How Does High-Speed Rail Impact the Industry Structure? Evidence from China

Ming-yu Hu1,2 · Jing Xu1

Received: 23 March 2022 / Revised: 21 June 2022 / Accepted: 16 August 2022 / Published online: 20 October 2022
© The Author(s) 2022

Abstract The economic implications of high-speed rail (HSR) cannot be overlooked in China. This paper studies the impact of HSR on the advancement and rationalization of industrial structure and the tertiary industry aggregation through theoretical derivation and multi-period difference-in-differences (DID) by improving the theoretical framework and empirical methods according to the characteristics of China’s HSR and economic development. From analyses on urban heterogeneity and inter-industry spillover effects, the transmission mechanism and expressions of the industrial structure are also discussed. The findings show that HSR promotes tertiary industry aggregation and contributes to the transformation of the industrial structure from the primary to secondary and tertiary industry sectors, as well as realizing industrial structure advancement but irrationalization. Furthermore, HSR has a more significant influence on tertiary industry aggregation in large cities and high-density cities. Additionally, the aggregation of the transportation, warehousing, and postal sectors has been reduced, with a significant spillover effect on neighboring cities, proving the siphon effect and conduction mechanism, resulting in a structural shift in the tertiary industry, from basic to advanced sectors. The movement of human resources is a key mediator in the economic impact of HSR.

Keywords High-speed rail · Industrial aggregation · Industrial structure advancement · Industrial structure rationalization · Siphon effect · Spillover effect

JEL Classification L16 · R40 · R41 · R48

1 Introduction

With a total length of 37,900 kilometers in 2020, high-speed railways (HSR) can be found in more than 80% of China’s cities. Since the opening of high-speed rail in China in 2008, the economy of the regions along the line has developed rapidly. Represented by Beijing-Guangzhou and Beijing-shanghai, as shown in Fig. 1 and Fig. 2, provinces along the Beijing-Shanghai and Beijing-Guangzhou HSR routes overlapped with provinces with high GDP in 2021, with the tertiary industry accounting for more than half of GDP. Spiekerman (1994) found that the space-time distance is compressed after the opening of HSR [1]. However, the HSR’s impact on the economy is not always positive. This might result in a siphon effect, suffocating the development of small and medium-sized cities [2]. The Beijing–Shanghai HSR, for example, establishes a 1-hour shuttle between Beijing, Tianjin, and Hebei, while simultaneously enabling Beijing to absorb Tianjin and Hebei’s vast resources. Without a doubt, HSR’s economic effect is significant, but how does HSR transform the industrial structure? What sectors will be affected as a result of this? Is HSR going to have a siphon effect? Theoretical and empirical studies are needed to better investigate these concerns.

With the advancement of modern transportation networks, more papers on the effect of transportation infrastructure on economic growth [3–5]. Since China’s HSR
launched late, related study on China has gradually gained significance until recent years. The summary of the literature on HSR’s impact on economic growth is shown in Appendix A. In reality, China’s HSR development and policy differ from those of other countries, and there is a knowledge gap that applies to China’s HSR. Few analyses of China’s HSR system take into consideration industry heterogeneity, which is the focus of research on industrial structure issues. Chandra and Thompson [6] found that certain industries profit from US interstates because of lower transportation costs, while others have migrated as a result of economic activity [6]. Holl [7] studied the effects of road investment and the aggregation effect on enterprises in 13 industrial and 9 service sectors in Portugal, and found significant differences across industries [7]. Western research material and methodologies about economic effects of HSR are numerous since it was the first to realize the industrial revolution [8–10]. Chen and Silva [11],

Fig. 1 GDP of provinces in China in 2021

Fig. 2 Proportion of tertiary industry of provinces in China in 2021
Ahlfeldt and Feddersen [12], Chen and Haynes [13], Guirao et al. [14], Guirao et al. [15], Cascetta et al. [16] studied the impact of HSR’s opening in Spain, Germany, China, France and Italy, and the positive impact is acknowledged. The recent papers on the economic consequences of HSR concentrate on individual industries and production factors including regional labor mobility, economic development, and urbanization [17–20]. Zheng et al. [21] and Pan et al. [22] measured the spatial spillover effect of HSR stations in China, and found that HSR spur the agglomeration of various economic activities from being near and far from the station. About empirical methods, DID (Differences-in-Differences) model is used widely. Wang et al. [23] used anti-gravity and gravity bias models to study how the urban structure of the Yangtze River Delta evolved after the opening of the HSR. Liu et al. [24] used a time-varying DID model to study the spatial integration and industrial development of the Yangtze River Economic Belt, and explained the relationship between the spatial integration of urban aggregations and the characteristics of industrial development after the opening of the HSR. Liang et al. [25] and Wang et al. [28] used the PSM-DID model to analyze the economic development along the HSR line based on the panel data of prefecture-level cities. Li et al. [26] used synthetic control methods (SCM) to investigate the economic effect of HSR, and the results show that the economic effect of HSR had strong disparity [26]. Li et al. [27] used both the DID model and threshold regression and found that the opening of HSR had a significant threshold effect on improving the efficiency of the service sector [27]. Meng et al. [29] constructs an HSR operation network (HSRON) model to study the impact of network position (NP) on service industry agglomeration (SIA) by employing complex network analysis and panel regression methods [29]. Melo et al. [30] discovered that estimates of transportation infrastructure’s productivity impact differed by major industry groups, with estimates for the US economy averaging higher than estimates for European countries, and estimates for highways averaging higher than estimates for other modes of transportation [30]. In the empirical analysis on HSR in China, there have been few papers focus on the quantification of transportation costs. It becomes clear that a well-established theoretical framework for analyzing the impact of HSR in China is desperately required. This paper is devoted to a more in-depth assessment of conditions based on previous research.

The contribution of this paper is to improve the theoretical and empirical framework regarding the economic impact of HSR in China on the basis of previous papers. On the one hand, the mathematical economic geography theory of the impact of the opening of the HSR on the industrial structure was proposed, which comprehensively complemented the hypotheses on the economic effects of HSR. The empirical analysis, on the other hand, was in line with the reality in China. In contrast to other studies, we chose appropriate control and treatment groups to avoid cross-influence of lines within the area. In addition to verifying the hypotheses, the mechanism and heterogeneity of HSR on the industrial structure were explored, revealing the manifestations of economic transformation in China.

The remainder of this paper is structured as follows. Section 2 presents the theoretical framework. Section 3 describes the data set. Our empirical results are presented in Sect. 4. Our main conclusions are then summarized in Sect. 5.

2 Theoretical and Analytical Framework

The enhancement of accessibility is a direct result of the opening of HSR [10, 31]. In addition to decreasing distances in time and space, the HSR indirectly increases the flow of factors such as labor, social capital, and technology, thereby reshaping the market structure [1, 5]. Moreover, the opening of HSR facilitates the establishment of economies of scale, which further alters market competition patterns, as seen by the clustering of urban tertiary sectors and the expansion of cities not on the route [8, 32]. The rationality of the industrial structure will be influenced by the advancement of the industrial structure [14]. As the cycle progresses, the advantageous industries will cluster, and the industrial structure will shift as well, both of which are manifestations of indirect impacts. The impact of the introduction of HSR on the industrial structure is seen in Fig. 3.

The theories and hypotheses proposed in this paper for the impact of HSR on the industrial structure are built using the core–periphery concept. 1 Area A and area B are the two economic zones. Assume there are no costs or restrictions to labor mobility in both places; consumers are rational, and they consume tradable product x and product y to maximize their utility with wages.

\[
U_A = C_A^u - \frac{\mu}{\sigma} C_A^1
\]

\[
S.t. P_A x C_A x + P_A y C_A y = W_A
\]

where \(0 < \mu < 1 < \sigma\), \(C_A x\) is the product x consumed in area A, \(C_A y\) is the product y consumed in area A, \(P_A x\) and \(P_A y\) are the corresponding prices, following the form of the D-S model. \(C_A\) is expressed by constant elasticity of substitution (CES), \(W_A\) is a consumer’s income in area A. If consumer indifference curves are continuous, then:

1 The core–periphery model is based on Nobel laureate Krugman’s Dixit-Stiglitz (D-S) model, which first proposed that the link between transportation costs and economies of scale is the basis of industrial agglomeration in 1991.
the industrial structure of the opening of HSR on urban rail transit (2022) 8(3-4):296–317 299

\[ C_{Ax} = \left( \int_{0}^{\alpha} c_{Ax}^{1-\left(1/\sigma\right)} d_{i} \right)^{1/(1-1/\sigma)} \]  

(3)

\[ P_{Ax} = (p_{i}^{1-\sigma} d_{i})^{(1-\sigma)} \]  

(4)

where \( p_{i} \) is the price of the \( i \)-th product. After optimization, the indirect utility function is:

\[ V_{A} = \mu^{(1 - \mu)} W_{A}/(P_{Ax}^{1 - \mu}) \]  

(5)

where \( \mu \) is the payment’s share for the \( i \)-th product. With the optimization conditions of suppliers, we define \( W_{B} \) to be the income of consumers in area \( B \), and \( s_{n} \) to be the proportion of product \( x \) in area \( A \) among all product \( x \) (\( S_{n} = n_{i}/n \) where \( n \) is the total of \( x \)). This paper introduces the theory of “iceberg transportation cost” proposed by Samuelson, which holds that there is a cost, represented by \( T \), and only \( 1/T \) of products can be reached in the process of transporting products from area \( A \) to area \( B \). Then, (4) can be transformed into the following:

\[ P_{Ax} = (s_{n} W_{A}^{1-\sigma} + (1 - s_{n})(W_{B} T)^{1-\sigma})^{1/(1-\sigma)} \]  

(6)

where \( \sigma \) is the elasticity of substitution, and \( T \) is the product’s transportation cost between area \( A \) and area \( B \). Substituting (6) into (5), we get:

\[ V_{A} = \mu^{(1 - \mu)} W_{A}/(P_{Ay}^{1 - \mu})^{1-\sigma} + (1 - s_{n})(W_{B} T)^{1-\sigma} \]  

(7)

where \( \sigma = \mu/(\sigma - 1) \), similarly, we get:

\[ V_{B} = \mu^{(1 - \mu)} W_{B}/(P_{By}^{1 - \mu})^{1-\sigma} + (1 - s_{n})(W_{A} T)^{1-\sigma} \]  

(8)

