DOI: 10.1111/rode.12642

REGULAR ARTICLE

# Does high-speed rail connection really promote local economy? Evidence from China's Yangtze River Delta†

Yanyan Gao[1] | Shunfeng Song[2,3] | Jun Sun[4] | Leizhen Zang[5]

[1]School of Economics and Management, Southeast University, Nanjing, China
[2]Department of Economics, University of Nevada, Nevada, USA
[3]Institute of Urban Development, Nanjing Audit University, Nanjing, China
[4]Business School, Jiangsu Ocean University, Jiangsu, China
[5]College of Humanities and Development Studies, China Agricultural University, Beijing, China

**Correspondence**
Shunfeng Song, Department of Economics, University of Nevada, Reno, NV 89557 USA.
Email: song@unr.edu

**Funding information**
The Major Project of China National Social Science Foundation, Grant/Award Number: 18ZDA077; Project of Humanities and Social Sciences, Grant/Award Number: 17YJC790040; Fundamental Research Funds for the Central Universities, Grant/Award Number: 2242018S20023

**Abstract**
High-speed rail (HSR) has led to a transportation revolution in China. This paper uses the county-level panel data of China's Yangtze River Delta to investigate the effect of HSR connection on local economy. To address the issue of endogenous HSR route placement, we use a straight-line strategy to construct potential HSR connection variables as instrumental variables of the actual HSR connection. Both the difference-in-differences and instrumental variable methods show that HSR connection impedes local economy, especially in peripheral regions. The impediment effect is channeled through population reallocation from peripheral to core areas and the restructuring of industries.

**KEYWORDS**
difference-in-differences, GDP per capita, high-speed rail, Yangtze River Delta

## 1 | INTRODUCTION

The railway has long played an important role in spatial resource allocation. China's railway network, even after increasing the speed of its rails six times, remains unable to fulfill the needs of modern

†Present address: Yanyan Gao, Mail box no. L10, Jingguan Building, Dongnandaxue Road 2, Jiangning District, Nanjing, 211189, China.

  

society for fast and comfortable transport. In recent years, the Chinese government has vigorously promoted the construction of high-speed rail (HSR) networks that, compared with traditional passenger railways, are much faster, more comfortable, and more punctual. The first HSR line in China, the Beijing–Tianjin intercity HSR, was operated on August 1, 2008. Through 2015, China constructed a "four vertical and four horizontal" HSR framework, with a total HSR line of over 20,000 km (Figure 1). As a result, China has the largest HSR network in the world. According to the promulgated mid- to long-term railway network plan approved by the State Council of China in July 2016, by the end of 2020, China's HSR networks are expected to reach 30,000 km, covering 80% of large cities with "eight vertical and eight horizontal" HSR channels.

With the hope of stimulating local economic development, Chinese local governments actively compete for HSR connection. However, two theories have provided different inferences on the effect of transportation improvement on local economic development. First, new economic geography (NEG) (Fujita, Krugman, & Venables, 2001; Helpman & Krugman, 1985; Krugman, 1991) argues that a reduction in transport costs will lead to trade integration and finally economic polarization. Some empirical evidence is consistent with this theory. For example, Fogel (1962) argued that the U.S. railway failed to promote local economic development compared with its waterways. The adverse effects of connecting to highway or railway on peripheral economic growth were also observed by Faber (2014), Vickerman (2015), and Hodgson (2018). Second, urban economics (Alonso, 1964;

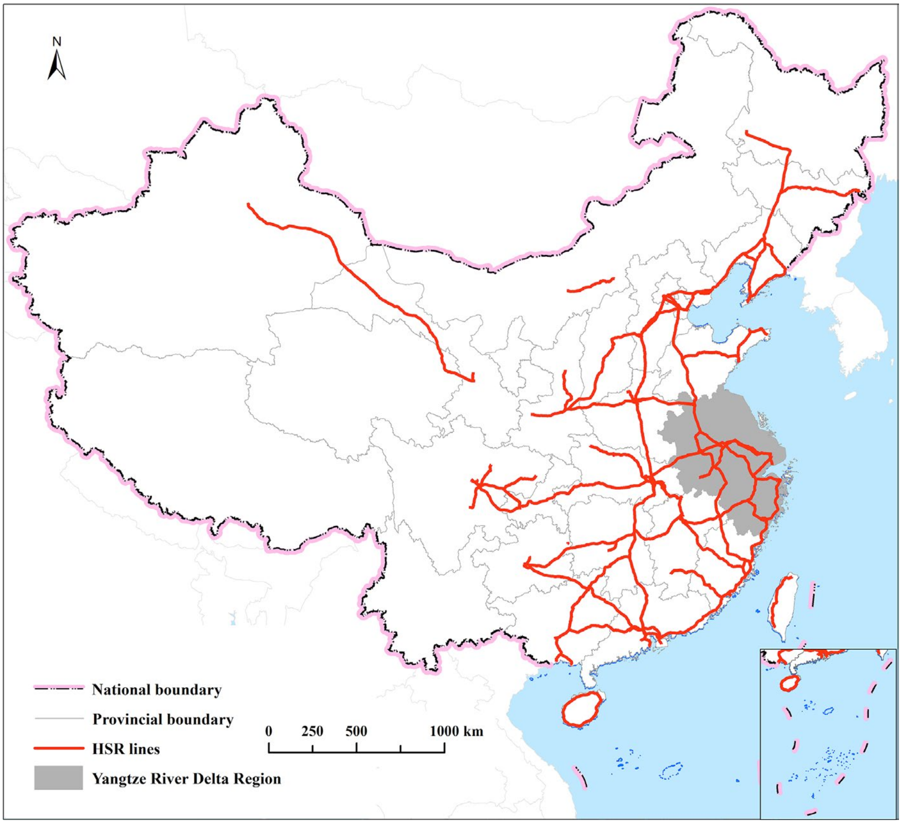


**FIGURE 1** China's HSR lines in 2016
*Source:* Authors' graph based on the HSR data from Li (2016).

Muth, 1969) predicts that the increase in accessibility from improving transport disperses economic activities from the central city to peripheral areas and thus produces effects locally, such as outflow of populations, dispersion of industries, and economic growth in connected areas (Baum-Snow, 2007; Baum-Snow, Brandt, Henderson, Turner, & Zhang, 2017; Chen & Hall, 2011; Deng, Wang, Yang, & Yang, 2019; Zheng & Kahn, 2013).

As a new mode of transport, HSR also leads to reduction in transport costs and improves accessibility. It can also be inferred from the two theories that HSR either polarizes or disperses regional economies. Extant empirical literature provides mixed evidence. On one hand, HSR is found to be beneficial to its connected cities by increasing accessibility (Sasaki et al., 1997), facilitating knowledge diffusion and idea spillovers (Dong, Zheng, & Kahn, 2018), increasing employment (Lin, 2017), promoting export and economic agglomeration (Ahlfeldt & Feddersen, 2018; Chen & Hall, 2011; Ke, Chen, Hong, & Hsiao, 2017; Shao, Tian, & Yang, 2017), and enhancing market integration and dispersing congested urban economies (Zheng & Kahn, 2013). On the other hand, evidence shows the economic polarization effects of HSR (Banister & Berechman, 2001; Qin, 2017; Vickerman, 2015). It has been shown that HSR improves the accessibility of both peripheral and core cities, and sometimes central cities benefit more (Sasaki et al., 1997). Such asymmetric accessibility improvement of HSR centralizes economies toward large cities (Banister & Berechman, 2001) and impedes economic growth in connected peripheral cities (Qin, 2017; Vickerman, 2015). Therefore, the actual effect of HSR connection on economic outcomes remains an empirical question that needs further investigation. This paper provides new evidence on the effect of HSR on local economic development, using county-level panel data from China's Yangtze River Delta (YRD).

Several empirical studies are highly relevant to our paper. They all use data from China though they obtained different results on the effect of HSR connection on economic growth. For example, Ke et al. (2017) use select prefectural-level data from 1990 to 2013, which consisted of 21 HSR cities along 4 HSR lines and 11 carefully chosen unconnected cities. They find with a counterfactual analysis that the effect of HSR on GDP per capita is heterogeneous over location, route, and region, ranging from 5% to 59%. They further find that coastal cities benefit more from HSR connection. Qin (2017) explores the two rounds of HSR upgrades in existing passenger railway lines as quasi-experiments to test the economic distribution effect of improvement on transport infrastructures. She finds that such upgrades decrease a county's total GDP and GDP per capita by 3%–5%. However, the HSR used by Qin (2017) is not the HSR defined by the International Union of Railways or the National Railway Administration of China (NRAC) in which an HSR refers to a rail with an operation speed of at least 200 km/hr and a designed speed of at least 250 km/hr. The data used by Qin (2017) are from 2002 to 2009, whereas the first HSR line recognized by NRAC is the Beijing–Tianjin intercity HSR that opened on August 1, 2008.

