Heliyon 9 (2023) e21088

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Analysis of the impact of high-speed rail on the spatio-temporal distribution of residential population and industrial structure

Xuanxuan Xia [a], Hongchang Li [a,*], Kun Wang [b], Yixian Liu [a]

[a] School of Economics and Management, Beijing Jiaotong University, Beijing, 100044, China
[b] Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, 100872, China

ARTICLE INFO

Keywords:
High-speed rail
Spatio-temporal distribution
Resident population
Industrial structure

ABSTRACT

With the development of high-speed rail, the spatiotemporal distance among regions has been shortened. However, the research on the spatiotemporal distribution and correlation of resident population and industrial structure caused by the opening of high-speed railway is relatively few. This study is aimed to explore the impact of high-speed rail on resident population and industrial structure, as well as the spatiotemporal agglomeration and evolution trend. The spatial-temporal distribution of the resident population, industrial structure, and the influence of high-speed rail was studied using panel data from 31 Chinese provinces from 2006 to 2018 using spatial autocorrelation analysis, standard deviation ellipse method, and DID model. The experimental results indicate that: (1) The resident population shows a significant positive spatial autocorrelation, while the industrial structure is vice versa. (2) The resident population and industrial structure are primarily centralized in the eastern and central areas, showing a "northeast-southwest" spatial distribution pattern. (3) High-speed rail and the resident population are negatively correlated, and there is heterogeneity. Although there is an association between high-speed rail and industrial structure that is positive, there is also a clear regional variability. This study contributes to providing the corresponding theoretical support and basis for the high-speed rail line planning in different regions and the relevant departments to formulate effective economic policies.

1. Introduction

Transportation infrastructure, as an important carrier for the spatial flow of labor, resources, and other production elements [1], plays a critical role in fostering local economic growth. The construction of transportation infrastructure not only changes the distribution and agglomeration pattern of the regional labor force but also affects the development of different industrial structures (hereafter IS) such as land value and real estate market. Transportation promotes the development of agglomeration economies in different regions [2]. China's urbanization level has continuously improved over the past few years, and impressive progress has been made in both new urbanization and rural-urban integration. China had a 914.25 million urban resident population (hereinafter RP) at the end of 2021, with a rate of urbanization of 64.72 %. China's industrial structure is also constantly optimizing and upgrading. According to relevant data, in 2021, 53.3 % of China's GDP consists of the added value of the tertiary industry and 39.4 % of the added value of the secondary industry. High-speed rail (hereafter HSR) is a modern kind of transportation facility that has altered the traditional concept of time and space and reduced production and transaction costs with the characteristics of convenience, efficiency,

* Corresponding author.
E-mail address: hchli@bjtu.edu.cn (H. Li).

https://doi.org/10.1016/j.heliyon.2023.e21088
Received 16 May 2023; Received in revised form 9 October 2023; Accepted 15 October 2023
Available online 17 October 2023

and punctuality. The way of regional economic growth in China is significantly affected by this [3,4]. Under the influence of the agglomeration effect and siphon effect, the impact of the opening of HSR on RP and IS has a certain spatial and temporal heterogeneity in different regions. The Outline of China's 14th Five-Year Plan has cut forward relevant policies such as accelerating the development of a potent transportation nation and essentially speeding up the "eight vertical and eight horizontal" HSR. China's operational railway network had a total length of 150,700 km as of the end of 2021, with HSRs accounting for more than 40,000 km, or more than two-thirds of the global total. Based on China's mid-long-term railway network plan, the long-term railway network is expected to be about 200,000 km by 2030, with around 45,000 km of HSR. The development of the HSR network promotes the flow of inter-city socio-economic activities and resource elements [2], which has an important impact on population flow and the upgrading of IS. It is of great guiding significance to grasp and explore the effect of the opening of HSR on RP and IS and the aggregation in time and space to promote the balanced development of the regional economy.

