LOSS		84.167***	74.781***
		(63.93)	(39.18)
VAR_RET		6.790*	8.119
		(1.81)	(1.23)
AGE		1.854	5.375*
		(1.39)	(1.93)
Top10		−0.127***	−0.170***
		(−6.79)	(−5.35)
Intan		10.214***	9.046
		(2.60)	(1.48)
Air		−0.589	0.066
		(−1.05)	(0.07)
Constant	24.915***	163.944***	121.902***
	(4.28)	(14.85)	(6.43)
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes
Analyst FE	No	No	Yes
N	170,132	170,132	170,132
Adj. R-Square	0.283	0.415	0.188

Notes: This table represents the coefficients of regressions which examine the effect of HSR opening on the forecast accuracy for the listed firms in China. The dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

leverage and older firms would experience a higher analyst forecast error. The coefficient of ***EV*** is 24.722, which is statistically significant at the 1% level, consistent with the results of Kross and Ro (1990), who suggest that firms with higher earnings volatility would have less accurate analyst forecasts.

4.2. Pre-trend check and dynamic DID

In this subsection, we test whether firms located in cities with or without connections to HSR experienced differential trends in ***FERROR*** before the initial HSR connection. In the dynamic DID estimation, we focus on two years before and five years after the initial connection year to illustrate the time-varying effects. $\boldsymbol{HSR^k}$ (k ≤ -2, k $= -1, 0, 1, 2$, k $=$ 3–5) are the dummy variables, which are equal to 1 if it is year k and the firm belongs to the treated group, and 0 otherwise. For example, $\boldsymbol{HSR^{\le -2}} = 1$ means a variable that takes 1 if the difference of sample year and the HSR opening year is less than or equal to -2 and this firm belongs to the treated group, and 0 otherwise, $\boldsymbol{HSR^1} = 1$ means that it is the first year after HSR opening and this firm belongs to the treated group, and 0 otherwise. $\boldsymbol{HSR^{3\text{–}5}} = 1$ means a variable that takes a value of 1 for firms from the third year following the HSR opening until to the fifth year and this firm belongs to the treated group, and 0 otherwise.

Panel A of Table 3 reports the estimated coefficients of the dynamic DID. Neither $\boldsymbol{HSR^{\le -2}}$ nor $\boldsymbol{HSR^{-1}}$ is significant in the specification, which suggests parallel trends in ***FERROR*** between the affected and unaffected firms before the initial HSR connection year. To analyze when the HSR opening begins to affect ***FERROR***, we evaluate the four dummies after the initial opening year. The coefficients of $\boldsymbol{HSR^0}$ and $\boldsymbol{HSR^1}$ are negative but statistically insignificant, whereas the coefficients of $\boldsymbol{HSR^2}$ and $\boldsymbol{HSR^{3\text{–}5}}$ are negative and both are significant at 5% level and 10% level. This finding indicates that the effect of ***HSR*** on ***FERROR*** appears from the second year after the initial opening year and the statistical significance is reduced from third year.

4.3. Placebo test

While we document that becoming connected to the HSR network is negatively related to ***FERROR*** for publicly listed firms, if the

Table 3
Dynamic DID and placebo test.

	Dep. Var.: *FERROR*				
	Panel A: Dynamic DID	Panel B: Placebo test			
	(1)	(2)	(3)	(4)	(5)
$HSR^{\le -2}$	−1.103				
	(−0.71)				
HSR^{-1}	0.115				
	(0.08)				
HSR^{0}	−2.066				
	(−1.61)				
HSR^{1}	−1.538				
	(−1.32)				
HSR^{2}	−2.250**				
	(−2.14)				
HSR^{3-5}	−1.479*				
	(−1.72)				
HSR_ p		0.806	0.428	−0.410	−0.629
		(1.26)	(0.73)	(−0.88)	(−1.40)
OP_DEBT	−2.786***	−2.785***	−2.788***	−2.765***	−2.773***
	(−4.82)	(−4.82)	(−4.82)	(−4.78)	(−4.80)
SIZE	−6.571***	−6.495***	−6.489***	−6.516***	−6.518***
	(−8.40)	(−8.31)	(−8.30)	(−8.33)	(−8.34)
LEV	22.682***	22.469***	22.479***	22.538***	22.608***
	(9.32)	(9.25)	(9.26)	(9.29)	(9.31)
EV	24.657***	24.713***	24.726***	24.719***	24.707***
	(21.56)	(21.62)	(21.64)	(21.61)	(21.61)
LOSS	74.795***	74.772***	74.792***	74.767***	74.774***
	(39.21)	(39.17)	(39.17)	(39.17)	(39.18)
VAR_RET	9.366	8.734	8.701	8.521	8.822
	(1.42)	(1.32)	(1.32)	(1.29)	(1.34)
AGE	6.021**	5.456*	5.265*	5.308*	5.416*
	(2.16)	(1.96)	(1.89)	(1.90)	(1.94)
Top10	−0.169***	−0.171***	−0.171***	−0.170***	−0.169***
	(−5.31)	(−5.37)	(−5.37)	(−5.36)	(−5.32)
Intan	9.082	9.070	9.018	8.993	8.919
	(1.48)	(1.48)	(1.47)	(1.47)	(1.45)
Air	−0.334	−0.068	−0.039	−0.032	−0.144
	(−0.37)	(−0.08)	(−0.04)	(−0.04)	(−0.16)
Constant	120.403***	120.265***	120.678***	121.224***	120.894***
	(6.35)	(6.34)	(6.36)	(6.39)	(6.37)
Year FE	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes
N	170,132	170,132	170,132	170,132	170,132
Adj.R-Square	0.188	0.188	0.188	0.188	0.188

Notes: This table represents the estimated results of dynamic DID and placebo test. The dependent variable is the forecast error (***FERROR***) multiplied by 100. The independent variables in column (1) is $\boldsymbol{HSR^k}$ ($k \le -2, -1, 0, 1, 2, k = 3$–5) are the dummy variables, which equal 1 if it is year k and the firm belongs to the treated group; otherwise, 0. For example, $\mathrm{HSR}^{\le -2} = 1$ means the difference between a specific year and the year of HSR opening is less than or equal to −2 and this firm belongs to the treated group, 0 otherwise. The column (2)–(5) reports the coefficients of placebo test we conduct that we randomly assign an opening year between 2006 and 2016 to each city in our sample. We measure the dummy variable ***HSR_p*** based on this randomly generated opening year and re-run the regressions in Table 2. In column (2) and column (5) we set the false time point two years before and after the actual point of HSR opening respectively. In column (3) and column (4) we set the false time point one year before and after the actual point of HSR opening respectively. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

HSR placement decision is based on the past and expected future economic growth of a city (Faber, 2014; Lin, 2017), our results may be driven by this pre-existing economic growth trend rather than the actual HSR connections. To disentangle the actual HSR connection effect from these pre-existing effects, we conduct a placebo test by randomly assigning an opening year between 2006 and 2016 to each connected city in our sample. Such placebo tests can help rule out alternative explanations (Roberts and Whited, 2013).