Our paper is also related to some previous studies that focus on the YRD and explore the HSR impacts. For example, using the difference-in-difference (DID) method, Shao et al. (2017) estimate the effect of HSR service intensity on city service industry in the YRD and find that HSR promotes the agglomeration of city service industry. Other studies using data from this region explore the accessibility-improvement impacts of HSR (Jin, Xu, Lu, & Huang, 2013; Wang, 2018; Wang & Duan, 2018). Although it is well agreed that HSR improves city accessibility, the enhancement is found producing spatially asymmetric effects across different levels of cities. Compared with peripheral cities, central cities benefit more from connecting to HSR in terms of accessibility improvement (Jin et al., 2013), population growth (Wang & Duan, 2018), housing price (Diao, Zhu, & Zhu, 2016), and economic activities regarding investment and consumption (Li, Huang, Li, & Zhang, 2016). However, to the best of our knowledge, a gap still exists in this strand of empirical literature in identifying the aggregate impact of HSR on local economy.

The core issue in disentangling the causality is that the placement of the HSR route is endogenous, leading to biased estimation. At the local or regional level, it is well observed in China that cities with a larger economy or greater political importance would be considered first for HSR connection. Also, local officials having a stronger desire to develop the economy or a closer connection with higher-level officials often lobby harder for their cities to be connected with HSR. Therefore, economic performance and HSR connection could mutually affect each other, suggesting a possible overestimate of the HSR role in promoting local economy. At the national level, to help reduce poverty and regional disparity, the central government may construct HSR lines that favor less-developed regions. This action, if not controlled for, tends to understate the effect of HSR on economy.

The DID method was commonly used in previous studies to estimate the effect of HSR connection on local economy (Ahlfeldt & Feddersen, 2018; Lin, 2017; Qin, 2017; Shao et al., 2017; Zhang, 2017). DID compares the difference of the outcome variable at two levels between the treatment and control groups, before and after the treatment, or both. A valid causality reference from DID, however, requires fulfilling the common trend assumption; that is, the two groups have the same trends over time (Angrist & Pischke, 2014). Because that assumption does not necessarily stand, the related literature also uses the instrumental variables (IVs) method to address the issue of endogenous route placement of transport lines.

To construct IVs for actual transport connections, previous studies either explored historical information (Baum-Snow et al., 2017; Donaldson, 2018; Michaels, 2008; Zheng & Kahn, 2013) or combined that information with a straight-line strategy (Atack, Bateman, Haines, & Margo, 2010; Banerjee, Duflo, & Qian, 2012; Faber, 2014; Hornung, 2015). Historical transport routes are connected with present ones, due to geographic condition, but are not correlated with present outcome determinants. Thus, they are considered qualified instruments for present transport networks. For example, Baum-Snow et al. (2017) use China's historical transportation network in 1962 as the IV of current highway and railway connections. They argue that transports in 1962 were targeted to deliver agricultural products out of rural areas, which are different from but connected to current transport plans. Similar practices are adopted by Michaels (2008), Zheng and Kahn (2013), and Dong et al. (2018).

The straight-line strategy is justified by the arguments that transports are targeted to connect large central cities, among which straight lines define the shortest distance. Places located either on the straight line or not are considered exogenous; they can be used to construct IVs for actual transport connection variable. For example, Atack et al. (2010) construct an IV for railway accessibility by drawing straight lines among U.S. terminal cities. A similar strategy is also used by Banerjee et al. (2012), Faber (2014), Hornung (2015), and Hodgson (2018).

We provide new evidence on the effect of HSR connection on local economy. We use DID as the basic tool for causal reference and construct IVs using the straight-line strategy to address the issue of endogenous route placement in HSR construction. The straight-line strategy is particularly suitable for our study because HSR is aimed to shorten the travel time among megalopolises such as Beijing and Shanghai. The counties located on those straight lines are considered potential counties connected by HSRs; they are used in our DID framework to construct IVs for the actual HSR connection.

In this paper, we use the county-level panel data, which are more exogenous to the omitted local growth determinants than those at the city or province level (Qin, 2017; Zhang, 2017). In China, HSR network plans are designed at the central government level. According to the latest mid- to long-term railway network plan announced by China National Development and Reform Committee in 2016, HSR network will connect all middle and large cities with a population of over 500,000 by the end of 2025. The first-phase connection targets are undoubtedly those megalopolises such as Beijing, Shanghai, and most provincial capitals. Therefore, county officials have little influence on choosing the location of the HSR system, resulting in a less concern on the issue of endogenous HSR route placement.

This paper focuses on the YRD, which is the largest regional economy and one of the most developed regions in China. Counties in this region have more complete official statistics and better HSR connections; they provide us with enough treatments at the county level and enable us to reliably estimate the effect of HSR connection. Equipped with developed infrastructure networks, economies within the YRD are highly integrated and interdependent on Shanghai, Nanjing, and Hangzhou. Before the HSR construction, counties in this region are more likely to share common development trends between HSR-connected places and places without HSR connection. How and what the HSR system brings to the local economy not only have important implications for evaluating existing HSR projects but also add new insights into HSR construction in the future and in other regions.

Our empirical results show that connecting to HSR produces a long-lasting negative effect on county GDP per capita, and this negative effect increases with the distance from provincial capitals, implying that peripheral regions suffer more from connecting to HSR. The results are robust when "contaminated" counties are omitted, using alternative variables or data and placebo tests. We find that HSR connections impede the economic development of counties through population migration from HSR-connected counties to central cities and spatial restructuring of industries, suggesting that the polarization effects of HSR dominate its dispersion effects. Thus, this paper also contributes to highlighting the channels that transportation improvement brought by HSR impedes local economy.

The rest of the paper is structured as follows: Section 2 discusses the empirical strategy. Section 3 reports the empirical results for the effect of HSR on GDP per capita. Section 4 tests the robustness of our results. Section 5 examines the mechanisms through which HSR connection affects local economy. Section 6 concludes the paper.

## 2 | EMPIRICAL STRATEGY

### 2.1 | Study Area and Data

We use China's YRD as a case to estimate the effect of HSR connection on local economy. The YRD is located in the coastal region alongside the east part of the world's third-longest river, Yangtze River. It includes all counties and cities in Shanghai, Jiangsu, Zhejiang, and Anhui provinces (see also Figure 1). Specifically, the YRD includes 1 central government municipality, 40 prefectural cities, and 305 county-level jurisdictions. As shown in Table 1, the total YRD's GDP in 2016 was 17,722.59 billion RMB (the average exchange rate in 2016 is 6.64 RMBs per U.S. dollar), accounting for 23.84% of China's total GDP. The YRD occupies only 3.65% of national land but accommodates 16.06% of China's population and contributes 35.93% of China's total international trade. Not surprisingly, the YRD has long been a focus of government policies and research interests (Shao et al., 2017; Wang, 2018).

The well-developed infrastructure system, which consists of dense networks of waterways, roadways, and railways, plays an important role in boosting local economy in the YRD. Yet, the existing railway system faces challenges to meet the increasing demand of modern economy and society for fast and comfortable transport tools like HSR. Because the Chinese government began to massively roll out HSR lines in the mid-2000s, the YRD has been one of the first HSR-targeted regions and now has the densest HSR networks in China. By the end of 2016, 14 HSR lines were opened in the YRD

**TABLE 1** Statistics of YRD and its national shares in 2016

| Area | GDP (billion RMB) | Population (10,000 persons) | Prefectures | Counties | Area (10,000 square km) | International trade (billion U.S. dollars) |
|---|---|---|---|---|---|---|
| Shanghai | 2,817.87 | 2,420 | 0 | 16 | 0.63 | 4,337.68 |
| Jiangsu | 7,738.83 | 7,999 | 13 | 96 | 10.26 | 5,092.96 |
| Zhejiang | 4,725.14 | 5,590 | 11 | 89 | 10.2 | 3,365.76 |
| Anhui | 2,440.76 | 6,196 | 16 | 105 | 13.97 | 444.13 |
| YRD | 17,722.59 | 22,205 | 40 | 306 | 35.06 | 13,240.53 |
| China | 74,358.55 | 1,38,271 | 334 | 2851 | 960 | 36,855.57 |
| Share of YRD (%) | 23.84 | 16.06 | 11.98 | 10.73 | 3.65 | 35.93 |

*Note:* The average exchange rate in 2016 is 6.64 RMBs per U.S. dollar.
*Source:* Author's calculations based on http://data.stats.gov.cn/easyquery.htm?cn=C01.

with more than 100 passenger stations (Wang, 2018). Thus, the YRD provides us enough HSR connections at the county level to evaluate their impact on local economy.

