No. 2208 [EN]

IMI Working Paper

Need for Speed: High-speed Rail and Firm Performance

Kuang Chun, Liu Zijie and Zhu Wenyu

INTERNATIONAL MONETARY INSTITUTE

Journal of Corporate Finance 66 (2021) 101830

Contents lists available at ScienceDirect

Journal of Corporate Finance

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

Need for speed: High-speed rail and firm performance☆

Chun Kuang [a], Zijie Liu [b], Wenyu Zhu [c,*]

[a] School of Banking and Finance, University of International Business and Economics, 10 Huixin East Street, Beijing 100029, China
[b] Simon Business School, University of Rochester, Rochester, NY 14627, USA
[c] School of Finance, Renmin University of China, 59 Zhongguancun Avenue, Mingde Building, Beijing 100872, China

ARTICLE INFO

JEL codes:
G30
D22
H54

Keywords:
Communication cost
High-speed rail
Firm location
Firm performance

ABSTRACT

Exploiting the staggered expansion of China's passenger-dedicated high-speed rail (HSR) network, we study the relationship between HSR connection and firm performance. By highlighting the importance of firm location and face-to-face interactions, we test the differential impact of HSR connection on the performance and growth of firms that differ in their dependence on communication and inter-city travel as production inputs. Our results confirm that firms in communication-intensive and travel-dependent industries benefit more from the operation of the HSR. Moreover, in examining the specific mechanisms at work, we find evidence that the HSR promotes firm performance and growth for the communication-intensive and travel-dependent through increased analyst attention, productivity boosts, and market expansions. Our findings imply that face-to-face interactions, through facilitating the exchange and acquisition of information, are potentially important for our understanding of "local (home) bias" of investment documented in the literature.

1. Introduction

In spite of astonishing advances in information technology and remote communication in recent decades, firms' geographical locations, and physical distances in general, remain crucial for firm growth. Lerner (1995) documents some of the earliest evidence on the importance of firm location in the finance literature, by showing that the venture capitalists' monitoring costs vary positively with

☆ This paper was previously circulated under the title "Public Capital and Firm Performance: Micro-Evidence from China's High-Speed Rail". The authors would like to thank Rémi Jedwab for his encouragement and advice on an earlier version of this paper, and are grateful for helpful comments from Douglas Cumming (the editor), three anonymous referees, Richard Arnott, Shiyu Bo, Shihe Fu, Yizhen Gu, Yatang Lin, Shimeng Liu, Yongzheng Liu, Wei Shi, Xuezhen Tao, Xue Wang, Tony Yezer, Chaoqun Zhan, and seminar participants at Jinan University, Xiamen University, Tsinghua University, and 2020 FMA Virtual Conference. Zhu acknowledges funding supported by fund for building world-class universities (disciplines) of Renmin University of China (KYGJC2020008). Kuang acknowledges funding supported by the fundamental research funds for the central universities in UIBE (19QN07).

* Corresponding author at: School of Finance, Renmin University of China, 59 Zhongguancun Avenue, Mingde Building, Beijing 100872, China.
E-mail addresses: chunkuang@uibe.edu.cn (C. Kuang), zliu88@simon.rochester.edu (Z. Liu), wenyuzhu@ruc.edu.cn (W. Zhu).

https://doi.org/10.1016/j.jcorpfin.2020.101830
Received 21 April 2020; Received in revised form 2 December 2020; Accepted 8 December 2020
Available online 15 December 2020
0929-1199/

distance.[1] Location remains important because face-to-face interactions are highly efficient. They are less subject to information friction and, therefore, cannot be replaced by virtual communications without some degree of information loss.[2] In most of the existing studies in corporate finance, firm locations are often treated as exogenous and fixed. However, major transportation infrastructure improvements could change the relative distances between firms and the rest of the market, by reducing the cost for people in different locations to interact, communicate, and exchange information face-to-face.

In this paper, we focus on a unique form of transportation system, the high-speed rail (HSR) network, and study its impact on firms' performance and growth. China's unprecedented HSR program entails a $300-billion investment to build a national network. The HSR network in China reached 35,000 km by the end of 2019, and carried 2.358 billions passengers in 2019 alone.[3] Unlike traditional railways, which transport both passengers and goods, HSR is almost entirely dedicated to passenger travel. By facilitating travel across cities, HSR prompts the flow of information and the exchange of ideas. The potential impact on face-to-face interactions across firms located in different cities is the object of this study. The differential effect across industries is hypothesized to be dependent on their communication and travel intensity. We proceed under the assumption that face-to-face interactions are productive and, more importantly, cannot be effectively replaced by remote virtual communications. Under these premises, we project that communication-intensive and travel-dependent firms that rely more heavily on face-to-face interactions and inter-city travels would benefit more from the operation of the HSR.

We construct an annual panel of firm-level data over the period of 2004 to 2017, augmented with information on individual firm's connection to the HSR network in each year. Our "treated" industries are defined as those that are both more communication-intensive and travel-dependent.[4] As there are no readily available firm or industry level measures of communication intensity or travel intensity, we first rank industries by the importance of interactive labor and inter-city travel inputs in their production. We adopt the non-routine interactive measures produced by Autor et al. (2003) for communication intensity, and the share of passenger travel input in total inputs according to the input-output tables for all industries to measure travel intensity. Industries that rank among the top half both in terms of interactive skills and passenger travel inputs make up the "treated" group in our sample. Communication-intensive and travel-dependent industries, such as education, and business services, are identified as our "treated" industries, while industries for which face-to-face communication is not as important a production input constitute the control group, including manufacturing, and information technology services.

Empirical issues such as the non-random placement of the HSR lines, possible endogenous entry or relacation of firms, and survival bias are dealt with using various sample restrictions targeted at each of these identification challenges, as we detail in Section 3.2. Exploiting the plausible randomness in the order in which cities became connected to the HSR network, we study the relationship between firm performance and HSR connection for industries that differ in terms of communication and travel intensity in a difference-in-differences (DiD) framework, with firm and year fixed effects as well as firm-level and city-level characteristics. Highlighting the unique feature of the HSR as a transportation network, we hypothesize that communication-intensive and travel-dependent firms benefit more from the operation of the HSR.

Our empirical results confirm that the HSR is particularly beneficial to the performance and growth of firms in the "treated" industries, compared to other industries, and that the effect is persistent over time. We demonstrate the robustness of our results (1) using a sub-sample with single-location firms, (2) by instrumenting the HSR connection with historical railways interacted with a hypothetical "connection order" based on relative elevations, as well as (3) with matching estimators. Results with an alternative measure of travel intensity constructed using travel expenses as a fraction of administrative expenses, and falsified "treated" industries demonstrate the robustness of our definition of the "treated" industries.

In uncovering the mechanisms at work, we directly test three potential mechanisms through which "treated" firms might benefit more from the HSR network. We first document the information channel of the HSR by showing that "treated" firms enjoy increased analyst attention upon being connected to the HSR network, especially those that are facing higher levels of information asymmetry (i.e., firms that are located in smaller cities or more financially constrained). Access to the HSR network also serves as a positive productivity shock to firms that are highly dependent on communication and inter-city travels in their daily operations, allowing employees in these firms to do more with the same or even fewer resources. Our results show that the "treated" group experiences larger productivity boosts than the control group after being connected to the HSR network. Moreover, "treated' firms, especially younger ones, see significant reduction in the share of their home province in overall revenue after being connected to the HSR network, indicating substantial market expansion beyond their home markets.

We further explore the heterogeneity of the HSR connection in terms of travel time savings from high-speed trains compared with regular-speed trains, and find that the improvements in firm performance and size growth for the "treated" firms are higher with larger savings in travel time. Among the "treated" firms connected to the HSR network, it is reasonable to expect those connected to mega

[1] In economics, there is a vast literature suggesting that freight rates on output influence market size and firm location. However, effects of changing passenger transportation cost are less well understood.

[2] We distinguish between routine versus non-routine interactions following Autor et al. (2003). We focus on non-routine interactions in this study and argue that routine face-to-face interactions are more likely to take place through remote communications. In other words, we assume that the degree of information loss over virtual communication is significantly larger for non-routine interactions than for routine interactions.

[3] See the 2019 Railway Safety Announcement released by the National Railway Administration of People's Republic of China (http://www.nra.gov.cn/jgzf/zfjg/zfdt/202003/t20200327_107025.shtml, accessed on April 5, 2020).

[4] We discuss in detail how we define the "treated" industries in Section 3.1. In Section 4.2.3, we show that our results are robust to an alternative proxy for travel intensity.

cities to perform even better, owing to the classic argument regarding agglomeration economies.[5] With hand collected data on the timing of direct HSR trips between each city and the four mega cities, we measure the number of mega cities that firms are connected via direct trips on the HSR for each year, as well as each firm's relative "distance" (measured in travel time) to their nearest mega cities. The results show that "treated" firms benefit from the HSR connection to more mega cities, confirming the importance of agglomeration economies. Among those connected to mega cities, "treated" firms within 3 to 6 hours away from the nearest mega cities benefit even more, consistent with more intensive utilization of the HSR in medium-range business travels.

This paper speaks to the corporate finance literature on the location of firms.[6] In most of the current literature, the relationship between firm location and other firm-level outcomes is treated as a fixed parameter. One of the notable exceptions is Giroud (2013), where the author studies the relation between headquarters' proximity to plants and plant-level investment, using the introduction of new airline routes as a key source of variation. Similar to this paper, we also highlight the importance of face-to-face interactions and travel distance in understanding the role of location and proximity in corporate development. Instead of focusing on the role of proximity in shaping firms' internal capital allocation, we examine the differential impact of a positive shock in transportation infrastructure that changes the relative distance between firms and the rest of the market. Our results demonstrate that communication-intensive and travel-dependent firms benefit significantly more from the operation of HSR, growing faster in profitability and scale upon the reduction of communication costs and travel distances.

Our paper is also related to the literature on geographical proximity and financial interactions, where "local bias" or "home bias" have been documented for both institutional and individual investors.[7] For example, Coval and Moskowitz (2001) and Sialm et al. (2020) find that locally biased mutual funds and funds of hedge funds enjoy abnormal returns in local stocks and fund investments. Ivković and Weisbenner (2005) and Cumming and Dai (2010) show that geographical proximity is important in shaping investment preferences for individual investors and venture capitalists (VCs) respectively, as they benefit from local knowledge. Bernstein et al. (2016) is among one of the few papers that document the travel time reducing effect of transportation on the relationship between VCs and their portfolio companies. Using survey data of VC investors, the authors find that the introduction of new airline routes facilitates the interactions between VCs and their portfolio companies, and in turn positively affects the innovation and success of these companies. Instead of looking at the effect of transportation on the relationship between VCs and their investing firms, we focus on its effect on the performance and growth of the more "affected" firms. In uncovering the underlying mechanisms at work, we find that communication-intensive and travel-dependent firms, especially those that are more likely to suffer from information asymmetries (i.e., those that are more financially constrained or located in non-mega cities) gain significantly more analysts' coverage upon being connected to the HSR network. Our findings imply that face-to-face interactions, by facilitating the exchange and acquisition of knowledge, are potentially important for our understanding of "local (home) bias".

This paper contributes firm-level evidence to the literature on the relation between transportation infrastructure and firm performance, by focusing on communication-intensive and travel-dependent firms and the HSR network in China.[8] Charnoz et al. (2018) and Bernard et al. (2019) document that firms are more likely to outsource their production to remote affiliates and to other firms, respectively, after being connected to the HSR. Nadiri and Mamuneas (1994), Shah (1992), and Li and Li (2013) study the impact of public infrastructure on the performance of manufacturing firms in the US, Mexico, and China, respectively. A recent paper by Bernard et al. (2019) takes a theoretical perspective in studying the impact on supply chains and firm performance in the context of high-speed rail in Japan. We highlight the importance of firm location and face-to-face interactions, and study the impact of the HSR on firms' performance and growth through substantial reductions in communication costs.

Finally, our study also adds to a growing literature on the spatial impacts of China's HSR network. Examining the impact of the HSR network on regional growth, Qin (2017) and Yu et al. (2018) document a decrease in GDP per capita in connected regions after being connected to the HSR network, while Zheng and Kahn (2013) and Ke et al. (2017) argue that smaller and more industrialized cities grow faster after being connected, respectively. Modeling sector specialization across cities, Lin (2017) finds that industries with a higher reliance on non-routine cognitive skills benefit more from HSR connection in terms of employment growth. Focusing on a a clearly defined set of communication-intensive and travel-dependent firms, our study complements her findings by providing detailed firm-level evidence on the underlying mechanisms at work as well as heterogeneity in the impact of HSR on firm performance and growth.

The remainder of the paper is structured as follows. Institutional background of the HSR planning is introduced in Section 2, along

[5] The four mega cities are Beijing, Shanghai, Guangzhou, and Shenzhen.

[6] There is a substantial body of literature on the potential impact of locational factors on various firm-level outcomes, including capital market performance (e.g., Baik et al., 2010; Bernile et al., 2015; Pirinsky and Wang, 2006), capital structure (e.g., Booth et al., 2001; Gao et al., 2011; Rajan and Zingales, 1995), financing activities (e.g., Arena and Dewally, 2012; Loughran, 2008), profitability (e.g., Carosi, 2016), ownership structure (e.g., Boubakri et al., 2016 geographic), dividend payout policy (e.g., John et al., 2011), and corporate governance (e.g., Chhaochharia et al., 2012).

[7] In addition, distance also plays a very important role in the bank-firm relationship literature. Firms that are more distant to their lenders have less access to credit and bear higher costs for loans. See, for example, Petersen and Rajan (2002), Degryse and Ongena (2005, 2007), and Sufi (2007).

[8] A substantial body of the literature has been devoted to examining the effect of transportation infrastructure systems, such as the highway system and the HSR network, on regional development, labor market outcomes, firms' location choices, business pattern across space, and interregional trade. Notable examples include the interstate highway system in the US (e.g., Michaels, 2008), the Chinese national highway network (e.g., Banerjee et al., 2020; Faber, 2014), the Indian railway network (e.g., Donaldson, 2018), the HSR networks in Europe, Japan, and more recently China (e.g., Charnoz et al., 2018). Redding and Turner (2015) provide a literature review on the impact of transportation cost on spatial economic activities.

with detailed data descriptions. Section 3 outlines our empirical framework, and describes our "treated" firms as well as our identification assumptions and strategy. The results are presented and discussed in Section 4, along with a series of robustness checks. Section 5 proposes and investigates three underlying mechanisms at work, and further demonstrate the relationship between HSR connection and firm performance with a series of extensional exercises. Section 6 concludes.

2. Institutional background and data

We start with a brief introduction of the institutional background of the HSR network - how it was initially proposed and how the purpose of the network evolved over time. Details about our data, sample, and measures are described afterwards, including information on the HSR network, as well as firm financials and city-level characteristics.

2.1. Institutional background of the HSR network

The construction of China's first high-speed passenger-dedicated railway, Qinhuangdao-Shenyang passenger railway, started on August 16, 1999, and was proposed in China's Ninth Five-Year Plan. Operating since October 12, 2003, it marks the transition in China's railway upgrading from raising speed along existing lines to constructing new lines for designated purposes (i.e., freight vs. passenger). The construction of a high-speed railway network was first proposed in the Medium- and Long-Term Railway Plan (MLTRP) in 2004, at a time when freight services were overloading the existing network. A passenger-dedicated HSR network, with four horizontal and four vertical corridors linking all major cities, was therefore proposed to transfer passengers to the new network, so as to increase the capacity for freight services on the original railways. Therefore, except a corridor along the southeast coast (i.e., the Hangzhou-Shenzhen corridor), all the other main HSR corridors in the original plan paralleled existing lines. When the MLTRP was further revised in 2016, it expanded the network structure to eight horizontal and eight vertical corridors, with a revised purpose to "promote regional connectivity and economic development" (Lawrence et al., 2019).

2.2. Data

Our empirical analysis relies on three sets of data: (1) on the construction and operation of the HSR network, (2) on firm-level measures, and (3) on city-level characteristics. HSR data are mainly from the China Research Data Service (CNRDS) database, firm-level data from the China Stock Market and Accounting Research (CSMAR) and the Wind databases, and city-level data are from the China City Census.[9]

We construct an annual panel of firm-level measures, augmented with information about the operation and expansion of the HSR network, to investigate the relationship between HSR connection and firm performance. Our original firm sample comprises all companies listed on the Shanghai and the Shenzhen Stock Exchanges in China. The study period is from 2004 to 2017, considering that 2003 is the year when the first HSR line, the Qinhuangdao-Shenyang Line, was put into use. We exclude from our analyses firms in industries that are directly related to HSR operations, including (1) construction, (2) real estate, (3) transportation, warehousing and postal services, and (4) hotel and catering services, to limit the possibility of confounding the HSR effects.[10]Table 1 reports the full list of industries and the number of firms in each industry in 2017.