The “accessibility” between locations has improved after the opening of the HSR. Assume \( H = e^{\tau t(H)} \), where \( H \) is a dummy variable for whether HSR is opened, \( \tau \) is time attenuation, and \( t \) is the average travel time from area \( A \) to area \( B \) after the opening of HSR. According to the location selection of the long-term equilibrium, the comparative utility function is as follows:

\[ S_{AB} = \frac{V_{A}}{V_{B}} \]

\[ = \frac{W_{A} (P_{Ax})^{\mu - 1} s_{n} W_{A}^{1-\sigma} + (1 - s_{n}) (W_{B} e^{\tau t(H)})^{1-\sigma} - \alpha}{W_{B} (P_{By})^{\mu - 1} s_{n} (W_{A} e^{\tau t(H)})^{1-\sigma} + (1 - s_{n}) W_{B}^{1-\sigma} - \alpha} \]

\[ = \left[ \frac{s_{n}}{1 - s_{n}} \left( \frac{W_{A}}{W_{B}} \right)^{1-\sigma} + \left( \frac{e^{\tau t(H)}}{r_{w}} \right)^{1-\sigma} + 1 \right] \]

(9)

Note \( (e^{\tau t(H)})^{a_{w}} = \lambda \). If we substitute the bivariable Taylor series expansion of \( \frac{s_{n}}{1 - s_{n}} \) , \( \frac{W_{A}}{W_{B}} \) and \( ln(1 - \frac{\lambda}{1 - \lambda^{2}}) \) into (9) and take the logarithm, we get:

\[ lnS_{AB} = ln \frac{W_{A}}{W_{B}} + (\mu - 1) ln \frac{P_{Ay}}{P_{By}} - \alpha ln \lambda - \frac{\lambda}{1 - \lambda^{2}} \frac{s_{n}}{1 - s_{n}} \]

(10)

The consumer utilities in two areas are equal in the long-term equilibrium, that is:

\[ S_{AB} = \frac{V_{A}}{V_{B}} = 1, \quad lnS_{AB} = 0 \]

(11)

Note \( \frac{s_{n}}{1 - s_{n}} = S \), then we get:

\[ S = \frac{s_{n}}{1 - s_{n}} \]

\[ = \frac{\lambda}{1 - \lambda^{2}} ln \lambda + \frac{\lambda}{\alpha(1 - \lambda^{2})} ln \frac{W_{A}}{W_{B}} + \frac{\lambda(\mu - 1)}{\alpha(1 - \lambda^{2})} ln \frac{P_{Ay}}{P_{By}} \]

(12)

\( \frac{\delta S}{\delta \alpha} \) is the impact of HSR on the industrial structure. Take the derivative of (12) with respect to \( H \):

\[ \frac{\delta S}{\delta H} = \frac{\delta \lambda}{\delta H} \left( 1 - 3 \lambda^{2} \right) \left| ln \frac{W_{A}}{W_{B}} + (1 - \mu) ln \frac{P_{Ay}}{P_{By}} \right| \alpha + \lambda^{2} \left( 1 - ln \lambda - \lambda^{2} \right) \]

(13)

It can be concluded that \( 0 < \frac{\delta \lambda}{\delta H} < 1 \). \( \frac{\delta S}{\delta H} \) can be determined by (13).
While $\lambda \to 0^+$ \[ \frac{\partial S}{\partial H} > 0 \]

While $\lambda \to 1/\sqrt{3}$ \[ \frac{\partial S}{\partial H} > 0 \]

While $\lambda \to 1^-$ \[ \frac{\partial S}{\partial H} > 0 \]

If $\frac{\partial S}{\partial H} > 0$, the behavior is “aggregation,” but if $\frac{\partial S}{\partial H} < 0$ then behaves as “dispersion.” Calculate the second derivative of (12) as follows:

\[
\frac{\partial^2 S}{\partial H^2} = \left(\frac{\partial}{\partial H}\right)^2 \left[ \frac{\beta}{\pi} \left( W_1 - W_2 \right) \left( 1 - \mu \right) \right] \left( \frac{4k^2 - 1 - 6k^2 \ln k - 2k^2 \ln 2k^2}{(\lambda - 1)^4} \right) \]

It is not difficult to determine that $\frac{\partial S}{\partial H} > 0$, and $\frac{\partial S}{\partial H} < 0$ within the domain. Next, we conclude the following:

1. $\left[ \ln \frac{W_1}{W_2} + (1 - \mu) \ln \frac{P_{tr}}{P_{fr}} \right] > 0$, then $\frac{\partial^2 S}{\partial H^2} < 0$, $\forall \lambda \in (0, 1)$ and $\frac{\partial S}{\partial H} > 0$, $\forall \lambda \in (0, 1)$. Fig. 4a shows that the opening of the HSR can boost the number of industries, but the pace of growth will slow at the periphery.

2. $\left[ \ln \frac{W_1}{W_2} + (1 - \mu) \ln \frac{P_{tr}}{P_{fr}} \right] < 0$. Fig. 4b shows $\frac{\partial^2 S}{\partial H^2} < 0$ at the beginning, $\frac{\partial S}{\partial H} > 0$ at the end, and only one zero point.

There are two possible situations. Within the domain, $\frac{\partial S}{\partial H} > 0$ is identical. The HSR is always in an “aggregating state” when it is opened; the pace of aggregation, however, varies. (2) $\frac{\partial S}{\partial H} > 0$ is observed. The aggregation effect of the HSR improves with time until it reaches its peak. The dispersion effect comes later, forcing suppliers with lower competitiveness to migrate to the outskirts. Finally, industries pool their resources to establish higher-quality tertiary industrial conglomerates.

The following hypotheses are proposed according to the derivation:

**Hypothesis I:** Once the HSR opens, the number of tertiary industry businesses will increase and higher-grade industries will continue to aggregate. In the short term, it encourages tertiary industry aggregation in cities along the route, but in the long run, the aggregation is dependent on the initial circumstances.

**Hypothesis II:** Once the HSR opens, the industrial structure advances steadily, resulting in a structural shift in the tertiary industry, from basic to advanced sectors.

### 3 Data

#### 3.1 Objects

This paper chooses Beijing–Shanghai and Beijing–Guangzhou HSR as the objects (see Appendix B for details). On the one hand, they are being built at an early stage. The Beijing–Shanghai HSR is China’s first dedicated long-distance passenger route, and it goes through some of China’s most densely inhabited and economically developed districts. The Beijing–Guangzhou HSR is a watershed moment in China’s HSR development and the world’s longest HSR. Their construction, on the other hand, is rather quick, which helps to eliminate cross-effects caused by other factors. The Beijing–Shanghai and Beijing–Guangzhou HSR run through a total of 131 prefecture-level cities and municipalities. We omitted cities with other HSR lines passing through, prefecture-level cities that were demoted and divided during the sample period, such as Chaohu, cities with poor accessible data, such as Enshi and Shennongjia forest area, and so on in order to obtain unbiased estimates. The control group consists of 40 cities, with a total of 82 prefecture-level cities as samples.

#### 3.2 Variables

#### 3.2.1 Dependent Variable

The dependent variables are the location entropy of the tertiary industry (LQ_third), the rationalization of

![Fig. 4 Forms of industrial aggregation and diffusion of HSR](image-url)
industrial structure (SR) and the advancement of industrial structure (SA). Location entropy is a commonly used index to measure the distribution of elements, and the location entropy of industry $i$ in area $j$ in the country $(LQ_{ij})$ is as follows:

$$LQ_{ij} = \frac{q_{ij}}{q}$$

(16)

where $q_{ij}$ is the indicator of industry $i$ in area $j$, $q$ is the indicator of all industries in the country; $q_i$ is the indicator of all industries in area $j$; $q_i$ is the indicator of industry $i$ in the country. The larger the $LQ_{ij}$ value, the higher the aggregation degree of the tertiary industry, implying that industry $i$ has a competitive advantage in area $j$.

The advancement of the industrial structure (SA) is affected by multiple factors, and its definition is currently neither standard nor rigid. We generated an index to quantify SA using the molar structure change value calculation technique. A set of three-dimensional vectors $X_0 = (x_{1,0}, x_{2,0}, x_{3,0})$ is constructed using the share of three industries in GDP as the spatial vector. The angles between $X_0$ and $X_1 = (1, 0, 0)$, $X_0$ and $X_2 = (0, 1, 0)$, $X_0$ and $X_3 = (0, 0, 1)$ are, respectively, $\theta_1, \theta_2$ and $\theta_3$.

$$\theta_j = \arccos \left[ \frac{\sum_{i=1}^{3} (x_{ij} \cdot x_{i0})}{\left( \sum_{i=1}^{3} x_{ij}^2 \right)^{\frac{1}{2}} \cdot \left( \sum_{i=1}^{3} x_{i0}^2 \right)^{\frac{1}{2}}} \right]$$

(17)

SA is thus defined as follows:

$$SA = \sum_{j=1}^{3} \sum_{i=1}^{3} \theta_j$$

(18)

The rationalization of the industrial structure (SR) measures whether the proportion of industries is appropriate. The following is a regularly used measuring form for assigning weights to the significance of each of the three industries:

$$SR = \sum_{i=1}^{3} \left( \frac{GDP_i}{GDP} \right) \left| \frac{GDP_i / L_i}{GDP / L} - 1 \right|$$

$$= \sum_{i=1}^{n} \left( \frac{GDP_i}{GDP} \right) \left| \frac{GDP_i / GDP}{L_i / L} - 1 \right|$$

(19)

where GDP is the city’s gross product, $GDP_i / L_i$ is labor productivity, $GDP_i / GDP$ is the output structure, and $L_i / L$ is the industry structure. Primary, secondary, and tertiary industries are represented by $i = 1, 2, 3$. The industrial structure deviation is 0 when the output structure matches the employment structure, which is the most appropriate situation. It should be noted that the lower this ratio is, the more the industrial structure has been rationalized.

3.2.2 Independent Variable

The treatment group (treated, set to 1) is cities along the HSR, whereas the control group (controlled, set to 0) is cities outside the HSR. There is also a time dummy variable, which is 0 before the HSR’s opening or 1 after the HSR’s opening. The core independent variable $HSR_i$ is the interaction term of the product of the aforementioned two. The Beijing–Shanghai HSR was opened in 2011, with data from 2001 to 2010 as pre-opening data, and data from 2011 to 2017 as post-opening data. The Beijing–Wuhan segment of the Beijing–Guangzhou HSR opened in 2012, while the Wuhan-Guangzhou section opened in 2009. The Beijing–Guangzhou HSR is valued in the same manner. The particular expression is as follows:

$$Y_{it} = \beta_0 + \beta_1 HSR_{it} + \beta_2 X_{it} + Z_i + year_t + \epsilon_{it}$$

(20)

where $Z_i$ denotes the city-level fixed effects that do not change over time, $X_{it}$ are the time-varying control variables, and $\epsilon_{it}$ is the residual term. The difference induced by $HSR$ is represented by $\beta_1$, and the difference between cities is represented by $\beta_2$.