Data used in this paper are from three provinces in the YRD. We collect the data from local official statistical yearbooks. Because most of HSR lines started to operate after 2009, we use the statistics data from 2006 to 2015. Shanghai is not included in our main analysis.[1] One consideration is that Shanghai is a national economic center that has a provincial-level city status, whereas all other cities and counties in our data are at lower levels. Meanwhile, all sub-level administrations of Shanghai are more integrated districts that usually have a higher administration level than those in other provinces. For example, Pudong New Area has an administration level just below a province, thereby making it incomparable with county-level administrations in other provinces. Another concern is about the endogeneity issue. Because Shanghai is one of the most important target cities of HSR, its HSR connection is much less exogenous. Including Shanghai in our analysis could weaken the quasi-experimental nature of our estimation. Our definition of HSR follows that of the National Railway Administration of China; that is, HSR refers to passenger transport railway with a designed speed of at least 250 km/hr and an operation speed of at least 200 km/hr at the initial stage. As a result, a total of 13 HSR lines are in our data, among which the first one is the Hefei–Nanjing HSR, opened on April 18, 2008, and the latest one is the Jinhua–Lishui–Wenzhou HSR, opened at the end of 2015. We define a county connected by HSR as one that has at least one HSR station in its jurisdiction, and for those counties connected by multiple HSR lines, we define their connection to HSR by the first year they were connected.[2] Our definition of connection using a dummy variable is similar to previous empirical works (Deng et al., 2019; Donaldson, 2018; Faber, 2014; Hornung, 2015; Shao et al., 2017). We do not use the distance measure of HSR connection because our identification strategy is based on a mutiple-period fixed-effects panel data model in which the effect of a time-invariant variable cannot be estimated. Moreover, HSR service intensity is more endogenous than HSR connection because stations with higher demand tend to have higher service intensity. Therefore, we only use a dummy variable to measure HSR connection that is similar to Hornung's (2015) definition of railway access and Deng et al.'s (2019) HSR variable. Following the HSR map by Li (2016, http://worldmap.harvard.edu/chinamap/), which provides detailed information on China's HSR lines and stations in 2011 and 2016, we manually construct a dummy variable panel data on whether a county is connected to HSR. Our data consist of 3 provincial capitals, 40 municipal districts of prefecture-level cities, and 171 counties and county-level cities.

## 2.2 | Methods

To determine the effect of HSR connection on local economic development, we first use the DID method and then use a simple straight-line strategy to construct IVs for HSR connection. The DID framework, based on panel data, is written as:

$$Ln\left(GDPPC_{it}\right)=\phi HSR_{i,t-1}+X_{it}^{'}\lambda+County_i+\theta_t+\varepsilon_{it} \tag{1}$$

where *GDPPC* denotes GDP per capita that measures the development level of local economy; *HSR* is a dummy variable regarding whether a county is connected by HSR; *X* denotes a vector of control variables; $County_i$ and $\theta_t$ denote county and year fixed effects, respectively; subscripts *i* and *t* denote county and year, respectively; $\varepsilon$ represents random disturbances; $\phi$ and $\lambda$ are parameters to be estimated. Because most HSR lines opened in the second half of the year, we use a 1-year lag for the HSR connection variable. To control heteroscedasticity and make the estimation robust, we cluster the standard errors at the county level.

To make a causality reference using the DID method, the common trends assumption must not be violated. As Qin (2017) and Gao et al. (2019) suggest, we test the common trend assumption with the idea proposed by Autor (2003), which is specified as the following DID model:

$$Ln\left(GDPPC_{it}\right)=\sum_{j=-m}^{q}\phi_j HSR_{i,k=o+j}+\varphi HSR_{i,t-q-1}+X_{it}^{'}\lambda+County_i+\theta_t+\varepsilon_{it} \tag{2}$$

where *o* is the year connecting to HSR, $q > 0$ and $m > 0$. When $j < 0$, $HSR_{i,k=o+j}$ is the *j*th year before HSR connection, and $\phi_j$ measures the *j*th year lead effect of HSR on GDP per capita; when $j > 0$, $HSR_{i,k=o+j}$ is the *j*th year after HSR connection and $\phi_j$ measures the *j*th year lag effect of HSR on GDP per capita; and when $j = 0$, $HSR_{i,k=o+j}$ is the year of connecting to HSR and $\phi_j$ is the effect of HSR on GDP per capita in the connection year. $HSR_{i,t-q-1}$ is $q + 1$ years lag variable of HSR and $\varphi$ is its estimate. If the common trends assumption is fulfilled, the leads will be statistically insignificant, whereas the lags will be significant.

Equation (2) can also be used to estimate the dynamic effect of the treatment variable on the outcome variable (Gao, Zang, Roth, & Wang, 2017). On one hand, HSR construction affects local investment and employment; local residents may also respond in advance of HSR operation. Therefore, HSR could cause lead effects on GDP per capita. On the other hand, HSR takes time to affect the local economy, producing lagged effects on GDP per capita. We estimate Equation (2) to empirically determine the dynamic effect of HSR on local economy.

To further address the endogeneity of HSR route placement, we use the IVs method. Following the literature (Atack et al., 2010; Banerjee et al., 2012; Faber, 2014; Hornung, 2015), we use simple straight-line strategies to construct the potential HSR connection variables as the IVs of actual HSR connection. Compared with highways and traditional railways, HSRs at an early stage aim particularly to reduce the travel time of targeted central cities such as Beijing, Shanghai, and Guangzhou. To best realize this aim, the designer could draw straight lines between targeted cities. Thus, the straight-line strategy produces valid IVs because counties lying on the straight line of two HSR-targeted cities are more likely to be actual HSR connection counties, but whether a county is on the line is somewhat exogenous. Figure 2 shows the straight lines between provincial capitals for all HSRs running within and through our study area.

The second IV is constructed by the straight-line strategy but further considers the HSR construction costs faced by that strategy. To reduce the construction costs of tunnels and bridges, potential

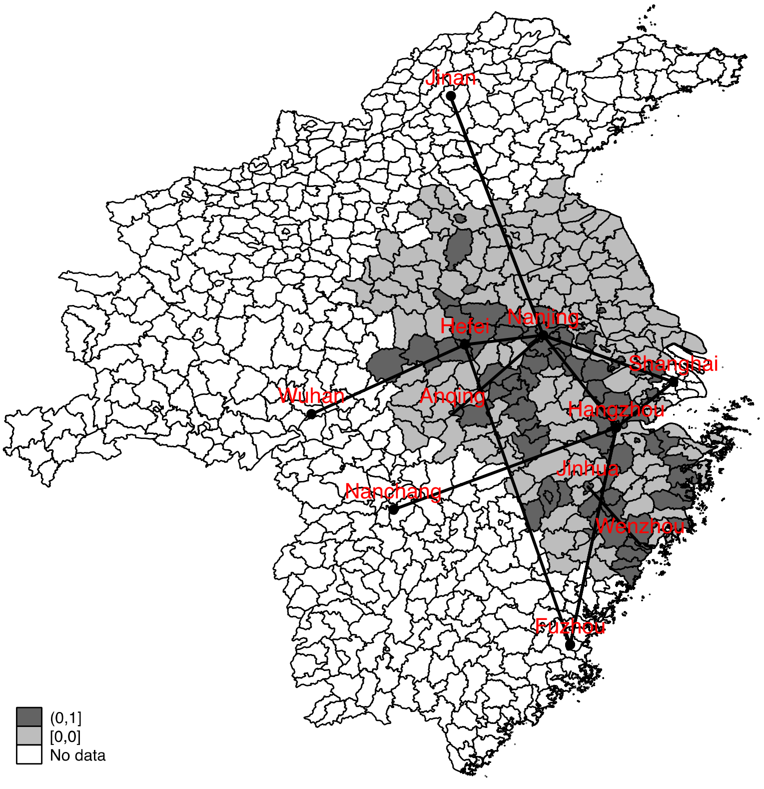


**FIGURE 2** The straight-line strategy for the instrumental variable. The blue lines are used for constructing IV for actual HSR connection, with 1 for counties and districts connected by HSR lines and 0 otherwise. Neighboring provinces of the YRD are marked as no-data.
*Source:* Authors' graph based on the county-level map data (version 2.8) from GADM at: http://gadm.org/download_country.html.

HSR lines in mountainous southwest Zhejiang and south Anhui must bypass large mountains, rivers, and lakes. Thus, we draw a broken line rather than a straight line to construct the IV for HSR lines that cross extreme geographic conditions. The second IV is closer to the actual HSR connections.

In China, central cities often refer to the four central government municipalities and provincial capitals. Because HSR aims to reduce the time of passenger transport among central cities, counties located on the straight lines and close to these cities are more likely to be connected by HSR networks. To some extent, the distance from a county to its own provincial capital is exogenous. We calculate the straight-line spherical distance of the longitude and latitude positions that are obtained from https://jingwei.supfree.net/. Therefore, we construct two additional IVs: the interactions of the straight-line distance of each county from its own provincial capital and previously constructed two IVs for the HSR connection variable.

The following two equations present the two-stage least square (2SLS) method using IVs to estimate the HSR effect on local economy:

$$Ln\left(GDPPC_{it}\right) = \phi HSR_{i,t-1} + X_{it}^{'}\lambda + County_i^2 + \theta_t^2 + \varepsilon_{it} \quad (3)$$

$$HSR_{i,t-1} = PHSR_{i,t-1}^{'}\rho + X_{it}^{'}\gamma + County_i^1 + \theta_t^1 + v_{it} \quad (4)$$

where *PHSR* is a vector of IVs, that is, potential HSR connection variables, $\rho$ and $\gamma$ are parameters to be estimated in the first stage, and $\nu$ is the random disturbances.