We use this randomly generated HSR opening year to redefine the DID term ***HSR*** (denoted as ***HSR_p*** here) and re-run the regressions as shown in Table 2. In columns (2) and (5) in Table 3, we set false time points two years before and after the actual point of HSR opening, respectively. In columns (3) and (4), we set false time points one year before and after the actual point of HSR opening, respectively. We also control for the firm characteristics and the fixed effects. Panel B of Table 3 shows that all of the

Table 4
Two-stage IV estimation.

	Dep. Var.: *FERROR*		Dep. Var.: *HSR*	Dep. Var.: *FERROR*
	OLS	*OLS*	First-stage estimation	Second-stage estimation
HSR	−3.326***	−2.492***		
	(−4.15)	(−3.97)		
HSR_IV			0.769***	
			(167.89)	
HSR_hat				−1.852**
				(−2.52)
OP_DEBT	−1.810**	−2.830***	0.016***	−2.718***
	(−2.25)	(−3.20)	(4.29)	(−4.62)
SIZE	−9.212***	−7.250***	−0.010***	−6.555***
	(−7.36)	(−7.49)	(−2.59)	(−8.46)
LEV	27.102***	23.987***	0.017	22.631***
	(7.72)	(7.50)	(1.40)	(9.51)
EV	20.472***	23.761***	0.010*	24.723***
	(12.92)	(18.36)	(1.66)	(21.39)
LOSS	72.213***	80.694***	0.008	74.781***
	(25.71)	(36.07)	(1.61)	(38.85)
VAR_RET	32.602***	8.048	−0.182***	8.141
	(3.46)	(0.90)	(−4.78)	(1.22)
AGE	2.867	18.586***	−0.070***	5.368**
	(0.65)	(3.07)	(−3.98)	(2.01)
Top10	−0.237***	−0.095**	−0.001***	−0.170***
	(−4.73)	(−2.37)	(−6.75)	(−5.40)
Intan	2.934	6.832	0.051	9.047
	(0.31)	(0.88)	(1.40)	(1.43)
Air	−1.062	−0.606	0.053***	0.062
	(−1.06)	(−0.53)	(7.39)	(0.07)
Constant	227.512***	111.722***		
	(7.77)	(4.63)		
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes
N	67,577	108,651	101,289	101,289
Adj.R-Square	0.193	0.198	0.652	0.188
First-stage F-statistic			28,185.975	

Notes: The dependent variable is the forecast error (***FERROR***) multiplied by 100 in column (1) and (2). ***HSR*** in column (1) and (2) is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. In column (1), we exclude the firms located in provincial capital, municipalities or cities with independent planning. In column (2), we exclude the firms going to public after 2008 (the first year of HSR opening). Column (3) and (4) of this table represents the estimates of two-stage least squares regression. The instrumental variable is constructed as follows. First, we connect the node cities on the eight planned HSR lines with straight line segments using Baidu Map API to form a theoretically least-cost network. Next, for each non-node city, we measure the shortest vertical Euclidean distance from to its nearby HSR line segments constructed previously and record the actual opening year of this nearest HSR line. The dummy variable ***HSR_IV*** is defined for each city, that equals 1 if the shortest vertical Euclidean distance is less than 50 km in the years after the recorded opening year, and 0 otherwise. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

coefficients of ***HSR_p*** are insignificant from zero, which suggests the lack of a pre-existing confounding factor to interfere with our results. This placebo test confirms that the decrease in ***FERROR*** is the result of becoming connected to the HSR network rather than other confounding factors.

4.4. 2SLS estimation

Another concern is that our results are potentially subject to the endogeneity problem. The underlying assumption of the OLS specifications is that the HSR connections among cities are randomly assigned. If the actual HSR placement between cities is not randomly assigned and our model omits any variable that affects the HSR placement and the forecast error for publicly listed firms, then the estimated coefficients from the OLS regressions are biased and inconsistent. Faber (2014) suggests that the highway system in China tends to cover politically important and economically prosperous regions and that local governments borrow against future revenues to finance its construction. Regional characteristics such as the level of economic development and political importance affect both HSR placements among cities and forecast error for publicly listed firms.

To mitigate these concerns, we exclude firms located in provincial capitals, municipalities and cities with independent planning and instead focus on firms located in small cities. These small cities have limited incentives and capacities to lobby the central planner to change the HSR routes. Thus, their connections to the HSR network are extremely exogenous. Column (1) of Table 4 presents the estimation results. The coefficient of ***HSR*** is −3.326, which is significant at the 1% level, as reported in column (1), when we exclude firms located in large cities. In addition, firms may tend to locate in cities with connections to HSR, which may make the geographical distribution of listed companies non-random. To further control this factor, we exclude firms going public after 2008, the first year of HSR opening. Column (2) of Table 4 presents the estimation results. The coefficient of ***HSR*** is −2.492, which is significant at the 1% level, as reported in column (2), when we exclude firms going public after 2008.

However, restricting our sample to small cities may not fully address the endogeneity problem because local government lobbying could also occur in small cities when the planned HSR lines are near these cities. For further identification, we apply the 2SLS regression using the instrument variable based on hypothetical least-cost HSR networks following Faber (2014) and Wang et al. (2019). We construct the instrumental variable that aligns with the *Mid-to-Long Term Railway Development Plan* proposed by MOR in 2004, which aimed to build four vertical and four horizontal HSR lines by the end of 2016. To construct the instrumental variable, we follow the methodology of Wang et al. (2019). In particular, we first connect the node cities on the eight planned HSR lines with straight line segments using the Baidu Map API to form a theoretically least-cost network. Second, we measure the shortest vertical Euclidean distance from each non-node city to its nearby HSR line segments and record the actual opening year of this nearest HSR line. Thereafter, a dummy variable ***HSR_IV*** is defined as 1 if the shortest vertical distance from this city to the nearby HSR line is within 50 km and this year is after the recorded opening year; otherwise, ***HSR_IV*** equals 0.

Whether a non-code city is connected to the HSR network is closely related to its actual geographical distance from the planned HSR routes, which satisfies the relevance of the instrumental variable. Although only a few non-node cities may slightly change the actual HSR route through lobbying and other methods, the theoretically least-cost networks connecting the 24 node cities are beyond their control. Therefore, our instrumental variable is unrelated to ***FERROR*** and other economic factors, which satisfies the exclusion restriction. Overall, we expect ***HSR_IV*** to affect the HSR connection of a city only through its geographical location and not through its economic level and political importance.

Columns (3) and (4) of Table 4 reports the estimated results of the 2SLS analysis based on the constructed instrumental variable. Column (3) presents the first-stage result that regresses ***HSR*** on the instrumental variable ***HSR_IV*** while controlling for other firm-level characteristics and fixed effects. The coefficient of ***HSR_IV*** is positive and significant at the 1% level. The instrumental variable, which also passes the relevance test as the F-statistic from the joint test of excluded instruments, is 28,185.98, which is significant at the 1% level. Column (4) shows the second-stage results. Column (4) shows that the coefficient of ***HSR_hat*** is approximately −1.852, which is statistically significant at the 1% level. We find a negative and significant effect of the exogenous HSR connections on ***FERROR***. Thus, we confirm the validity of the ***HSR*** impact on the forecast accuracy.