2.2.1. HSR data

Our HSR data are mainly from the CNRDS database, which covers information about HSR lines, stations, and trains. According to the official guidelines published by the Ministry of Railways of China, HSR lines are newly built railway lines designated for passenger travel, running at a speed of 250 km/h or above.[11] The first HSR line, the Qinhuangdao-Shenyang Line, became operational in 2003. Construction of the HSR network began to boom in 2008, after four years of dormancy. By the end of 2017, the HSR network had 96 completed lines and segments, with a total mileage exceeding 20,000 km. Table 2 reports the number of new lines and segments completed, and new HSR stations opened each year from 2003 to 2017. Firms in our sample are identified by the cities in which they are headquartered, while each city is connected to the HSR network by the HSR stations in the city. The key explanatory variables in this study, $HSR_{c,t}$, indicates whether city c in year t was connected to the HSR network. To measure the duration of HSR connection over time, we construct $HSR^{Year}_{c,t}$, which indicates the number of years city c has been connected to the HSR network by year t. The Appendix provides more details on the construction of our HSR measures.

[9] For some measures, such as the number of mega cities each firm is connected to every year, we manually check via web search from other sources. A more detailed description is provided in our Appendix.

[10] The HSR network may benefit these industries directly through channels other than the ones studied in this paper. For example, the hotel and catering services industry in connected cities may benefit from the increasing number of tourists because of greater accessibility via the HSR. Similarly, the construction of HSR is also associated with rising house prices in nearby cities (Zheng and Kahn, 2013), which could in turn boost the construction and real estate sectors.

[11] Information is obtained from http://www.gov.cn/flfg/2013-02/20/content_2334582.htm. There are three types of HSR trains, denoted by G, D, and C. G refers to high-speed electric multiple unit (EMU) trains, D refers to general EMU trains and C refers to intercity trains. We do not distinguish between them in this study, as they are generally collectively referred to as high-speed rail trains due to their high travel speed.

Table 1
Industry Classification.

Code	Industry	Treated	# of Firms in 2017
A	Agriculture, Forestry, Livestock Farming, Fishery	0	27
B	Mining and Quarrying	0	39
C	Manufacturing	0	1237
D	Production and Supply of Electricity, Heat, Gas, and Water	0	73
E	Construction*	(0)	(57)
F	Wholesale and Retail Trades	1	118
G	Transportation, Warehousing, and Post*	(0)	(68)
H	Hotels and Catering Services*	(0)	(7)
I	Information Transmission, Software, and Information Technology Services	0	146
J	Finance and Insurance	1	7
K	Real Estate*	(0)	(75)
L	Leasing and Business Services	1	26
M	Scientific Research and Technical Services	1	18
N	Management of Water Resources, Environment, and Public Utility	1	24
O	Household Services, Repair, and Other Services	0	0
P	Education	1	1
Q	Public Health and Social Services	0	4
R	Culture, Sports, and Entertainment	1	26
S	Public Administration, Social Security, and Social Organizations**	(1)	–

Notes: The original sample includes industries A-S. *We exclude firms belonging to industries directly related to HSR operations, including (1) construction, (2) transportation, warehousing, and post, (3) hotels and catering services, and (4) real estate. **We exclude firms belonging to industry "S" due to discrepancies in the definitions for industry "S" from the National Bureau of Statistics of China and the China Securities Regulatory Commission.

Table 2
Expansion of the HSR Network.

Year	# of Newly Constructed		# of Newly Connected		
	Segments	Stations	Mega Cities	Provincial Capitals*	Other Cities
2003	1	9	0	1	5
2004	0	0	0	0	0
2005	0	0	0	0	0
2006	0	0	0	0	0
2007	0	0	0	0	0
2008	3	20	1	4	4
2009	4	48	1	5	14
2010	8	79	1	5	19
2011	4	45	1	1	14
2012	7	50	0	1	23
2013	12	76	0	1	20
2014	16	142	0	5	39
2015	19	147	0	0	31
2016	12	108	0	1	11
2017	10	53	0	1	6

Notes: Cities and HSR stations on multiple line segments are only counted in the year of establishment. *The city of Guangzhou, although a provincial capital, is only included under Mega Cities to avoid double counting.

2.2.2. Firm-level data

We obtain firm-level financial data as well as information on firm headquarters locations from the CSMAR and the Wind databases. Each firm reports two addresses in its annual report, a business address and a registered address. When the two are different, we use business addresses to locate firms.[12] We focus on two aspects of firms' performance, namely profitability and size. Our profitability measures include (1) earnings before interest and tax (*EBIT*) ratio, which is calculated as EBIT over total assets, (2) return on assets (*ROA*), and (3) return on equity (*ROE*).[13] Firm size is measured by the natural logarithm of total assets (*lnAssets*) or number of employees (*lnEmployees*). We also control for a rich set of firm-level financial variables in our regressions. *lnAssets* as defined above is

[12] Only 2.18% of firm-year observations reported different cities for business and registration addresses. Our results are robust to using registered addresses instead. For the purpose of our research design, we prefer to use business address as it is where a firm's daily operations are carried out. We also collect information on the locations of firms' subsidiaries for robustness checks presented in Section 4.2.1. We describe in detail how the subsidiaries' locations are identified in the Appendix.

[13] We use the one-year lags of total assets as denominators for *EBIT* and *ROA* calculations. Similarly, the one-year lag of shareholders' equity is used to calculate *ROE*.

included except when the outcome variable is firm size. The leverage ratio (*Leverage*) is defined as the ratio of total debt to total assets, and is introduced as a control for firms' capital structure. *CashFlow*, which is the net operating cash flow normalized by total assets in the previous year, is included as a measure of firm liquidity. We also control for the market estimation of the firms' long-term growth opportunity using their market-to-book ratios, *q*. All firm-level financials are winsorized at the 1^{st} and 99^{th} percentiles of their empirical distributions of each year. Table 3 presents detailed definitions of all variables employed in the regressions, while Table 4 presents the summary statistics for the estimation sample. Firm-level data are merged with the HSR data using the locations of firms' headquarters to measure each firm's access to the HSR network as well as connection to other cities. Table 5 demonstrates the share of firms and cities being connected to the HSR network by year. The HSR network experienced a significant expansion between 2008 and 2011, growing from only 1.76% of firms being connected to the HSR in 2007 to more than 70% in 2011.

2.2.3. City-level data

We recognize the importance of city-level characteristics in relation to the staggered expansion of the HSR network, and control for a multitude of city-level attributes in our model. To control for local socioeconomic environment, which is crucial in determining the priority a city receives in the HSR network planning, we include real GDP (*lnRealGDP*) and population (*lnPopulations*). Given our focus on communication-intensive and travel-dependent industries, we control for local industry composition with the size of employment of the tertiary industry (*lnTertiaryEmployment*). Considering the potential substitution between air and HSR travel, we include the number of passengers in each airport each year, using data collected from the Civil Aviation Administration of China. *lnAirportPassengers* is the natural logarithm of 1 plus the total passengers.[14] The number of households with broadband connection (*lnInternetUsers*) is also introduced to capture penetration of communication technology across cities over time. We report in Table 6 summary statistics of all city-level characteristics both together for all cities in the estimation sample, and separately for mega and non-mega cities.

3. Empirical strategy

We describe our empirical framework and identification strategy in detail in this section. Admittedly, all firms in cities with HSR connections potentially benefit from the operation of the transportation network. However, since the HSR is designated to transport people across cities, rather than to ship freight, the potential impact of the HSR on industries varies widely with their dependence on face-to-face interactions and inter-city travel opportunities. We begin by describing how we identify industries that are most likely to benefit significantly from the HSR network (i.e., "treated" industries). As an important part of our identification strategy, we discuss our sample restrictions and working hypotheses, before formally describing the difference-in-differences empirical framework.

3.1. Identifying "treated" industries

In the spirit of Fernald (1999), who studies the effect of highways on the productivity of vehicle-intensive firms, we hypothesize that our "treated" firms should be those who most heavily utilize the HSR network in their daily operations. However, it is difficult to measure the importance of inter-city travel via the HSR at the firm level as no such measures are readily available.[15] Therefore, we rank industries by the importance of (1) non-routine face-to-face interactions, as well as (2) inter-city travels by rail, to identify industries that are the most likely to benefit significantly from the HSR operation.

We define our "treated" industries as those that are both more communication-intensive and travel-dependent than other industries. For communication intensity, we use the industry-level non-routine interactive skill intensity measure constructed by Autor et al. (2003). They measure different task intensities at the occupation level in the following five categories based on the Dictionary of Occupational Titles (DOT): routine manual, routine cognitive, non-routine manual, non-routine cognitive/interactive, and non-routine cognitive/analytical.[16] We manually match our industries to the industry-level DOT measures in Autor et al. (2003), which they aggregated from the occupation level.[17] To measure travel dependence, we rank industries by the share of rail passenger travel input in total inputs, using information reported in the input-output table for all industries published by the National Bureau of Statistics of China in 2017.[18] We refer to industries that rank among the top half in terms of rail passenger input simply as "travel-dependent", as air passenger input yields the same ranking. Such overlap between rail and air passenger input rankings confirms that our measure identifies demand for inter-city travel as a production input, rather than travel demand induced by the operation of the HSR, at the

[14] By construction, *lnAirportPassengers* for cities with no airport is zero.

[15] As a robustness check, in Section 4.2.3 we construct an alternative proxy for travel intensity using the ratio of travel expenses over administrative expenses at the firm level.

[16] According to their definition, routine tasks are those that can be completed by following clearly defined rules, while non-routine tasks involves more "problem-solving and complex communication activities." Since we only focus on non-routine interactive skills in this study, we use "interactive" instead of "non-routine interactive" for brevity.

[17] We manually match our industries to the consistent 60–90 NAICS codes in Author et al. (2003). All five DOT measures at the occupation level are aggregated to industry level where higher values indicate more intensive use of the task.

[18] The reporting format of industry inputs varies year to year. Freight shipping input and rail passenger travel input were separately reported only in 2002 and 2017, not only for rail but also for air and land. In other years, either these two were reported as one rail transportation input, or all transportation inputs by road, rail, air and marine were reported as one. Our results are robust to using the 2002 data as well as the average of the 2002 and 2017 data.

Table 3
Variable Definitions.

Variable	Definition
$COMM^{High}$	indicator for communication- and travel-intensive (i.e. "treated") industries, as listed in Table 1
HSR	indicator for being connected to the HSR network
HSR^{Year}	number of years of being connected to the HSR network
$HSR^{TimeSaving}$	average savings in travel time from HSR compared with regular speed train (in hours)
HSR^{Mega}	indicator for being connected to at least one mega city via HSR with direct trip
$HSR^{\#Mega}$	number of mega cities connected to via HSR with direct trip
	Firm-level Characteristics
EBIT	earnings before interest and taxes scaled by total asset
ROA	return on assets, calculated as the ratio of net profit over total assets
ROE	return on equity, calculated as the ratio of net profit over equities
lnAssets	logarithm of total assets
lnEmployees	logarithm of the number of employees
Leverage	ratio of liabilities over total assets
CashFlow	ratio of net operating cash flow over total assets
q	ratio of market capitalization over total assets
lnAnalystCoverage	logarithm of one plus the number of analysts covering a given firm*
TFP^{LP}	TFP estimated using Levinsohn and Petrin (2003) procedure with a value-added production function
TFP^{ACF}	TFP estimated using Ackerberg et al. (2015) procedure with a value-added production function
ProfitPerCapita	profit per capita, calculated as net profit (in millions CNY) over the number of employees
%HomeRevenue	ratio of revenue from home province over total revenue
$\%TravelExpense^{ave}$	multi-year average of ratios of travel expenses over administrative expenses**
	City-level Characteristics
lnRealGDP	logarithm of real GDP
lnPopulations	logarithm of population
lnTertiaryEmployment	logarithm of employment in the tertiary sector
lnAirportPassengers	logarithm of one plus the average city-level passenger flow*
lnInternetUsers	logarithm of households with broadband connection

Notes: Total assets used in calculation of *EBIT*, *ROA*, and *CashFlow* are lagged by one year. Total shareholders' equity used to calculated *ROE* is also lagged by one year. Mega cities include Beijing, Shanghai, Guangzhou, and Shenzhen. In constructing HSR^{Mega} and $HSR^{\#Mega}$, Guangzhou and Shenzhen are treated as one metropolis as they are only about 80 miles apart. **lnAnalystCoverage* and *lnAirportPassengers* equal ln(1) or 0 for cities with no analyst coverage or an airport. **Multi-year averages, rather than actual travel expenses ratios, are used due to unsatisfactory data quality before 2010.

Table 4
Summary Statistics of Firm Financial Data.

	All firms					Treated firms				
	Mean	SD	p(25)	p(50)	p(75)	Mean	SD	p(25)	p(50)	p(75)
EBIT	0.07	0.10	0.03	0.06	0.10	0.07	0.08	0.04	0.06	0.09
ROA	0.05	0.08	0.01	0.04	0.07	0.05	0.06	0.02	0.04	0.07
ROE	0.09	0.17	0.03	0.07	0.13	0.10	0.14	0.04	0.09	0.15
lnAssets	21.99	1.25	21.12	21.82	22.64	22.27	1.48	21.26	22.03	22.92
lnEmployees	7.77	1.23	6.98	7.73	8.53	7.61	1.40	6.67	7.54	8.47
Leverage	0.46	0.20	0.31	0.46	0.61	0.53	0.20	0.39	0.54	0.68
CashFlow	0.06	0.14	0.01	0.05	0.10	0.05	0.25	0.01	0.05	0.11
q	2.47	1.86	1.37	1.90	2.89	2.18	1.53	1.30	1.73	2.56
lnAnalystCoverage	1.45	1.15	0.00	1.39	2.40	1.50	1.23	0.00	1.39	2.64
TFP^{LP}	14.15	1.53	13.11	13.81	15.06	16.30	1.28	15.51	16.28	17.17
TFP^{ACP}	9.86	3.28	8.91	9.53	10.54	14.66	1.68	13.78	14.61	15.60
ProfitPerCapita	0.12	0.86	0.01	0.04	0.10	0.19	0.59	0.02	0.07	0.19
$\%TravelExpense^{ave}$	2.30	1.45	1.27	1.97	2.93	2.35	1.50	1.26	2.03	3.39
%HomeRevenue	0.68	0.31	0.42	0.77	0.97	0.78	0.27	0.65	0.92	0.98
Number of observations			13,726					1852		

Notes:EBIT is the sum of net profit, corporate income tax, and financing cost, divided by total assets. *ROA* is the ratio of net profit over total assets. *ROE* is the ratio of net profit over total shareholders' equity. *lnAssets* and *lnEmployees* are natural logarithms of firms' total assets and employment size respectively. *Leverage* is the ratio of total debt to total assets. *CashFlow* is cash flow normalized by total assets in the year before. *q* is calculated as the market-to-book ratio. *lnAnalystCoverage* is the logarithm of one plus the number of analysts covering a given firm. TFP^{LP} and TFP^{ACF} are estimated with procedures outlined in Levinsohn and Petrin (2003) and Ackerberg et al. (2015) using a value-added production function. *ProfitPerCapita* is calculated as net profit (in millions) over total number of employees. $\%TravelExpense^{ave}$ is the average of the ratios of travel expenses over total administrative expenses. *%HomeRevenue* is the share of revenue from home province in total operating revenue in a given year. All variables are winsorized at the 1^{st} and 99^{th} percentiles of their empirical distributions of each year.

Table 5
HSR Connnectedness Over Time.

Year	Cities			Firms		
	% Connected to		*N*	% Connected to		*N*
	HSR Network	Mega Cities		HSR Network	Mega Cities	
2003	2.79	0.00	215	2.26	0.00	620
2004	2.79	0.00	215	2.17	0.00	644
2005	2.79	0.00	215	1.98	0.00	706
2006	2.79	0.00	215	1.93	0.00	724
2007	2.79	0.00	215	1.81	0.00	775
2008	6.98	0.47	215	21.05	1.78	841
2009	16.28	4.65	215	34.61	7.52	864
2010	27.91	7.91	215	60.61	17.52	919
2011	35.35	18.14	215	72.74	53.87	1086
2012	46.51	31.63	215	77.39	62.56	1234
2013	56.28	49.30	215	85.19	78.74	1317
2014	76.74	63.72	215	93.34	85.10	1322
2015	91.16	73.02	215	96.29	89.92	1320
2016	96.74	79.53	215	99.41	92.98	1354
2017	100.00	84.65	215	100.00	94.76	1354

Notes: The estimation sample includes only cities and firms connected to the HSR network by the end of 2017. Connection to mega cities measures the share of cities or firms with connection via HSR to at least one mega city.

Table 6
Summary Statistics of City-Level Characteristics.