## 2.3 | Variables

Following the literature (Banerjee et al., 2012; Faber, 2014; Ke et al., 2017; Qin, 2017), we use GDP per capita to measure the level of a local economy. The HSR connection variable is a dummy variable, which takes 1 for a county in the years it is connected to HSR and 0 otherwise. Figure 3 depicts the per capita GDP trends between the HSR-connected group and the control group without HSR connection. It shows that since the year 2008, when the first HSR line opened in the YRD, an obvious decrease is seen in the per capita GDP of the treatment group, whereas a decrease is not observed in the control group. We also see that the two groups have the same trends before 2008. However, the trends should

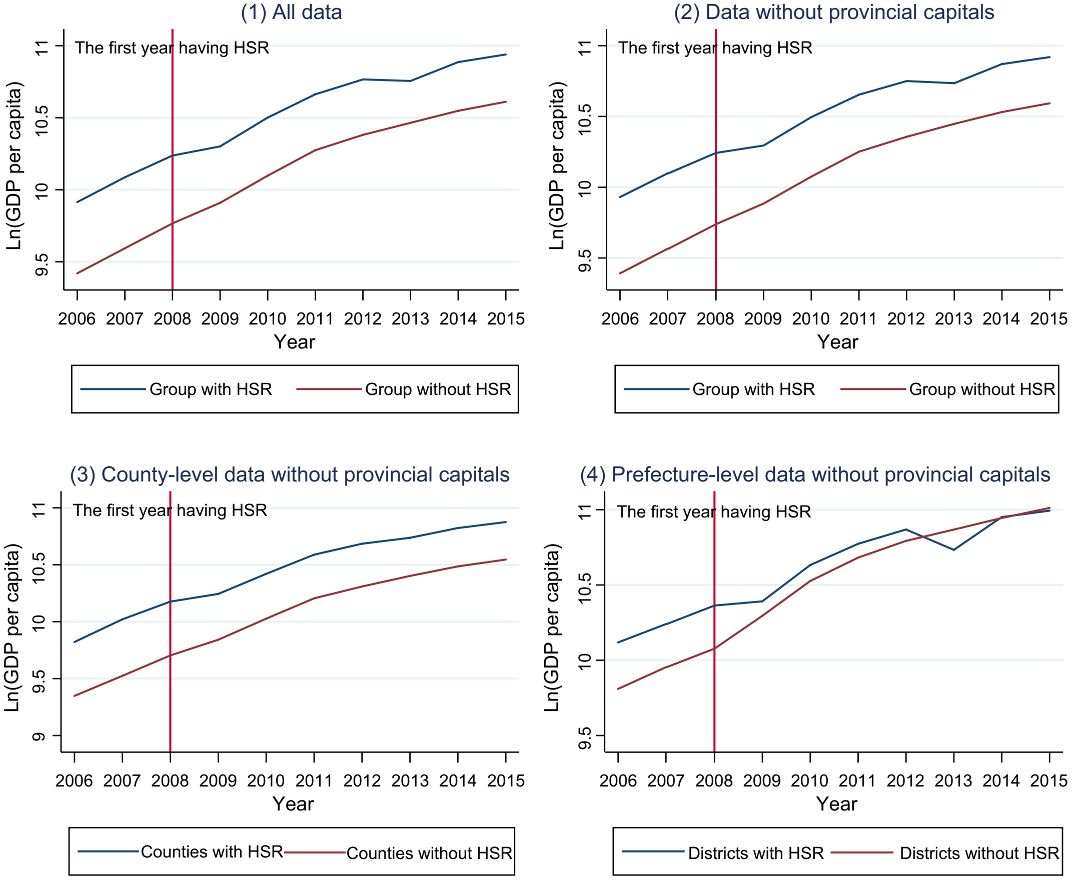


**FIGURE 3** GDP per capita by counties with and without HSR connections. The sharp reduction in GDP per capita of prefectural districts in 2013 is due partly to transforming some counties into urban districts. For example, Jiangyan, a county-level city in Taizhou, Jiangsu Province, was transformed into a municipal district of Taizhou City at the end of 2012; Shaoxing and Shangyu, both of which are county-level administrations in Shaoxing, Zhejiang Province, were also transformed into municipal districts of Shaoxing City in October 2013.
*Sources:* The data on GDP per capita are from official statistics websites of Anhui, Jiangsu, and Zhejiang; the data on HSR connection are manually constructed by the authors.

not be used as strong evidence of the negative effect of HSR connection on economic growth because not all HSR lines were opened in 2008, and policies to combat the world economic crisis with mass investments by the Chinese government are not controlled for.

Control variables are selected based on the related literature (e.g., Faber, 2014; Ke et al., 2017; Qin, 2017) and the data availability at the county level. These variables include investments, public expenditures, economic vitality, consumption, international trade, and resource endowments. Investment is measured by per capita fixed asset investment because a well-recognized discount rate is unavailable at the county level. Public expenditures are those from local budgets, which are also in per capita. Economic vitality is reflected in the size of the private sector. Based on the data available from official statistics yearbooks, we use the share of the number of urban and township nonprivate sector employees to total population to measure this variable. Thus, a larger size of the nonprivate sector, or a smaller private sector, implies a lower level of economic vitality. Consumption is measured by total social retail consumption, which can also measure the market size. International trade is measured by exports per capita. Resource endowments are measured by population density, that is, the number of people per square kilometer. Population density is also related to market size and thus the potential for labor division and specialization. Due to the lack of a county-level Consumer Price Index (CPI) and the fact that our panel data model controls for both year and county-level fixed effects, we use nominal values of all variables. In section 4, we will report the empirical result using provincial CPI to deflate all nominal variables.

Table 2 reports the summary statistics of all variables for all counties, the treatment group, and the control group.[3] We find that, on average, HSR-connected counties and districts have higher levels

**TABLE 2** Variables and summary statistics

| | All counties | | HSR-unconnected counties | | HSR-connected counties | | |
|---|---|---|---|---|---|---|---|
| **Variables** | **Obs.** | **Mean** | **Obs.** | **Mean** | **Obs.** | **Mean** | **Difference in means** |
| Per capita GDP | 2,062 | 10.273 | 1,308 | 10.102 | 754 | 10.57 | –0.468*** |
| Per capita fixed asset investment | 2,068 | 9.601 | 1,307 | 9.407 | 761 | 9.935 | –0.528*** |
| Per capita public expenditure | 2,069 | 8.227 | 1,308 | 8.113 | 761 | 8.421 | –0.308*** |
| Economic vitality | 2,069 | 24.799 | 1,308 | 25.997 | 761 | 22.74 | 3.257*** |
| Population density | 2,069 | 6.199 | 1,308 | 6.132 | 761 | 6.314 | –0.183*** |
| Per capita retail consumption | 2,065 | 9.175 | 1,306 | 8.954 | 759 | 9.556 | –0.602*** |
| Per capita export | 1,573 | 6.635 | 950 | 6.171 | 623 | 7.341 | –1.170*** |
| Potential HSR connection 1 (IV1) | 2,110 | 0.095 | 1,340 | 0.051 | 770 | 0.17 | –0.119*** |
| Potential HSR connection 2 (IV2) | 2,110 | 0.116 | 1,340 | 0.051 | 770 | 0.23 | –0.179*** |
| Distance to provincial capital | 1,880 | 1.591 | 1,170 | 1.783 | 710 | 1.273 | 0.510*** |
| HSR connection | 2,110 | 0.154 | 1,340 | 0 | 770 | 0.422 | –0.422*** |

*Notes:* All variables except for HSR connection variables, economic vitality, and distance to provincial capitals are in natural logarithm; the last column reports the *t* test on the mean difference between counties connected and not connected by HSR; *** denotes the significance level of 1%.

of GDP per capita, fixed asset investment, population density, retail consumption and exports, longer distance from the provincial capitals, but a lower level of economic vitality. Because other things are not necessarily equal, a simple comparison between the treatment and the control groups with respect to GDP per capita does not indicate a real causality from HSR connection to the level of economic development.

# 3 | EMPIRICAL RESULTS

## 3.1 | Results from the DID Method

We report results from the DID method in Table 3. The results show that connecting to HSR has a negative effect on GDP per capita, which decreases from 0.188 to below 0.09 as more variables are controlled for (columns [1]–[4]). A further control of exports leads to a larger effect of HSR connection on GDP per capita but also leads to fewer observations due to a lack of export statistics for many counties in Anhui Province (column [5]). However, once prefecture-level specific trends are controlled for, the effect of HSR on GDP per capita is similar to that without controlling for exports and retail consumption (column [6]). To keep as many observations as possible, we do not control for exports and retail consumption in the following estimations except for IV estimation where we want to see how the estimate would change when all controls are included. Among the control variables, we observe that public expenditures have a positive statistically significant effect on GDP per capita, indicating the government-led economy in China.