4.5. Excluding alternative explanations

In this subsection, we attempt to exclude alternative explanations by addressing the effects related to changes of the population mobility of cities connection to the HSR network. Lin (2017) suggests an HSR connection increases city-wide passenger flows by 10% and employment by 7%. In this case, the increase in population mobility may help information dissemination of firms connected to HSR, which could provide more inputs for analysts to make earnings forecasts. To mitigate concerns that the main source of variation in forecast accuracy stems from other confounding factors (the mobility of people) rather than the HSR connection, we split the sample firms into different groups according to the mobility of people and to check the differences of impacts of HSR on analyst forecasts.

If the connections to HSR increases the forecast accuracy by influencing cities' population mobility, we might expect different effects of HSR on the forecast accuracy between firms located in high mobility cities and low mobility cities.[6] For example, the increase in population mobility is associated with the mobility of analysts or other informed individuals, who could provide more inputs for analysts located in cities with net increase of population migration. To address this concern, we split the sample into two subgroups according to population mobility. We define firms located in high mobility cities if the cities' total population of immigration and emigration (***Mob_tt***) is above the sample median in the year, and low mobility cities otherwise. Similarly, we define firms located in high immigration/emigration cities if the cities' population of immigration (***Mob_in***)/ emigration (***Mob_out***) is above the sample median in the year, and low mobility cities otherwise. Table 5 reports estimated results. In Column (1) and (2), both coefficients of HSR are significant and with similar magnitude (−6.098 and − 5.586). Once again, our results show that there are no significantly different effects of high-speed rail (HSR) on the forecast accuracy between firms located in high mobility cities and low mobility cities. Results are qualitatively similar when we split the sample firms into different groups according to population of immigration and emigration as reported in Column (3), (4), (5) and (6). Note that ***Air*** is omitted because of collinearity in Column (1) and Column (3) because in high mobility groups of Column (1) and Column (3), the air openings taking account for 99.72% and 100%. Thus, population mobility is less likely to be a mechanism through which the HSR increases forecast accuracy for firms located in the cities connected to HSR.

[6] We hand-collect historical population mobility data from Beijing, Shanghai, and Guangdong Statistical Yearbook, provided by Beijing, Shanghai, and Guangdong Provincial Statistics Bureau, respectively (Beijing: http://www.bjstats.gov.cn/tjsj/; Shanghai: http://www.stats-sh.gov.cn/html/sjfb/; Guangdong: http://www.gdstats.gov.cn/tjsj/gdtjnj/).

Table 5
Heterogeneity: high mobility cities vs. low mobility cites.

	Dep. Var.: *FERROR*					
	High_Mob	Low_Mob	High_In	Low_In	High_Out	Low_Out
HSR	−6.098**	−5.586***	−10.318*	−5.297***	−4.112*	−6.064***
	(−2.16)	(−4.60)	(−1.74)	(−4.46)	(−1.78)	(−4.58)
OP_DEBT	−1.820	−1.476	−7.693***	0.353	−1.808	−1.624
	(−1.13)	(−1.23)	(−3.03)	(0.30)	(−1.20)	(−1.30)
SIZE	3.787**	−7.555***	4.416*	−6.539***	4.162***	−8.891***
	(2.52)	(−6.56)	(1.83)	(−5.87)	(2.93)	(−7.41)
LEV	18.331***	21.539***	−17.719**	31.441***	18.286***	22.883***
	(3.53)	(5.28)	(−2.33)	(7.98)	(3.74)	(5.30)
EV	23.627***	25.949***	36.277***	24.216***	22.674***	25.700***
	(9.73)	(17.08)	(11.28)	(15.90)	(10.02)	(16.41)
LOSS	68.761***	72.381***	73.002***	72.283***	69.983***	71.276***
	(29.61)	(44.77)	(24.23)	(45.65)	(31.82)	(43.30)
VAR_RET	−32.412**	30.288***	−86.897***	32.507***	−8.645	24.665**
	(−2.07)	(2.67)	(−3.90)	(2.85)	(−0.60)	(2.08)
AGE	−2.766	16.953***	−14.082*	11.155**	−12.709**	19.172***
	(−0.52)	(3.31)	(−1.74)	(1.98)	(−2.54)	(3.60)
Top10	−0.262***	−0.185***	−0.299***	−0.181***	−0.173**	−0.217***
	(−3.47)	(−3.98)	(−3.33)	(−3.28)	(−2.51)	(−4.49)
Intan	1.972	7.886	19.893	−0.952	8.315	7.881
	(0.13)	(0.64)	(0.96)	(−0.08)	(0.60)	(0.61)
Air		0.220		−0.781	14.086	0.327
		(0.06)		(−0.22)	(0.68)	(0.09)
Constant	−76.857**	94.555***	−13.343	23.930	−90.897**	124.047***
	(−2.11)	(3.31)	(−0.22)	(0.77)	(−2.29)	(4.20)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes	Yes	Yes	Yes
Difference of HSR	*P*-value = .907		P-value = .904		P-value =.626	
N	22,178	34,951	16,619	40,510	25,120	32,009
Adj. R-Square	0.157	0.195	0.185	0.186	0.151	0.201

Notes: The dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. And we differentiate the firms with the population mobility. In column (1) and (2), we define firms located in high mobility cities if the cities' ***Mob_tt*** is above the sample median in the year, and low mobility cities otherwise. Here, ***Mob_tt*** is the total population of immigration and emigration. In column (3) and (4), we define firms located in high immigration cities if the cities' ***Mob_in*** is above the sample median in the year, and low mobility cities otherwise. Here, ***Mob_in*** is the population of immigration. In column (5) and (6), we define firms located in high emigration cities if the cities' ***Mob_out*** is above the sample median in the year, and low mobility cities otherwise. Here, ***Mob_out*** is the population of emigration. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

4.6. Further test

Following Dhaliwal et al. (2012), we further consider the long horizontal impacts of ***HSR*** on forecast accuracy. We define ***FERROR(1)*** and ***FERROR(2)*** as the firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t + 1 and year t + 2 respectively. Table 6 reports the estimated results. The coefficient of ***HSR*** on ***FERROR (1)*** is −2.123 and significant at the 5% level as reported in column (1), and the coefficient of ***HSR*** on ***FERROR (2)*** is −4.55 and significant at the 5% level, as reported in column (2).

5. Possible mechanism and heterogeneity

5.1. Possible mechanism

Thus far, we have obtained evidence that the HSR opening significantly decreases ***FERROR***. However, the reason why HSR results in decreasing forecast error remains to be determined. We predict that the HSR opening decreases ***FERROR*** for following three non-mutually exclusive reasons in subsection 5.1.1, 5.1.2, and 5.1.3:

Table 6
Further tests.