4 Summary Statistics

Openness ($lfdi$), Economy ($lgdp average$), Wage ($lwage average$), Government interference ($lfee$), Human capital ($lhuman capital$), Informatization ($ltele pop$), Transportation ($lroad average$) and Infrastructure ($lbooks average$ and $lhospital num$) are the control variables. The sample period is 2003–2017, and the data come from the China City Statistical Yearbook and the China Statistical Yearbook. We use the logarithm of the data throughout the paper and provide the summary descriptive statistics in Table 1. Obviously, with the opening of the HSR, the location entropy of the tertiary industry and the advancement and rationalization of industrial structure are greater, and more empirical study is necessary.

5 Empirical Strategy

5.1 Time-Varying DID Regression

5.1.1 Parallel Trend Test

The parallel trend test, as described by Shao [17], is used to determine whether trends in the treatment and control groups are consistent:
\[ Y_{it} = \beta_0 + \sum_{m=1}^{3} \beta_m FirstHSR_{it-m} + \sum_{n=0}^{5} \beta_n FirstHSR_{i,t+n} + \beta_2 X_{it} + Z_i + year_t + \epsilon_{it} \]  

(21)

where \( t \) is the first HSR year of operation. The year before the opening is \( m(m = 1, 2, 3) \), while the year following is \( n(n = 0, 1, 2, 3, 4, 5) \). \( FirstHSR_{i,t} \), \( FirstHSR_{i,t-m} \), \( FirstHSR_{i,t+n} \) are dummy variables that return 1 if city \( i \)'s HSR is operational in the specified year. The initial 4 years (year \( = -4, -3, -2, -1 \)) and the latter 4 years (year \( = 5, 6, 7, 8 \)) are both significant, as shown in Fig. 5. As a result, the opening of HSR has a large delayed impact on the aggregation of tertiary industries, and this conclusion is consistent with Shao [17]. And before year \( = 4 \), \( \beta_m \) increases year by year, which can be explained by the fact that the corporation made strategic planning after getting the news ahead of time [5, 33]. It may be inferred that there is no significant difference in trend between the treatment and control groups before HSR’s opening. By repeating the steps, it is confirmed that both industrial structure advancement and rationalization satisfy the hypothesis of a parallel trend.

5.1.2 Baseline Regression

The opening of HSR can be regarded as a quasi-natural experiment. Multi-period DID is used as the identification model, in the form:

\[ Y = \beta_0 + \delta du + \beta_1 dt + \delta_1 du \cdot dt + \epsilon_{it} \]  

(22)

where \( du \) is the entity fixed effect, \( dt \) is the time fixed effect, \( \delta_1 \) is the coefficient of double differences, which is also the value of interest in this paper. The coefficients of \( HSR \) remain significant and positive when control variables from models (1) to (6) are included in Table 2. Specifically, ceteris paribus, \( LQ_{third} \) and \( SA \) increased by 0.029 and 0.021 respectively after HSR opened. This supports Hypothesis I and Hypothesis II of this paper, namely, that the HSR has a significant positive impact on tertiary industry aggregation and industrial structure advancement.

---

Footnote 2 continued

According to the principles of China’s railway network planning, the location of HSR is determined by the “natural force” of the geographical location, the “government force” of the HSR strategic planning, and the competition and cooperation of local governments and the “market force” of the first-mover advantage of developed regions. As a result, local governments have little control in whether HSR lines go through their jurisdictions.

Footnote 3 continued

A series of tests shown in Appendix C conducted before regression revealed no multicollinearity, heteroscedasticity, or model identification error.
The deviation from equilibrium in industrial structure is shown by the significant positive coefficient of SR. The findings also show that the opening of the HSR causes an imbalance in the local industrial structure and a lack of coordination in resource allocation. Wang (2019) studied the Yangtze River Delta region and discovered that HSR boosted the proportion of tertiary industry added value and improved the quality of urbanization [21]. Liang (2020) performed research on the Guangdong–Guangxi–Guizhou HSR and discovered that by altering the industrial structure of the area, the HSR may promote the development of undeveloped regions [25]. Wang (2019) and Liang (2020) came to similar conclusions as this paper’s findings.

5.2 Robustness Test

5.2.1 Changes in the Sample

The robustness test includes sample change, instrumental variables, and a placebo test. The changes in the sample come from adjusting the period scope [Model (1) to (3)] and excluding provincial capitals and municipalities [Model (4) to (6)]. The HSR coefficients remain significantly positive in Table 3, except for model (5). Model (5)’s insignificance might be due to endogenous recognition across cities, and a heterogeneous effect will be discussed later. As a result, the positive effect of HSR is consistent across samples, and the conclusion in this paper is robust.

5.2.2 Instrumental Variables (IV)

To test for endogeneity owing to omitted variables, instrumental variables (IV) was used in this paper. The IV of transportation infrastructure commonly mentioned in previous papers are landslides, geographic slope, ancient postal services and historical passenger traffic, historical planning information, and so on [34–36]. Referred to Dong

| Variables         | (1)    | (2)    | (3)    | (4)    | (5)    | (6)    |
|-------------------|--------|--------|--------|--------|--------|--------|
| LQ_third          | 0.013* | 0.029*** | 0.041*** | 0.021*** | 2.718*** | 3.077*** |
| SR                | (1.74) | (3.84) | (8.00) | (4.40) | (3.18) | (3.19) |
| lgdp_average      | −0.031*** | −0.046*** | 0.042*** | 0.038*** | 1.094* | 0.717 |
| (−5.05)           | (−4.82) | (7.74) | (5.19) | (1.94) | (0.77) |        |
| lwage_average     | 0.025*** | 0.029*** | 0.003*** | 0.0000 | 0.0900 | 0.1100 |
| (14.16)           | (15.86) | (3.48) | (0.67) | (0.51) | (0.65) |        |
| lfdi              | −0.017*** | −0.013*** | 0.007*** | 0.004** | 0.963*** | 0.887*** |
| (−5.83)           | (−4.57) | (3.57) | (2.03) | (3.53) | (3.20) |        |
| lhuman_capital    | 0.051** | 0.035 | −0.024 | −0.023 | −4.672*** | −3.925*** |
| (2.11)            | (1.53) | (−1.63) | (−1.61) | (−3.40) | (−2.74) |        |
| fee               | 0.000** | 0.000*** | 0.0000 |        |        |        |
| (0.28)            | (5.08) | (−3.61) |        |        |        |        |
| lroad_average     | −0.022*** | 0.019*** | 2.624*** |        |        |        |
| (−2.71)           | (3.09) | (2.59) |        |        |        |        |
| ltele_pop         | 0.001 | −0.002*** | −0.004 |        |        |        |
| (0.96)            | (3.68) | (0.05) |        |        |        |        |
| lbooks_average    | 0.034*** | −0.020*** | −1.256** |        |        |        |
| (7.05)            | (−6.39) | (−1.97) |        |        |        |        |
| lhospital_num     | 0.015 | −0.009* | 3.216** |        |        |        |
| (1.47)            | (−1.79) | (2.48) |        |        |        |        |
| Constant          | 0.225 | 0.538 | 2.067*** | 2.172*** | 55.863** | 27.983 |
| (0.55)            | (1.32) | (8.22) | (8.42) | (2.27) | (1.05) |        |
| City fixed        | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    |
| Time fixed        | Yes    | Yes    | Yes    | Yes    | Yes    | Yes    |
| Observations      | 1,226  | 1,209  | 1,226  | 1,209  | 1,226  | 1,209  |
| R-squared         | 0.847  | 0.861  | 0.893  | 0.911  | 0.445  | 0.452  |

Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively.

4 The finding is still robust (P>|t|=0.125) if the significance level is increased to 20%.
this paper uses China’s historical railway network in 1961 as the IV of the opening of the HSR. The Beijing–Shanghai and Beijing–Guangzhou HSR construction is based on the railways constructed in 1961, and railways built in 1961 have no direct impact on the present industrial structure. Exogeneity and correlation are both satisfied using the railway network in 1961 as IV. The results in Table 4 show that the significant effect of HSR on SH is 0.0822, which is consistent with the baseline regression in Table 2. The correlation between instrumental and explanatory variables is also verified (old railway on HSR is 0.2198), and the assumption of weak instrumental variables was rejected (F/C0Statistic [10]).

5.2.3 Placebo Test

To ensure that the results are not the result of chance or randomness, a placebo test is utilized, in which the treatment group is generated at random. The estimated coefficient of interaction in (22) is as follows:

$$\tilde{\beta}_1 = \beta_1 + \gamma \frac{\text{cov}(\text{treated}_c \times \text{time}_t, z)}{\text{var}(\text{treated}_c \times \text{time}_t[z])}$$

where z is the control variables. When γ = 0, the estimator \(\tilde{\beta}_1\) is unbiased. If \(\text{treated}_c \times \text{time}_t\) is replaced by other variables that do not affect explained variables (\(\beta_1 = 0\)), and \(\tilde{\beta}_1 = 0\) is obtained by estimation, then γ = 0 can be realized. Following this line of reasoning, we make the event of the opening of HSR random, so it has no effect on LQ_third, SA, and SR, i.e., \(\beta_{\text{random}} = 0\). The distribution of \(\tilde{\beta}_{\text{random}}\) is obtained by repeating the above preceding technique as shown in Fig. 6, and t-statistics are distributed U-shaped, with peaks around zero.