Table 4 reports the results based on Equation (2), where HSR connection ($o + j$) refers to the $j$th year before connecting to HSR (when $j < 0$), the $j$th year after connecting to HSR (when $j > 0$), or exactly the year of connecting to HSR (when $j = 0$), and HSR connection ($t - q$) is $q$ years lag of HSR connection.[4] Results in columns (1) and (3) show that in the years before HRS connections there are no significant differences in GDP per capita between the two groups connected and not connected to HSR, verifying that the common trends assumption is fulfilled. Regarding dynamic effects, the results in columns (1) to (3) show no lead effect of HSR connection on GDP per capita. However, columns (4) and (5) show that there are significantly negative lag effects of HSR connection on GDP per capita. We observe that this negative effect does not decrease over time, remaining –0.133 three years later, which indicates that connecting to HSR reduces GDP per capita by 13.3% three years later. Such an increasing trend of the adverse HSR impact is further confirmed by a dynamic analysis that includes both leads and lags of HSR connection in the same model. As shown in the last column of Table 4, at the 5% significance level, HSR produces a persistent negative effect on GDP per capita in years after a county is connected to HSR.

## 3.2 | Results from IVs Methods

Table 5 reports the results from the IVs method. The first-stage estimation (Panel B) shows that IVs as a whole are highly correlated with endogenous HSR connection because of high $F$-values, indicating that they are not weak IVs, although some IVs are not statistically significant. To test whether the exclusion restriction can be fulfilled, that is, potential HSR connection variables affect GDP per capita only through actual HSR connections, we follow the practice adopted by Faber (2014). As argued by Faber (2014), if the exclusion restriction is fulfilled after controlling for time-varying variables, time-fixed effect, and individual-fixed effect, there is not a significant

**TABLE 3** Results from the difference-in-differences method

| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| HSR | –0.188*** | –0.113** | –0.085** | –0.085** | –0.106*** | –0.087* |
| connection ($t$ – 1) | (0.057) | (0.045) | (0.035) | (0.035) | (0.036) | (0.050) |
| Fixed asset | | 0.031** | 0.021 | 0.021 | 0.020 | 0.028** |
| investment | | (0.013) | (0.014) | (0.014) | (0.016) | (0.014) |
| Public | | 0.550*** | 0.511*** | 0.520*** | 0.571*** | 0.378*** |
| expenditures | | (0.078) | (0.063) | (0.074) | (0.079) | (0.059) |
| Economic | | | 0.003** | 0.003** | 0.003*** | 0.001 |
| vitality | | | (0.001) | (0.001) | (0.001) | (0.001) |
| Population | | | –0.299** | –0.299** | –0.262** | –0.138 |
| density | | | (0.133) | (0.134) | (0.129) | (0.091) |
| Retail | | | | –0.020 | –0.122 | 0.033 |
| consumption | | | | (0.095) | (0.105) | (0.141) |
| Exports | | | | | 0.033* | 0.015 |
| | | | | | (0.018) | (0.013) |
| Constant | 9.627*** | 5.372*** | 7.579*** | 7.685*** | 7.962*** | –18.793 |
| | (0.016) | (0.518) | (0.750) | (1.051) | (1.093) | (12.717) |
| Year fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| County fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Prefecture specific trends | No | No | No | No | No | Yes |
| Observations | 2,062 | 2,061 | 2,060 | 2,056 | 1,561 | 1,561 |
| Within $R$-squared | 0.824 | 0.861 | 0.869 | 0.869 | 0.840 | 0.874 |
| Number of counties | 211 | 211 | 211 | 211 | 167 | 167 |

*Notes:* The dependent variable is Ln (GDP per capita); robust standard errors clustered at the county level are reported in parentheses; all results are from the DID method in a fixed effects model; *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively.

relationship between the IVs and time-varying variables as well as the outcome variable in years before the treatments occurred. We also test the effect of the potential HSR connection variable on residuals from column (3) of Table 3. Our test results, which are available upon request, show that the potential HSR connection variables are not significantly correlated with other control variables before connecting to HSR, after controlling for fixed effects. Moreover, Table 3 shows that after controlling for fixed asset investment, public expenditures, economic vitality, and population density, more controls lead to small changes in the effects of HSR connections on GDP per capita. Therefore, the exclusion restriction is not violated when the straight-line strategy is used to construct IVs for actual HSR connections.

The results in columns (1) and (2) of Table 5 are with two IVs, the potential HSR connection variables with and without consideration of construction costs. Columns (3) and (4) present results with

**TABLE 4** Test on common trend assumption and dynamic effects

| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| HSR connection ($o - 3$) | 0.020 | | | | | –0.006 |
| | (0.024) | | | | | (0.014) |
| HSR connection ($o - 2$) | –0.011 | –0.022 | | | | –0.044* |
| | (0.015) | (0.015) | | | | (0.023) |
| HSR connection ($o - 1$) | –0.042 | –0.054 | –0.043 | | | –0.078* |
| | (0.029) | (0.037) | (0.031) | | | (0.046) |
| HSR connection ($o$) | –0.063* | –0.077* | –0.065* | –0.046* | –0.048* | –0.103* |
| | (0.036) | (0.043) | (0.036) | (0.024) | (0.025) | (0.053) |
| HSR connection ($o - 1$) | –0.122** | –0.142** | –0.125** | | | |
| | (0.053) | (0.067) | (0.058) | | | |
| HSR connection ($o + 1$) | | | | –0.079*** | –0.082*** | –0.138** |
| | | | | (0.030) | (0.031) | (0.059) |
| HSR connection ($t - 2$) | | | | –0.109** | | |
| | | | | (0.048) | | |
| HSR connection ($o + 2$) | | | | | –0.064** | –0.126** |
| | | | | | (0.027) | (0.057) |
| HSR connection ($t - 3$) | | | | | –0.133** | –0.220** |
| | | | | | (0.062) | (0.104) |
| Observations | 2,060 | 2,060 | 2,060 | 2,060 | 2,060 | 2060 |
| Within $R$-squared | 0.870 | 0.870 | 0.870 | 0.870 | 0.870 | 0.871 |
| Number of counties | 211 | 211 | 211 | 211 | 211 | 211 |

*Notes:* Other controls are the same to the column (3) of Table 3; $o$ denotes the year of connecting to HSR; for $j = –3, –2, –1, 1$, and 2, HSR connection ($o + j$) refers to the $j$th lead ($j < 0$) or lag ($j > 0$) year of HSR connection; for $q = 1, 2$, and 3, HSR connection ($t–q$) is $q$ years lag term of HSR connection; other notes are the same as those in Table 3; *, **, and ***, denote the significance level of 10%, 5%, and 1%, respectively.

two additional IVs, namely the interactions of the distance to the provincial capital with two potential connection variables. Tests reported at the bottom of Table 5 support the use of these IV variables for the actual HSR connection. Again, we find that HSR connection has a statistically significant and negative effect on GDP per capita. The estimated coefficient, which is about –0.1, indicates that connecting to HSR leads to a reduction in GDP per capita by about 10%. The 2SLS estimation with first two IVs, that is, potential HSR connection variables under the straight-line strategy with or without considering construction costs (see columns [1] and [2]), tends to produce a larger negative effect than the simple DID method as shown in columns (3) and (5) of Table 3. This 1–2 percentage point larger negative effect from the IVs method implies the existence of the endogeneity of HSR routes, but it also indicates that the straight-line rule is largely followed by the designers to shorten the traveling distance between central cities.[5]

**TABLE 5** Results from instrumental variables method

| Variables | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Panel A: Second-stage results | | | | |
| HSR connection ($t$ – 1) | –0.113** | –0.106** | –0.092** | –0.100** |
| | (0.058) | (0.053) | (0.045) | (0.048) |
| Observations | 2,060 | 1,561 | 2,060 | 1,561 |
| Number of counties | 211 | 167 | 211 | 167 |
| Over-identification test | 0.013 | 0.745 | 4.276 | 1.854 |
| Panel B: First-stage results | | | | |
| IV 1 | –0.0692 | –0.170** | 0.133 | –0.092 |
| | (0.073) | (0.066) | (0.128) | (0.107) |
| IV 2 | 0.629*** | 0.734*** | 0.51*** | 0.689*** |
| | (0.073) | (0.061) | (0.13) | (0.096) |
| IV 1 × distance to provincial capital | | | –0.171** | –0.062 |
| | | | (0.083) | (0.072) |
| IV 2 × distance to provincial capital | | | 0.076 | 0.025 |
| | | | (0.067) | (0.044) |
| $F$-value | 15.91*** | 17.43*** | 14.78*** | 16.18*** |

*Notes:* The over-identification test reports Sargan-Hansen statistics; other controls in odd- and even-numbered columns are the same to the column (3) and (5), respectively; other notes are the same as those in Table 3; ** and *** denote the significance level of 5% and 1%, respectively.