	(1)	(2)
	FERROR(1)	*FERROR(2)*
HSR	−2.123**	−4.550**
	(−2.22)	(−2.57)
OP_DEBT	−5.343***	−12.261***
	(−4.74)	(−4.61)
SIZE	5.999***	27.508***
	(4.21)	(9.72)
LEV	6.021	−65.431***
	(1.35)	(−7.53)
EV	−16.275***	13.167***
	(−7.92)	(3.46)
LOSS	−19.517***	9.898
	(−5.68)	(1.40)
VAR_RET	−21.355	−217.861***
	(−1.59)	(−7.79)
AGE	9.339*	25.979**
	(1.72)	(2.54)
Top10	−0.222***	0.044
	(−4.00)	(0.41)
Intan	11.599	66.047***
	(0.95)	(2.95)
Air	2.411	9.666***
	(1.21)	(2.58)
Constant	−104.613***	−389.809***
	(−3.04)	(−5.26)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Stock FE	Yes	Yes
Analyst FE	Yes	Yes
N	144,222	102,071
Adj. R-Square	0.035	0.044

Notes: This table represents further test for the main results. In column (1), the dependent variable is ***FERROR(1)*** multiplied by 100, which is the firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t + 1. In column (2), the dependent variable is ***FERROR(2)*** multiplied by 100, which is the firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t + 2. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

5.1.1. Does HSR increase the probability of company visits?

If HSR increases the probability of analysts' visit of listed firms, we would expect to see an increase in the probability of analyst visits after the HSR opening. Table 7 shows the consistent results of this prediction. We define ***HSR_D*** as 1 if a direct high-speed railway has been opened between cities where the firm located and cities where brokers' firm located, and 0 otherwise. ***Local*** is indicator variable for whether the analyst and firm are within 200 km of each other, and non-local otherwise. ***Near*** is indicator variable for whether the distance between analyst and firm is below the non-local sample median in the year, and ***Far*** otherwise. Panel A of Table 7 shows that the impact of HSR opening on analysts' forecast accuracy is different for firms at different distances to broker firms. Intuitively, the local companies would not be influenced by the HSR opening because local analysts could visit the listed firms whether HSR was open or not (as reported in column (1)). For firms located near their analysts, the coefficient of ***HSR*** is −3.126, which is significant at the 1% level as reported in column (2). For firms located far from their analysts, the coefficient of ***HSR*** is −1.945, which is significant at the 5% level, as reported in column (3), which shows that the impact is less than the impact when the analyst is moderately far from the firm. Panel A suggests that it be an inverted U for the impact of HSR on forecast accuracy by the distance. These results suggest that the benefits from the HSR are diminishing as firms get farther away from the analysts.

Furthermore, as shown in Panel B, the impact of direct HSR opening on ***FERROR*** exists only for firms far from their broker firms. The coefficient of ***HSR_D*** is −1.87, which is significant at the 5% level, as reported in column (3). Based on the results mentioned above, we further analyze the relationship between HSR opening and visit probability by splitting the sample firms into different groups according to their distance to broker firms, as reported in Panel C. Consistent with the results in Panel A, we find that the HSR increases the probability of analysts' visit of listed firms in non-local firms. For firms located near their analysts, the coefficient of ***HSR*** is 0.251 and significant at the 1% level, as reported in column (3). For firms located far from their analysts, the coefficient of ***HSR*** is 0.207, which is significant at the 1% level, as reported in column (4). Furthermore, in Panel D of Table 7, we split the sample firms

Table 7
HSR and firm-site visit.

Panel A: *HSR* and *FERROR* in different distances.

	Dep. Var.: *FERROR*		
	(1)	(2)	(3)
	Local	Near	Far
HSR	−1.038	−3.126***	−1.945**
	(−0.75)	(−3.50)	(−2.32)
OP_DEBT	−2.394*	−3.277***	−2.745***
	(−1.66)	(−3.48)	(−3.00)
SIZE	−2.816	−8.069***	−5.668***
	(−1.32)	(−6.38)	(−4.51)
LEV	21.975***	16.350***	26.399***
	(3.23)	(4.17)	(6.99)
EV	17.310***	20.198***	29.689***
	(5.47)	(11.67)	(16.01)
LOSS	65.934***	75.772***	75.145***
	(12.66)	(25.22)	(26.16)
VAR_RET	−50.814***	37.063***	7.204
	(−2.73)	(3.53)	(0.68)
AGE	7.627	10.110**	1.841
	(0.90)	(2.30)	(0.42)
Top10	−0.215***	−0.105**	−0.222***
	(−2.70)	(−2.08)	(−4.21)
Intan	−1.130	−3.661	20.410**
	(−0.07)	(−0.37)	(2.15)
Air	−2.297	−1.032	2.360
	(−0.92)	(−0.87)	(1.33)
Constant	39.655	140.374***	115.762***
	(0.88)	(4.23)	(3.86)
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes
N	24,003	73,085	73,044
Adj. R-Square	0.149	0.186	0.206

Panel B: HSR_D and FERROR in different distances.

	Dep. Var.: *FERROR*		
	(1)	(2)	(3)
	Local	Near	Far
HSR_D	−0.735	−0.585	−1.870**
	(−0.40)	(−0.74)	(−2.21)
OP_DEBT	−2.395*	−3.345***	−2.775***
	(−1.66)	(−3.56)	(−3.04)
SIZE	−2.787	−8.035***	−5.575***
	(−1.31)	(−6.35)	(−4.44)
LEV	21.828***	16.545***	26.279***
	(3.22)	(4.23)	(6.96)
EV	17.279***	20.277***	29.672***
	(5.46)	(11.72)	(16.01)
LOSS	65.918***	75.785***	75.088***
	(12.65)	(25.22)	(26.13)
VAR_RET	−50.804***	37.942***	8.434
	(−2.73)	(3.60)	(0.80)
AGE	7.603	9.937**	2.288
	(0.89)	(2.25)	(0.52)
Top10	−0.218***	−0.101**	−0.221***
	(−2.73)	(−2.02)	(−4.18)
Intan	−1.248	−3.488	20.558**
	(−0.08)	(−0.35)	(2.16)
Air	−2.199	−1.044	1.993
	(−0.86)	(−0.89)	(1.12)

(*continued on next page*)

Table 7 (*continued*)

Panel B: HSR_D and FERROR in different distances.

	Dep. Var.: *FERROR*		
	(1)	(2)	(3)
	Local	Near	Far
Constant	39.315	139.275***	112.739***
	(0.88)	(4.20)	(3.77)
Year FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Stock FE	Yes	Yes	Yes
Analyst FE	Yes	Yes	Yes
N	24,003	73,085	73,044
Adj. R-Square	0.149	0.185	0.206

Panel C: *HSR_D* and *Visit* in different distances.