5.3 Heterogeneity Test

5.3.1 Socioeconomic Characteristics Among Cities

The effect of HSR is varied across different regions due to differences in endowment [37–39]. This paper takes 3

| Variables                  | (1) | (2) | (3) | (4) | (5) | (6) |
|----------------------------|-----|-----|-----|-----|-----|-----|
| LQ_third                   | 0.023*** | 0.014*** | 2.513** | 0.022*** | 0.011 | 3.490*** |
| SA                         | (2.95) | (3.11) | (2.44) | (2.85) | (1.54) | (2.95) |
| SR                         | 0.003 | 0.057*** | 1.632** | –0.037*** | 0.034*** | 1.357 |
| lwdi                       | (0.27) | (7.05) | (2.12) | (−3.78) | (3.61) | (1.25) |
| lhuman_capital             | 0.029*** | −0.000 | 0.087 | 0.026*** | −0.001 | 0.094 |
| fee                        | (18.55) | (−0.58) | (0.59) | (15.02) | (−1.01) | (0.48) |
| lroad_average              | −0.013*** | −0.000 | 0.518* | −0.013*** | 0.002 | 0.950*** |
| iltele_pop                 | (−4.29) | (−0.16) | (1.85) | (−4.31) | (1.34) | (3.11) |
| lbooks_average             | 0.062** | −0.019 | −3.131* | 0.082*** | 0.002 | −1.828 |
| ffee                       | (2.21) | (−1.17) | (−1.71) | (3.28) | (0.17) | (−0.90) |
| lfdi                       | 0.000 | 0.000*** | −0.000*** | 0.000*** | 0.000*** | −0.000 |
| lhuman_capital             | (1.49) | (4.94) | (−3.09) | (3.70) | (4.01) | (0.93) |
| lroad_average              | −0.009 | −0.017** | −0.277 | −0.30*** | 0.002 | 2.688*** |
| iltele_pop                 | (−0.84) | (−2.37) | (−0.30) | (−3.42) | (0.32) | (2.42) |
| lbooks_average             | −0.000 | −0.001* | 0.060 | 0.002* | −0.001*** | 0.028 |
| fee                        | (−0.47) | (−1.96) | (0.67) | (1.67) | (−2.87) | (0.30) |
| lroad_average              | 0.002 | −0.009*** | 1.619* | 0.031*** | −0.017*** | −1.733*** |
| iltele_pop                 | (0.22) | (−2.10) | (1.72) | (5.85) | (−3.97) | (2.13) |
| lbooks_average             | −0.014 | 0.001 | 6.144*** | 0.019* | −0.008 | 3.970*** |
| Constant                   | (−1.39) | (0.13) | (3.30) | (1.67) | (−1.37) | (2.57) |
| lhospital_num              | −0.294 | 1.953*** | 0.954 | −0.341 | 1.793*** | −15.811 |
| Constant                   | (−0.59) | (6.43) | (0.03) | (−0.78) | (6.35) | (0.43) |
| City fixed                 | Yes  | Yes  | Yes  | Yes  | Yes  | Yes  |
| Time fixed                 | Yes  | Yes  | Yes  | Yes  | Yes  | Yes  |
| Observations               | 889  | 889  | 889  | 1063 | 1063 | 1063 |
| R-squared                  | 0.900 | 0.943 | 0.511 | 0.764 | 0.860 | 0.446 |

Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively.
The impact of HSR implying a link between socioeconomic characteristics and statistically more significant than those of other cities, respectively. Table 5 shows that the economic effects of the spatial correlation test, this paper obtains the spatially relevant industries as follows: mining (LQ_2), electricity, gas, and water production and supply (LQ_1), transportation warehousing and postal (LQ_7), accommodation and catering (LQ_6), financial (LQ_10), real estate (LQ_11), scientific research and technical services (LQ_13), education (LQ_16), health and social security (LQ_17) and public management and social organizations (LQ_19).

Spatial lag models (SAR), spatial error models (SEM), and spatial Durbin models (SDM) are examples of spatial econometric models. SAR and SEM assume spatial autocorrelation between dependent variables and error terms, respectively, whereas SDM considers both. To obtain SAR and SEM, the Kronecker product is employed to integrate the spatial matrix across time. The following are their expressions:

Spatial lag model (SAR):
\[ Y_{it} = \beta_0 + \rho W_{ij} Y_{it} + \beta X_{it} + \gamma_i + \epsilon_{it} \]  
(27)

Spatial error model (SEM):
\[ Y_{it} = \beta_0 + \beta X_{it} + \epsilon_{it} + \gamma_i + \mu_{it}, \mu_{it} = \lambda W_{ij} + \epsilon_{it} \]  
(28)

Spatial Durbin model (SDM):
\[ Y_{it} = \beta_0 + \rho W_{ij} Y_{it} + \theta W_{ij} Y_{it} + \beta X_{it} + \gamma_i + \epsilon_{it} \]  
(29)

where \( \beta_0 \) is a constant, \( \beta \) is the matrix of the variable coefficients, \( X \) is the matrix of the independent variables, and \( W_{ij} \) is the weight matrix. \( \rho \) is the spatial autoregressive coefficient, \( \lambda \) is the spatial autocorrelation coefficient, and \( \theta \) is the spatial spillover effect. \( \gamma_i \) and \( \epsilon_{it} \) are used to measure spatial fixed effects and time fixed effects, respectively. \( \epsilon_{it} \) is the error term that is subject to normal distribution. This paper uses the approach described by Pace (2009) to separate the estimation of direct and indirect effects from SAR and SDM [40], and it takes the following form:

\[ Y_{it} = (I - \rho W_{ij})^{-1}(\beta X_{it} - W_{ij} X_{i0}) \theta + (I - \rho W_{ij})^{-1} \gamma_i + (<br>\[ \lambda_i + (I - \rho W_{ij})^{-1} \epsilon_{it} \]  
(30)

5 The reference is the scale classification standard of the State Council of China for large cities (type I) and small and medium-sized cities (type II), and the population density standard of the State Council of China for the construction of metro cities.

6 The economic effects of the high-speed rail’s opening are numerous, but it is impossible to cover everything at this time. We correctly decrease the significance limit to 20% and delete Moran’s I insignificant indications.
The partial derivatives with regard to the $k^{th}$ independent variable are as follows from area $1$ to area $N$:

\[
\begin{bmatrix}
\frac{\partial Y}{\partial X_{1k}} & \cdots & \frac{\partial Y}{\partial X_{nk}} \\
\frac{\partial Y}{\partial X_{1k}} & \cdots & \frac{\partial Y}{\partial X_{nk}} \\
\vdots & \ddots & \vdots \\
\frac{\partial Y}{\partial X_{1k}} & \cdots & \frac{\partial Y}{\partial X_{nk}}
\end{bmatrix} = (I - \rho W_q)^{-1}
\]

\[
= \begin{bmatrix}
\beta_k & w_{12} \theta_k & \cdots & w_{12} \theta_k \\
\vdots & \ddots & \ddots & \vdots \\
w_{N1} \theta_k & w_{N2} \theta_k & \cdots & \beta_k
\end{bmatrix}
\]

where $w_{ij}$ is the element $(i,j)$ of the matrix $W_{ij}$. The direct effect is the average of the sum of the diagonal elements of the matrix (32). The indirect effect is the average of the sum of all row and column elements of the non-diagonal elements, which is also the spillover effect. The spatial panel models are estimated by maximum likelihood estimation (MLE) to avoid biased estimators. When LM-err is significant, the criterion for deciding the optimal spatial model is SEM, and when LM-lag is significant, SAR. The robustness of LM-err and LM-lag are compared if they are both substantial. The results are shown in the Tables 14 to 16 in Appendix D, indicating that SDM is the model with the best explanatory power. The aggregation of the transportation, warehousing, and postal sectors is significantly reduced as a consequence of the HSR’s opening, as can be seen in Table 6, and they have a significant spillover impact on neighboring cities. Because HSR boosts urban housing costs, relatively low-end transportation and warehousing in the tertiary sector will be shifted to distant locations, with nearby cities being suitable places to accept them, leading to an increase in the industrial aggregation of these industries. This phenomenon is called the siphon effect.

Table 5 Heterogeneity test of socio-economic characteristics among cities

| Variables | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  | (8)  | (9)  | (10) | (11) | (12) |
|-----------|------|------|------|------|------|------|------|------|------|------|------|------|
| LQ_third  |      |      |      |      |      |      |      |      |      |      |      |      |
| SA        |      |      |      |      |      |      |      |      |      |      |      |      |
| SR        |      |      |      |      |      |      |      |      |      |      |      |      |
| Megalopolis | 0.027*** | 0.016*** | 1.535** | 0.051* | -0.002  | 14.403*** |
| Observations | (3.16) | (3.25) | (2.16) | (1.96) | (-0.13) | (2.94) |
| Adjusted R-squared | 0.852 | 0.915 | 0.454 | 0.908 | 0.920 | 0.510 |
| Small and medium-sized cities | 0.034*** | 0.017*** | 3.038*** | 0.029* | 0.013*** | 2.122 |
| Observations | (3.50) | (3.06) | (2.77) | (1.77) | (1.51) | (1.41) |
| Adjusted R-squared | 0.888 | 0.929 | 0.482 | 0.788 | 0.842 | 0.419 |
| Density>|0.05 | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Variable control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| City fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Time fixed | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively.
effect [2]. Furthermore, since the opening of HSR, there has been a structural shift in the tertiary industry, from basic to advanced sectors. By analyzing the heterogeneity of HSR in the industrial structure from the effects of HSR on the industry, HSR has an overall positive effect on most sectors, showing that the opening of HSR has fostered the aggregation of the tertiary industry as a whole. Debrezion (2007) [41], He (2020) [42], Huang (2020) [43], Shao (2017) [17], and Wang (2018) [44] studied the impact of HSR on real estate, automobiles, services, and finance, respectively, and found heterogeneity in the impact of HSR across industries, which is consistent with the findings in this paper.

5.4 Human Capital’s Mediation Effect

Following the demonstration of the heterogeneity in the impact of HSR on various cities, it should be determined whether the impact is produced by the flow of human resources. In this paper, the number of college students is utilized as a variable to quantify human resources. If the coefficients of the three regressions are significant but \( c \) is minor, then the human resource has an intermediary effect on the impact of HSR on industrial structure. Using the steps below:

\[
Y = cX + e_1 \tag{32}
\]

\[
M = aX + e_2 \tag{33}
\]

\[
Y = c'X + bM + e_3 \tag{34}
\]

where \( X \) is HSR, \( Y \) are dependent variables, and \( M \) is the mediator variable. The significance of HSR’s coefficients in Models (1), (3), and (5) in Table 7 demonstrates its explanatory power for \( Y \) and \( M \). In Model (2), the partial intermediary effect of human resources is shown with a coefficient of \( ab/c = 0.0723 \) for \( LQ_{third} \) as the explanatory variable. Because the \( b_{wrdxs} \) coefficients in Model (4) are insignificant, the bootstrap test is required to obtain a distribution that is close to the population, and the results for SA are reported in Table 8. The sign of the coefficient in Table 8 is significantly positive, suggesting that human resources have only an indirect effect. Combining Tables 7 and 8, it can be concluded that human resources serve as a complete intermediary in the impact of HSR on the industrial structure.