## 3.3 | Heterogeneity

One may wonder whether the negative effect of HSR on GDP per capita is heterogeneous between counties and districts of provincial capitals or prefectures. To test these heterogeneities, we report the results with subsamples in Table 6. Because HSR routes aim to shorten the travel time between central cities, we

**TABLE 6** Heterogeneity analysis by subsamples

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | DID | DID-IV | DID | DID-IV | DID | DID-IV |
| Variables | Removing the data from provincial capitals | | Counties | | Districts in prefectures | |
| HSR connection | –0.116*** | –0.127** | –0.086*** | –0.080** | –0.167* | –0.225 |
| ($t$ – 1) | (0.037) | (0.054) | (0.022) | (0.037) | (0.091) | (0.141) |
| Observations | 1,921 | 1,921 | 1,559 | 1,559 | 362 | 362 |
| Within $R$-squared | 0.866 | 0.866 | 0.947 | 0.947 | 0.612 | 0.611 |
| Number of counties | 197 | 197 | 160 | 160 | 37 | 37 |
| Over-identification test | – | 2.254 | – | 2.046 | – | 2.184 |

*Notes:* The IVs used in even-numbered columns are the same to column (3) of Table 5; other controls are the same to those in column (3) of Table 3; all columns have removed observations from provincial capitals; other notes are the same as those in Table 3; *, **, and *** denote the significance level of 10%, 5%, and 1%, respectively.

first omit the data from the districts or counties of provincial capitals and see if the adverse effect of HSR remains the same. The results from DID and IV based on DID (columns [1] and [2], respectively) show that the negative effect is about 3 percentage points larger than that without omitting data from provincial capitals (see column [3] of Table 3 and column [3] of Table 5). A larger estimate when the data are removed from provincial capitals may imply that central cities gain or suffer less from HSR connection. Columns (3) through (6) of Table 6 report the results from counties and districts of prefectural cities without observations from provincial capitals. Again, we find a negative effect of HSR connection on GDP per capita, which are larger at the district level but are more statistically significant at the county level.

We also consider whether the negative effect of HSR on GDP per capita is heterogeneous regarding the distance to central cities. Table 7 reports the heterogeneity with respect to the coordinate distance from counties or districts to two types of central cities: their own provincial capital and the nearest other provincial capital. The results show that the farther a county or a prefecture district is from its own provincial capital, the larger the negative effect is from connecting to HSR. As shown in column (1), with an average distance of 159 km in our data, we can calculate that the aggregate average effect of HSR connection on GDP per capita is about –0.11, a size similar to those in Table 3. However, we do not observe a similar heterogeneity regarding the distance to the nearest other provincial capital, as shown in column (2) of Table 7, consistent with the fact that local economies are more tied to their own' political centers in China.

## 4 | ROBUSTNESS CHECKS

In this section, we provide several robustness checks by omitting some counties that could "contaminate" DID estimation, using alternative variables, including Shanghai, and conducting placebo tests.

### 4.1 | Omitting "Contaminated" Counties

This research uses data from 2006 to 2015. One concern is that constructions of some HSR routes were started during our study period, but the HSR routes were opened in 2016 or later. Such construction activities certainly could promote local economic growth before the HSR route's operation. As a result, counties that have experienced such construction activities, included in the control group in

**TABLE 7** The role of distance to central cities

| Variables | (1) | (2) |
|---|---|---|
| HSR connection ($t$ – 1) | 0.033 | –0.074 |
| | (0.039) | (0.073) |
| HSR connection ($t$ – 1) × the distance to the provincial capital | –0.092** | |
| | (0.045) | |
| HSR connection ($t$ – 1) × the distance to the nearest other provincial capital | | –0.004 |
| | | (0.039) |
| Observations | 1,655 | 1,655 |
| Within $R$-squared | 0.836 | 0.834 |
| Number of counties | 188 | 188 |

*Notes:* Same notes as those in Table 4.

our DID framework, are "contaminated" and omitted in our robustness check. Similarly, we also omit those counties that have experienced other kinds of railway construction. In total, we omit 14 counties and districts to re-estimate the effect of HSR connection on GDP per capita. Columns (1) and (2) of Table 8 show the results. By comparing with those in column (3) of Table 3 and column (3) of Table 5, respectively, we observe that omitting such potentially "contaminated" counties in the control group change little the results whether the IVs method is used.

We further omit "contaminated" data in the treatment group, considering that counties connecting to HSR in 2014 or 2015 may either have abnormal economic growth or the time is too short to observe the effect of HSR connection on local economy. The results are reported in columns (3) and (4) of Table 8. We find that the negative effect of HSR connection remains, although it is smaller than those observed in columns (1) and (2).

## 4.2 | Using Alternative Variables and Data

Variables used in previous estimations are in nominal values. One reason is because we do not have county-level deflators, and the provinces in the study area do not differ much in this regard. Another reason is that we control for year fixed effects in our DID framework, which helps to reduce the impact of inflation. If inflation across counties in YRD has a similar movement, which is likely because YRD is one of the more integrated regions, it can be largely addressed by year fixed effects. To check if using real values would change our results, we use 2005 as the base year and deflate the nominal variables with each province's CPI. The results are reported in columns (1) and (2) of Table 9. We find that the estimate of HSR connection is almost unchanged (all are the same up to the third decimal point, compared to column [3] of Table 3 and column [3] of Table 5, respectively). Table 9 also confirms the role of year fixed effects in reducing the impact of inflation. Therefore, in our following analysis and discussions, we continue to use the nominal values.

In this paper, we use GDP per capita to measure the level of economic development and examine HSR impact on it. We understand that both level and rate variables have been used in the empirical growth literature (Acemoglu, Naidu, Restrepo, & Robinson, 2019; Banerjee et al., 2012; Faber, 2014; Ke et al., 2017; Qin, 2017). The level variable measures the changes in economic development level,

**TABLE 8** Results after omitting "contaminated" counties

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Removing counties under HSR construction | | Further removing counties connecting to HSR in 2014 and 2015 | |
| Variables | DID | DID-IV | DID | DID-IV |
| HSR connection | –0.087** | –0.090** | –0.073** | –0.069* |
| ($t$ – 1) | (0.035) | (0.044) | (0.036) | (0.042) |
| Observations | 1,928 | 1,928 | 1,680 | 1,680 |
| Within $R$-squared | 0.864 | 0.864 | 0.853 | 0.853 |
| Number of counties | 197 | 197 | 172 | 172 |
| Over-identification Test | | 4.442 | | 5.472 |

*Notes:* The IVs and controls are the same to column (3) of Table 5; other notes are the same as those in Table 5; * and ** denote the significance level of 10% and 5%, respectively.

**TABLE 9** Robustness checks using alternative variables or including Shanghai

| Variables | (1) Using real values in 2005 price level DID | (2) Using real values in 2005 price level DID-IV | (3) Using growth of GDP per capita DID | (4) Using growth of GDP per capita DID-IV | (5) Including Shanghai DID | (6) Including Shanghai DID-IV |
|---|---|---|---|---|---|---|
| HSR connection (*t* – 1) | –0.085** | –0.092** | | | –0.086** | –0.092** |
| | (0.035) | (0.045) | | | (0.035) | (0.045) |
| HSR connection | | | –0.015 | –0.021 | | |
| | | | (0.010) | (0.017) | | |
| Observations | 2,060 | 2,060 | 1,847 | 1,651 | 2,070 | 2,060 |
| Within *R*-squared | 0.809 | | 0.963 | | 0.869 | |
| Number of counties | 211 | 211 | 211 | 188 | 212 | 211 |
| Over-identification test | | 4.242 | | 4.028 | | 4.276 |

*Notes:* The IVs used are the same to column (3) of Table 5; controls in all columns are the same to those in column (3) of Table 3, except for columns (3) and (4) where 1-year lag of Ln (GDP per capita) is also controlled; other notes are the same as those in Table 3; ** denotes the significance level of 5%.

whereas the rate variable measures the changes in growth rate of GDP per capita. Here, we conduct a robustness check where the dependent variable is the growth rate of GDP per capita. The results are reported in columns (3) and (4) of Table 9, which are statistically insignificant. We would not interpret that these results are inconsistent with our baseline results where the level variable is used. Because the purpose of our paper is to estimate the impact of HSR connection on local economy by using a DID framework, we think it is more appropriate to use the level variable (really, the difference of the natural logarithm of GDP, i.e., the percentage changes of GDP per capita) than the growth rate variable.

In Tables 3–8, we exclude Shanghai from our dataset for reasons mentioned earlier in section 2. However, to check if Shanghai would change our results, we conduct a robustness check that includes the data from Shanghai. The results are reported in columns (5) and (6) of Table 9. Comparing them with those in column (3) of Table 3 and column (3) of Table 5, we find no qualitative change in our results.

## 4.3 | Performing Placebo Test

We finally perform some placebo tests by constructing fake HSR connection variables. In detail, we move the time connecting to HSR forward by 4–7 years. Because such HSR connections do not exist in reality, we expect the estimates to be statistically insignificant. It is indeed the case as we can see from Table 10 that, by confining the observations without HSR connection, none of the fake HSR connection variables have a significant effect on GDP per capita.

The preceding robustness checks all confirm our finding that HSR connection undermines local economy in terms of GDP per capita.[6] Our result is consistent with some studies with data from other traditional transports such as truck highway (Faber, 2014), railroads (Hodgson, 2018), and HSR (Banister and Berechman, 2001; Qin, 2017), providing new evidence that transport improvements lead to polarization of regional economy, that is, an agglomeration shadow termed by Hodgson (2018).