	Dep. Var.: *Visit*			
	(1)	(2)	(3)	(4)
	Total	Local	Near	Far
HSR	0.246***	0.087	0.251***	0.207***
	(7.86)	(0.89)	(5.18)	(4.48)
TREAT	0.359***	0.968***	0.160**	0.500***
	(8.26)	(4.42)	(2.50)	(7.79)
OP_DEBT	−0.006	0.223***	−0.024	−0.092**
	(−0.24)	(3.50)	(−0.56)	(−2.26)
SIZE	−0.075***	−0.175***	−0.058***	−0.052***
	(−7.07)	(−6.21)	(−3.46)	(−3.24)
LEV	−0.344***	0.314*	−0.427***	−0.509***
	(−5.31)	(1.84)	(−4.17)	(−5.17)
EV	0.114***	0.273**	−0.060	0.196***
	(2.72)	(2.49)	(−0.88)	(3.20)
LOSS	−0.323***	−0.776***	−0.320***	−0.179**
	(−6.63)	(−5.40)	(−4.22)	(−2.51)
VAR_RET	0.119	0.355	0.198	−0.055
	(0.56)	(0.63)	(0.60)	(−0.16)
AGE	−0.238***	−0.393***	−0.240***	−0.168***
	(−10.48)	(−7.03)	(−6.54)	(−4.99)
Top10	0.003***	0.004**	0.002**	0.004***
	(4.83)	(2.28)	(2.11)	(3.89)
Intan	−0.346*	−1.487***	−0.162	−0.117
	(−1.77)	(−2.95)	(−0.50)	(−0.41)
Air	0.112***	0.147**	0.066**	0.127***
	(5.01)	(2.31)	(2.05)	(3.49)
Constant	−0.560**	2.453***	−0.720*	−1.370***
	(−2.32)	(3.48)	(−1.91)	(−3.71)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
N	69,924	10,368	29,687	29,852
Pseudo R2	0.0528	0.0534	0.0574	0.0529

Panel D: Heterogeneity: *Visit* vs. *Non_Visit.*

	Dep. Var.: *FERROR*	
	(1)	(2)
	Visit	*Non_Visit*
HSR	−4.000*	0.962
	(−1.73)	(0.55)
OP_DEBT	−3.653**	0.245
	(−2.40)	(0.19)
SIZE	−8.234***	−5.781***
	(−3.96)	(−2.94)

(*continued on next page*)

Table 7 (*continued*)

Panel D: Heterogeneity: *Visit* vs. *Non_Visit*.

	Dep. Var.: *FERROR*	
	(1)	(2)
	Visit	*Non_Visit*
LEV	3.933	24.984***
	(0.59)	(3.91)
EV	26.166***	27.930***
	(5.95)	(9.05)
LOSS	74.472***	70.196***
	(12.08)	(18.55)
VAR_RET	63.081***	49.702***
	(3.13)	(3.32)
AGE	18.030*	13.343*
	(1.79)	(1.77)
Top10	−0.305***	0.204**
	(−3.29)	(2.53)
Intan	23.500	23.180
	(1.25)	(1.36)
Air	−4.904**	3.184
	(−2.10)	(1.20)
Constant	120.730**	−12.121
	(2.38)	(−0.25)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Stock FE	Yes	Yes
Analyst FE	Yes	Yes
N	27,310	42,620
Adj. R-Square	0.225	0.234

Notes: This table represents the impact of HSR connection on company visits. In Panel A and Panel B, the dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. ***HSR_D*** is a dummy variable equals to one if a direct high-speed railway has been opened between cities where the firm located and cities where brokers' firm located, and zero otherwise. Fixed effects of year, industry, firm and analyst are controlled. In Panel C, the dependent variable is ***Visit***. ***Visit*** is a dummy variable equals one if a visit occurred for the (analyst-listed company) pair in a given year t, and zero otherwise. According to Pouget et al. (2017), discrete choice models, such as logit and probit, become unidentified when a large number of fixed effects are included. Therefore, we run a logit model in Panel C without including stock and analyst fixed effects. ***Local*** is indicator variable for whether the ***Dis_Broker*** (geographic distance in kilometres between analyst and firm based on latitudes and longitudes of their city locations) are within 200 km of each other, and non-local otherwise. ***Near*** is indicator variable for whether the ***Dis_Broker*** is below the non-local sample median in the year, and ***Far*** otherwise. In Panel D, the dependent variable is ***FERROR*** multiplied by 100. We split the sample firms into different groups according to whether the analysts visit the firm or not. Fixed effects of year, industry, firm and analyst are controlled. All variables are defined in section 2 and Appendix 1. For Panel A, B and, C and D the t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

into different groups according to whether the analysts visit the firm or not.[7] The coefficient of HSR is −4 and significant at the 10% level as reported in Column (1), while the coefficient of HSR is 0.962 but insignificant as reported in Column (2), which is consistent with the findings in Han et al. (2018). These findings are consistent with the reason that connection to HSR could increase the probability of analysts visiting the listed firms and improve analysts' forecast accuracy, which help verify the potential mechanism underlying our results.

5.1.2. Does the HSR increase analysts' coverage?

If HSR attracts more analysts, we would expect to see an increase in analyst coverage after the HSR opening. We define ***Coverage*** as natural logarithm of a company's analyst coverage plus one and ***D_Coverage*** as a dummy variable equals to one if the firms have more than one analyst, and zero otherwise. We handle data at the firm-year level in Table 8. Table 8 shows the inconsistent results for this prediction. The coefficient of ***HSR*** in Column (1) is −0.012 and that in Column (2) is 0.084; neither is significant.

[7] We retrieve the visit data from CSMAR and the data is limited to firms listed in Shenzhen Stock Exchange from 2012.

Table 8
HSR and analyst coverage/ report quality.

	Dep. Var.: *Coverage*	Dep. Var.: *D_ Coverage*	Dep. Var.: *DA*	Dep. Var.: *Big4*
	(1)	(2)	(3)	(4)
HSR	−0.012	0.084	−0.003	−0.088
	(−0.49)	(1.47)	(−1.54)	(−0.82)
TREAT		0.112		1.615***
		(1.63)		(7.13)
OP_DEBT	0.133***	0.653***	−0.056***	0.379***
	(4.68)	(8.50)	(−14.07)	(2.65)
SIZE	0.471***	1.004	0.001	1.287***
	(24.31)	(40.74)	(0.52)	(33.10)
LEV	−0.583***	−1.885***	−0.009	−2.085***
	(−7.54)	(−15.36)	(−1.07)	(−7.76)
EV	0.023	0.356***	0.039***	−0.423***
	(0.60)	(3.59)	(9.67)	(−2.68)
LOSS	−0.212***	−0.842***	−0.001	0.253
	(−10.40)	(−13.19)	(−0.59)	(1.56)
VAR_RET	1.476***	−0.138	0.126***	−6.520***
	(6.40)	(−0.28)	(5.89)	(−5.29)
AGE	−0.233**	−0.861***	−0.010	0.482***
	(−2.17)	(−12.39)	(−1.04)	(5.54)
Top10	0.008***	0.022***	0.001***	0.024***
	(7.06)	(15.40)	(8.36)	(9.64)
Intan	−0.216	0.053	−0.105***	3.393***
	(−1.03)	(0.15)	(−5.61)	(5.92)
Air	−0.047	−0.122***	−0.000	0.324***
	(−1.04)	(−2.70)	(−0.05)	(3.22)
Constant	−8.869***	−19.422***	0.031	−34.988***
	(−19.94)	(−34.89)	(0.63)	(−35.11)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Stock FE	Yes	No	Yes	No
N	20,712	20,687	20,712	20,034
Adj.R2 / Pseudo R2	0.206	0.239	0.079	0.345