Along with the ongoing development and extension of HSR, which not only improve the overall transport capacity of the main line but also improved the accessibility of regional transportation, promote element flow and concentration by compression of time and space, greatly influenced the urban labor and service markets, as well as space structure, IS, inject new vitality for economic growth. However, as the regional economy continues to grow, a series of urban development problems also arise, such as the relatively low level of urbanization, the uncoordinated growth of the urban RP [5], the unreasonable IS and the low level of comprehensive transportation infrastructure. According to the National New Urbanization Plan (2014–2020), the urban agglomeration's ability to gather economy and population should be enhanced, coordinated regional development should be promoted, and the supporting role of comprehensive transport networks should be strengthened. It's important to coordinate the growth of the industrial economy while also absorbing the new urban RP, regulating population size, realizing the positive interaction between IS and population structure, fostering regional economic integration, and accelerating the urbanization process. Therefore, how will the HSR's construction and operation impact urban RP and IS? The impact of the HSR's opening on the RP and IS in different regions is spatial heterogeneity is of key significance for promoting the growth of urbanization. For logical HSR construction planning, IS optimization, and urbanization promotion, a thorough and profound understanding of the interaction between the construction of HSR and urban RP and IS is crucial.

This paper includes the following parts. Section 2 reviews the literature on HSR and urban economic development. Section 3 introduces the experimental data and models (including the spatial autocorrelation model, standard deviation ellipse model, and PSM-DID model). Making the empirical analysis is section 4. The main findings and policy recommendations are presented in section 5 as a conclusion.

2. Literature review

HSR is a transport mode that can quickly cross regions, shorten distances in both space and time, and vastly enhance the accessibility and connection of regional cities [6]. The rapid expansion of the HSR network in China has a significant impact on the spatial structure of urban agglomerations, regional integration, economic and social development, and other sides [7,8]. Researchers have investigated the effects of HSR opening on the degree of economic growth, regional spatial pattern, and industrial aggregation using a variety of study approaches based on different research perspectives.

2.1. Effect of HSR on the degree of economic development

Scholars continue to differ on the research of HSR on local economic growth as a result of the various research objects and research methods. Both the impacts of HSR on balanced regional economic growth and the impacts of HSR on unbalanced regional economic growth are covered in the current study on the impacts of HSR on regional economic growth.

Most academic studies show that the development of HSR not only improves regional traffic conditions but also optimizes the allocation efficiency of labor, capital, and natural resources between regions, which is strongly tied to economic growth, population flow, and tourism [9,10]. Using the DID model, Fang et al. (2016) analyzed how the opening of HSR will impact regional economic growth and discovered that it may greatly accelerate per-capita GDP growth and support the economic expansion of the Yangtze River Delta urban agglomeration [11]. Jiang et al. (2017) studied China's 2012 input-output table and found that every 100 million yuan investment in HSR construction boosted total output by 372 million yuan, GDP increased by 121 million yuan and created 1084 jobs, which could inject strong impetus to the national economy in a short period. In addition, the operation of HSR has largely facilitated the development of the residential property market in cities along the route [12]. Liu et al. (2018) discovered that the opening of HSR has a considerable beneficial effect on house prices in the cities along the route, and these increases are greater in the cities that are located in the center of the urban economic groups [13].

Against the background of China's economic slowdown, Wang and Nian (2014) questioned whether the HSR had driven regional economic growth, and believed that the operation of the HSR could not drive local economic growth in a short period [14]. Some scholars have found that the impacts of HSR on local economic development vary, revealing the phenomena of imbalanced development during the construction of HSR networks [15,16]. HSR widens the gap between regions by quickening the differentiation and movement of elements. In particular, western and northeast regions with relatively poor economic foundations are likely to be marginalized, and prefectures close to regional central cities will be more negatively affected [17,18]. In addition, some scholars point out that the construction and opening of HSR and other infrastructure may cause central cities to over-absorb mobile factors from medium and small cities, which will have an adverse spillover influence on the economy of medium and small cities. From the perspective of how HSR affects the population, Lin et al. (2015) conducted a talent attraction test on the station cities on the Wuhan-Guangzhou HSR and found that the non-station cities had weak talent attraction and the possibility of brain drain [19].

Similarly, Huang and Han (2021) believed that HSR would have a siphon effect on the population of medium-sized cities and enhance the population adsorption capacity of megacities [20]. Zhang and Chen (2022) confirmed the existence of agglomeration shadow and found that after the operation of HSR, the per capita GDP of the county decreased by 2.6 %, but considering the change of RP, this negative effect may be weakened [21].