	Mean	SD	p(25)	p(50)	p(75)
Panel A: All Cities					
lnRealGDP	24.70	0.89	24.07	24.58	25.29
lnPopulations	15.24	0.67	14.84	15.32	15.71
lnTertiaryEmployment	12.33	0.83	11.82	12.24	12.74
lnAirportPassengers	7.27	7.26	0.00	9.50	14.52
lnInternetUsers	12.98	1.17	12.22	12.99	13.75
Number of observations			1974		
Panel B: Mega Cities					
lnRealGDP	26.81	0.28	26.58	26.82	27.07
lnPopulations	15.85	0.71	15.48	16.12	16.42
lnTertiaryEmployment	14.59	0.67	14.04	14.51	15.23
lnAirportPassengers	17.57	0.57	17.13	17.59	18.06
lnInternetUsers	15.33	0.58	14.84	15.34	15.66
Number of observations			56		
Panel C: Non-Mega Cities					
lnRealGDP	24.63	0.83	24.06	24.55	25.22
lnPopulations	15.22	0.66	14.84	15.31	15.69
lnTertiaryEmployment	12.27	0.74	11.80	12.22	12.68
lnAirportPassengers	6.97	7.14	0.00	0.00	14.15
lnInternetUsers	12.91	1.11	12.19	12.97	13.68
Number of observations			1918		

Notes: The estimation sample includes only cities that were connected to the HSR network by the end of 2017. Mega cities include Beijing, Shanghai, Guangzhou, and Shenzhen.

industry level. We rank industries along these two dimensions and identify those that rank among the top half in both measures as our "treated" group. In other words, we identify firms that are more dependent on non-routine interactions and inter-city travels in their production process as firms that benefit more from being connected to the HSR network. The industry-level rankings are presented in Fig. 1, and those that rank among the top 10 both in terms of interactive skills and passenger travel inputs (i.e., in the upper right corner) make up the "treated" group in our sample. Ranked among the top both in terms of interactive skill and inter-city travel inputs are industries such as education, public administration, finance & insurance, technical services, and business services. Industries that are in the control group include agriculture, mining, manufacturing, social services, household services, and information technology services. We use $COMM^{High}$ as an indicator for firms that are more communication-intensive and travel-dependent in our empirical analysis.

Figs. 2 and 3 show plots of the distributions of firms in "treated" versus control industries across cities separately, along with the HSR network in operation by the end of 2016. While the number of firms in the control group is larger than that of the "treated" group, the two groups do not seem to differ in terms of their spatial distribution across locations. In Table 4, we report summary statistics for

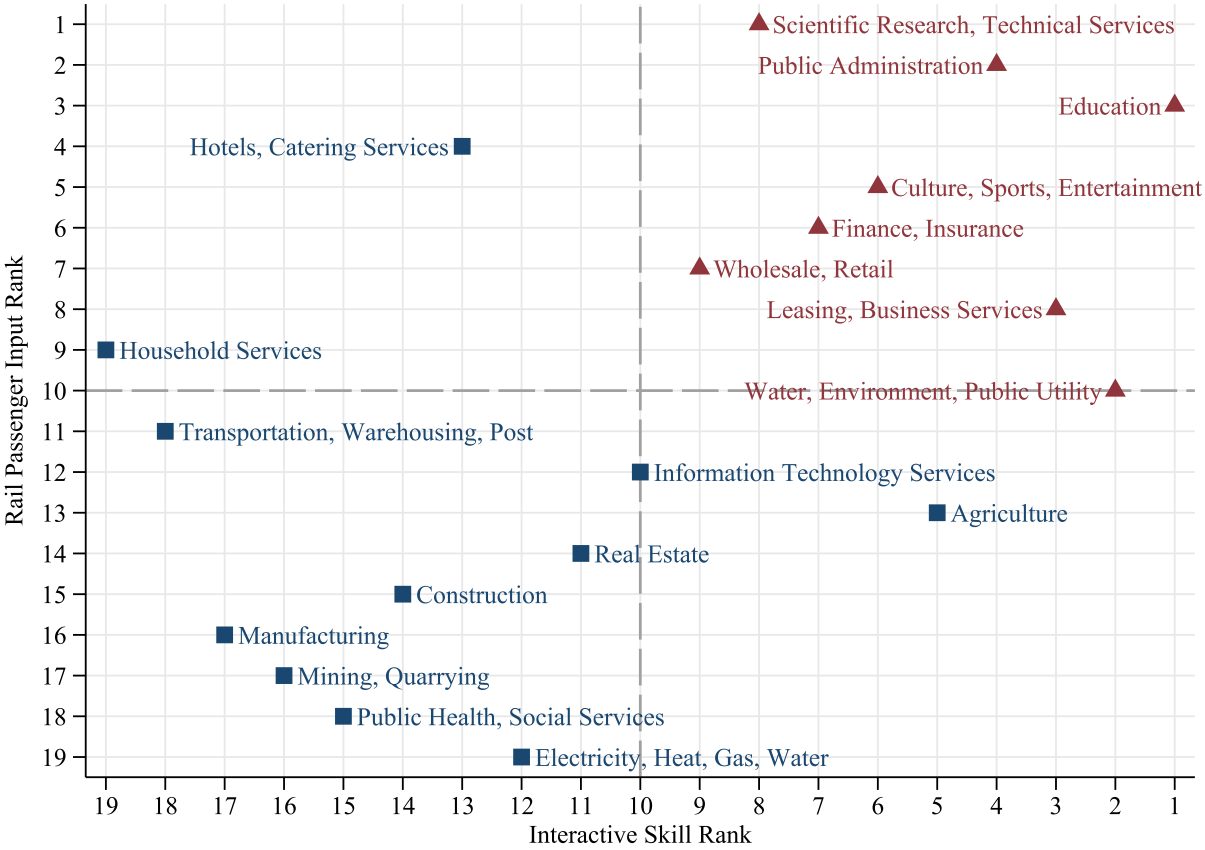

Fig. 1. Industry Rankings of Communication and Travel Intensity.
Notes: All industries are ranked by (1) non-routine interactive labor intensity (Autor et al., 2003, 2) share of passenger travel input. Firms in industries that rank among the top half for both criteria are identified as our "treated" group (red triangles). The rest of the industries form the control group (blue squares). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

our "treated" firms. Comparing with the full sample summary statistics, firms in the "treated" industries are similar in terms of profitability, but are slightly larger in asset size but slightly smaller in employment size.[19]

3.2. *Sample restrictions and identification assumptions*

As a first step to address endogeneity concerns due to the non-random placement of HSR lines and to ensure comparability of firms across locations, we restrict our empirical sample to firms headquartered in cities that were connected to the HSR network by the end of our sample period.[20] This restriction is important because it is possible that firms located in cities that are never connected to the HSR network might be inherently different from those with HSR connections. Given this sample restriction, our identifying variation comes from differences in the order in which a city is connected, rather than whether a city is connected, to the HSR network.

We believe there to be some plausible randomness in the connection order of cities. It is worth pointing out that the actual connection order of a city is not the same as the level of priority that it received in the design the network, as factors such as engineering difficulties also contribute to determining the order in which a city was actually connected. Should the order in which cities were connected reflect the importance of regional economies across the country, bilateral connections between the four mega cities would enjoy the highest priority. Yet, the four mega cities were connected in 2008 (Beijing), 2009 (Guangzhou), 2010 (Shanghai), and 2011 (Shenzhen), respectively.

First, the HSR network was designed to connect large cities, but mechanically it has to cross and connect many smaller cities along the way. In other words, the order in which cities were connected cannot be perfectly predicted by city characteristics such as the level

[19] To ensure that the two groups of firms are more comparable and that the unbalanced sample size of these two groups does not affect our results, we use the matching estimators with the difference-in-differences framework as one of our robustness checks in Section 4.2.5.

[20] In practice, we restrict our estimation sample to firms that are connected to the HSR network by the end of 2017, since the HSR connection indicators enter the regressions as one-year lags.

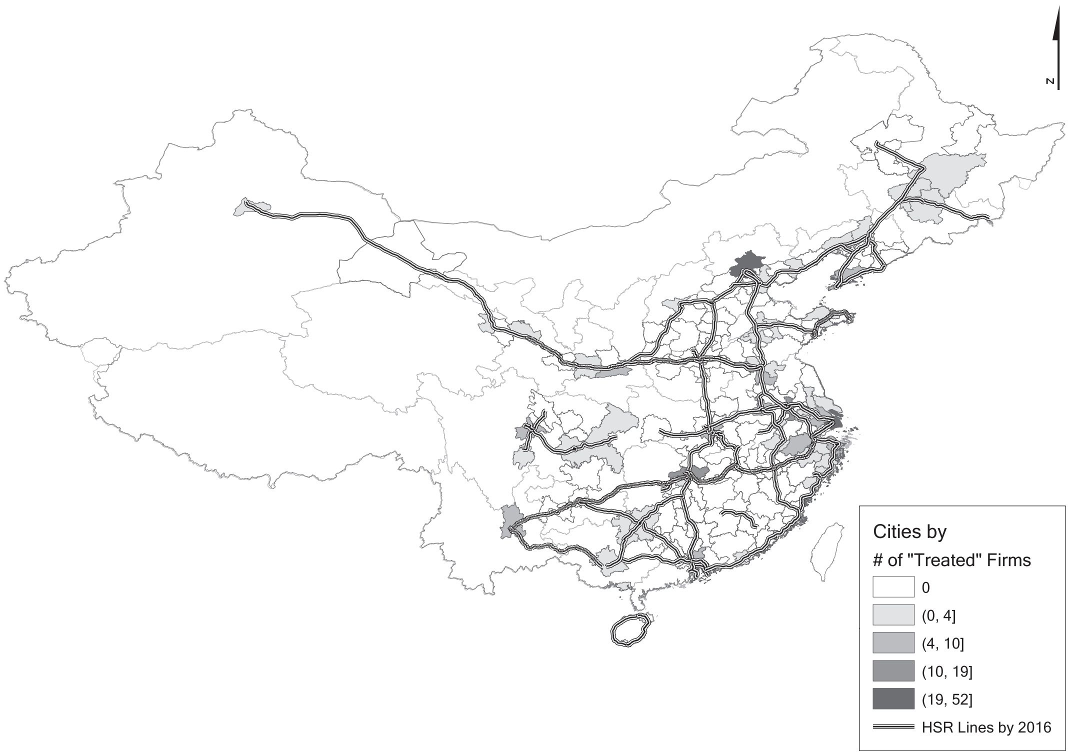

Fig. 2. Distribution of "Treated" Firms Across Cities.
Notes: The graph plots the distribution of the number of "treated" firms across cities identified by their headquarter locations. "Treated" firms refer to those in industries that rank in the top halves in communication and travel intensity in Fig. 1. The map of HSR lines is modified based on Li (2016).

or growth of local economic development and city size (i.e., population). We report year-by-year the number of newly connected cities in different categories (i.e., mega cities, provincial capitals, and other cities) in the right panel of Table 2. The gradual expansion of the HSR network can be seen from the table - not only large cities such as provincial capitals and mega cities, but also smaller cities in between were added to the network each year. Second, we plot the distribution of real GDP for newly connected cities by connection year, and do not observe a systematic pattern between the local economic development level and the order in which a city was connected. Fig. 4 presents the box plot of the distribution of real GDP for newly connected cities by connection year. The horizontal lines in the middle of the boxes correspond to the median real GDP among cities being connected in a given year, and the height of the boxes correspond to the interquartile ranges (i.e., IQR, the differences between 25th and 75th percentiles). Both the median and IQR of real GDP of the newly connected cities appear to be fluctuating over time, rather than trending monotonically up or down.

Firm locations need to be treated carefully, as they reflect, among other things, firms' preference for access to the HSR network. In an effort to eliminate such influence on our results, we first try to make firm location a pre-determined characteristic, by restricting our empirical sample to firms that were established before 2003 - one year before the HSR network was first proposed in 2004. However, any changes in firm locations could still lead to reverse causality. That is if more productive or profitable firms tend to adjust their location more responsively to changes in local shocks such as infrastructure improvements, then the observed positive association between firm performance and the HSR connection could be due to reverse causality. To guard against this possibility, we then manually check firms' business addresses in our sample period and exclude firms that have ever changed headquarter cities. This exclusion criteria is not trivial to our identification, as our results should not pick up the fact that firms respond to the productivity shock associated with the HSR operation by moving to a different location.

Our main empirical framework is a difference-in-differences analysis with firm and year fixed effects. With the previously discussed restrictions and assumptions, the identifying variations in our model are the differences in the timing of being connected to the HSR network across cities. Under these assumptions, the two layers of differences in our model are the differences between the communication-intensive and travel-dependent firms and those in the control industries, and the differences between firms before versus after they become connected to the HSR network. We believe that the within-(HSR-connected)-city cross-firm variations are just as important as the cross-city variations for identifying the impact of HSR connection on firm performance and growth, if not more so.

3.3. HSR connection and firm performance

We test the main hypothesis that firms in communication-intensive and travel-dependent industries benefit more from HSR con-

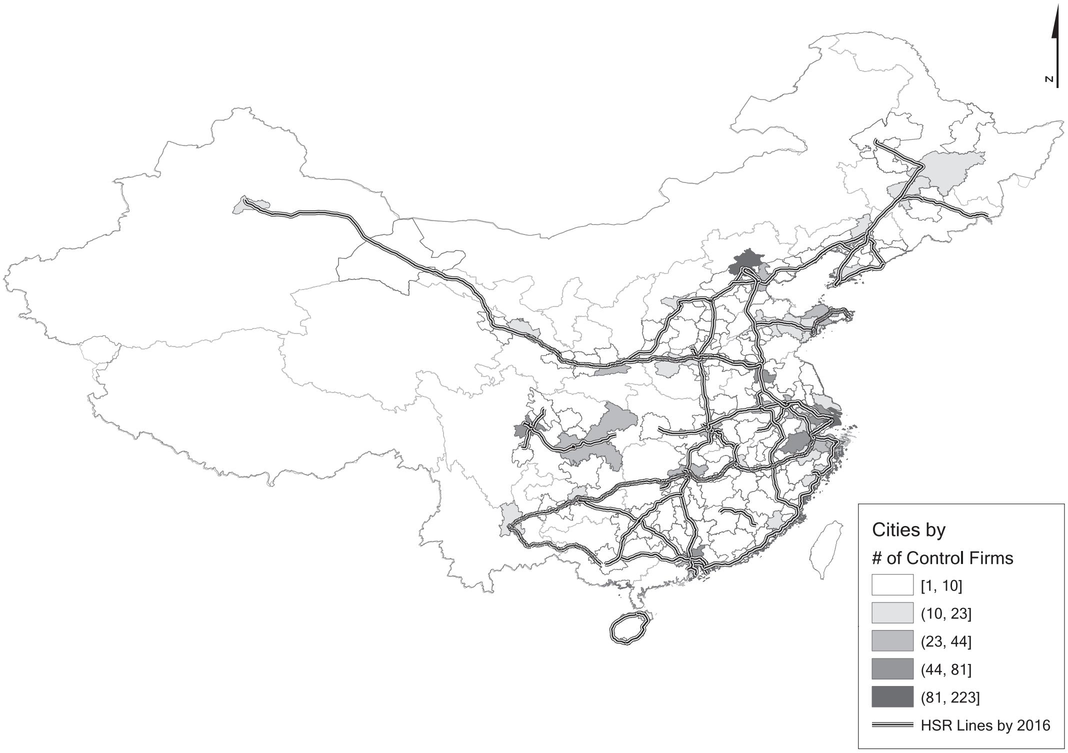

Fig. 3. Distribution of Control Firms Across Cities.
Notes: The graph plots the distribution of the number of control firms across cities identified by their headquarter locations. Control firms refer to those in industries that rank in the bottom halves in communication and travel intensity in Fig. 1. The map of HSR lines is modified based on Li (2016).

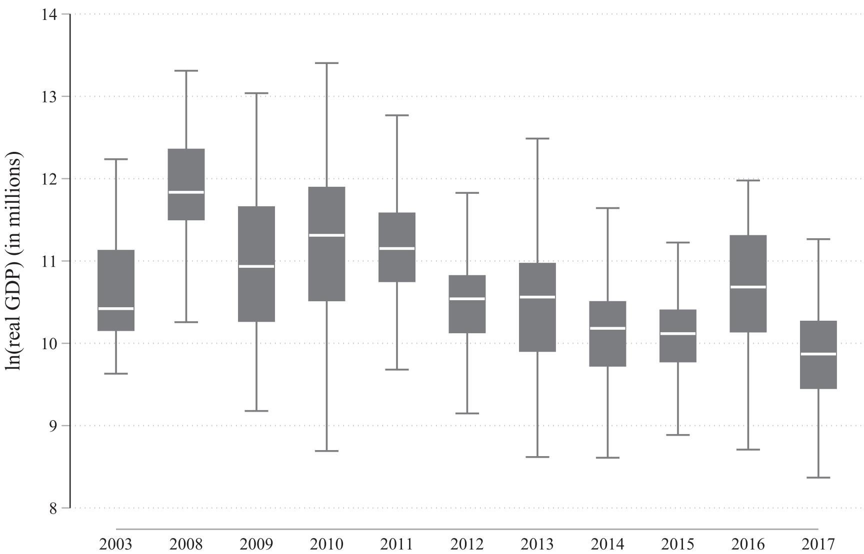

Fig. 4. Distributions of Real GDP of Connected Cities by Connection Year.
Notes: The graph presents the box plot of the distribution of real GDP for newly connected cities by connection year. The horizontal lines in the middle of the boxes correspond to the median real GDP among cities being connected in a given year, and the height of the boxes correspond to the interquartile ranges (i.e., IQR, the differences between 25th and 75th percentiles). The two ends of the line segments represent the maximum and minimum values of the distribution, excluding outliers.

nections, compared to firms in other industries, due to the significant savings on travel time and communication cost brought by the HSR network. Empirically, we estimate our baseline regression in the following form:

$$Y_{ic,t} = \alpha + \beta^{*}HSR_{c,t-1} + \delta^{*}HSR_{c,t-1}{}^{*}COMM_{i}^{High} + \gamma^{*}X_{i,t-1} + \lambda^{*}Z_{c,t-1} + u_i + v_t + \varepsilon_{ic,t}, \tag{4.11}$$

where $Y_{ic,\ t}$ is the outcome variable of firm i headquartered in city c in year t, including three profitability measures (i.e., *EBIT*, *ROA*, and *ROE*), and two size measures (i.e., *lnAssets* and *lnEmployees*); the key explanatory variable, $HSR_{c,\ t-1}$, is an indicator for connection to the HSR network for firm's headquarter city c in year $t-1$; $X_{i,\ t-1}$ is a vector of one-year lags of firm financials, including firm total assets, leverage ratio, cash flow and q; $Z_{c,\ t-1}$ is a vector of one-year lags of city-level attributes, including GDP, population, industry mix, volume of air travel, and boardband penetration; while u_i and v_t are vectors of firm fixed effects and year fixed effects, respectively.[21] Year fixed effects are included in the model to strip away any system-wide variations in firm performance and growth each year. It is possible that firms headquartered in cities that are connected to the HSR network first differ from firms that are connected later along some unobservable dimensions, such as firms' innate ability. If such unobservable attributes are permanent, our firm-level fixed effects, to a large extent, alleviate this concern. The coefficient on the interaction between *HSR* and $COMM^{High}$, δ, captures the differences in firm performances between communication-intensive and travel-dependent firms and other firms before versus after they are connected to the HSR network.[22]

We further hypothesize that the effect of the HSR on "treated" industries is persistent over time. This hypothesis is rooted in the differences between the HSR network and other types of infrastructure projects such as the highway systems in terms of the degree to which they are subject to congestion.[23] While the HSR network is not congestion-free, it is much less susceptible to congestion in the same fashion as the highway system by design.[24] On the contrary, as the HSR network expands, not only are new cities being connected to the network, but cities connected earlier also enjoy substantial enhancement in their accessibility to the rest of the network. Therefore, we project that with the expansion of the HSR network and the improvement of on-board services, the effect of HSR connection on firm performance for communication-intensive and travel-dependent industries would not be a one-time shock.