Notes: This table represents the coefficients of regressions which examine the second and third mechanism of the impact of HSR opening on the forecast accuracy. We examine the second mechanism in columns (1) and (2) and third mechanism in columns (3) and (4). The dependent variable in column (1) ***Coverage*** is natural logarithm of a company's analyst coverage plus one. The dependent variable in column (2) ***D_Coverage*** is a dummy variable equals to one if the firms have more than one analyst, and zero otherwise. The dependent variable in column (3) ***DA*** is absolute value of a company's abnormal accruals in year t, using a modified Jones model. The dependent variable in column (4) ***Big4*** is dummy variable that takes the value of 1 (0) if the firm is audited by a Big 4 (non-Big 4) auditor firms. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. Fixed effects of year, industry, firm and analyst are controlled in column (1) and column (3). According to Pouget et al. (2017), discrete choice models, such as logit and probit, become unidentified when a large number of fixed effects are included. Therefore, we run a logit model in column (2) and column (4) without including stock and analyst fixed effects. All variables are defined in section 2 and Appendix 1. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

5.1.3. Does the HSR improve firms' report quality?

If HSR increases firms' report quality, we would expect to see an increase in the quality of the report after the HSR opening. We define ***DA*** as absolute value of a company's abnormal accruals in year t, using a modified Jones model and ***Big4*** as dummy variable that takes the value of 1 (0) if the firm is audited by a Big 4 (non-Big 4) auditor firms following Gul et al. (2010) and Yang et al. (2019). Columns (3) and (4) of Table 8 show the inconsistent results of this prediction. The coefficient of ***HSR*** in Column (3) is −0.003 and that in Column (4) is −0.088; neither is significant.

In conclusion, a plausible mechanism is that HSR increases the probability of analysts' visit of listed firms rather than via an increase in analyst coverage or the improvement of the firm information environment.

5.2. Cross-sectional differences in the improvements of forecast accuracy

5.2.1. Do analysts benefit more from higher earning quality firms after HSR opening?

If, as our findings suggest, analysts' visit is a significant factor affecting the analyst's forecast accuracy of HSR-connected firms instead of the increase in analyst coverage or the improvement of the firm information environment, firms with lower information asymmetry could benefit more for they could provide high quality information for analysts. To further demonstrate the mechanism that HSR decreases FERROR for firms, we split the sample firms into different groups and provide evidence on heterogeneity. In particular, we analyze whether the impact of HSR differs for firms with different earnings quality, firms audited by different auditors, and firms with different distances to information centers. We conjecture that analysts are more sensitive to HSR-connected firms'

Table 9
Heterogeneity: high vs. low earnings quality.

	Dep. Var.: *FERROR*	
	(1) Firms with high earnings quality	(2) Firm with low earnings quality
HSR	−2.668***	−1.262
	(−3.07)	(−1.38)
OP_DEBT	−2.322*	−3.695***
	(−1.74)	(−3.82)
SIZE	−10.570***	−1.705
	(−7.66)	(−1.34)
LEV	36.501***	15.203***
	(8.16)	(3.87)
EV	22.916***	25.342***
	(11.79)	(13.91)
LOSS	71.765***	66.883***
	(23.10)	(21.32)
VAR_RET	−22.591**	32.625***
	(−2.00)	(2.95)
AGE	5.538	−2.025
	(1.28)	(−0.43)
Top10	−0.013	−0.316***
	(−0.26)	(−5.99)
Intan	11.200	33.118***
	(1.21)	(3.15)
Air	1.391	−1.079
	(0.95)	(−0.78)
Constant	190.687***	11.800
	(5.84)	(0.36)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Stock FE	Yes	Yes
Analyst FE	Yes	Yes
N	85,163	84,969
Adj. R-Square	0.182	0.166

Notes: The dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. And we differentiate the firms with earnings quality. We define it as a high earnings quality firm if its ***DA*** is below the sample median in the year, and a low earnings quality firm if it is above the sample median in the year. Here, ***DA*** is absolute value of a company's abnormal accruals in year t, using a modified Jones model. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

information and are more likely to have higher forecast accuracy when information asymmetry is lower.

Research suggests that earnings management is related to information asymmetry and that high earnings management corresponds to high information asymmetry (Cormier et al., 2013). For example, Francis et al. (2005) suggest that earnings management could increase information asymmetry in the stock market. Firms with higher earnings quality have lower information asymmetry, which is good for analyst forecasts. We define it as a high earnings quality firm if its DA is below the sample median in the year, and a low earnings quality firm otherwise. We estimate how the HSR connection affects high earnings quality and low earnings quality firms differently and report the results in Table 9. The coefficient of ***HSR*** is −2.668 and significant at the 1% level as reported in Column (1), while the coefficient of ***HSR*** is −1.262 but insignificant as reported in Column (2). Table 9 shows that the opening of HSR mainly decreases ***FERROR*** for firms with higher earnings quality.

5.2.2. Do analysts benefit more from firms audited by big 4 auditors after HSR open?

Numerous studies have documented that firms audited by big auditor firms are of higher quality and lower information asymmetry because big auditor firms can constrain aggressive earnings management and the annual reports they audit could obtain higher investor confidence by investors (Francis and Wang, 2008). The existing studies find that in emerging market the Big 4 auditors could help promote the flow of more credible information to the market and play a corporate governance role (Gul et al., 2010; Ke et al., 2015). Thus, firms audited by big auditor firms would have lower information asymmetry, which is good for analyst forecasts. Following Gul et al. (2010) and Yang et al. (2019), we define firms are audited by Big4 if the auditors are one of the joint ventures of international Big 4 firms (E&Y, KPMG, Deloitte, and PWC), and Non-Big 4 otherwise. We estimate how the HSR connection affects firms audited by Big 4 audit firms and Non-Big 4 firms differently and report the results in Table 10. The coefficient of ***HSR*** is −9.696 and significant at the 1% level as reported in Column (1), while the coefficient of ***HSR*** is −0.783 and insignificant as reported in Column (2). Table 10 shows that the opening of HSR mainly decreases ***FERROR*** for firms audited by Big 4 auditor firms.

Table 10
Heterogeneity: Big4 vs. Non-Big4 audit firms.