### Table 6  Effect decomposition of different industries along Beijing-Guangzhou and Beijing-Shanghai lines

| Variables | SEM | SAR | SDM |
|-----------|-----|-----|-----|
|           | Direct | Direct | Indirect | Total | Direct | Indirect | Total |
| \( LQ_{_2} \) | -0.1519*** | -0.1473*** | -0.0115 | -0.1587*** | -0.0184*** | 0.0703 | -0.1141 |
|           | (-3.8694) | (-3.8387) | (-1.2294) | (-3.7030) | (-3.9172) | (0.7359) | (-1.4300) |
| \( LQ_{_4} \) | -0.1187*** | -0.0861*** | -0.0305*** | -0.1167*** | 0.1771*** | 0.0292*** | 0.1150* |
|           | (-4.7052) | (-3.7432) | (-2.8607) | (-3.5925) | (-6.7065) | (4.3179) | (1.8330) |
| \( LQ_{_5} \) | -0.1134*** | -0.1035*** | -0.0159*** | -0.0119*** | -0.1448*** | 0.1377*** | -0.0071 |
|           | (-6.4077) | (-6.1867) | (-2.2107) | (-5.6810) | (-7.2599) | (3.2819) | (-0.2058) |
| \( LQ_{_7} \) | -0.0368 | -0.0378 | -0.0021 | -0.0398 | 0.0426 | -0.2154 | -0.1728 |
|           | (-0.6570) | (-0.6762) | (-0.3953) | (-0.6791) | (0.6282) | (-1.6232) | (-1.5984) |
| \( LQ_{_9} \) | -0.0885*** | -0.0684*** | -0.0407*** | -0.1091*** | -0.0982*** | 0.1247* | 0.0265 |
|           | (-3.6840) | (-3.0821) | (-2.6752) | (-2.9964) | (-3.9350) | (1.8093) | (0.4055) |
| \( LQ_{_11} \) | 0.1157*** | 0.1056*** | 0.0271*** | 0.1328*** | 0.1344*** | -0.0296 | 0.1048 |
|           | (3.8898) | (3.7482) | (2.7698) | (3.7653) | (4.0432) | (-0.3969) | (1.5386) |
| \( LQ_{_13} \) | -0.0146 | -0.0145 | -0.0013 | -0.0159 | -0.0119 | 0.0122 | 0.0002 |
|           | (-0.7204) | (-0.7200) | (-0.6322) | (-0.7318) | (-0.4943) | (0.2581) | (-0.0196) |
| \( LQ_{_15} \) | -0.0199 | -0.0105 | -0.0084 | -0.0190 | -0.0317 | 0.0093 | -0.0223 |
|           | (-1.0490) | (-0.6076) | (0.6041) | (0.6017) | (-1.6394) | (0.1534) | (-0.3620) |
| \( LQ_{_17} \) | -0.0465*** | -0.0276 | -0.0253 | -0.0530 | -0.0601*** | 0.1428** | 0.0827 |
|           | (-2.5483) | (-1.6468) | (-1.5662) | (-1.6128) | (-1.6383) | (0.1509) | (-0.3600) |
| \( LQ_{_19} \) | -0.0104 | -0.0018 | -0.0016 | -0.0035 | -0.0263 | 0.0267 | 0.0004 |
|           | (-0.5588) | (-0.1174) | (-0.1239) | (-0.1113) | (-1.3891) | (0.4233) | (-0.0122) |

Bold indicates SDM is the model with the best explanatory power
Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively
With China’s fast expansion of HSR, how to accurately quantify the impact of HSR on the industrial structure is of great concern to many scholars. In light of China’s current state of the economy, the theoretical framework and hypothesis of the impact of HSR on the industrial structure are derived using the core–periphery model. This demonstrates that the impact of HSR on industrial structure aggregation includes decreasing-speed and U-shaped-speed aggregation, while the impact of HSR on industrial structure rationalization is uncertain. A series of empirical studies are based on the three hypotheses given in this paper.

The findings show that, first, HSR promotes tertiary industry aggregation and contributes to the transformation of China’s industrial structure. This is evident in the significant positive coefficients of HSR in the regression models for tertiary industry aggregation. The coefficients of HSR are significant at the 1% level in models (3) and (4), indicating a robust and significant impact. Additionally, the coefficients of HSR in models (1) and (2) are positive but not statistically significant, suggesting that the impact of HSR on tertiary industry aggregation is not consistent across different samples or time periods.

In summary, the empirical evidence supports the hypothesis that HSR promotes tertiary industry aggregation. This finding is consistent with the theoretical framework and hypothesis derived from the core–periphery model. The results provide valuable insights for policy makers in China, as they consider ways to further enhance the impact of HSR on China’s industrial structure, particularly in promoting the development of the tertiary sector.

### Table 7: Intermediary effect test of human capital

| Variables | (1) LQ_third | (2) LQ_third | (3) SA | (4) SA | (5) lwrdks |
|-----------|-------------|-------------|-------|-------|------------|
| lwrdks    | -0.028***   | -0.026***   | 0.021*** | 0.019*** | -0.067*** |
|           | (-2.96)     | (-2.98)     | (4.53) | (4.12) | (-2.63)   |
| HSR       | 0.029***    | 0.026***    | 0.038*** | 0.038*** | 0.320***  |
|           | (3.62)      | (3.31)      | (7.19) | (6.84) | (11.09)   |
| lgdp_average | -0.046***   | -0.038***   | 0.000  | 0.001  | 0.011*    |
|           | (-5.12)     | (-3.98)     | (7.19) | (6.84) | (11.09)   |
| lwage_average | 0.029***    | 0.029***    | 0.000  | 0.024* | 0.332***  |
|           | (14.36)     | (14.53)     | (0.42) | (0.44) | (1.70)    |
| lfdi      | -0.013***   | -0.012***   | 0.004** | 0.004** | 0.015     |
|           | (-4.17)     | (-4.17)     | (2.07) | (2.13) | (1.56)    |
| lhuman_capital | 0.035*      | 0.045**     | -0.023* | -0.024* | 0.332***  |
|           | (1.65)      | (2.06)      | (-1.84)| (-1.93)| (4.83)    |
| fee       | 0.000       | -0.000      | 0.000*** | 0.000*** | -0.000*** |
|           | (0.37)      | (-0.00)     | (6.05) | (6.38) | (-4.24)   |
| lroad_average | -0.022***   | -0.013      | 0.019*** | 0.022*** | 0.325***  |
|           | (-2.79)     | (-1.51)     | (4.10) | (4.40) | (12.62)   |
| ltele_pop | 0.001       | 0.001       | -0.002*** | -0.002*** | -0.012*** |
|           | (0.98)      | (0.62)      | (-3.46)| (-3.54)| (-3.92)   |
| lbooks_average | 0.034***    | 0.033***    | -0.020*** | -0.020*** | -0.011    |
|           | (7.19)      | (7.15)      | (-7.46)| (-7.52)| (-0.71)   |
| lhospital_num | 0.015       | 0.013       | -0.009*  | -0.010*  | -0.079*** |
|           | (1.59)      | (1.34)      | (-1.73)| (-1.88)| (-2.64)   |
| Constant  | 0.538       | 0.420       | 2.172*** | 2.221*** | -4.006*** |
|           | (1.47)      | (1.12)      | (10.12)| (10.22)| (-3.35)   |
| City fixed | Yes         | Yes         | Yes     | Yes     | Yes       |
| Time fixed | Yes         | Yes         | Yes     | Yes     | Yes       |
| Observations | 1,209       | 1,206       | 1,209  | 1,206  | 1,206     |
| R-squared  | 0.861       | 0.861       | 0.911  | 0.912  | 0.951     |

Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively.

### Table 8: Bootstrap test of SA

| Variables | Coefficient | Bootstrap standard error | Z     | P>|Z|  | Normal-based [95% confidence interval] |
|-----------|-------------|--------------------------|-------|------|-----------------------------|
| Indirect influence | 0.010***    | 0.002                    | 4.49  | 0.005| 0.014                       |
| Direct influence   | -0.002      | 0.006                    | -0.26 | 0.000| -0.014                      |

Note: *, **, and *** indicate significant levels at 10%, 5%, and 1%, respectively.

### 6 Discussion and Conclusion

With China’s fast expansion of HSR, how to accurately quantify the impact of HSR on the industrial structure is of great concern to many scholars. In light of China’s current state of the economy, the theoretical framework and hypothesis of the impact of HSR on the industrial structure are derived using the core–periphery model. This demonstrates that the impact of HSR on industrial structure aggregation includes decreasing-speed and U-shaped-speed aggregation, while the impact of HSR on industrial structure rationalization is uncertain. A series of empirical studies are based on the three hypotheses given in this paper.

The findings show that, first, HSR promotes tertiary industry aggregation and contributes to the transformation of China’s industrial structure. This is evident in the significant positive coefficients of HSR in the regression models for tertiary industry aggregation. The coefficients of HSR are significant at the 1% level in models (3) and (4), indicating a robust and significant impact. Additionally, the coefficients of HSR in models (1) and (2) are positive but not statistically significant, suggesting that the impact of HSR on tertiary industry aggregation is not consistent across different samples or time periods.

In summary, the empirical evidence supports the hypothesis that HSR promotes tertiary industry aggregation. This finding is consistent with the theoretical framework and hypothesis derived from the core–periphery model. The results provide valuable insights for policy makers in China, as they consider ways to further enhance the impact of HSR on China’s industrial structure, particularly in promoting the development of the tertiary sector.
of the industrial structure from primary to the secondary and tertiary industry sector, as well as realizing the industrial structure advancement but not rationalization. Next, the impact of HSR on tertiary industry aggregation in major cities and high-density cities is greater than that in other cities, whereas the impact on the industrial structure advancement is smaller. Moreover, following the HSR’s opening, the aggregation of the transportation, warehousing, and postal sectors has been greatly reduced, with a significant spillover effect on neighboring cities, proving the siphon effect and conduction mechanism of the HSR on industrial structure. There has also been a structural shift in the tertiary industry, from basic to advanced sectors. Finally, it has been confirmed that HSR decreases human resource flow costs and plays a partial intermediate function in the aggregation of tertiary industry, and the advanced industrial structure’s intermediary role is clearer.

The primary contribution of this paper is the selection of appropriate research objects. The impacts of newly built stations and the rehabilitation of existing stations will overlap if all HSR in the area is evaluated, and the cross-effects will lead to skewed results. This study determines the suitable control and treatment groups for improving recognition accuracy after extensive comparison. The second contribution is that this research creates a mathematical economic geography model of the influence of HSR opening on the industrial structure based on Nobel laureate in Economics Krugman’s core–periphery model. Samuelson’s iceberg transportation cost notion is introduced, which is supplemented to provide a reference for follow-up research. The third contribution is that, unlike previous research, this work takes into account the heterogeneity of HSR’s economic impacts across sectors, examines the direct and indirect heterogeneity of HSR on various industries in the city, and investigates the “siphon effect” of HSR.