2.2. *Effect of HSR on regional spatial pattern*

As the product attributes of HSR itself have obvious spatiotemporal compression characteristics, mainly reflected in improving regional urban accessibility and connectivity, breaking the original regional spatial pattern. Scholars often use spatial models to examine how HSR affects urbanization. It mainly includes the mixed spatial correlation logit model [22,23], spatial random utility model [24,25], generalized spatial correlation logit model [26], dynamic spatial multinomial probability model [27], and p-space model [28,29]. In addition, common econometric models in the existing literature include the social network analysis method [30,31], coefficient analysis [32], integrated evaluation method [33], and different-in-different model [34,35].

The construction and growth of the regional spatial framework depend greatly on HSR [36]. Dong proposed that the construction and operation of HSR will create opportunities for economic development in the station cities, such as Japan's Urasa-town and France's Lille, which have both become a new generation of important cities with the rapid development of characteristic industries. In addition, it can somewhat lower production costs for businesses in towns along the route, change the original industrial agglomeration's spatial structure, and eventually create a new growth pole [9]. The HSR network also connects the center metropolis with nearby satellite cities, promoting the formation of metropolitan areas [13]. In the discussion of the effect of HSR on the evolution of the megalopolis structure in the Yangtze River Delta, Wang et al. (2019) found that following the HSR's opening, 75 % of the urban areas linked to the HSR network began to shift toward HSR stations [35]. The HSR project contributes to increasing the density of spatial connectivity inside the urban agglomeration and stabilizing its spatial framework, according to Zhang et al. (2020), who used the Beijing-Tianjin-Hebei urban agglomeration as their research subject [37].

However, some scholars have found that the extension of HSR lines may lead to the polarization of economic spatial distribution, and economic resources are concentrated in hub cities. Cities with HSR stations will benefit more from the significant improvement in accessibility, while cities without HSR will have an "island effect". Monzón et al. (2013) proposed the influence of HSR expansion on efficiency and spatial equity [38], and Xu and Song (2019) believed that the opening and operation of the HSR somewhat weakened the economic linkages between cities of the same level [39].

2.3. *Influence mechanism of HSR on urban population-industry*

The RP and IS are the micro-expression subjects of socio-economic activities. During the urbanization process, the agglomeration and diffusion degree of population and industry will directly affect the urbanization level [40,41]. On the one hand, the construction and operation of HSR produce a spatial-temporal compression effect, enhance the effectiveness of labor and other resource allocation among regions, and encourage population mobility between areas. The high-frequency migration of the population will inevitably lead to the change of urban RP, and the human capital agglomeration it brings about helps raise the degree of urbanization [42]. On the other hand, the continuous improvement of the HSR network enhances the independent research and development ability of the HSR industry and accordingly drives the steel, locomotive manufacturing, and other related industries, which is conducive to the growth of the tertiary industry and the secondary industry At the same time, it provides opportunities for the transformation of traditional industries to emerging industries along the route and promotes the expansion of industrial scale and optimization of IS to support the urbanization's qualitative development [41]. Therefore, under the background of vigorously developing HSR network construction, an extensive investigation of the influence of HSR on the RP and IS has certain guiding significance for rational planning of HSR construction, adjusting the economic structure, and supporting the high-quality development of urbanization.

Previous literature has focused on the impact of HSR on the level of regional economic development and spatial patterns. Although previous studies have achieved good research results, there are still a series of problems that need to be further solved and thought about. However, relatively few studies have analyzed the spatial aggregation and correlation of the opening of HSR on RP and IS under the spatiotemporal dimension. HSR construction from line to network, forming a criss-cross HSR network, urban development is impacted by both the construction of an HSR line and additional HSR lines that travel through the region If only a certain HSR line is used as the research object, the conclusion may be biased. Firstly, we quantitatively study the agglomeration and diffusion degree of RP and IS in different regions using spatial autocorrelation analysis and standard deviation ellipse method, and visualize the spatial agglomeration form and evolution distribution of population and IS before and after the opening of HSR. Secondly, the PSM-DID model was utilized to verify the influence of HSR construction on RP and IS, and to investigate the spatial heterogeneity of the influence from the holistic and local aspects. According to the results, on the one hand, the relevant theories of HSR on urban economic development factors are enriched. On the other hand, relevant departments can make corresponding decisions on the overall construction of the HSR network, rationally plan the spatial layout of HSR lines, and provide theoretical support for the sustainable development of the regional economy.