To test this hypothesis, we construct a cumulative measure $HSR^{Year}{}_{c,\ t-1}$, indicating the number of years that city c is connected to the HSR network in year $t-1$ and substitute it for $HSR_{c,\ t-1}$ in eq. (4.11):

$$Y_{ic,t} = \alpha + \beta^{*}HSR_{c,t-1}^{Year} + \delta^{*}HSR_{c,t-1}^{Year}{}^{*}COMM_{i}^{High} + \gamma^{*}X_{i,t-1} + \lambda^{*}Z_{c,t-1} + u_i + v_t + \varepsilon_{ic,t}. \tag{4.12}$$

The persistent effect of HSR is captured by δ, which is the coefficient on the interaction between HSR^{Year} and the "treated" firm indicator. The estimated coefficient on this interaction is the yearly average of the differences in firm performance improvement and growth in firm size between "treated" and control firms after being connected to the HSR network, and is hypothesized to be positive. To examine the dynamics of this effect over time, we also estimate a model with multiple indicators for the n^{th} year (i.e. $t-2$, $t-1$, t, $t+1$, …,$t+5^{+}$) that a firm is connected to the network and visualize how the effect of HSR evolves over time. The indicator for $t+5^{+}$ takes the value of 1 for the fifth-year of HSR connection and beyond.

4. Empirical results

In this section, we present estimation results for the impact of HSR on the performance and growth of firms in communication-intensive and travel-dependent industries, and discuss the robustness of our results.

4.1. HSR connection and firm performance

We report in Table 7 the estimated results of the HSR connection on firm profitability and firm size, with column titles indicating the specific outcome variables. Estimates from eq. (4.11) are reported in Panel A, while results from estimating eq. (4.12) are in Panel B. All regressions include firm fixed effects and year fixed effects. Standard errors are clustered at the firm level.[25] In Panel A, the results reported from the profitability regressions show that "treated" firms see more significant improvement in profitability after being connected to the HSR network than firms in other industries. The combined estimated differences are 1.6 percentage points for *EBIT*, 1.3 percentage points for *ROA*, and 3.0 percentage points for *ROE*, respectively. Given that the sample averages of *EBIT*, *ROA*,

[21] *lnAssets* is excluded from the specification where the outcome variables are firm size measures.

[22] In the current study, we are not able to rule out the possibility that average firm performance could increase due to rising competition (HSR-induced or not) at the industry level. However, we believe that it would only be problematic if the increase in competition is significantly more concentrated in our "treated" industries. We also collect data from the Statistical Yearbook of China and check how the number of all registered firms in the "treated" versus control industries change over time. We do not find the level of competition to be increasing more significantly in the "treated" industries than in the control industries, suggesting that increase in competition is likely not a key driver of our results.

[23] An empirically important feature of the highway system is that it is highly subject to congestion, as an additional user might slow others down under sufficiently high traffic flow. Fernald (1999) finds the productivity boost from the completion of the highway system in the 1950's and 1960's to be a one-time, unrepeatable productivity shock for vehicle-intensive industries, as congestion became empirically important after the interstate system was completed.

[24] Strictly speaking, even though those traveling on the HSR are unlikely to experience congestion the same way as those driving on the highway do, the availability of HSR tickets is indeed decreasing due to the increase in the number of travelers.

[25] Our results are robust to clustering standard errors at the city level.

Table 7
HSR Connection and Firm Performance.

	Profitability			Size	
	EBIT	*ROA*	*ROE*	*lnAssets*	*lnEmployees*
	(1)	(2)	(3)	(4)	(5)
Panel A: HSR Connection					
HSR_{t-1}	−0.001	0.000	0.004	−0.060***	−0.028
	(0.003)	(0.003)	(0.007)	(0.020)	(0.024)
$COMM^{High} * HSR_{t-1}$	0.016***	0.013***	0.030***	0.129***	0.174**
	(0.005)	(0.004)	(0.009)	(0.049)	(0.070)
$lnAssets_{t-1}$	−0.031***	−0.023***	−0.055***		
	(0.004)	(0.003)	(0.008)		
$Leverage_{t-1}$	0.016*	−0.017**	0.147***	0.764***	0.550***
	(0.009)	(0.007)	(0.022)	(0.077)	(0.096)
$CashFlow_{t-1}$	0.027	0.029*	0.073**	0.021	0.194***
	(0.018)	(0.016)	(0.035)	(0.088)	(0.064)
q_{t-1}	0.014***	0.012***	0.023***	−0.067***	−0.066***
	(0.002)	(0.002)	(0.004)	(0.008)	(0.010)
Adjusted R^2	0.259	0.301	0.185	0.898	0.838
Panel B: HSR Connection in Years					
${HSR^{Year}}_{t-1}$	0.001	0.001	0.004*	−0.003	0.015
	(0.001)	(0.001)	(0.002)	(0.009)	(0.011)
$COMM^{High} * {HSR^{Year}}_{t-1}$	0.003***	0.002***	0.005***	0.018*	0.025*
	(0.001)	(0.001)	(0.002)	(0.010)	(0.014)
$lnAssets_{t-1}$	−0.031***	−0.023***	−0.055***		
	(0.004)	(0.003)	(0.008)		
$Leverage_{t-1}$	0.016*	−0.017**	0.146***	0.766***	0.547***
	(0.009)	(0.007)	(0.022)	(0.077)	(0.096)
$CashFlow_{t-1}$	0.027	0.029*	0.072**	0.019	0.192***
	(0.018)	(0.016)	(0.035)	(0.088)	(0.064)
q_{t-1}	0.014***	0.012***	0.023***	−0.067***	−0.066***
	(0.002)	(0.002)	(0.004)	(0.008)	(0.009)
Observations	13,726	13,726	13,726	13,726	13,726
Adjusted R^2	0.259	0.302	0.185	0.897	0.838

Notes: The estimation sample includes all firms connected to the HSR network by the end of 2017. Firms that are established in 2003 or later are excluded from the estimation sample. Dependent variables are indicated by column titles in italics. All firm-level and city-level controls including HSR connection are lagged by one year. City-level characteristics are controlled for in all regressions, including logarithm of real GDP, population, tertiary employment, air passengers, and internet households. All regressions include firm fixed effects and year fixed effects. Robust standard errors are clustered at the firm level.

and *ROE* for firms in the "treated" industries are 7%, 5%, and 9%, respectively, these estimated differences in the improvement in profitability correspond to increases in the magnitude of 22.9%–33.3%. In terms of firm size, we find that firms in the "treated" industries on average see an additional 6.9% increase in asset size and a 17.4% increase in the number of employees after being connected to the HSR network, compared with firms in other industries.[26]

Results in Panel B of Table 7 suggest that the impact of HSR on firm performance and growth in the communication-intensive and travel-dependent industries are persistent over time. For instance, for "treated" firms, being connected to the HSR network for one more year, on average, is associated with an additional increase in *ROA* of 0.2 percentage points. In terms of firm size, the differential effect of connecting "treated" firms to the HSR network for one more year, on average, is associated with an additional difference in *lnAssets* of 1.8% and *lnEmployees* of 2.5%. Fig. 5 plots the estimated coefficients on indicators of the n^{th} year of being connected to the HSR network (i.e. $t-2, t-1, t, t+1, \ldots, t+5^{+}$) in the profitability and size regressions, respectively. There is a visible upward shift in the magnitude of estimated coefficients on our focal interaction terms at $t+1$ for profitability. The estimates show that instead of a one-time boost, the improvement in "treated" firms' performance and growth appears to persist after being connected to the HSR network for several years. These plots also confirm that before being connected to the HSR network, firms in our "treated" industries appear to be largely similar to firms in other industries, consistent with the parallel trend assumption of the difference-in-differences framework. In sum, results in Table 7 and Fig. 5 suggest that firms that are communication-intensive and travel-dependent benefit more from being connected to the HSR network and that such benefits appears to persist over time.

[26] In column (4) of Panel A in Table 7, the main effect of HSR connection on asset size on all firms is estimated to be −0.060, and the additional change in the "treated" industries is estimated to be 0.129, resulting in a combined differential increase of 6.9% of firms in the "treated" industries after being connected to the HSR network.

4.2. Robustness checks

We begin by presenting sub-sample results using single-location firms to address concerns over the precision of our HSR measure. Results using historical railways interacted with city-level relative elevations as instruments for the HSR placement, and those with alternative measures for travel intensity are also presented. We also demonstrate the robustness of our results using falsified "treated" industries and matching estimators to ease concerns over the imbalanced sample sizes between the "treated" and control groups.

4.2.1. Single-location firms

A natural concern for studies about firm location is the location measure itself. We construct our HSR measures based on the connectedness of firms' headquarter cities. However, it is possible for firms to have subsidiaries at multiple locations. In that case, our HSR measures would be imperfect measures of the connectedness of firm location, if firms are operating across different cities. To address this concern, we collect information on firms' subsidiaries, including location, amount of registered capital, and industry of operation in each accounting period from the CNRDS database.[27] We then measure the relative size of each subsidiary using its registered capital and the holding company's book value of equity in the corresponding accounting period. Using each subsidiary's location, we identify whether a subsidiary is at different locations from their holding companies. To separate firms that have subsidiaries at multiple locations (i.e., multi-location firms) from those with no subsidiary or those that only have subsidiaries at the same location as their holding companies (i.e., single-location firms), we calculate the total number of subsidiaries and their total size in terms of registered capital for each firm in each year. In our estimation sample, the medium number of subsidiaries is 5, while the medium number of subsidiaries that are at different locations from their holding companies is 0. We also calculate the relative size of all subsidiaries for each firm-year. For the subset of firms with at least one subsidiary, the medium share of all other-location subsidiaries' registered capital to the holding company's equity is only 0.94%, confirming that multi-location firms are not common and that other-location subsidiaries are often small.[28]

To demonstrate the robustness of our empirical results, we estimate our baseline regressions with firms that have no subsidiary in other locations.[29] Results are reported in Panel A of Table 8. The estimated coefficients from the single-location sub-sample are similar to that in Table 7, with the magnitude of the coefficients being slightly larger for all outcome variables. By removing multi-locational factors that could potentially contaminate our results, we improve the precision of our HSR measure and find larger effect of the HSR connection for communication-intensive firms with single operational locations.

4.2.2. Instrumenting for HSR connection

Given that our estimation sample includes only firms (cities) that are connected to the HSR network by the end of 2017, the selection bias between "the connected" and "the never-connected" cities are excluded from our results by design. However, if cities connected first versus those connected later are not comparable in some attributes that might affect firm performances, relying simply on the differences in the timing of being connected to the HSR network would not give us a casual relationship between the HSR connection and firm performance for the communication-intensive and travel-dependent industries. Ideally, we need an instrumental variable that only affects the order in which a city is connected without influencing firm performance.

We follow Baum-Snow et al. (2017) and instrument the HSR network connection with the 1962 railway network in China. Fig. 6 plots the 1962 railroad network as in Baum-Snow et al. (2017).[30] At the time, railways were constructed primarily to transport raw materials and manufacturing goods to provincial capitals to support the industrial development for national defense purposes in China's interior regions, in accordance with the national and provincial five-year plans (Baum-Snow et al., 2017). Since then, not only has the role of modern railroad evolved, but the HSR network has become even more distinct in that most of the network is designated solely for passenger travel. For this instrument to be valid in our analysis, the 1962 railway network should affect today's firm performance only through their effects on the HSR network configuration, and more specifically, the order in which cities in our sample were connected to the HSR network, conditional on all the firm-level and city-level controls in the model. Given that the initial purpose of the HSR network was to serve passengers so as to free up capacity in the original network for freight service, most of the HSR lines designed in the original plan ran parallel to the then-existing railway lines. We believe that the configuration of the 1962 railway network is highly relevant to the layout of the HSR network, satisfying the relevance assumption of an instrument. The distinct purposes for railway construction between the 1962 and the current networks give us confidence that it is excludable from the main model of interest.

[27] A few filters for data quality control are applied when constructing measures of firm subsidiaries. The specifics of the filters are described in the Appendix.

[28] Detailed summary statistics of firms' subsidiaries are reported in Table A1 in the Appendix. We separately report statistics for (1) all subsidiaries, (2) subsidiaries that are at other locations, and (3) subsidiaries that are at other locations and in the same sector as their holding companies. Panel A reports the number of firms' subsidiaries for all firms in our sample, and Panel B reports, for the subset of firms with at least one subsidiary, the relative sizes of all subsidiaries by calculating the percentage shares of their registered capital in their holding companies' equity.

[29] We believe that this is the strictest criterion since we also exclude firm-year observations with very small subsidiaries in other locations. We report more robustness test results in the Appendix with more relaxed criteria by including firms that have very small subsidiaries in other locations. Results are reported in Panel A and Panel B of Table A2. We report estimation results with the sub-sample of firms that have subsidiaries at other locations but have a combined equity share lower than 1.41%(75^{th} percentile) or 9.65%(90^{th} percentile) in Panels A and B, respectively.

[30] The GIS shapefile used to produce this map is generously shared by those authors online.

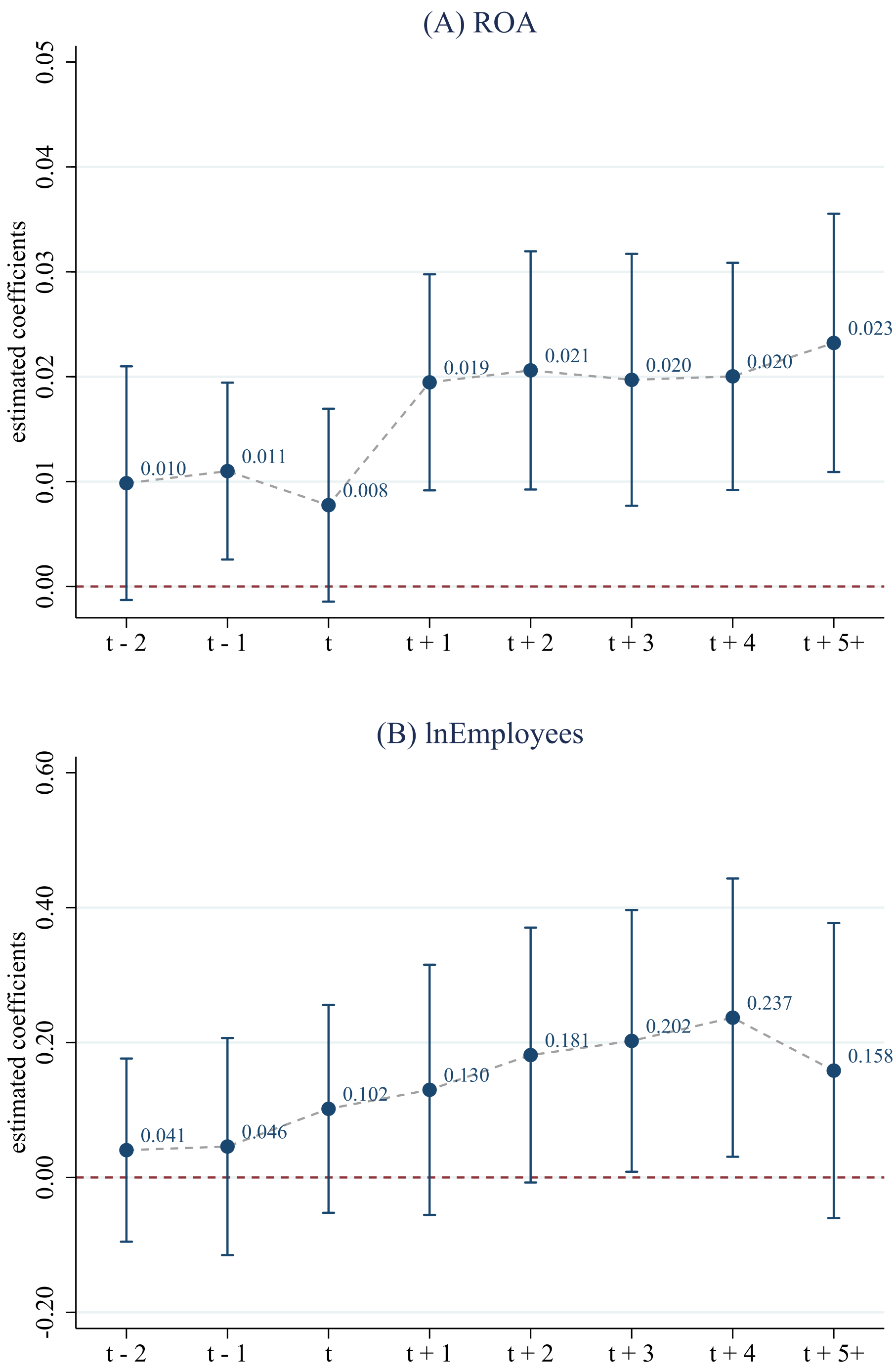

Fig. 5. Year-by-Year Estimates of HSR Connection on Firm Performance and Size.
Notes: The graph plots the estimated coefficients on indicators of various years before and after being connected to the HSR network in profitability and size regressions, respectively. Similar to Model (4.12), all firm-level and city-level controls are lagged by one year. All regressions include firm fixed effects and year fixed effects. Robust standard errors are clustered at the firm level.