As urbanization and the establishment of HSR progress, China should continue to boost investment in HSR development and make active use of resource allocation tools to aid in the transformation of the industrial structure. Simultaneously, it should not pursue tertiary industry growth blindly, in order to prevent the establishment of the siphon effect. The shortcomings of this paper, on the one hand, the data used are prefecture-level, not county-level units, potentially resulting in inadequate precision. This work, on the other hand, uses the whole railway line as the research object to prevent cross-influence, potentially resulting in a self-selection dilemma, which could be solved by segmenting the line to form a control group.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40864-022-00175-w.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Appendix

Appendix A: The Review of the Literature on High-Speed Rail and Economic Growth

See Table 9.

| Literature | Research object | Hierarchy | Sample period | Identification method | Notes |
|------------|-----------------|-----------|---------------|-----------------------|-------|
| Hall (2011) | UK | Nation | 1976–2006 | Statistics | – |
| Levinson (2012) | America | Nation | 2012 | Case analysis | – |
| Silva (2014) | Spain | Nation | 1990–2010 | SEM, MLE | – |
| Ahlfeldt (2015) | Germany | Area | 1995–2009 | DID | Cologne–Frankfurt (HSR) |
| Pugliara (2017) | Italy | Nation | 2006–2013 | GLM | Impact on tourism |
| | China | Nation | 2000–2014 | FE | Colonial railway |
Table 9 continued

| Literature | Research object | Hierarchy | Sample period | Identification method | Notes |
|------------|-----------------|-----------|---------------|----------------------|-------|
| Haynes (2017) | Guirao (2017) | Spain | Area | 2004–2015 | FE | Labor migration in Madrid |
| Juan (2018) | Cascetta (2020) | Italy | Nation | 2008–2018 | Case analysis | Reduced reachability fairness by 11% |
| Shao (2017) | Dong (2018) | China | Yangtze River Delta region | 1995–2014 | Multi-period DID | Train frequency in the service industry |
| Meng (2018) | Zheng (2019) | China | Nation | 98 cities | 2004–2013 | Multi-period DID | Calculation of spatial spillover effects of high-speed railway stations in China |
| Wang (2019) | Wang (2019) | China | Yangtze River Delta region | 2000–2013 | Gravity model | Research on the evolution law of hierarchical structure and physical structure |
| Pan (2019) | Liu (2020) | China | Yangtze River Delta region | 2008–2010 | SARAR | Fine-grained nighttime light levels represent the level of economic growth |
| Liang (2020) | Li (2020) | China | Guangdong, Guangxi, Guizhou | 2012–2017 | PSM-DID | Using NPP-VIIRS remote sensing data to represent the regional economic growth |
| Li (2020) | Wang (2020) | China | 40 prefecture-level cities | 2001–2017 | SCM | Using a synthetic control method |
| Li (2020) | Wang (2020) | China | Nation | 2004–2013 | PSM-DID | Using DMSP/OLS and night lighting data |
| Meng (2021) | Li (2020) | China | 61 prefecture-level cities | 2008–2017 | DID, threshold effect model | Research on the relationship between HSR frequency and economic efficiency |

Appendix B: The list of HSR and cities along and outside the Beijing–Shanghai and Beijing–Guangzhou HSR

See Tables 10, 11.

Table 10 A list of the history of the opening of some high-speed rail

| High-speed rail line | Mileage (km) | Speed (km/h) | Opening time |
|----------------------|-------------|--------------|--------------|
| Beijing–Tianjin intercity railway | 115 | 350 | 2008/08/01 |
| Wuhan–Guangzhou passenger dedicated line | 968 | 300 | 2009/12/26 |
| Zhengxi passenger dedicated line | 484 | 300 | 2010/02/06 |
| Changjiu intercity railway | 131 | 300 | 2010/09/20 |
| Hainan east ring railway | 308 | 250 | 2010/12/30 |
| Changji intercity railway | 111 | 250 | 2010/12/30 |
| Beijing–Shanghai high-speed railway | 1318 | 325 | 2011/06/30 |
Table 10 continued

| High-speed rail line | Mileage (km) | Speed (km/h) | Opening time |
|---------------------|--------------|--------------|--------------|
| Hebong passenger dedicated line | 131 | 350 | 2012/10/16 |
| Harbin-Dalian passenger line | 904 | 300 | 2012/12/01 |
| Shiwei passenger dedicated line | 841 | 300 | 2012/12/26 |
| Hangyong passenger dedicated line | 150 | 300 | 2013/07/01 |
| Nanjing-Hangzhou passenger dedicated line | 249 | 300 | 2013/07/01 |
| Panying passenger dedicated line | 90 | 300 | 2013/07/01 |
| Hengliu railway | 498 | 250 | 2013/12/28 |
| Liunan intercity railway | 226 | 250 | 2013/12/28 |
| Wuxian intercity railway | 90 | 300 | 2013/12/28 |
| Xibao passenger dedicated line | 148 | 250 | 2013/12/28 |
| Qinfang railway | 63 | 250 | 2013/12/30 |
| Yongbei line | 199 | 250 | 2013/12/30 |
| Nanguan railway | 577 | 250 | 2014/04/18 |
| Wuhuang intercity railway | 133 | 250 | 2014/06/18 |
| Daxi passenger dedicated line | 859 | 250 | 2014/07/01 |
| Hangzhou-Changzhou passenger dedicated line | 924 | 300 | 2014/09/16 |
| Changkun passenger dedicated line | 706 | 300 | 2014/12/16 |
| Chengmian passenger dedicated line | 319 | 250 | 2014/12/20 |
| Guiguan passenger dedicated line | 861 | 250 | 2014/12/26 |
| Lanzin railway second double line | 1787 | 250 | 2014/12/26 |
| Qingrong intercity railway | 299 | 250 | 2014/12/28 |
| Hefu passenger dedicated line | 808 | 300 | 2015/06/28 |
| Haqi passenger dedicated line | 286 | 250 | 2015/08/17 |
| Shendan passenger line | 208 | 250 | 2015/09/01 |
| Jihun passenger dedicated line | 378 | 250 | 2015/09/20 |
| Jinbin intercity railway | 45 | 350 | 2015/09/20 |
| Lanzhong intercity railway | 63 | 250 | 2015/09/30 |
| Ning’an passenger dedicated line | 257 | 250 | 2015/12/06 |
| Nankun passenger dedicated line | 710 | 250 | 2015/12/11 |
| Chengdu-Chongqing passenger line | 305 | 300 | 2015/12/26 |
| Jinwen railway capacity expansion and reconstruction project | 188 | 250 | 2015/12/26 |
| Lanzhou-Chongqing railway | 352 | 250 | 2015/12/26 |
| Jinbao railway | 145 | 250 | 2015/12/28 |

Table 11 List of cities along and outside the Beijing–Shanghai and Beijing–Guangzhou high-speed rail

| Cities along the line | Beijing–Shanghai high-speed rail: |
|----------------------|----------------------------------|
| Beijing (1)          | Beijing                           |
| Tianjin (1)          | Tianjin                           |
| Shanghai (1)         | Shanghai                          |
| Hebei (2)            | Cangzhou, Langfang                |
| Jiangsu (6)          | Nanjing, Wuxi, Suzhou, Zhenjiang, Xuzhou, Changzhou |
| Anhui (3)            | Suzhou, Chuozhou, Bengbu          |
| Shandong (5)         | Jinan, Zhaozhuang, Tai’an, Dezhou, Jining |

| Cities along the line | Beijing–Guangzhou high-speed railway: |
|----------------------|---------------------------------------|
| Beijing (1)          | Beijing                               |
Appendix C: Tests in heteroscedasticity, multicollinearity, and RESET

See Table 12.

Table 12 Variance inflation factors

| Variables     | VIF | 1/VIF |
|---------------|-----|-------|
| lgdp_average  | 5.82| 0.1720|
| road_average  | 3.45| 0.2897|
| lbooks_average| 2.77| 0.3610|
| lfdi          | 2.54| 0.3937|
| lhospital_num  | 2.39| 0.4190|
| lhuman_capital| 2.27| 0.4405|
| fee           | 1.74| 0.5759|
| ltele_pop     | 1.71| 0.5802|
| hsr           | 1.37| 0.7277|
| lwage_average | 1.35| 0.7407|
| Mean          | 2.54|       |

Heteroscedasticity Test

Considering the possible influence of heteroscedasticity, the BP test was performed on the heteroscedasticity of the data. The basic principle of the BP test is to test whether there is heteroscedasticity through auxiliary regression, as shown in Fig. 7. The results show that the null hypothesis is accepted, indicating that there is no obvious heteroscedasticity effect in the data set, and further discussion can be continued. The verified industry-advanced SH and industrial structure rationalization SR have also passed the BP test.

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

Ho: Constant variance
Variables: fitted values of LQ_third

\[
F(1, 1207) = 0.56
\]
\[
Prob > F = 0.4551
\]

Fig. 7 The results of the BP test
**Multicollinearity Test**

The variance inflation factor is a measure of the severity of multicollinearity in a multiple linear regression model, and it represents the ratio of the variance of the regression coefficient estimator compared with the variance assuming non-linear correlations among the independent variables. A rule of thumb for judging whether there is multicollinearity is that the VIF of the independent variable should not exceed 10, otherwise multicollinearity may become a problem. B.1 shows that the VIF of all variables does not exceed 10, that is, there is no obvious multicollinearity for the variables in this paper.

**RESET Test**

Considering the effect of possible omitted interactions of higher-order terms and other variables, a regression set error test (RESET) was performed. It is a general method for testing functional form in multiple regression models. Its rationale is a joint significance F-test of squared, cubic, and possibly higher powers of the fitted values derived from the original OLS estimate. Fig. 8 Accepting the null hypothesis means that the model does not obviously miss the need to include complex interaction terms between all independent variables and higher-order terms of the respective variables, and the next in-depth analysis can be carried out.

**Appendix D: Division standard of industries and test results of spatial correlation**

See Tables 13, 14, 15, and 16.