**TABLE 10** Placebo test

| Variables | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| HSR connection ($t + 7$) | 0.031 | | | |
| | (0.021) | | | |
| HSR connection ($t + 6$) | | –0.022 | | |
| | | (0.018) | | |
| HSR connection ($t + 5$) | | | –0.007 | |
| | | | (0.015) | |
| HSR connection ($t + 4$) | | | | 0.012 |
| | | | | (0.016) |
| Observations | 537 | 706 | 864 | 1,015 |
| Within $R$-squared | 0.928 | 0.866 | 0.892 | 0.915 |
| Number of counties | 190 | 190 | 190 | 190 |

*Notes:* Same notes as those in Table 4.

## 5 | MECHANISMS

This section examines two mechanisms through which HSR connection impedes local economy: migration of (rural or peripheral) population and restructuring of industries. China has a long history of a dual socioeconomic structure with huge urban–rural development gaps (Gao, Zheng, & Bu, 2014). A sharp reduction in costs accessing central urban cities, due to HSR, might lead to a fast flow of information and knowledge diffusion, and an increase in the outflow of labor toward core areas, given that the latter has taken on most advanced resources such as education, economy, and culture. Because outflowing labor usually has higher levels of human capital than those left behind, HSR connection could reduce GDP per capita in rural and peripheral areas.

Industries are also restructured along with the migration of population and information brought by HSR connection. For example, Shao et al. (2017) find with data from China's YRD region that HSR promotes agglomeration of the producer service industry, particularly with medium- and small-size cities located on the rail lines. Using China's national trunk highway data, Faber (2014) also finds that highway network connection reduces peripheral counties' GDP growth, which is driven by the reduction in industrial output growth. Qin (2017) examines China's rail upgrades and finds that the reduction in the transport cost of people due to increases in railway speed diverts economic activities from peripheral counties to central cities. Therefore, industrial restructuring may also channel the economy-impediment effect of HSR.

Panel A in Table 11 reports the effect of HSR connection on population growth, that is, the first-order difference of the natural logarithm of population. We observe significant negative effects of HSR connection on population growth, which is larger at the county level than that at the prefectural municipal district level. Specifically, HSR connection decreases population growth in connected counties by 6.3% and in prefectural districts by 5.6%. A slightly larger effect of HSR connection from all data than from the data excluding provincial capitals also indicates that the population polarization effect of HSR connection outweighs its population diffusion effect. This channel is consistent with empirical studies for China's municipal districts (Deng et al., 2019) and the Shinkansen in Japan (Sasaki et al., 1997).

**TABLE 11** Mechanisms by which HSR connection affects GDP per capita

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | **All data** | **Removing the data from provincial capitals** | **Data from counties** | **Data from districts of prefectures** |
| Panel A: | Population growth | | | |
| HSR connection ($t$ – 1) | –0.060*** | –0.061** | –0.063* | –0.056* |
| | (0.023) | (0.024) | (0.036) | (0.030) |
| Observations | 1,856 | 1,731 | 1,399 | 332 |
| Number of counties | 211 | 197 | 160 | 37 |
| Over-identification test | 3.365 | 3.283 | 1.798 | 3.283 |
| Panel B: | Share of the secondary industry | | | |
| HSR connection ($t$ – 1) | –1.836 | –2.608** | –2.759** | –1.495 |
| | (1.118) | (1.069) | (1.315) | (1.544) |
| Observations | 2,067 | 1,928 | 1,559 | 369 |
| Number of counties | 211 | 197 | 160 | 37 |
| Over-identification test | 4.222 | 4.352 | 3.525 | 2.345 |
| Panel C: | Urbanization | | | |
| HSR connection ($t$ – 1) | –0.518 | –0.502 | –0.968 | 0.275 |
| | (1.162) | (1.361) | (1.697) | (1.391) |
| Observations | 1,504 | 1,385 | 1,037 | 348 |
| Number of counties | 158 | 146 | 110 | 36 |
| Over-identification test | 1.271 | 0.860 | 1.555 | 1.854 |

*Notes:* Other controls include fixed asset investment and public expenditures in Panel A and Panel C, and also economic vitality and population density in Panel B; the data in columns (3) and (4) do not include observations from provincial capitals; other notes are the same as those in Table 3; *, **, and *** denote the significance level of 10%, 5%, and 1%, respectively.

Panel B presents the results of HSR connection effect on the industrial restructure. We find that HSR connection reduces the share of the secondary industry at the county level by about 2.8%. This finding is consistent with the results in the literature (Faber, 2014; Percoco, 2016; Shao et al., 2017), suggesting that HSR connection indeed restructures industries by decreasing the secondary industry and increasing other industries. We also estimate the effect of HSR connection on the share of other two industries, agricultural and service industries, and find that it increases the share of agricultural industry at the prefectural municipal district level and the share of service industry at the country level (details are available upon request). Peripheral areas, after having access to large central cities through HSR, can develop their advantages in primary industries as well as transport-related or producer services (Percoco, 2016; Shao et al., 2017). Because China's GDP has long been dominated by the secondary industry, such a restructuring process contributes to a reduction in GDP per capita.

In Panel C, we test if HSR connection leads to economic dispersion through local effects, as argued by urban economists (Alonso, 1964; Baum-Snow, 2007; Muth, 1969; Zheng & Kahn, 2013), whereby improvements in transport accessibility contribute to urbanization and house prices in peripheral areas. The results show that connecting to HSR does not significantly increase urbanization at all levels. Our results are consistent with that of Faber (2014) in showing that local effects in the form of fast urbanization are not observed as a result of HSR connection.

## 6 | CONCLUSIONS

This paper has used the county-level panel data from China's YRD to determine the effect of HSR connection on local economy. The DID and IVs methods both show that connecting to HSR results in a long-lasting reduction in GDP per capita by about 10%. We further find that peripheral areas suffer a larger reduction in GDP per capita from HSR connection. Our results provide new evidence that supports the new economic geography, which predicts polarization effects of the reduction in transport cost. This paper has also confirmed two mechanisms through which HSR connection impedes local economy. We find that HSR connection facilitates migration from peripheral areas to core cities and decreases the share of the secondary industry in HSR-connected counties, both impeding local economy. Therefore, this paper has provided additional evidence on how fast transports reshape the spatial economy, which confirms that polarization effects outweigh the dispersion effects.

Several caveats are worth mentioning. First, HSRs are targeted to shorten the travel time among large cities. Many HSR trains do not stop at county stations. In this paper, we do not analyze the impact of the frequency of HSR trains stopped in county HSR stations. Second, HSR is more expensive than other passenger trains. Therefore, it is relatively more attractive for skilled workers than to unskilled workers to migrate from rural counties to central cities. We do not examine the effect of HSR on sub-populations based on their skill levels. Third, HSR improves access to central cities and thus possibly depresses local consumption. It, for example, enables county residents to enjoy advanced medical cares at modern hospitals in central cities, lowering the demand for local medical services. Fourth, HSR replaces other kinds of passenger transports such as buses, which are run by local enterprises that create jobs and generate income for the local economy. Such substitution effects could be substantial because HSR lines are operated by agencies of National Railway Administration in central cities. Fifth, although our results suggest that central cities such as provincial capitals in the YRD region could benefit more from HSR connection or suffer less from it, whereas peripheral areas suffer from it, we are unable to conclude that HSR makes the central cities stronger because our data are from the YRD region that includes only one megacity (Shanghai) and three central cities (Hangzhou, Hefei, and Nanjing). To have a valid reference about HSR's role in promoting central city's economy, we will need a much larger dataset that consists of more central cities and counties connected and not connected by HSR. Finally, due to the lack of availability of long-term data, this paper does not provide evidence on long-term effects of HSR connection on GDP per capita. Moreover, more specific micro-level data are also needed to evaluate how HSR connection sorts population as well as industries between central and peripheral areas. The aforementioned caveats are topics for future studies. Nevertheless, the main findings of this study provide an alert that local officials at peripheral areas should not count on connecting to HSR to boost local economy, at least in the short run.


### CONFLICT OF INTEREST STATEMENT

None of the authors has any conflict of interest.

### ACKNOWLEDGMENT

The authors thank for the helpful comments from Shuai Shao, Jian Lu, Ruixiang Xiong, Yangfei Xu, Shilin Zheng, and discussants at the first Camphor Economics Circle Seminars at Xiangtan University (2017) and Harbin Engineering University (2018), the 2018 CES China Conference (Hefei), the 2018 CESA Annual Conference (Sydney), the research assistance from Zeqian Lin, Kaini Wang, Ziqi Xu, and Qingrong Zhi, and the fund supports from the Major Project of China National Social Science Foundation (Grant No. 18ZDA077), the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Grant No. 17YJC790040) and the Fundamental Research Funds for the Central Universities (Grant No. 2242018S20023).

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID

*Yanyan Gao* https://orcid.org/0000-0003-0326-504X
*Shunfeng Song* https://orcid.org/0000-0002-3624-3093

## ENDNOTES

[1]To use all data from the YRD and not neglect the fact that Shanghai is one dispensable economic and political center of the YRD, we conduct a robustness check including the data from Shanghai.