Given that the 1962 railway instrument is static, while our HSR connection indicator is time-varying, we augment the 1962 railway IV with a hypothetical "connection order" predicted using city-level elevations. The idea is that the construction of HSR lines in cities with lower elevations is less likely to face engineering difficulties, allowing these cities to be connected before areas with higher elevations. More specifically, we rank cities in our sample based on their average elevations and divide them equally into 14 groups so that for each year between 2003 and 2016, an equal number of cities become hypothetically "connected" to the HSR network. The

Table 8
HSR Connection and Firm Performance: Robustness Checks.

	Profitability			Size	
	EBIT (1)	*ROA* (2)	*ROE* (3)	*lnAssets* (4)	*lnEmployees* (5)
Panel A: Single-Location Firms					
HSR_{t-1}	−0.001	0.000	0.002	−0.054**	−0.039
	(0.004)	(0.003)	(0.009)	(0.022)	(0.027)
$COMM^{High} * HSR_{t-1}$	0.017***	0.015***	0.033***	0.122**	0.248***
	(0.006)	(0.004)	(0.011)	(0.052)	(0.075)
Observations	8818	8818	8818	8818	8818
Adjusted R^2	0.276	0.331	0.174	0.898	0.845
*Panel B: IV Estimator**					
HSR_{t-1}	−0.031	−0.017	0.016	−0.156	−0.281
	(0.027)	(0.021)	(0.056)	(0.170)	(0.198)
$COMM^{High} * HSR_{t-1}$	0.023***	0.016***	0.034**	0.207***	0.195**
	(0.007)	(0.005)	(0.014)	(0.065)	(0.099)
Weak identification test					
Kleibergen-Paap rk Wald F statistic		43.41		43.36	
Stock-Yogo critical values (10%)		7.03		7.03	
First-stage F statistic (HSR)		45.68		45.46	
First-stage F statistic (COMM*HSR)		977.99		988.89	
Observations	13,726	13,726	13,726	13,726	13,726
Panel C: Alternative Measure					
HSR_{t-1}	−0.006	−0.005	−0.011	−0.117***	−0.100**
	(0.006)	(0.005)	(0.012)	(0.042)	(0.047)
$\%TravelExpense^{ave} * HSR_{t-1}$	0.004**	0.004**	0.011***	0.034**	0.042**
	(0.002)	(0.002)	(0.004)	(0.015)	(0.019)
Observations	9822	9822	9822	9822	9822
Adjusted R^2	0.234	0.274	0.163	0.875	0.814
*Panel D: Matching Estimator***					
HSR_{t-1}	−0.004	−0.004	−0.004		
	(0.008)	(0.006)	(0.014)		
$COMM^{High} * HSR_{t-1}$	0.018*	0.015**	0.037**		
	(0.009)	(0.007)	(0.018)		
Observations	3423	3423	3423		
Adjusted R^2	0.317	0.375	0.250		

Notes: The estimation sample includes all firms connected to the HSR network by the end of 2017. Firms that are established in 2003 or later are excluded from the estimation sample. Dependent variables are indicated by column titles in italics. All regressions include the same firm-level and city-level control variables as in Table 7, as well as firm fixed effects, and year fixed effects. Robust standard errors are clustered at the firm level. *We use the Kleibergen-Paap rk Wald F statistic to test for weak identification of the two endegenous variables jointly (Baum et al., 2007). The critical value compiled by Stock and Yogo (2005) is 7.03. Our results are also robust to the "rule of thumb" of Staiger and Stock (1997), which requires the first stage F statistic to be larger than 10. **Matching estimator for firm sizes are omitted due to the fact that lnAssets itself was included as a matching criterion when searching for matched control sample.

validity of this strategy depends on (1) the predicting power of relative city-level elevations and the order in which cities were connected to the HSR network, and (2) the extent to which the relative elevations of cities can be excluded from a model of firm performance or growth with the firm-level and city-level characteristics in our baseline model. Our instrument is therefore the interaction between the 1962 railway indicator and the hypothetical "connection" predicted using relative elevations. The estimation results along with the 1^{st} stage F-statistics are reported in Panel B of Table 8. With the HSR connection indicator instrumented with a hybrid of the 1962 railway connection and relative elevations of cities, our baseline results remain qualitatively the same, with slightly larger point estimates.[31]

4.2.3. *Alternative proxy for travel intensity*

Another question to the robustness of our empirical findings is how closely the "treated" firm proxy captures firms' communication

[31] We understand that perfect exogeneity is unlikely to hold exactly, not only under our context but in general as well. More specifically, it remains possible that the historically connected cities are larger and more agglomerated today than the historically non-connected cities, thus allowing historical railway connection to contribute directly to contemporary firm performance. Similar concerns could exist for relative elevations. Nonetheless, we do believe that the direct impact of the instrument is secondary to the impact of HSR connection on firm performance, which allows us to have confidence in the overall IV strategy (Jiang, 2017).

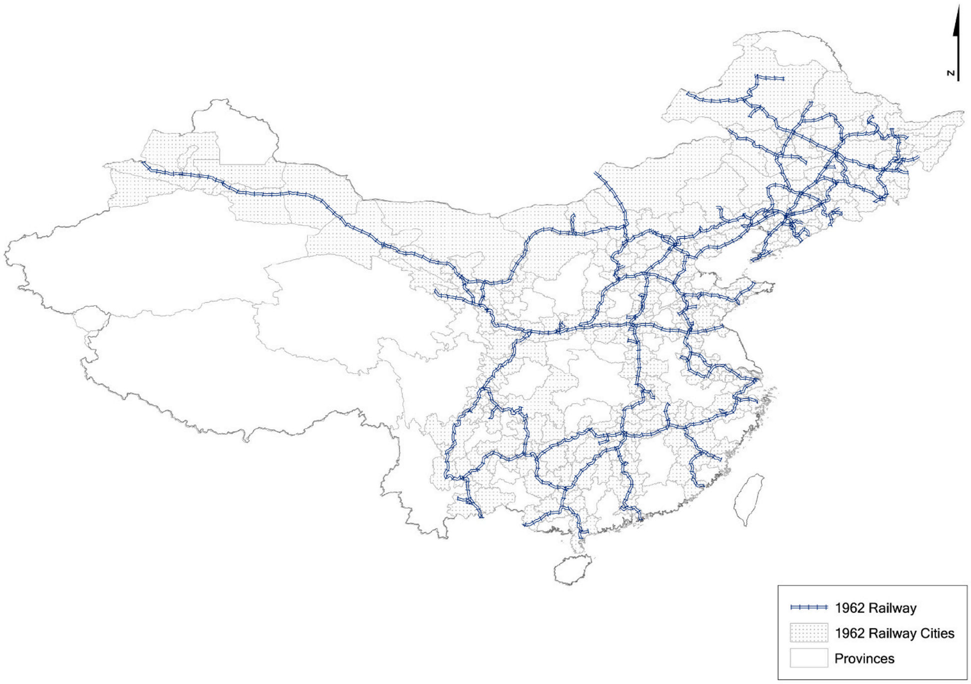

Fig. 6. Instrumental Variable: Railways in 1962.
Notes: The graph plots the 1962 railways in China generously shared by the authors of Baum-Snow et al. (2017) online.

and travel intensity that we set out to measure. In addressing this concern to the best of our ability, we look for firm-level information on travel intensity. The intuition is simple. Under the assumption that the HSR network is exceptionally productive for communication-intensive and travel-dependent firms, such firms ought to have employees traveling more frequently via the HSR network to firms and consumers in other cities. This argument follows naturally from Fernald (1999) where it is hypothesized that vehicle-intensive firms use highways more intensively.

Unlike Fernald (1999) where "vehicle-intensity" is measured directly, neither direct measures for communication intensity nor travel intensity at the firm level are available. An ideal travel intensity index would capture the frequency and number of trips for each firm at different points in time, with some information on the destination. The only relevant information we are able to find in firms' annual reports are the total travel expenses for firms' administrative personnel. However, this information is only reported regularly across firms after 2010, and furthermore, it does not measure quantities (i.e. travel demand), but values (i.e., travel expenses). The limitations of such measures include the loss of separability in travel intensity due to price change versus quantity adjustment. The direction of the potential bias on our estimates is not clear. If the average cost of trips increased over time, this measure could overstate the level of travel intensity. However, with the introduction of the HSR services, it is highly possible that the average cost decreased in recent years, which could lead to an increase in travel demand. In this case, the changes in travel expenses would understate the changes in travel intensity. Nonetheless, such information is the only direct measures available at the firm level, reported in the notes of the financial statements. Due to data availability, we take the multi-year averages of the ratios of travel expenses over total administrative expenses ($\%TravelExpense^{ave}$) as a proxy for travel intensity. As previously discussed, we reproduce our baseline results of model (4.11) by replacing $COMM^{High}$ with $\%TravelExpense^{ave}$. The results are in Panel C of Table 8, and show that travel-intensive firms benefit more from the HSR connections, consistent with our baseline findings.

4.2.4. Falsification tests

Identifying the "treated" firms in this study is challenging in the sense that firms in cities that are connected to the HSR network are affected to some extent by the operation of the HSR. We next use a set of falsification tests to confirm our inferences about the effect of HSR on communication-intensive and travel-dependent firms' performance and growth is through significant reduction of communication cost. Instead of using the industry-level ranks of nonroutine interactive labor inputs and rail passenger travel inputs, we construct two pseudo treatment identifiers, by mixing one true input (i.e., the same as our baseline "treatment" measure) with one false input ranking. The first measure, "Interactive-Freight", keeps non-routine interactive labor inputs, and substitutes rail freight shipping inputs for rail passenger travel input in our "treated" measure. Similarly, the second measure, "Manual-Passenger", keeps the rail passenger travel inputs, while replacing the non-routine interactive skills with manual skills in our "treated" measure. The underlying rationale behind these measures is that while these firms might benefit from the HSR network, they should not benefit from the operation of HSR as much as our "treated" firms would. Therefore, we expect to see the coefficients of the interaction terms between the pseudo treatment measures and the HSR to have smaller magnitude and lower significance level than those in our baseline models. For the sake of brevity, we perform the falsification tests with one performance measure and one firm size measure, namely *ROA* and *lnEmployees*.[32] Two coefficients of the pseudo treatment interactions are plotted in Fig. 7, along with the coefficients in our baseline models. Consistent with our prediction, pseudo treatment groups do not benefit from the HSR network as much as the treatment group identified in our baseline models as the coefficients are mostly not statistically different from zero.

4.2.5. *Matching estimators*

The last set of robustness tests we perform is to mitigate potential estimation bias due to the unbalanced sample sizes between the "treated" versus control groups. As reported in Table 1, the sample size of our "treated" group is much smaller than that of the control group. To address this issue, we combine our difference-in-differences model with a matching estimator as an additional robustness check for our empirical results.[33] Specifically, we use the Abadie and Imbens (2006) estimator, which minimizes the Mahalanobis distance between a vector of observed covariates across "treated" and "control" firms. By minimizing the Mahalanobis distance, "treated" observations are matched with untreated ones based on all firm-level controls, including asset size, leverage, cash flow, and q. We use the nearest-neighbor matching criteria with $n = 3$, which matches each "treated" firm with three untreated firms. The matching process is done with replacement.[34] The same quasi-difference-in-differences specification as eq. (4.11) is employed to test our baseline hypothesis with the matched sample. The empirical results are reported in Panel D of Table 8. We find similar results as in our baseline regressions in terms of both the coefficients' magnitude and levels of significance. Firm size measures are omitted in this analysis as it is used as a matching criterion.

5. Discussion and extensions

To further demonstrate the impact of a massive inter-city passenger travel network on the performance and growth of communication-intensive and travel-dependent firms, we discuss in detail three possible channels through which communication-intensive and travel-dependent firms might benefit from the HSR network. We also perform a series of extensional analyses to show that the potential gains from being connected to the HSR network (1) increase with savings in travel time compared to regular speed rail, (2) are larger for firms connected to a mega city, and (3) especially for those within 3–6 hours from the nearest mega city.

5.1. *Mechanisms*

We start by investigating the mechanisms through which the identified "treated" firms benefit more from their connection to the HSR network relative to firms in other industries. To better understand the process through which the HSR promotes the performance of firms in the "treated" industries, we propose and test three mechanisms at work: analyst attention, productivity boost, and market expansion.

5.1.1. *Analyst attention*

The "local (home) bias" literature find that investors have a strong preference for local firms, as firms located far away are at a disadvantage in terms of information communication. We hypothesize that HSR connection could, to some extent, alleviate the information asymmetries between "treated" firms and their potential investors, and look at analyst coverage as the outcome variable in verifying this information channel of HSR connection. Chang et al. (2006) and Derrien and Kecskés (2013) both show that analyst coverage could alleviate information asymmetry and lower level of analyst coverage is associated with lower financing and investment. We expect to see "treated" firms to have more analyst coverage after being connected to the HSR network compared to the firms in the control industries.

We further hypothesize the information channel of HSR to be stronger for firms with higher levels of information asymmetries, due to the possibility that communication frictions and travel distances matter even more for these types of firms (Sufi, 2007). Information asymmetries are proxied for along two dimensions: firms' degree of financial constraint and their headquarter location. We adopt two of the most commonly used financial constraint indexes: (1) Kaplan-Zingales (KZ) index following Kaplan and Zingales (1997) and Lamont et al. (2001,2) Whited-Wu (WW) index following Whited and Wu (2006).[35] Financial constraint is a natural proxy for information asymmetries, as the more financially constrained firms are likely to suffer from higher degrees of information asymmetries.

[32] Using other outcome variables, as in our baseline models, produces similar results.

[33] We are aware that matching estimators alone do not solve the endogeneity problem as the method itself also relies on the conditional independence assumption (Roberts and Whited, 2013). This method is widely used in the empirical corporate finance literature as a useful tool to mitigate bias from self-selection. See, for example, Campello et al. (2010) and Almeida et al. (2012).

[34] We carefully check the balancing of matching estimators by directly testing the differences in the means of the covariates we consider between our "treated" and control groups. Balancing test results are reported in Appendix Table A3 and Figure A2. Our empirical results remain similar if we use propensity score matching instead. We choose Mahalanobis distance matching based on the percentage reduction in bias. Our results are also robust to matching without replacement.

[35] Since the dividend policy of listed firms in China could sometimes be noisy in predicting firms' level of financial constraint, we construct an additional measure for each of the index by excluding the dividend component. Results are robust to using alternative financial constraint measures that exclude the dividend component. For brevity, we do not report this set of results in our paper.