### Table 13 Division standard of three industries

| LQ1 Primary industry | Agriculture, forestry, animal husbandry, and fishery |
|----------------------|-----------------------------------------------------|
| LQ2 Secondary industry | Mining industry                                     |
| LQ3                   | Manufacturing                                        |
| LQ4                   | Production and supply of electricity, gas, and water |
| LQ5                   | Construction industry                                |
| LQ6 Tertiary industry | Wholesale and retail trade                          |
| LQ7                   | Transportation, warehousing, and postal industry     |
| LQ8                   | Accommodation and catering                           |
| LQ9                   | Information transmission, computer                   |
| LQ10                  | Financial industry                                   |
| LQ11                  | Real estate                                          |
| LQ12                  | Leasing and business services                        |
| LQ13                  | Scientific research and technical services           |
| LQ14                  | Water conservancy, environment and public facilities management industry |
| LQ15                  | Residential services, repairs and other services     |
| LQ16                  | Education industry                                   |
| LQ17                  | Health, social security                               |
| LQ18                  | Culture, sports, and entertainment industries        |
| LQ19                  | Public administration and social organization        |
### Table 14 Test results of spatial correlation of LQ_third

| Year | Spatial error: Moran's I | Spatial error: Lagrange multiplier | Spatial error: Robust Lagrange multiplier | Spatial lag: Lagrange multiplier | Spatial lag: Robust Lagrange multiplier |
|------|--------------------------|-----------------------------------|------------------------------------------|----------------------------------|----------------------------------------|
| 2003 | 2.424                    | 2.604                             | 3.318                                    | 0.004                            | 0.719                                  |
|      | 0.015                    | 0.107                             | 0.069                                    | 0.949                            | 0.397                                  |
| 2004 | 2.908                    | 4.515                             | 3.914                                    | 0.652                            | 0.051                                  |
|      | 0.004                    | 0.034                             | 0.048                                    | 0.419                            | 0.821                                  |
| 2005 | 2.916                    | 4.583                             | 4.671                                    | 0.263                            | 0.351                                  |
|      | 0.004                    | 0.032                             | 0.031                                    | 0.608                            | 0.553                                  |
| 2006 | 2.722                    | 3.929                             | 0.207                                    | 0.207                            | 0.339                                  |
|      | 0.006                    | 0.047                             | 0.649                                    | 0.649                            | 0.560                                  |
| 2007 | 1.397                    | 0.571                             | 1.121                                    | 0.139                            | 0.690                                  |
|      | 0.162                    | 0.450                             | 0.290                                    | 0.709                            | 0.406                                  |
| 2008 | 1.875                    | 1.490                             | 1.840                                    | 0.005                            | 0.354                                  |
|      | 0.061                    | 0.222                             | 0.175                                    | 0.946                            | 0.552                                  |
| 2009 | 5.010                    | 15.688                            | 13.507                                   | 2.354                            | 0.173                                  |
|      | 0.000                    | 0.000                             | 0.000                                    | 0.125                            | 0.678                                  |
| 2010 | 3.373                    | 6.862                             | 4.873                                    | 2.023                            | 0.034                                  |
|      | 0.001                    | 0.009                             | 0.027                                    | 0.155                            | 0.853                                  |
| 2011 | 3.717                    | 8.065                             | 4.989                                    | 3.453                            | 0.376                                  |
|      | 0.000                    | 0.005                             | 0.026                                    | 0.063                            | 0.540                                  |
| 2012 | 3.176                    | 5.432                             | 2.876                                    | 3.581                            | 1.025                                  |
|      | 0.001                    | 0.020                             | 0.090                                    | 0.058                            | 0.311                                  |
| 2013 | 3.078                    | 5.158                             | 2.784                                    | 3.466                            | 1.093                                  |
|      | 0.002                    | 0.023                             | 0.095                                    | 0.063                            | 0.296                                  |
| 2014 | 1.335                    | 0.508                             | 0.041                                    | 1.975                            | 1.508                                  |
|      | 0.182                    | 0.476                             | 0.840                                    | 0.160                            | 0.219                                  |
| 2015 | 1.078                    | 0.245                             | 0.190                                    | 0.060                            | 0.005                                  |
|      | 0.281                    | 0.620                             | 0.663                                    | 0.806                            | 0.945                                  |
| 2016 | 0.295                    | 0.067                             | 0.067                                    | 0.002                            | 0.001                                  |
|      | 0.768                    | 0.795                             | 0.795                                    | 0.968                            | 0.969                                  |
| 2017 | 0.276                    | 0.058                             | 0.049                                    | 0.009                            | 0.000                                  |
|      | 0.783                    | 0.809                             | 0.824                                    | 0.923                            | 0.983                                  |

### Table 15 Test results of spatial correlation of LQ2–LQ10

| Year | LQ_2 | LQ_3 | LQ_4 | LQ_5 | LQ_6 | LQ_7 | LQ_8 | LQ_9 | LQ_10 |
|------|------|------|------|------|------|------|------|------|-------|
| 2003 | 1.673| 0.877| 5.814| 1.298| 3.443| 2.453| 1.779| 0.731| 1.203 |
|      | 0.094| 0.620| 0.000| 0.194| 0.001| 0.014| 0.075| 0.465| 0.229 |
| 2004 | 2.296| 0.324| 6.472| 0.617| 1.564| 1.147| 0.684| 0.955| 2.104 |
|      | 0.022| 0.254| 0.000| 0.537| 0.118| 0.251| 0.494| 0.339| 0.035 |
| 2005 | 2.304| 0.235| 3.895| 1.149| 0.812| 0.812| 0.958| 1.578| 1.414 |
|      | 0.021| 0.186| 0.000| 0.251| 0.417| 0.417| 0.338| 0.115| 0.157 |
| 2006 | 2.090| 0.223| 4.608| 1.766| 0.953| 1.591| 3.420| 1.456| 2.304 |
|      | 0.037| 0.176| 0.000| 0.077| 0.340| 0.111| 0.001| 0.146| 0.021 |
| 2007 | 1.837| 0.562| 2.740| 2.777| 1.095| 1.025| 3.405| 0.877| 2.762 |
|      | 0.066| 0.574| 0.006| 0.005| 0.274| 0.305| 0.001| 0.380| 0.006 |
| 2008 | 3.946| 1.400| 3.314| 1.241| 1.173| 1.101| 3.405| 0.931| 2.658 |
|      | 0.000| 0.162| 0.001| 0.215| 0.241| 0.271| 0.001| 0.352| 0.008 |
Table 15  continued

| Year | LQ_2  | LQ_3  | LQ_4  | LQ_5  | LQ_6  | LQ_7  | LQ_8  | LQ_9  | LQ_10 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2009 | 3.178 | 0.596 | 3.134 | 0.426 | 1.241 | 0.588 | 4.575 | 1.609 | 2.416 |
| 2010 | 4.175 | 0.235 | 4.405 | 1.450 | 1.422 | 0.306 | 2.214 | 0.638 | 2.047 |
| 2011 | 3.687 | 0.827 | 4.135 | 0.743 | 0.895 | 0.542 | 0.714 | 0.684 | 1.483 |
| 2012 | 3.885 | 0.984 | 2.695 | −0.099 | 1.401 | 0.594 | 1.475 | 0.896 | 2.007 |
| 2013 | 2.779 | 0.660 | 2.177 | 2.390 | 1.834 | 0.672 | 1.936 | 0.681 | 2.159 |
| 2014 | 2.281 | 0.230 | 2.988 | 2.441 | 1.040 | 0.326 | 4.533 | 0.503 | 2.402 |
| 2015 | 1.607 | 1.040 | 3.391 | 2.842 | 1.903 | 0.490 | 3.481 | 1.000 | 2.288 |
| 2016 | 1.923 | 1.261 | 4.473 | 3.335 | 2.221 | 0.279 | 3.346 | 0.534 | 1.997 |
| 2017 | 1.935 | 1.309 | 3.982 | 4.331 | 1.537 | 1.196 | 0.751 | 0.941 | 1.037 |

Table 16  Test results of spatial correlation of LQ11–LQ19

| Year | LQ_11 | LQ_12 | LQ_13 | LQ_14 | LQ_15 | LQ_16 | LQ_17 | LQ_18 | LQ_19 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2003 | 0.001 | 0.001 | 1.614 | 5.205 | 0.818 | 2.098 | 4.191 | 0.212 | 2.029 |
| 2004 | 0.509 | 0.300 | 3.128 | 3.568 | 1.339 | 2.238 | 4.058 | 0.529 | 1.514 |
| 2005 | 0.611 | 0.764 | 0.002 | 0.000 | 0.181 | 0.025 | 0.000 | 0.597 | 0.130 |
| 2006 | 1.050 | 0.582 | 4.056 | 2.640 | 1.557 | 2.186 | 2.179 | 1.591 | 1.632 |
| 2007 | 0.294 | 0.439 | 0.000 | 0.008 | 0.199 | 0.029 | 0.029 | 0.112 | 0.103 |
| 2008 | 1.828 | 0.747 | 4.073 | 1.936 | 2.083 | 2.450 | 0.965 | 1.847 | 1.865 |
| 2009 | 0.068 | 0.545 | 0.000 | 0.053 | 0.037 | 0.014 | 0.334 | 0.065 | 0.062 |
| 2010 | 1.316 | 0.262 | 3.003 | 1.415 | 1.606 | 0.170 | 0.563 | 1.783 | 1.140 |
| 2011 | 0.188 | 0.207 | 0.003 | 0.157 | 0.108 | 0.865 | 0.573 | 0.075 | 0.254 |
| 2012 | 1.151 | 0.208 | 2.842 | 2.384 | 1.115 | 1.568 | 0.291 | 1.925 | 0.996 |
| 2013 | 0.250 | 0.165 | 0.004 | 0.017 | 0.265 | 0.117 | 0.771 | 0.054 | 0.319 |
| 2014 | 1.070 | 0.524 | 2.033 | 0.370 | 1.700 | 0.541 | 0.149 | 1.223 | 1.294 |
| 2015 | 0.285 | 0.399 | 0.042 | 0.712 | 0.089 | 0.588 | 0.118 | 0.221 | 0.196 |
| 2016 | 1.869 | 0.181 | 3.039 | 0.816 | 0.973 | 2.861 | 1.645 | 0.222 | 1.249 |
| 2017 | 0.062 | 0.856 | 0.002 | 0.414 | 0.331 | 0.004 | 0.100 | 0.176 | 0.212 |
| 2018 | 0.078 | 0.171 | 4.988 | 1.538 | 1.752 | 1.734 | 0.921 | 0.374 | 1.223 |
| 2019 | 0.938 | 0.864 | 0.000 | 0.124 | 0.080 | 0.083 | 0.357 | 0.708 | 0.221 |
| 2020 | 1.550 | 1.870 | 5.417 | 0.405 | 1.053 | 1.556 | 2.338 | 1.786 | 2.482 |
| 2021 | 0.121 | 0.939 | 0.000 | 0.315 | 0.293 | 0.120 | 0.019 | 0.074 | 0.013 |
| 2022 | 2.888 | 1.614 | 5.835 | 0.345 | 0.066 | 1.312 | 3.114 | 2.216 | 2.529 |
| 2023 | 0.004 | 0.893 | 0.000 | 0.270 | 0.995 | 0.189 | 0.002 | 0.027 | 0.011 |
| 2024 | 2.858 | 1.387 | 5.831 | 0.862 | 1.597 | 0.692 | 2.412 | 3.207 | 2.996 |
| 2025 | 0.004 | 0.834 | 0.000 | 0.611 | 0.110 | 0.489 | 0.016 | 0.001 | 0.003 |
Table 16 continued

|      | LQ_11 | LQ_12 | LQ_13 | LQ_14 | LQ_15 | LQ_16 | LQ_17 | LQ_18 | LQ_19 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 2016 | 3.100 | 0.360 | 5.678 | 0.858 | 1.573 | 1.186 | 2.767 | 3.827 | 3.355 |
|      | 0.002 | 0.719 | 0.000 | 0.609 | 0.116 | 0.236 | 0.006 | 0.000 | 0.001 |
| 2017 | 3.178 | 0.714 | 3.899 | 1.004 | 2.310 | 1.992 | 3.306 | 4.529 | 3.439 |
|      | 0.001 | 0.475 | 0.000 | 0.315 | 0.021 | 0.046 | 0.001 | 0.000 | 0.001 |

Author Contributions Conceptualization, MH and JX; methodology, MH; software, MH; validation, MH and JX; formal analysis, MH and JX; investigation, MH and JX; resources, MH; data curation, MH and JX; writing—review and editing, JX and MH; visualization, MH; supervision, JX. All authors have read and agreed to the published version of the manuscript.