	Dep. Var.: *FERROR*	
	(1) Firms audited by Big4	(2) Firms audited by Non-Big4
HSR	−9.696***	−0.783
	(−6.35)	(−1.35)
OP_DEBT	0.128	−3.012***
	(0.04)	(−5.09)
SIZE	−0.938	−7.276***
	(−0.34)	(−8.63)
LEV	17.199*	22.949***
	(1.82)	(8.95)
EV	27.506***	24.782***
	(11.54)	(19.00)
LOSS	69.360***	75.164***
	(13.74)	(35.91)
VAR_RET	10.585	4.474
	(0.49)	(0.64)
AGE	−12.055**	8.246***
	(−2.06)	(2.60)
Top10	0.166**	−0.182***
	(2.18)	(−5.14)
Intan	−7.053	7.376
	(−0.37)	(1.12)
Air	5.037**	−0.114
	(2.11)	(−0.12)
Constant	24.389	129.695***
	(0.37)	(6.53)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Stock FE	Yes	Yes
Analyst FE	Yes	Yes
N	19,095	151,037
Adj. R-Square	0.204	0.188

Notes: The dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. And we differentiate the firms with audit firm. We define it as a ***Big4*** audit firm if the auditors are one of the joint ventures of international Big 4 firms (E&Y, KPMG, Deloitte, and PWC), and ***Non-Big 4*** otherwise. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

5.2.3. Do analysts benefit more from firms located near information centers after HSR open?

Loughran (2007) shows that more investors are familiar with urban firms than rural ones and have access to informal information about them because urban firms are located near more sophisticated money managers and potential investors. Boubakri et al. (2016) suggest foreign investors are less likely to buy shares of the firms faraway to domestic financial centers. Hochberg et al. (2007) find that the network centrality of VC has a positive and significant effect on the probability. In China, most fund companies—one type of sophisticated institutional investors—are located in large cities such as Beijing, Shanghai, Shenzhen, and Guangzhou. In addition, most broker firms, which serve as important information intermediaries, are also located in these large cities. Thus, we treat those four cities as information centers and suppose that the firms located nearer information centers would have lower information asymmetry. Specifically, we compute the minimal geographic distance (***Dis_Center***) in kilometres between information center (Beijing, Shanghai, Guangzhou and Shenzhen) and firm based on latitudes and longitudes of their city locations. Then we define a firm as near to information centers firm if the distance between the firm to information centers is below the sample median in a specific year; otherwise, we define a firm as far to information centers firm. Then, we estimate how the HSR connection affects firms located near to and far from information centers differently and report the results in Table 11. The coefficient of ***HSR*** is −2.672 and significant at the 1% level as reported in Column (1), while the coefficient of ***HSR*** is −1.054 and insignificant as reported in Column (2). Table 11 shows that the opening of HSR mainly decreases ***FERROR*** for firms located nearby the information centers.

Overall, the evidence presented in this section confirms our proposed mechanism. After the opening of HSR, firms with lower information asymmetry are more likely to provide good input for analysts making earnings forecasts, and therefore they would be influenced more significantly. The opening of HSR improves forecast accuracy by increasing the probability of analysts' visit of listed firms instead of via the improvement of firm information environment and increased analyst coverage. Those effects mainly exist in firms with high earnings quality, firms audited by Big 4 firms, and firms located near an information center.

Table 11
Heterogeneity: near vs. far to information centers.

	Dep. Var.: *FERROR*	
	(1) Near to information centers	(2) Far to information centers
HSR	−2.672***	−1.054
	(−3.70)	(−1.27)
OP_DEBT	−2.570***	−2.360***
	(−3.24)	(−2.78)
SIZE	−4.312***	−7.962***
	(−4.04)	(−6.88)
LEV	22.592***	22.425***
	(6.88)	(6.19)
EV	21.453***	27.313***
	(13.97)	(16.57)
LOSS	67.822***	79.630***
	(22.27)	(32.44)
VAR_RET	6.665	4.370
	(0.71)	(0.48)
AGE	1.593	14.875***
	(0.43)	(3.47)
Top10	−0.227***	−0.113**
	(−5.66)	(−2.23)
Intan	4.942	10.262
	(0.66)	(1.09)
Air	−2.333*	0.572
	(−1.84)	(0.45)
Constant	69.538***	160.021***
	(2.78)	(5.68)
Year FE	Yes	Yes
Industry FE	Yes	Yes
Stock FE	Yes	Yes
Analyst FE	Yes	Yes
N	85,156	84,976
Adj. R-Square	0.158	0.217

Notes: The dependent variable is the forecast error (***FERROR***) multiplied by 100. ***HSR*** is a dummy variable equals to one if the HSR opens in the city where the firm located, and zero otherwise. And we differentiate the firms with the distance between firms and information centers. We define a firm as near to information centers firm if the distance between the firm to information centers is below the sample median in a specific year; otherwise, we define a firm as far to information centers firm. All variables are defined in section 2 and Appendix 1. Fixed effects of year, industry, firm and analyst are controlled. The t-statistics reported in parentheses are heteroscedasticity robust. Significance at the 10%, 5%, and 1% level is indicated by *, **, ***, respectively.

6. Conclusions

This study provides evidence linking China's HSR to analyst forecast accuracy. Exploiting the quasi-random variation in a city's connection to the HSR network, we use the DID methodology to analyze the impact of ***HSR*** on ***FERROR***. We find that the HSR connection could decreases ***FERROR*** significantly. Our findings are robust to different specifications and addressing endogeneity. We argue that HSR improves the forecast accuracy of analysts by increasing the probability of analysts' visit of listed firms rather than via an increase in analyst coverage or an improved firm information environment. We further verify the mechanism by providing evidence on other outcomes and the heterogeneity. The opening of HSR has its largest effect on firms with high earnings quality, firms audited by Big 4 firms, and firms located near an information center.

Appendix A. Variable definitions

Variables	Definitions
Main variables	
HSR	A dummy variable *HSR* as 1 for the treatment group after the HSR opening year and 0 before the opening year; *HSR* constantly takes 0 for the control group. Specifically, if opening time is in the first half of year t (before 1st July), we treat the HSR opening year a year t, t + 1 otherwise
FERROR	The firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t
FERROR(-1)	The firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t + 1
FERROR(-2)	The firm's absolute value of analyst forecast errors for forecasts made in year t for the earnings of year t + 2
HSR_D	A dummy variable *HSR_D* as 1 for the treatment group after the cities firm located and cities of the broker opening direct HSR and 0 before the opening year; *HSR_D* constantly takes 0 for the control group