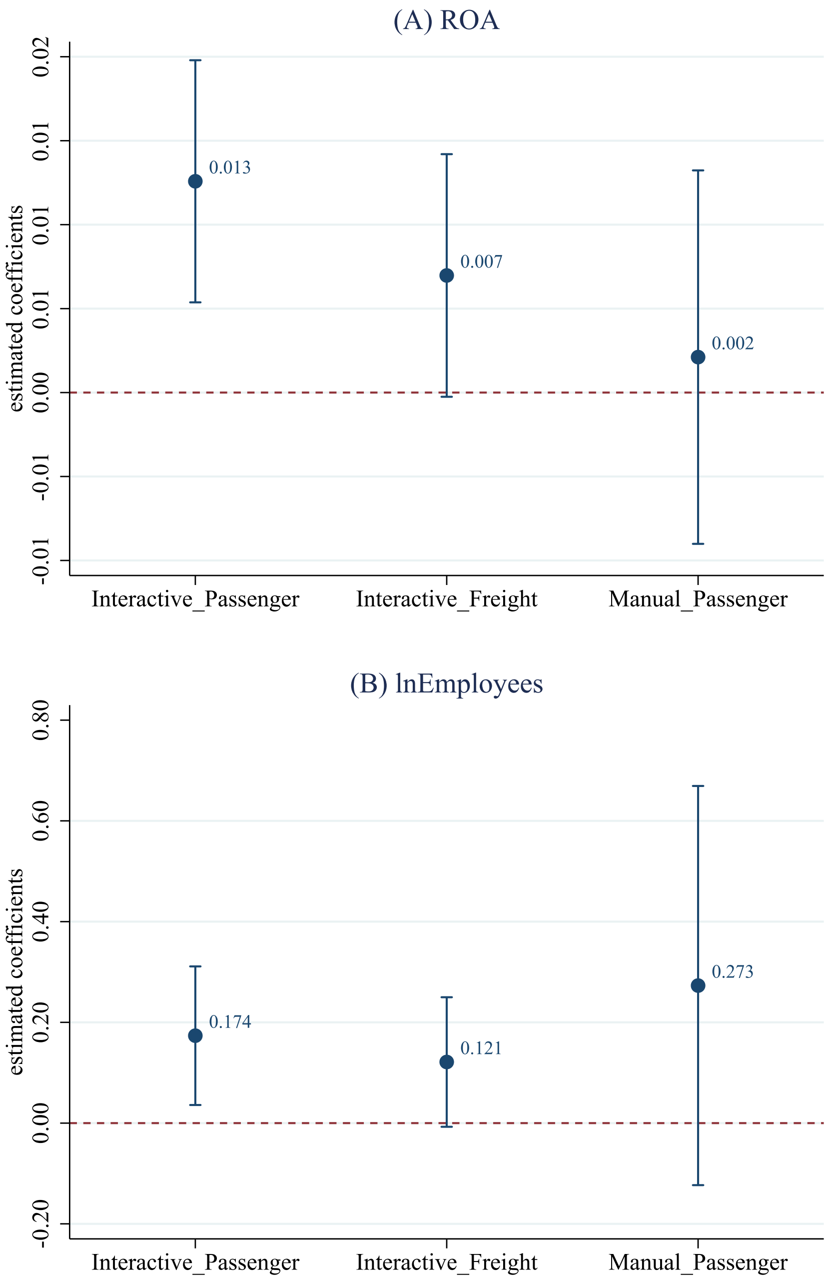

Fig. 7. Falsification Tests.
Notes: The graph plots the estimated coefficients and the confidence intervals of $COMM^{High} * HSR$ in our baseline model and interactions of *HSR* and coefficients in the falsification models as described in Section 4.2.4. "Interactive-Freight" identifies industries ranked on the top halves in non-routine interactive skills and rail freight shipping inputs. "Manual-Passenger" identifies industries ranked on the top halves in manual skills and rail passenger travel inputs.

A firm is identified as constrained if it is in the top tertile of the distribution of KZ or WW index in a given year, while firms in the bottom tertile are identified as unconstrained. We also differentiate firms' level of information asymmetries using their headquarter locations (i.e., mega city vs. non-mega city). Besides the meaningful differences in local economic development levels between mega and non-mega cities, more importantly for our analyst coverage results, more than half of the institutional investors in China are headquartered in mega cities, with the largest 15 all in mega cities adding up to a combined assets of about 60% of the industry total in 2017.[36] We therefore split our estimation sample using the three information asymmetry proxies discussed above, and test in each subsample if the "treated" firms benefit more from the HSR connection in terms of analyst coverage compared to firms in other industries.

The information channel results are reported in Table 9. The first column shows the full sample results, where firms that are communication-intensive and travel-dependent attract more analyst attention after being connected to the HSR network relative to other firms. Results in columns (2–5) suggest that such differential increases in analyst coverage are only observed for the more financially constrained firms. Similar patterns are also found in columns (6, 7), in that "treated" firms that are located in smaller cities (i.e., non-mega cities) enjoy significant increase in analyst attention upon being connected to the HSR network, compared to those located in mega cities. Consistent with our hypotheses and existing findings in the literature, we find the information channel of HSR connection to be significantly stronger for "treated" firms that suffer from higher levels of information asymmetries. This indicates that being connected to the HSR network enhances the performance and growth of communication-intensive and travel-dependent firms through increased analyst attention, potentially facilitating capital flows from investors and lowering their cost of capital. This set of findings suggest that face-to-face interactions, through facilitating the exchange and acquisition of knowledge and information, are potentially important for our understanding of "local (home) bias" of investments documented in the literature.

5.1.2. *Productivity boost*

We argue that being connected to the HSR network significantly reduces the cost of face-to-face communication, and therefore increases the operating efficiency of the "treated" firms. Access to the HSR network serves as a positive productivity shock to firms that are highly dependent on communication and inter-city travels in their daily operations, allowing employees in these firms to produce more with the same or even fewer resources. For instance, with significant reduction in travel time between cities via the HSR, workers in the "treated" firms are able to meet with more potential customers (in multiple cities) to provide their services or to market their products. Higher productivity at the individual employee level gives firms the option to expand their operation with minimal increases in their operating costs.

We directly test this hypothesis using three productivity measures as the outcome variables: (1) two estimates of total factor productivity (TFP), estimated following procedures outlined in Levinsohn and Petrin (2003) (i.e., LP) and Ackerberg et al. (2015) (i.e., ACF) with a value-added production function, denoted by TFP^{LP} and TFP^{ACF} respectively[37]; and (2) profit per capita (*ProfitPerCapita*), which is defined as net income (in millions) divided by the number of employees. Model (4.11) is estimated with TFP^{LP}, TFP^{ACF}, and *ProfitPerCapita* as dependent variables. A positive coefficient on the interaction between HSR connection and the "treated" firm indicator would be consistent with our hypothesis that HSR boosts the level of operating efficiency by lowering the communication cost for face-to-face interactions across cities, and therefore improves the productivity of "treated" firms connected to the HSR network. Results are reported in columns (1–3) of Table 10. For both the TFP (columns (1–2)) and profit per capita (column (3)), "treated" firms on average see a larger increase in productivity than other firms after being connected to the HSR network. Given the fact that "treated" firms on average have higher levels of productivity than other firms in our sample period (see Table 4), being connected to the HSR network further increases the difference in productivity between the "treated" and other firms.

5.1.3. *Market expansion*

In addition to boosts in the "treated" firms' productivity, "treated" firms could benefit from the HSR network is through another direct channel: by expanding business more efficiently into regions that firms are not currently connected to via the HSR. For firms in traditional industries such as agriculture and manufacturing, the expansion of the highway network is likely more beneficial than the HSR network. This is because business expansion for these firms relies more heavily on shipping their products out, whereas the HSR is dedicated for passenger travel. In contrast, firms that are more communication-intensive and travel-dependent would benefit more directly from significant savings on travel time and costs when fostering a relationship with new costumers and entering new territories in the domestic market. To test the effect of the HSR network on the regional distribution of firms' sales revenues, we collect data on the revenue share across different regions from firms' annual reports for each firm in each year. We construct a measure for the share of sales revenues in firms' home market (i.e., headquarter location) and test if the home shares decrease after firms are connected to the HSR network.

Ideally, we would prefer to define "home market" as the headquarter city, rather than the headquarter province. This is because as an inter-city passenger travel network, the HSR facilitates just as much travel within each province, as it does across provinces, if not more. For firms that originally operate mainly within their headquarter province, it is more likely that other cities in the same province would be among their top contenders from the perspective of market expansion. Unfortunately, the smallest geographical unit reported in the notes of firms' financial reports are often provinces rather than cities. Thus, in our data, we are not able to observe a more

[36] Source of information: Securities Association of China (SAC), which functions under the supervision of the China Securities Regulatory Commission (CSRC); accessed on November 7, 2020.

[37] The TFPs are estimated without information on communication and travel inputs, as they are impossible to measure given our data. Therefore, the increased usage of these inputs in the production process, due to the decrease of cost to communicate and travel, are potentially reflected in the TFP estimates.

Table 9
Mechanism Test: Analyst Attention.

	lnAnalystCoverage						
	Full Sample	KZ Index		WW Index		Firm Location	
		NFC	FC	NFC	FC	Mega	Non-Mega
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HSR_{t-1}	−0.029	−0.133*	−0.003	−0.046	−0.007	0.135**	−0.062*
	(0.032)	(0.073)	(0.048)	(0.046)	(0.047)	(0.069)	(0.036)
$COMM^{High} * HSR_{t-1}$	0.173***	0.045	0.378***	0.103	0.177**	0.075	0.221***
	(0.053)	(0.084)	(0.106)	(0.079)	(0.083)	(0.099)	(0.059)
Observations	13,726	3745	4768	4757	4077	3735	9991
Adjusted R^2	0.675	0.735	0.659	0.686	0.586	0.708	0.662

Notes: The estimation sample includes all firms connected to the HSR network by the end of 2017. Firms that are established in 2003 or later are excluded from the estimation sample. Dependent variables are *lnAnalystCoverage*. All firm-level controls including HSR connection are lagged by one year. All regressions include firm-level and city-level controls, firm fixed effects, and year fixed effects. Robust standard errors are clustered at the firm level. In columns (2–5), the estimation sample is divided into financially constrained (FC) and unconstrained (NFC) firms using Kaplan and Zingales (KZ) index (Kaplan and Zingales, 1997; Lamont et al., 2001), or Whited and Wu (WW) index (Whited and Wu, 2006). A firm is identified as constrained if it is in the top tertile of the distribution of KZ or WW index in a given year, while firms in the bottom tertile are identified as unconstrained. In columns (6–7), the estimation sample is split based on firms' headquarter locations (i.e., mega versus non-mega cities).

Table 10
Mechanism Test: Productivity Boost and Market Expansion.

	Productivity			Market Expansion		
	TFP^{LP}	TFP^{ACF}	*ProfitPerCapita*	*%HomeRevenue*		
	All Firms			All Firms	Young Firms	Old Firms
	(1)	(2)	(3)	(4)	(5)	(6)
HSR_{t-1}	−0.044*	−0.019	−0.052*	0.005	0.047	0.003
	(0.023)	(0.053)	(0.031)	(0.017)	(0.036)	(0.019)
$COMM^{High} * HSR_{t-1}$	0.497***	0.588***	0.148**	−0.027	−0.120**	−0.027
	(0.075)	(0.170)	(0.070)	(0.019)	(0.056)	(0.020)
Observations	13,718	13,718	13,726	2146	200	1944
Adjusted R^2	0.895	0.882	0.284	0.853	0.827	0.855

Notes: The estimation sample includes all firms connected to the HSR network by the end of 2017. Firms that are established after in 2003 or later are excluded from the estimation sample. Dependent variables are indicated by column titles in italics. All firm-level controls including HSR connection are lagged by one year. All regressions include firm fixed effects and year fixed effects. Robust standard errors are clustered at the firm level. Columns (1)–(4) are estimated using all firms in the estimation sample. Columns (5, 6) are estimated based on young firms and old firms, respectively, where firms established before 1999 (sample median) are identified as old.

gradual expansion of territory for each firm over time. However, in the annual reports, we are able to observe the top 5 regional revenues, with reporting format varying wildly across firms and over time.[38]

Model (4.11) is estimated with *%HomeRevenue* as the dependent variable. Column (4) of Table 10 shows that, on average, the home market revenue shares of "treated" firms decrease by 2.7 percentage points after being connected to the HSR network, compared to firms in the control group, although not statistically significant at conventional levels. Inspired by the literature on firm age and growth, we also estimate this market expansion mechanism separately for younger versus older firms in our sample (Dunne and Hughes, 1994; Hart and Oulton, 1996; Mansfield, 1962). Given that the median firms in our sample were founded in 1999, we identify those founded after 1999 as younger firms. On average, we find a much larger difference between the "treated" and control firms in the reduction of home market revenue share (12.0 percentage points) among younger firms after they are connected to the HSR network (column (5)). As for the older firms, we do not observe statistically significant difference between firms in the "treated" and control industries with HSR connection, as is shown in column (6) of Table 10. Our finding that young firms (are able to) expand more aggressively is consistent with the established finding in the literature on the negative association between firm age and growth. In sum, we conclude that the HSR network helps communication-intensive and travel-dependent firms expand their operation network more rapidly, especially for younger firms.

[38] For example, some firms report regions as provinces, while others report southeastern versus northwestern China. We treat data as missing when the regions reported are too large for individual provinces to be identified (e.g., overseas versus domestic markets). Due to inconsistencies in reporting format, we are only able to observe the home province shares in sales revenues for a sub-set of firms in our sample.

5.2. Extensions

Given that cities were connected via regular speed rail network before the construction of HSR, we directly test the how the impact of HSR connections changes with savings in travel time in this section. We further explore the heterogeneity of connected "treated" firms' connection to mega cities to examine the role of agglomeration economies in the relationship between HSR connection and the performance and growth of communication-intensive and travel-dependent firms.

5.2.1. Travel time savings

As discussed previously, "treated" firms are able to significantly increase their operating efficiency and in turn improve their performance and growth rate after being connected to the HSR network. One of the most important forms of efficiency gain is the cost and time savings from inter-city travels. Before the high-speed trains, the fastest trains were *Z*-series trains, which were non-stop express trains with a maximum speed of 160 km/h and an average speed of 71.6 km/h.[39] Since cities were also inter-connected via the regular speed rail network before high-speed trains becomes available, it is reasonable to expect that the greater the travel time savings from the HSR compared to the regular-speed rail, the larger the improvement on the performance for the communication-intensive and travel-dependent firms when connected to the HSR network.[40]

Using information on the trip lengths for both regular and high-speed trains between all city pairs in our sample, we measure the average travel time reductions for each city due to the HSR connection. Specifically, we first calculate the differences in travel time between each city pair in our sample between regular- and high-speed trains.[41] We then calculate the average travel time savings for each city,$HSR^{TimeSaving}$, from high-speed trains compared with regular-speed trains. This variable is set to zero before the city is connected to the HSR network. Results are reported in Panel A of Table 11. Columns (1–5) report coefficients for the $HSR^{TimeSaving}$ and its interaction with the indicator for our "treated" firms. Similar to our baseline analyses, we test two sets of outcome variables, firm performance (columns (1–3)) and firm size (columns (4–5)). Results show that the improvement of firm performance and size growth associated with HSR connection for "treated" firms increases with travel time savings.[42]

5.2.2. Agglomeration economies

There is a large body of literature on the productiveness of cities. To begin with, a concentration of human capital facilitates learning and information sharing, and thus promotes productivity (Marshall, 1890). If communication-intensive and travel-dependent firms in the largest cities, such as Beijing and Shanghai, are highly agglomerated, we would expect "treated" firms in other cities that are connected to these large cities to enjoy even larger benefits from the HSR network. In addition, the HSR connection enables "treated" firms in smaller cities to access consumers in these large cities to market their own products and services without having to pay the high rents for physically operating in these high-cost locations. To directly test the role of agglomeration economies associated with large cities in the relationship between HSR connection and the performance of the communication-intensive and travel-dependent firms, we augment our baseline model with manually collected information that measures, for each firm in each year, their degree of connectedness to each of these mega cities.

Mega-City Connections. Mega cities, such as Beijing, Shanghai, Guangzhou, and Shenzhen, are large densely populated urban metropolises with great economic and political importance.[43] We focus on our "treated" firms and examine the differences in the HSR effects due to differences in firms' accessibility to their regional mega cities. Table 5 presents the share of firms connected to mega cities via the HSR from 2003 to 2017. It is evident from the table that by 2010, more than half of our estimation sample was connected to at least one of the four mega cities in China. Therefore, we distinguish connected "treated" firms by the extent of their connection to mega cities and estimate a model similar to eq. (4.12), where $HSR^{Year}{}_{c,\,t-1}$ is replaced with $HSR^{Mega}{}_{c,\,t-1}$ or $HSR^{\#Mega}{}_{c,\,t-1}$. $HSR^{Mega}{}_{c,\,t-1}$ is an indicator for mega city connection by direct trips via HSR, and $HSR^{\#Mega}{}_{c,\,t-1}$ is the number of mega cities that city *c* is connected to via the HSR in year $t-1$, and can range from 0 to 3. We exclude firms headquartered in the four mega cities in this analysis in order (1) to avoid mixing mega-to-mega and non-mega-to-mega connections, and (2) to understand the impact of HSR connections to mega cities on firms in smaller cities.

To ensure that our mega city connection variables faithfully reflect the actual year in which each city is connected to different mega cities, we manually check the train information on the official railway service website to pinpoint all the stations in between the origin city and the destination mega city.[44] We also check news reports for the actual opening time for direct trips between different city pairs. A positive estimate would confirm the role of agglomeration economies in the HSR connection on "treated" firms' performance and growth. Panels B and C of Table 11 present the estimation results using $HSR^{Mega}{}_{c,\,t-1}$ and $HSR^{\#Mega}{}_{c,\,t-1}$, respectively. Results show that

[39] Other regular-speed trains include T-series trains (express trains, with a maximum speed of 140 km/h), K-series trains (fast trains, with a maximum speed of 120 km/h), and regular-speed trains (four-digit-numbered trains, with a maximum speed of 120 km/h or lower).

[40] We control for the volume of air passenger for each city in each year with *lnAirportPassengers* in all regressions. As for the highway, we assume that travel time savings would be similar to that of the regular trains when compared to the HSR, as the average travel speed on the highway is similar to that of the regular-speed trains, which is much lower than the speed of high-speed trains.