Data Availability The data were collected from the China Statistical Yearbook and the China City Statistical Yearbook.

References

1. Spiekermann K, Wegener M (1994) The shrinking continent: new time–space maps of Europe. Environ Plann B Plann Des 21(6):653–673. https://doi.org/10.1068/b210653
2. Niu F, Xin Z, Sun D (2021) Urban land use effects of high-speed railway network in China: a spatial spillover perspective. Land Use Policy 105. https://doi.org/10.1016/j.landusepol.2021.105417
3. Donaldson D (2010) Railroads of the raj: estimating the impact of transportation infrastructure. Lse Res Online Doc Econ 32(2):16487. https://doi.org/10.1257/aer.20101199
4. Alexander K (2015) Improving transport infrastructure in Russia. OECD Econ Dep Work Pap 1193:1–22. https://doi.org/10.1787/5js4hmcs3mxp-en
5. Fageda X, Gonzalez-Aregall M (2017) Do all transport modes impact on industrial employment? Empirical evidence from the Spanish regions. Transp Policy 55:70–78. https://doi.org/10.1016/j.tranpol.2016.12.008
6. Chandra A, Thompson E (2000) Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. Reg Sci Urban Econ 30(4):457–490. https://doi.org/10.1016/S0166-0462(00)00040-5
7. Holl A (2004) Transport infrastructure, agglomeration economies, and firm birth: empirical evidence from Portugal. J Reg Sci 44(4):693–712. https://doi.org/10.1111/j.0022-4146.2004.00354.x
8. Aschauer DA (1989) Is public expenditure productive? J Monet Econ 23(2):177–200. https://doi.org/10.1016/0304-3932(89)90047-0
9. Chen CL, Hall P (2011) The impacts of high-speed trains on British economic geography: a study of the UK’s intercity 125/225 and its effects. J Transp Geogr 19(4):689–704. https://doi.org/10.1016/j.trangeo.2010.08.010
10. Levinson DM (2012) Accessibility impacts of high-speed rail. J Transp Geogr 22:288–291. https://doi.org/10.1016/j.trangeo.2012.01.029
11. Chen G, Silva J (2014) Estimating the provincial economic impacts of high-speed rail in Spain: an application of structural equation modeling. Soc Behav Sci 111:157–165. https://doi.org/10.1016/j.sbspro.2014.01.048
12. Ahlfeldt GM, Feddersen A (2017) From periphery to core: measuring agglomeration effects using high-speed rail. Serc Discuss Pap. https://doi.org/10.1093/jeg/lbx005
13. Chen Z, Haynes KE (2017) Impact of high-speed rail on regional economic disparity in China. J Transp Geogr 65:80–91. https://doi.org/10.1016/j.jtrangeo.2017.08.003
14. Guirao B, Lara-Galera A, Campa JL (2017) High speed rail commuting impacts on labour migration: the case of the concentration of metropolis in the Madrid functional area. Land Use Policy Int J Cover All Aspects Land Use 66:131–140. https://doi.org/10.1016/j.landusepol.2017.04.035
15. Guirao B, Campa JL, Casado-Sanz N (2018) Labour mobility between cities and metropolitan integration: the role of high-speed rail commuting in Spain. Cities 78:140–154. https://doi.org/10.1016/j.cities.2018.02.008
16. Cascetta E, Carteni A, Ih A, Henke I (2020) Economic growth, transport accessibility and regional equity impacts of high-speed railways in Italy: ten years ex post evaluation and future perspectives. Transp Res Part A; Policy Pract 139:419–428. https://doi.org/10.1016/j.tra.2020.07.008
17. Shao S, Tian Z, Yang L (2017) High speed rail and urban service industry agglomeration: evidence from China’s Yangtze River Delta region. J Transp Geogr 64:174–183. https://doi.org/10.1016/j.jtrangeo.2017.08.019
18. Dong X (2018) High-speed railway and urban sectoral employment in China. Transp Res Part A Policy Pract 116:603–621. https://doi.org/10.1016/j.tra.2018.07.010
19. Meng X, Lin S, Zhu X (2018) The resource redistribution effect of high-speed rail stations on the economic growth of neighbouring regions: evidence from China. Transp Policy 68:178–191. https://doi.org/10.1016/j.tranpol.2018.05.006
20. Wang F, Wei X, Liu J, He L, Gao M (2019) Impact of high-speed rail on population mobility and urbanisation: a case study on Yangtze river delta urban agglomeration, China. Transp Res Part A; Policy Pract 127(SEP.):99–114. https://doi.org/10.1016/j.tra.2019.06.018
21. Zheng L, Long F, Chang Z, Ye J (2019) Ghost town or city of hope? the spatial spillover effects of high-speed railway stations in China. Transp Pol 81(SEP.):230–241. https://doi.org/10.1016/j.tranpol.2019.07.005
22. Pan H, Cong C, Zhang X (2020) How do high-speed rail projects affect the agglomeration in cities and regions? Transp Res Part D; Transp Environ 88:102561. https://doi.org/10.1016/j.trd.2020.102561
23. Wang S, Wang J, Liu X (2019) How do urban spatial structures evolution in the high-speed railway era? Case study of Yangtze River Delta, China. Habitat Int 93:102051. https://doi.org/10.1016/j.habitatint.2019.102051
24. Liu Y, Zhang X, Pan X, Ma X, Tang M (2020) The spatial integration and coordinated industrial development of urban agglomerations in the Yangtze River economic belt China. Cities 104. https://doi.org/10.1016/j.cities.2020.102801
25. Liang Y, Zhou K, Li X, Zhou Z, Sun W, Zeng J (2020) Effectiveness of high-speed railway on regional economic growth for less developed areas. J Transp Geogr. https://doi.org/10.1016/j.jtrangeo.2019.102621

26. Li X, Wu Z, Zhao X (2020) Economic effect and its disparity of high-speed rail in China: a study of mechanism based on synthesis control method. Transp Policy 99:262–274. https://doi.org/10.1016/j.tranpol.2020.09.003

27. Li Y, Chen Z, Wang P (2020) Impact of high-speed rail on urban economic efficiency in China. Transp Policy 97:220–231. https://doi.org/10.1016/j.tranpol.2020.08.001

28. Wang C, Meng W, Hou X (2020) The impact of high-speed rail on urban economy: an investigation using night lighting data of Chinese cities. Res Transp Econ. https://doi.org/10.1016/j.retrec.2020.100819

29. Meng TA, Ti A, Xy B, Hz A, Xia MA (2021) The impact of high-speed rail on service industry agglomeration in peripheral cities. Transp Res Part D: Transp Environ. https://doi.org/10.1016/j.trd.2021.102745

30. Melo PC, Graham DJ, Brage-Ardao R (2013) The productivity of transport infrastructure investment: a meta-analysis of empirical evidence. Reg Sci Urban Econ 43(5):695–706. https://doi.org/10.1016/j.regsciurbeco.2013.05.002

31. Hansen WG (1959) How accessibility shapes land-use. J Am Inst Plann 25(1):73–76

32. Guy M, Ferdinand R, Redding SJ (2008) Urbanization and structural transformation. Quart J Econ 2:535–586. https://doi.org/10.1016/10.2139/qrsm.1281691

33. Pagliara F, Mauriello F, Garofalo A (2017) Exploring the interdependences between high-speed rail systems and tourism: some evidence from Italy. Transp Res Part A: Policy Pract 106:300–308. https://doi.org/10.1016/j.tra.2017.09.022

34. Duflo E, Pande R (2007) Dams. Q J Econ 122(2):601–646. https://doi.org/10.1162/qjec.122.2.601

35. Duranton G, Turner MA (2011) The fundamental law of road congestion: evidence from US cities. Am Econ Rev 101(6):2616–2652. https://doi.org/10.1257/aea.101.6.2616

36. Turner G (2012) Urban growth and transportation. Cepr Discuss Pap 79(4):1407–1440. https://doi.org/10.2307/23355076

37. Melitz MJ (2003) The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71(6):1695–1725. https://doi.org/10.1111/1468-0262.00467

38. Richard EB et al (2006) Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. J Econ Geogr 6(3):323–346. https://doi.org/10.1093/jegf/ibf020

39. Melitz MJ, Ottaviano G (2008) Market size, trade, and productivity. Rev Econ Stud 75(1):295–316. https://doi.org/10.1111/j.1467-937x.2007.00463.x

40. Pace RK, Lesage JP (2009) A sampling approach to estimate the log determinant used in spatial likelihood problems. J Geogr Syst 11(3):209–225. https://doi.org/10.1007/s10109-009-0087-7

41. Debrezion G, Pels E, Rietveld P (2007) The impact of railway stations on residential and commercial property value: a meta-analysis. J Real Estate Financ Econ 35(2):161–180. https://doi.org/10.1007/s11146-007-9032-z

42. He SY (2020) Regional impact of rail network accessibility on residential property price: modelling spatial heterogeneous capitalisation effects in Hong Kong. Transp Res Part A Policy Pract. https://doi.org/10.1016/j.jtra.2020.01.025

43. Huang Z, Du X (2020) How does high-speed rail affect land value? Evidence from China. Land Use Policy 101. https://doi.org/10.1016/j.landusepol.2020.105068

44. Wang L, Yuan F, Duan X (2018) How high-speed rail service development influenced commercial land market dynamics: a case study of Jiangsu Province, China. J Transp Geogr 72:248–257. https://doi.org/10.1016/j.jtrangeo.2018.09.010