[2]One exception is Fengyang, a county in Anhui Province, which is not connected to HSR but is treated as connecting to the Beijing–Shanghai HSR in 2011 because the Bengbu South HSR station is just about 10 km away from the center of Fengyang and about 5 km from its jurisdiction. However, empirical results are not sensitive to such treatment.

[3]We also calculated the correlation matrix of these variables and test the multicollinearity with variance inflation factor (VIF). The results suggest that multicollinearity is not a severe issue, because the VIFs of all explanatory variables are less than 10, a threshold value used by empirical literature to tell the existence of multicollinearity. We observed that retail consumption and export are moderately correlated with fixed asset investment and public expenditure. In our analysis (Table 4 through Table 11), we excluded these variables because controlling for them changes our empirical results little but leads to about 500 missing observations. Details on correlation coefficients and VIFs are available upon request.

[4]To see the difference between HSR connection ($o + j$) and HSR connection ($t$–1), let us consider the case that $j$ equals to –1 and $q$ equals to 1, that is, HSR connection ($o$–1) and HSR connection ($t$–1). The former defines the first year before connecting to HSR. It takes a value of 1 for counties having HSR connection in the first year before that connection and 0 otherwise. The latter is a lag variable of HSR connection that moves the original HSR connection variable backward by 1 year. Moreover, HSR connection ($o$) refers to the year connecting to HSR that takes a value of 1 in that year and 0 otherwise; HSR connection ($t$) takes a value of 1 in all years connecting to HSR and 0 in years before that connection.

[5]The Hausman tests conducted between simple DID model and IVs method also support this argument by showing that the null hypothesis of non-systematic difference in coefficients cannot be rejected.

[6]We also conducted additional robustness checks with alternative measures of HSR connection. Results are consistent with previous findings. Due to word limit, results are not reported in the paper but available upon request.

## REFERENCES

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, *127*(1), 47–100.

Ahlfeldt, G. M., & Feddersen, A. (2018). From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, *18*(2), 355–390. Retrieved from https://doi.org/10.1093/jeg/lbx005

Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Cambridge, MA: Harvard University Press.

Angrist, J. D., & Pischke, J.-S. (2014). *Mastering 'metrics: The path from cause to effect*. New Jersey: Princeton University Press.

Atack, J., Bateman, F., Haines, M., & Margo, R. A. (2010). Did railroads induce or follow economic growth? *Social Science History*, *34*(2), 171–197.

Autor, D. H. (2003). Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics*, *21*(1), 1–42.

Banerjee, A., Duflo, E., & Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China (MIT Department of Economics Working Paper No. 12–06). MIT. Retrieved from https://ssrn.com/abstract=2018637

Banister, D., & Berechman, Y. (2001). Transport investment and the promotion of economic growth. *Journal of Transport Geography*, *9*(3), 209–218.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, *122*(2), 775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads and decentralization of Chinese cities. *Review of Economics and Statistics*, *99*(3), 435–448.

Chen, C.-L., & Hall, P. (2011). The impacts of high-speed trains on British economic geography: a study of the UK's InterCity 125/225 and its effects. *Journal of Transport Geography*, *19*(4), 689–704.

Deng, T., Wang, D., Yang, Y., & Yang, H. (2019). Shrinking cities in growing China: Did high speed rail further aggravate urban shrinkage? *Cities*, *86*, 210–219. https://doi.org/10.1016/j.cities.2018.09.017

Diao, M., Zhu, Y., & Zhu, J. (2016). Intra-city access to inter-city transport nodes: The implications of high-speed-rail station locations for the urban development of Chinese cities. *Urban Studies*, *54*(10), 2249–2267. https://doi.org/10.1177/0042098016646686

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, *108*(4–5), 899–934.

Dong, X., Zheng, S., & Kahn, M. E. (2018). The Role of Transportation Speed in Facilitating High Skilled Teamwork (NBER Working Paper No. 24539). Retrieved from http://www.nber.org/papers/w24539

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System. *Review of Economic Studies*, *81*(3), 1046–1070.

Fogel, R. W. (1962). A quantitative approach to the study of railroads in American economic growth: a report of some preliminary findings. *The Journal of Economic History*, *22*(2), 163–197.

Fujita, M., Krugman, P. R., & Venables, A. J. (2001). *The spatial economy: cities, regions and international trade*. Cambridge, MA: MIT Press.

Gao, Y., Su, W., & Wang, K. (2019). Does High-Speed Rail Boost Tourism Growth? New Evidence from China. *Tourism Management*, *72*, 220–231.

Gao, Y., Zang, L., Roth, A., & Wang, P. (2017). Does democracy cause innovation? An empirical test of the popper hypothesis. *Research Policy*, *46*(7), 1272–1283.

Gao, Y., Zheng, J., & Bu, M. (2014). Rural-Urban Income Gap and Agricultural Growth in China——An Empirical Study on Provincial Panel Data, 1978–2010. *China Agricultural Economic Review*, *6*(1), 92–107. Retrieved from https://doi.org/10.1108/CAER-02-2012-0016

Helpman, E., & Krugman, P. R. (1985). *Market structure and foreign trade: Increasing returns, imperfect competition, and the international economy*. Cambridge, MA: MIT press.

Hodgson, C. (2018). The effect of transport infrastructure on the location of economic activity: Railroads and post offices in the American West. *Journal of Urban Economics*, *104*, 59–76. https://doi.org/10.1016/j.jue.2018.01.005

Hornung, E. (2015). Railroads and growth in Prussia. *Journal of the European Economic Association*, *13*(4), 699–736.

Jin, C., Xu, J., Lu, Y., & Huang, Z. (2013). The impact of Chinese Shanghai-Nanjing high-speed rail on regional accessibility. *Geografisk Tidsskrift-Danish Journal of Geography*, *113*(2), 133–145. https://doi.org/10.1080/00167223.2013.836787

Ke, X., Chen, H., Hong, Y., & Hsiao, C. (2017). Do China's high-speed-rail projects promote local economy?—New evidence from a panel data approach. *China Economic Review*, *44*, 203–226. https://doi.org/10.1016/j.chieco.2017.02.008

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), 483–499. https://doi.org/10.1086/261763

Li, X., Huang, B., Li, R., & Zhang, Y. (2016). Exploring the impact of high speed railways on the spatial redistribution of economic activities-Yangtze River Delta urban agglomeration as a case study. *Journal of Transport Geography*, *57*, 194–206.

Li, Y. (2016). China high speed railways and stations (2016). Harvard Dataverse. https://doi.org/10.7910/DVN/JIISNB

Lin, Y. (2017). Travel costs and urban specialization patterns: Evidence from China's high speed railway system. *Journal of Urban Economics*, *98*, 98–123. https://doi.org/10.1016/j.jue.2016.11.002

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, *90*(4), 683–701.

Muth, R. (1969). *Cities and housing: The spatial patterns of urban residential land use*. Chicago: University of Chicago Press.

Percoco, M. (2016). Highways, local economic structure and urban development. *Journal of Economic Geography*, *16*(5), 1035–1054. Retrieved from https://doi.org/10.1093/jeg/lbv031

Qin, Y. (2017). 'No county left behind?' The distributional impact of high-speed rail upgrades in China. *Journal of Economic Geography*, *17*(3), 489–520. Retrieved from https://doi.org/10.1093/jeg/lbw013

Sasaki, K., Ohashi, T., & Ando, A. (1997). High-speed rail transit impact on regional systems: does the Shinkansen contribute to dispersion? *The Annals of Regional Science*, *31*(1), 77–98.

Shao, S., Tian, Z., & Yang, L. (2017). High speed rail and urban service industry agglomeration: Evidence from China's Yangtze River Delta region. *Journal of Transport Geography*, *64*, 174–183. https://doi.org/10.1016/j.jtrangeo.2017.08.019

Vickerman, R. (2015). High-speed rail and regional development: the case of intermediate stations. *Journal of Transport Geography*, *42*(Supplement C), 157–165. https://doi.org/10.1016/j.jtrangeo.2014.06.008

Wang, L. (2018). High-speed rail services development and regional accessibility restructuring in megaregions: A case of the Yangtze River Delta, China. *Transport Policy*, *72*, 34–44. https://doi.org/10.1016/j.tranpol.2018.09.015

Wang, L., & Duan, X. (2018). High-speed rail network development and winner and loser cities in megaregions: The case study of Yangtze River Delta, China. *Cities*, *83*, 71–82. https://doi.org/10.1016/j.cities.2018.06.010

Zhang, J. (2017). High-speed Rail Construction and County Economic Development: The Research of Satellite Light Data. *China Economic Quarterly*, (4), 1533–1562.

Zheng, S., & Kahn, M. E. (2013). China's bullet trains facilitate market integration and mitigate the cost of megacity growth. *Proceedings of the National Academy of Sciences*, *110*(14), E1248–E1253.

**How to cite this article:** Gao Y, Song S, Sun J, Zang L. Does high-speed rail connection really promote local economy? Evidence from China's Yangtze River Delta. *Rev Dev Econ*. 2019;00:1–23. https://doi.org/10.1111/rode.12642