[41] Travel time savings are equal to the differences in travel time by high-speed trains compared to the fastest available regular-speed trains (measured in hours). Further details on the measure construction is provided in the Appendix.

[42] Results remain quantitatively similar if we use percentage of travel time saved with the HSR instead.

[43] Guangzhou and Shenzhen are only about 80 miles apart. Therefore, we assume that cities are automatically connected to both of them when they have access to one of them by the HSR, treating them as one metropolis. Our results remain unchanged if we measure them separately.

[44] The official railway service website is www.12306.cn.

Table 11
HSR Connection and Firm Performance: Extensions.

	Profitability			Size	
	EBIT	*ROA*	*ROE*	*lnAssets*	*lnEmployees*
	(1)	(2)	(3)	(4)	(5)
Panel A: Travel Time Savings of HSR					
$HSR^{TimeSaving}{}_{t-1}$	0.000	0.000	0.002**	−0.003	0.001
	(0.000)	(0.000)	(0.001)	(0.003)	(0.004)
$COMM^{High} * HSR^{TimeSaving}{}_{t-1}$	0.002***	0.001***	0.003***	0.012**	0.014*
	(0.001)	(0.000)	(0.001)	(0.006)	(0.008)
Observations	13,726	13,726	13,726	13,726	13,726
Adjusted R^2	0.259	0.302	0.185	0.897	0.838
Panel B: Connection to Mega Cities					
$HSR^{Mega}{}_{t-1}$	−0.006	−0.003	−0.002	−0.057**	−0.004
	(0.004)	(0.003)	(0.006)	(0.026)	(0.028)
$COMM^{High} * HSR^{Mega}{}_{t-1}$	0.015***	0.011**	0.034***	0.145**	0.353***
	(0.006)	(0.005)	(0.012)	(0.062)	(0.097)
Observations	9977	9977	9977	9977	9977
Adjusted R^2	0.281	0.337	0.206	0.876	0.841
Panel C: Number of Mega Cities Connected					
$HSR^{\#Mega}{}_{t-1}$	−0.000	0.001	0.004	−0.005	0.023
	(0.002)	(0.001)	(0.003)	(0.013)	(0.015)
$COMM^{High} * HSR^{\#Mega}{}_{t-1}$	0.005**	0.004*	0.011**	0.049*	0.142***
	(0.002)	(0.002)	(0.005)	(0.027)	(0.041)
Observations	9977	9977	9977	9977	9977
Adjusted R^2	0.281	0.337	0.206	0.876	0.841

Notes: The estimation sample in Panel A includes all firms connected to the HSR network by the end of 2017, and the estimation sample in Panels B and C excludes firms headquartered in the four mega cities. Firms that are established in 2003 or later are excluded from the estimation sample. Dependent variables are indicated by column titles in italics. All regressions include the same firm-level and city-level control variables as in Table 7, as well as firm fixed effects, and year fixed effects. Robust standard errors are clustered at the firm level.

"treated" firm benefit more from being connected to mega cities via the HSR, and confirm the existence of agglomeration economies in the relationship of the HSR and firm performance for the communication-intensive and travel-dependent. Moreover, the "treated" firms benefit more from being connected to more mega cities through the HSR in the form of improved profitability and further size growth. For example, "treated" firms that are connected to the HSR network on average have 0.5 percentage point higher EBIT, for being connected to an additional mega city, compared to firms in the control group (column (1) of Panel C).

Time to Nearest Mega City. We next test if distance to mega city, as measured in travel time, matters for the effect of HSR connection on firm performance in the communication-intensive and travel-dependent industries. Our hypothesis is that the firms within the medium range to their nearest mega cities would benefit more from being connected to the HSR network. For short travels, highways may be more flexible and the time-saving might be limited, if we compare the highway and the HSR.[45] However, for long trips, air travel may be a more attractive option. In this case, we argue that the HSR is likely more frequently used for medium-range business travel, and hypothesize that the "treated" firms that are within medium-range distances to mega cities might benefit the most from their connection to large cities.

Specifically, we exclude firms in mega cities from our estimation sample in this exercise. We measure firms' travel time to their nearest mega cities in three-hour intervals, 1–3 hoursh, 3–6 hours, and 6 hours or more. The heterogeneous effects of the HSR connection to mega cities on the performance of the communication-intensive and travel-dependent firms are estimated by replacing $HSR_{c,\,t-1}$ in our baseline model (4.11) with three indicators of travel time to the nearest mega cities. Firm-year observations before the HSR connections make up the reference group in this specification. The coefficients of the three interactions terms between our "treated" indicator and the three travel-time indicators are plotted in Fig. 8 separately. Consistent with our hypothesis, we find more significant results for the "treated" firms located 3–6 hours away from their nearest mega cities. Our empirical findings are similar for our firm performance and size measures. This evidence confirms the importance of firm locations, especially for firms that have higher rail passenger input and are more dependent on interactive non-routine works in their daily operation.

6. Conclusions

In this paper, we focus on the HSR network in China, and study its impact on firms' performance and growth through substantial

[45] For example, traveling from the center of Beijing to the center of a nearby city, Tianjin, takes about 2 h by car, and about 1 h and 50 min by the HSR, taking into account time spent traveling to and from the HSR stations within city.

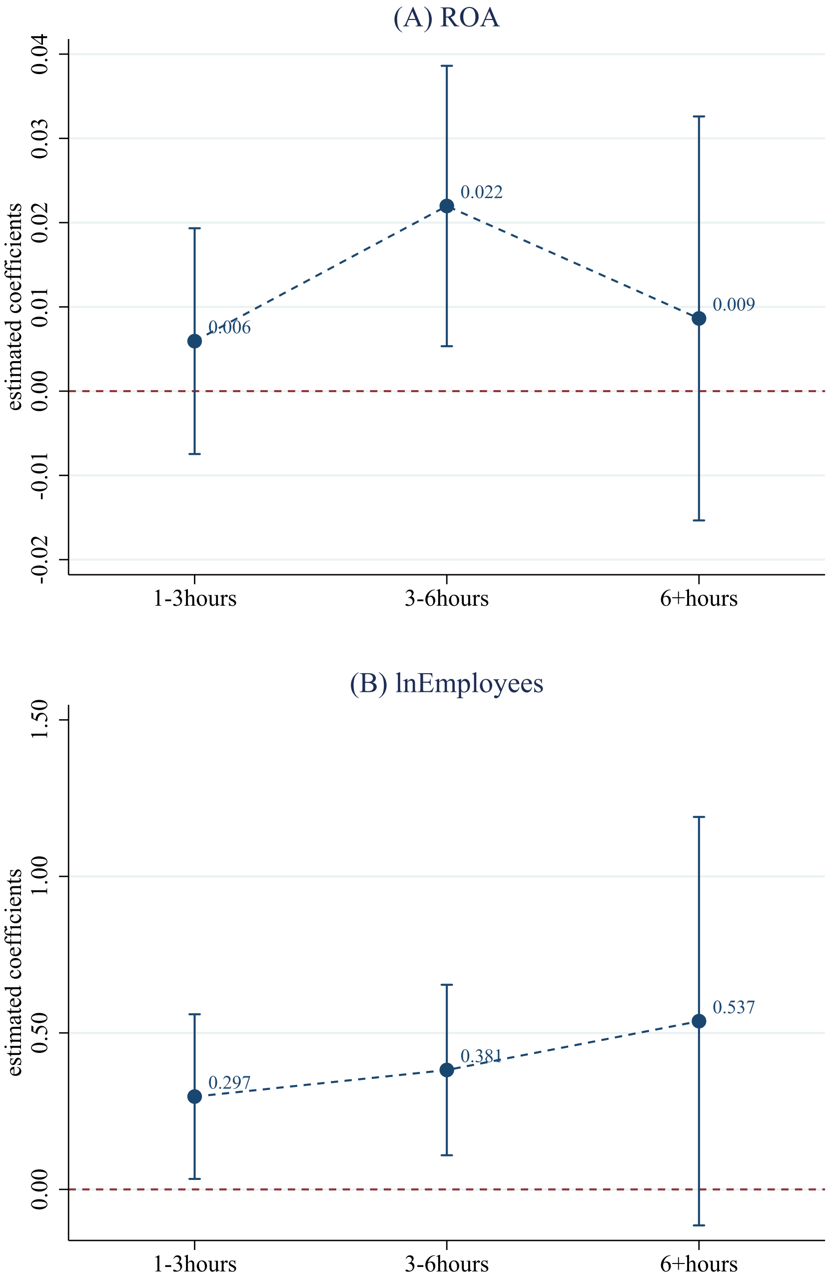

Fig. 8. HSR Connection and Firm Performance: Travel Time to Nearest Mega City.
Notes: The graph plots the estimated coefficients and the confidence intervals of the interaction terms between the "treated" firm indicator and three-hour travel time intervals between firms and their nearest mega cities.

reductions in communication costs. Using an annual panel of firm-level data for the period of 2004–2017, we test the differential impact of HSR connection on firm performance and size growth for firms that differ in their dependence on communication and inter-city travel as production inputs. Our empirical strategy is a difference-in-differences framework with firm and year fixed effects, and firm-level financials and city-level characteristics as controls. The results confirm our hypotheses that firms in communication and travel intensive industries benefit more from the operation of the HSR, and that the effect is persistent over time. This finding is robust to (1) restricting to a sub-sample of single-location firms, (2) instrumenting the HSR connection with historical railways, (3) alternative measures of travel intensity, and (4) using matching estimators.

Moreover, in examining the specific mechanisms at work, we find evidence supporting the argument that HSR connection promotes the performance and size growth of firms in communication-intensive and travel-dependent industries through increased analyst attention, productivity boosts, and market expansions. We also find evidence of agglomeration economies in that communication-intensive and travel-dependent firms that are connected to (multiple) mega cities perform better and grow faster. Lastly, we find that more intensive utilization of HSR travel are associated with more sizable benefits enjoyed by the "treated" firms. This is consistent with the findings that the impact of the HSR on the communication-intensive and travel-dependent firms increases with the savings on travel time, and that firms located about 3–6 hoursh away from their nearest mega cities benefit more.

By exploiting the staggered expansion of China's passenger-dedicated HSR network, we highlight the importance of firm location and face-to-face interactions. Our results suggest that a massive inter-city travel network like this changes the relative distances between firms and the rest of the market, potentially making up for some of the disadvantages associated with firm location which were assumed to be permanent. Our findings also imply that face-to-face interactions, by facilitating the exchange and acquisition of knowledge, are potentially important for our understanding of "local (home) bias" of investments.

Acknowledgements

Zhu acknowledges funding supported by fund for building world-class universities (disciplines) of Renmin University of China (KYGJC2020008). Kuang acknowledges funding supported by the fundamental research funds for the central universities in UIBE (19QN07).

Appendix. Data and Measure Details

1. Multiple segments of the same line open at different times

Some HSR lines have several segments completed at different times. Cities along different segments of the same HSR line that are finished in different years are only connected when all segments are completed. For example, the Beijing–Guangzhou HSR line (or Jing-Guang HSR line for short) has three segments. The Wuhan–Guangzhou segment was finished in 2009, but the Beijing–Zhengzhou and Zhengzhou–Wuhan segments were completed in 2012. In this case, although Wuhan and Guangzhou were connected by the HSR in 2009, Beijing and Guangzhou were not connected by the HSR until 2012. We manually check the segment information on the same HSR lines and update the HSR measure for each city year by year to reflect the reality of HSR connection for each city.

2. Multiple stations in the same city

There might be several stations in one city. For instance, there are three HSR stations in Beijing, namely Beijing South, Beijing West, and Yizhuang. In this case, the earliest of all stations in a city determines the time when said city became connected to the HSR network. We do not differentiate cities with different numbers of HSR stations in this study.

3. Manual check of mega city connections.

To truthfully reflect the status of the mega cities connections of each firm in our sample over time, we manually check for each headquarter city the year in which a direct HSR train connecting said city and each of the four mega cities became operational. Only direct trips without transfer between each origin city and one of the mega cities are counted.[46] Fig. A1 demonstrates the construction of $HSR^{\#Mega}$ using the city of Nanjing as an example. In 2010, of the four mega cities, Nanjing is only connected to Shanghai via HSR by the Shanghai-Nanjing Line. $HSR^{\#Mega}$ for Nanjing is recorded as 1 in 2010. In 2011, the Beijing-Shanghai Line was completed, and Nanjing was connected to Beijing along this line. Nanjing was directly connected to Guangzhou in 2012, using different segments of two HSR lines, the Nanjing–Chengdu Line (from Nanjing to Yichang) and the Beijing–Guangzhou Line (from Yichang to Guangzhou). It is worth noting that although the trip from Nanjing to Guangzhou consists of different segments of two HSR lines, it is indeed a direct trip without connection. By 2012, the only mega city that Nanjing did not have direct connection to via HSR was Shenzhen. However, since Shenzhen is in the same province and is only 80 miles away from Guangzhou, it is reasonable to assume that people can travel relatively easily between the two mega cities. Nonetheless, we manually check that in 2013, Nanjing was connected to Shenzhen directly via the HSR, through the Nanjing–Hangzhou Line (from Nanjing to Hangzhou) and the Hangzhou–Shenzhen Line (from Hangzhou to Shenzhen).

4. Measuring firm subsidiaries.

A few filters for data quality control are applied when constructing measures of firm subsidiaries. More specifically, we deleted subsidiaries-year observations (1) that were flagged as "closed"; (2) that were marked as "moved away"; (3) whose registered capital were recorded in foreign currencies; (4) whose registered capital information was missing; (5) that had registered capital exceeding the shareholders' equity of the holding companies in the same year; and (6) for which location cannot be identified.[47] For about 22% observations missing currency information for registered capital, we assume that the currency used was RMB.

5. Measuring travel time savings.

Travel time savings are calculated for each city-pair in our sample, and is equal to the differences in travel time by high-speed trains compared to the fastest available regular speed trains (measured in hours or percentages). For city-pairs with no information on regular-speed trains, it is possible that (1) there was no regular speed train before the two cities were connected by the HSR, or (2) the regular speed train was suspended after high-speed trains became available. For these observations, we project the travel time for regular speed trains as the ratio of the railway travel distance between the two cities over the average travel speed of regular speed trains.

[46] In our sample, in addition to direct connection between cities that are on the same HSR line, there can also be indirect connections through different HSR lines. For instance, the trip from Shanghai to Shenzhen could go through three different HSR lines, the Shanghai–Kunming Line, the Beijing–Guangzhou Line, and the Beijing–Hongkong Line. The train departs from Shanghai Hongqiao Station to Changsha South Station via the Shanghai–Kunming HSR line, then goes towards Guangzhou South via the Beijing–Guangzhou HSR, finally arriving at Shenzhen North through the Beijing–Hongkong Line.

[47] 12,785 out of a total of 156,691 observations were deleted as a result of these filters, amounting to 8.16% of the full subsidiary sample.

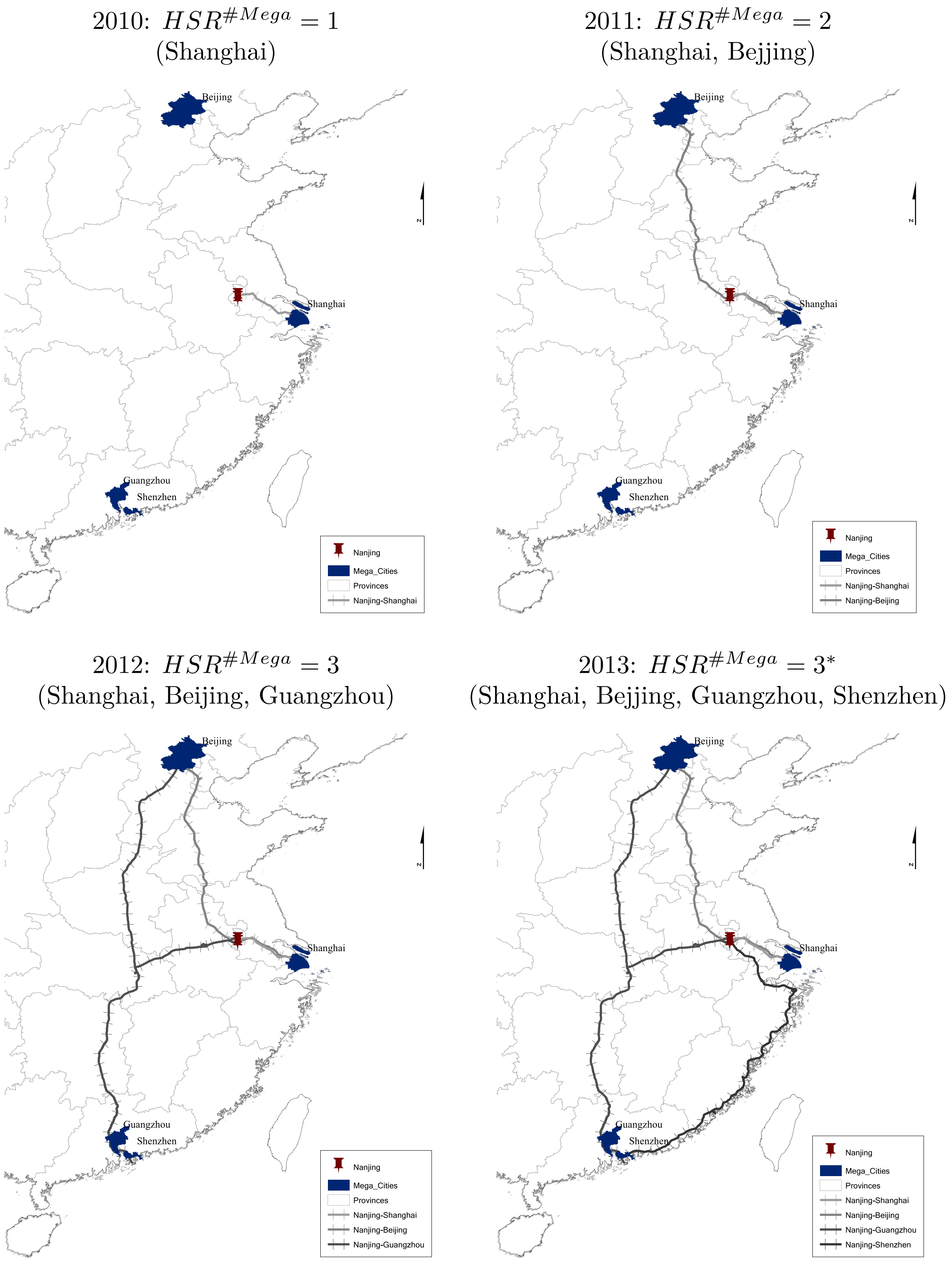

Fig. A1. Demonstrating the Construction of $HSR^{\#Mega}$ using Nanjing as an Example.
Notes: *Because Guangzhou and Shenzhen are only 80 miles apart, we assume automatic connection between Guangzhou and Shenzhen, and record $HSR^{\#Mega}$ for Nanjing in 2012 and 2013 as 3, treating both cities as one metropolis. Our results are highly similar when measuring them separately.