TREAT	A dummy variable *TREAT* as 1 if the cities firm located are connected to the HSR network during the sample period and 0 otherwise.
Dis_Broker	Geographic distance in kilometres between analyst and firm based on latitudes and longitudes of their city locations
Visit	Indicator variable, which takes a value of one if a visit occurred for the (brokerage firm listed company) pair in a given year t, and 0 otherwise
Coverage	Natural logarithm of a company's analyst coverage plus one
DA	DA is absolute value of a company's abnormal accruals in year t, using a modified Jones model (Jones, 1991; Dechow et al., 1995)
Big 4	Dummy variable that takes the value of 1 (0) if the firm is audited by a Big 4 (non-Big 4) auditor firms
Dis_Center	The minimal geographic distance in kilometres between information centre (Beijing, Shanghai, Guangzhou and Shenzhen) and firm based on latitudes and longitudes of their city locations
Mob_tt	The total population of immigration and emigration of cities firm located
Mob_in	The population of immigration of cities firm located
Mob_out	The population of emigration of cities firm located
Control variables	
OP_DEBT	Company's operating cash flow divided by total debt
SIZE	Natural logarithm of a company's total assets
LEV	Debt-to-assets ratio of a company
EV	The standard deviation of actual earnings per share for the last 3 years
LOSS	Indicator variable, which equals one if the company's net profit is negative, and zero otherwise
VAR_RET	The standard deviation of stock returns over the past 36 months
AGE	Natural logarithm of the number of years since the firm is established
Top10	The shareholding ratio of the top ten shareholders
Intan	The ratio of intangible assets to total assets
Air	A dummy variable, Air as 1 if the city where a firm is located has at least one airport in use; otherwise, Air takes 0

References

Boubakri, N., Guedhami, O., Saffar, W., 2016. Geographic location, foreign ownership, and cost of equity capital: evidence from privatization. J. Corp. Finan. 38, 363–381.
Brown, L.D., Call, A.C., Clement, M.B., Sharp, N.Y., 2015. Inside the 'Black Box' of sell-side financial analysts. J. Account. Res. 53 (1), 1–47.
Campante, Filipe, Yanagizawa-Drott, David, 2018. Long-range growth: economic development in the global network of air links. Q. J. Econ. 133 (3), 1395–1458.
Chen, C., Hall, P., 2011. The impacts of high-speed trains on British economic geography: a study of the UK's InterCity 125/225 and its effects. J. Transp. Geogr. 19 (4), 689–704.
Cormier, D., Houle, S., Ledoux, M., 2013. The incidence of earnings management on information asymmetry in an uncertain environment: some Canadian evidence. J. Int. Account. Audit. Tax. 22 (1), 26–38.
Coval, J., Moskowitz, T., 1999. Home bias at home: local equity preference in domestic portfolios. J. Financ. 54 (6), 2045–2073.
Dechow, P., Sloan, R., Sweeney, A., 1995. Detecting earnings management. Accounting Review 70, 193–226.
Demers, E., 2002. Discussion of high-technology inatangibles and analysts' forecasts. J. Account. Res. 40 (2), 313–319.
Dhaliwal, D.S., Radhakrishnan, S., Tsang, A., Yang, Y.G., 2012. Nonfinancial disclosure and analyst forecast accuracy: international evidence on corporate social responsibility disclosure. Account. Rev. 87 (3), 723–759.
Faber, B., 2014. Trade integration, market size, and industrialization: evidence from china's national trunk highway system. Rev. Econ. Stud. 81 (3), 1046–1070.
Francis, J.R., Wang, D., 2008. The joint effect of investor protection and big 4 audits on earnings quality around the world. Contemp. Account. Res. 25 (1), 157–191.
Francis, J., LaFond, R., Olsson, P., Schipper, K., 2005. The market pricing of accruals quality. J. Account. Econ. 39 (2), 295–327.
Giroud, X., 2013. Proximity and investment: evidence from plant-level data. Q. J. Econ. 128 (2), 861–915.
Gul, F.A., Kim, J.B., Qiu, A.A., 2010. Ownership concentration, foreign shareholding, audit quality, and stock price synchronicity: evidence from China. J. Financ. Econ. 95 (3), 425–442.
Han, B., Kong, D., Liu, S., 2018. Do analysts gain an informational advantage by visiting listed companies? Contemp. Account. Res. https://doi.org/10.1111/1911-3846.12363.
Hochberg, Y.V., Ljungqvist, A., Lu, Y., 2007. Whom you know matters: venture capital networks and investment performance. J. Financ. 62 (1), 251–301.
Holl, A., 2016. Highways and productivity in manufacturing firms. J. Urban Econ. 93, 131–151.
Ke, B., Lennox, C.S., Xin, Q., 2015. The effect of China's weak institutional environment on the quality of big 4 audits. Account. Rev. 90 (4), 1591–1619.
Jones, J., 1991. Earnings management during import relief investigations. Journal of Accounting Research 29, 193–228.
Ke, X., Chen, H., Hong, Y., Hsiao, C., 2017. Do china's high-speed-rail projects promote local economy?—new evidence from a panel data approach. China Econ. Rev. 44, 203–226.
Kross, W., Ro, B., 1990. Earnings expectations: the analysts' information advantage. Account. Rev. 65 (2), 461–476.
Lang, M., Lundholm, R.J., 1996. Corporate disclosure policy and analyst behavior. Account. Rev. 71 (4), 467–492.
Lin, Y., 2017. Travel costs and urban specialization patterns: evidence from china's high speed railway system. J. Urban Econ. 98, 98–123.
Loughran, T., 2007. Geographic dissemination of information. J. Corp. Finan. 13 (5), 675–694.
Malloy, C.J., 2005. The geography of equity analysis. J. Financ. 60 (2), 719–755.
Mayer, Thierry, Trevien, Corentin, 2017. The impact of urban public transportation evidence from the Paris region. J. Urban Econ. 102, 1–21.
Merkley, K., Michaely, R., Pacelli, J., 2017. Does the scope of the sell-side analyst industry matter? An examination of bias, accuracy, and information content of analyst reports. J. Financ. 72 (3), 1285–1334.
O'Brien, Patricia C., Tan, Hongping, 2015. Geographic proximity and analyst coverage decisions: Evidence from IPOs. Journal of Accounting and Economics 59 (1), 41–59.
Platikanova, P., Mattei, M.M., 2016. Firm geographic dispersion and financial analysts' forecasts. J. Bank. Financ. 64 (3), 71–89.
Pouget, S., Sauvagnat, J., Villeneuve, S., 2017. A mind is a terrible thing to change: confirmatory bias in financial markets. Rev. Financ. Stud. 30 (6), 2066–2109.
Qin, Y., 2017. 'No county left behind?'the distributional impact of high-speed rail upgrades in China. J. Econ. Geogr. 17 (3), 489–520.
Roberts, M.R., Whited, T.M., 2013. Endogeneity in empirical corporate finance1. In: Handbook of the Economics of Finance. 2. pp. 493–572.
Sabherwal, S., Smith, S.D., 2008. Concentrated shareholders as substitutes for outside analysts. Corp. Gov. 16 (6), 562–577.
Soltes, E., 2014. Private interaction between firm management and sell-side analysts. J. Account. Res. 52 (1), 245–272.
Wang, Y., Liang, S., Kong, D., Wang, Q., 2019. High-speed rail, small city, and cost of debt: firm-level evidence. Pac. Basin Financ. J. 57, 1–19.
Yang, X., Jin, Z., Tan, J., 2019. Foreign residency rights and companies' auditor choice. China J. Account. Res. 12 (1), 93–112.
Zhang, L., Wei, M., 2020. How analyst recommendations respond to corporate uncertainty caused by investment behavior. China Finance Rev. Int. https://doi.org/10.1108/CFRI-05-2019-0046. (forthcoming).