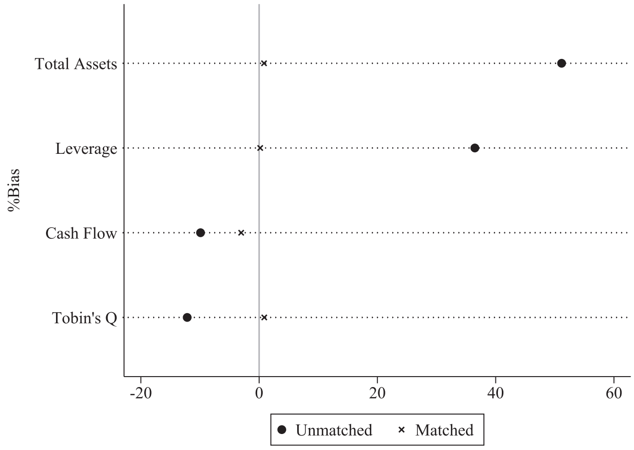

Fig. A2. Standardized % Bias Reduction from Matching.
Notes: The graph plots the standardized percentage bias before and after matching. The standardized % bias is the percentage difference of the sample means in the "treated" versus control sub-samples as a percentage of the square root of the average of the sample variances in the "treated" and control groups (Rosenbaum and Rubin, 1985).

Table A1
Summary Statistics of Subsidiaries.

	Mean	p(25)	p(50)	p(75)
Panel A: All Firms				
Number of				
– Subsidiaries	8.64	3	5	10
– Subsidiaries at Other Locations	2.67	0	0	2
– Subsidiaries at Other Locations in Same Industry	1.10	0	0	0
# of Firms		1508		
# of Firm-Year Observations		13,726		
Panel B: All Subsidiaries				
(%) Registered Capital/Holding Company Equity				
– All Subsidiaries	4.21	0.29	1.15	3.97
– Subsidiaries at Other Locations	3.34	0.23	0.94	3.13
– Subsidiaries at Other Locations in Same Industry	3.81	0.37	1.30	3.84
# of Subsidiaries		29,415		
# of Firm-Subsidiary-Year Observations		119,496		

Notes: Panel A is reported using the full set of firms in our estimation sample, while Panel B includes only firms with at least one subsidiary. "Subsidiaries at Other Locations" refers to subsidiaries at locations different from the holding companies' headquarters. "Subsidiaries at Other Locations in Same Industry" are subsidiaries in the same industries but different locations as the holding companies.

Table A2
HSR Connection and Firm Performance: Robustness Checks.

	Profitability			Size	
	EBIT	*ROA*	*ROE*	*lnAssets*	*lnEmployees*
	(1)	(2)	(3)	(4)	(5)
Panel A: Other-Location Subsidiaries' Total Equity Share ≤ p(75)					
HSR_{t-1}	−0.001	0.000	0.002	−0.057***	−0.037
	(0.004)	(0.003)	(0.008)	(0.022)	(0.024)
$COMM^{High} * HSR_{t-1}$	0.015***	0.013***	0.030***	0.124**	0.203***
	(0.005)	(0.004)	(0.010)	(0.053)	(0.072)
Observations	10,195	10,195	10,195	10,195	10,195
Adjusted R^2	0.284	0.340	0.185	0.901	0.849
Panel B: Other-Location Subsidiaries' Total Equity Share ≤ p(90)					
HSR_{t-1}	0.000	0.001	0.005	−0.056***	−0.026
	(0.003)	(0.003)	(0.007)	(0.020)	(0.023)

(continued on next page)

Table A2 (*continued*)

	Profitability			Size	
	EBIT	*ROA*	*ROE*	*lnAssets*	*lnEmployees*
	(1)	(2)	(3)	(4)	(5)
$COMM^{High} * HSR_{t-1}$	0.015***	0.012***	0.029***	0.121**	0.159**
	(0.005)	(0.004)	(0.009)	(0.049)	(0.070)
Observations	13,021	13,021	13,021	13,021	13,021
Adjusted R^2	0.273	0.317	0.196	0.902	0.842

Notes: The estimation sample includes firms that have subsidiaries at other locations with a combined equity share less than the 75^{th} percentile (=1.41%) or 90^{th} percentile (=9.65%) of the baseline estimation sample respectively. Dependent variables are indicated by column titles in italics. All regressions include the same control variables, firm fixed effects, and year fixed effects as in Table 7. Robust standard errors are clustered at the firm level.

Table A3
Balancing Tests on Matched Sample.

Covariates	Sample	Mean		t-test	
		Treated	Control	*t*	$p >\| t\|$
Total Assets	Unmatched	22.507	21.799	18.86	0.000
	Matched	22.507	22.496	0.18	0.860
Leverage	Unmatched	0.525	0.452	12.07	0.000
	Matched	0.525	0.525	0.04	0.966
Cash Flow	Unmatched	0.0377	0.060	−5.22	0.000
	Matched	0.037	0.044	−0.70	0.483
Tobin's Q	Unmatched	2.271	2.488	−3.78	0.000
	Matched	2.271	2.255	0.22	0.826

Notes: The table shows the t-test results for firm covariates used as matching criteria both before and after the Mahalanobis distance matching procedure between the treated and control sub-samples. *t*-tests are performed based on a regression of the variable on a treatment indicator using the on-support sample.

References

Abadie, Alberto, Imbens, Guido W., 2006. Large sample properties of matching estimators for average treatment effects. Econometrica 74 (1), 235–267.
Ackerberg, Daniel, Caves, Kevin, Frazer, Garth, 2015. Identification properties of recent production function estimators. Econometrica 83 (6), 2411–2451.
Almeida, Heitor, Campello, Murillo, Laranjeira, Bruno, Weisbenner, Scott, et al., 2012. Corporate debt maturity and the real effects of the 2007 credit crisis. Critical Finance Review 1 (1), 3–58.
Arena, Matteo, Dewally, Michaël, 2012. Firm location and corporate debt. J. Bank. Financ. 36 (4), 1079–1092.
Autor, David H., Levy, Frank, Murnane, Richard J., 2003. The skill content of recent technological change: An empirical exploration. Q. J. Econ. 118 (4), 1279–1333.
Baik, Bok, Kang, Jun-Koo, Kim, Jin-Mo, 2010. Local institutional investors, information asymmetries, and equity returns. J. Financ. Econ. 97 (1), 81–106.
Banerjee, Abhijit, Duflo, Esther, Qian, Nancy, 2020. On the road: Access to transportation infrastructure and economic growth in China. Journal of Development Economics 102442.
Baum, Christopher, Schaffer, Mark E., Stillman, Steven, 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. Stata J. 7 (4), 465–506.
Baum-Snow, Nathaniel, Brandt, Loren, Henderson, J. Vernon, Turner, Matthew A., Zhang, Qinghua, 2017. Roads, railroads, and decentralization of Chinese cities. Rev. Econ. Stat. 99 (3), 435–448.
Bernard, Andrew, Moxnes, Andreas, Saito, Yukiko U., 2019. Production networks, geography, and firm performance. J. Polit. Econ. 127 (2), 639–688.
Bernile, Gennaro, Kumar, Alok, Sulaeman, Johan, 2015. Home away from home: Geography of information and local investors. Rev. Financ. Stud. 28 (7), 2009–2049.
Bernstein, Shai, Giroud, Xavier, Townsend, Richard R., 2016. The impact of venture capital monitoring. J. Financ. 71 (4), 1591–1622.
Booth, Laurence, Aivazian, Varouj, Demirguc-Kunt, Asli, Maksimovic, Vojislav, 2001. Capital structures in developing countries. J. Financ. 56 (1), 87–130.
Boubakri, Narjess, Guedhami, Omrane, Saffar, Walid, 2016. Geographic location, foreign ownership, and cost of equity capital: evidence from privatization. J. Corp. Finan. 38, 363–381.
Campello, Murillo, Graham, John R., Harvey, Campbell R., 2010. The real effects of financial constraints: Evidence from a financial crisis. J. Financ. Econ. 97 (3), 470–487.
Carosi, Andrea, 2016. Do local causations matter? The effect of firm location on the relations of ROE, R&D, and firm SIZE with MARKET-TO-BOOK. J. Corp. Finan. 41, 388–409.
Chang, Xin, Dasgupta, Sudipto, Hilary, Gilles, 2006. Analyst coverage and financing decisions. J. Financ. 61 (6), 3009–3048.
Charnoz, Pauline, Lelarge, Claire, Trevien, Corentin, 2018. Communication costs and the internal organisation of multi-plant businesses: Evidence from the impact of the French high-speed rail. Economic Journal 128 (610), 949–994.
Chhaochharia, Vidhi, Kumar, Alok, Niessen-Ruenzi, Alexandra, 2012. Local investors and corporate governance. J. Account. Econ. 54 (1), 42–67.
Coval, Joshua, Moskowitz, Tobias J., 2001. The geography of investment: Informed trading and asset prices. J. Polit. Econ. 109 (4), 811–841.
Cumming, Douglas, Dai, Na, 2010. Local bias in venture capital investments. J. Empir. Financ. 17 (3), 362–380.
Degryse, Hans, Ongena, Steven, 2005. Distance, lending relationships, and competition. J. Financ. 60 (1), 231–266.
Degryse, Hans, Ongena, Steven, 2007. The impact of competition on bank orientation. J. Financ. Intermed. 16 (3), 399–424.
Derrien, François, Kecskés, Ambrus, 2013. The real effects of financial shocks: evidence from exogenous changes in analyst coverage. J. Financ. 68 (4), 1407–1440.
Donaldson, Dave, 2018. Railroads of the raj: estimating the impact of transportation infrastructure. Am. Econ. Rev. 108 (4–5), 899–934.
Dunne, Paul, Hughes, Alan, 1994. Age, size, growth and survival: UK companies in the 1980s. Journal of Industrial Economics 115–140.
Faber, Benjamin, 2014. Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. Review of Economic Studies 81 (3), 1046–1070.
Fernald, John, 1999. Roads to prosperity? Assessing the link between public capital and productivity. Am. Econ. Rev. 89 (3), 619–638.

Gao, Wenlian, Ng, Lilian, Wang, Qinghai, 2011. Does corporate headquarters location matter for firm capital structure? Financ. Manag. 40 (1), 113–138.
Giroud, Xavier, 2013. Proximity and investment: evidence from plant-level data. Q. J. Econ. 128 (2), 861–915.
Hart, Peter, Oulton, Nicholas, 1996. Growth and size of firms. Economic Journal 106 (438), 1242–1252.
Ivković, Zoran, Weisbenner, Scott, 2005. Local does as local is: information content of the geography of individual investors' common stock investments. J. Financ. 60 (1), 267–306.
Jiang, Wei, 2017. Have instrumental variables brought us closer to the truth. Rev. Corp. Financ. Stud. 6 (2), 127–140.
John, Kose, Knyazeva, Anzhela, Knyazeva, Diana, 2011. Does geography matter? Firm location and corporate payout policy. J. Financ. Econ. 101 (3), 533–551.
Kaplan, Steven, Zingales, Luigi, 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? Q. J. Econ. 112 (1), 169–215.
Ke, Xiao, Chen, Haiqiang, Hong, Yongmiao, Cheng, Hsiao, 2017. Do China's high-speed-rail projects promote local economy? New evidence from a panel data approach. China Econ. Rev. 44, 203–226.
Lamont, Owen, Polk, Christopher, Saaá-Requejo, Jesús, 2001. Financial constraints and stock returns. Rev. Financ. Stud. 14 (2), 529–554.
Lawrence, Martha, Bullock, Richard, Liu, Ziming, 2019. China's High-Speed Rail Development. The World Bank.
Lerner, Josh, 1995. Venture capitalists and the oversight of private firms. J. Financ. 50 (1), 301–318.
Levinsohn, James, Petrin, Amil, 2003. Estimating production functions using inputs to control for unobservables. Rreview of Economic Studies 70 (2), 317–341.
Li, Han, Li, Zhigang, 2013. Road investments and inventory reduction: firm level evidence from China. J. Urban Econ. 76, 43–52.
Li, Yifan, 2016. China High Speed Railways and Stations (2016). In: Data retrieved from Harvard Dataverse. https://doi.org/10.7910/DVN/JIISNB.
Lin, Yatang, 2017. Travel costs and urban specialization patterns: evidence from China's high speed railway system. J. Urban Econ. 98, 98–123.
Loughran, Tim, 2008. The impact of firm location on equity issuance. Financ. Manag. 37 (1), 1–21.
Mansfield, Edwin, 1962. Entry, Gibrat's law, innovation, and the growth of firms. Am. Econ. Rev. 52 (5), 1023–1051.
Marshall, Alfred, 1890. Principles of Economics. Macmillan.
Michaels, Guy, 2008. The effect of trade on the demand for skill: evidence from the interstate highway system. Rev. Econ. Stat. 90 (4), 683–701.
Nadiri, M. Ishaq, Mamuneas, Theofanis P., 1994. The effects of public infrastructure and R&D capital on the cost structure and performance of U.S. manufacturing industries. Rev. Econ. Stat. 76 (1), 22–37.
Petersen, Mitchell, Rajan, Raghuram G., 2002. Does distance still matter? The information revolution in small business lending. J. Financ. 57 (6), 2533–2570.
Pirinsky, Christo, Wang, Qinghai, 2006. Does corporate headquarters location matter for stock returns? J. Financ. 61 (4), 1991–2015.
Qin, Yu, 2017. No county left behind? The distributional impacts of high-speed rail upgrades in China. J. Econ. Geogr. 17 (3), 489–520.
Rajan, Raghuram, Zingales, Luigi, 1995. What do we know about capital structure? Some evidence from international data. J. Financ. 50 (5), 1421–1460.
Redding, Stephen J., Turner, Matthew A., 2015. Transportation costs and the spatial organization of economic activity. In: Handbook of Regional and Urban Economics, Vol 5. Elsevier, pp. 1339–1398.
Roberts, Michael R., Whited, Toni M., 2013. Endogeneity in empirical corporate finance. In: Handbook of the Economics of Finance, Vol. 2. Elsevier, pp. 493–572.
Rosenbaum, Paul, Rubin, Donald B., 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am. Stat. 39 (1), 33–38.
Shah, Anwar, 1992. Dynamics of public infrastructure, industrial productivity and profitability. Review of Economics and Statistics 28–36.
Sialm, Clemens, Sun, Zheng, Zheng, Lu, 2020. Home bias and local contagion: evidence from funds of hedge funds. Rev. Financ. Stud. 33 (10), 4771–4810.
Staiger, Douglas, Stock, James H., 1997. Instrumental Variables Regression with Weak Instruments. Econometrica 65 (3), 557–586.
Stock, James H., Yogo, Motohiro, 2005. Testing for Weak Instruments in Linear IV Regression. In: Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, p. 80.
Sufi, Amir, 2007. Information asymmetry and financing arrangements: evidence from syndicated loans. J. Financ. 62 (2), 629–668.
Whited, Toni, Wu, Guojun, 2006. Financial constraints risk. Rev. Financ. Stud. 19 (2), 531–559.
Yu, Feng, Lin, Faqin, Tang, Yihong, Zhong, Changbiao, 2018. High-Speed Railway to Success? The Effects of High-Speed Rail Connection on Regional Economic Development in China. Journal of Regional Science.
Zheng, Siqi, Kahn, Matthew E., 2013. China's bullet trains facilitate market integration and mitigate the cost of megacity growth. Proc. Natl. Acad. Sci. 110 (14), E1248–E1253.