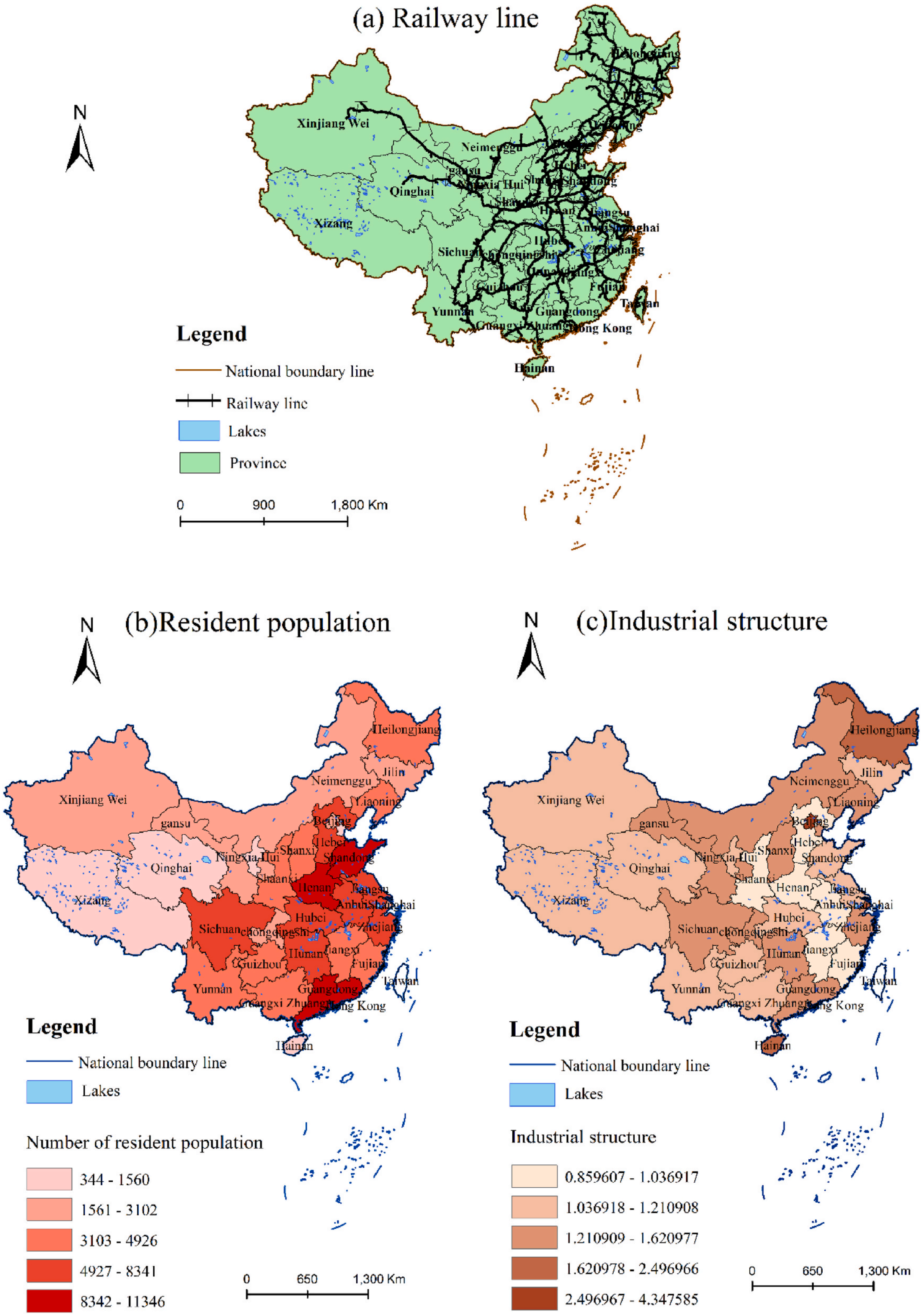

Fig. 1. The spatial distribution of railway, RP, and IS.

3. Data and methodology

3.1. Study field

By May 2018, HSR lines had been operational in 28 provincial capitals, over 180 prefecture-level cities, and more than 370 county-level cities. Taking into account various types of stations, there were more than 700 HSR stations nationwide, basically covering major population centers. According to Fig. 1, the research area of this paper covers 31 Chinese provinces, which excludes Hong Kong and Macao.

Fig. 1 (a) shows the railway distribution in China. We can see that the railway distribution in China is relatively perfect, but the HSR in China is relatively lacking. Fig. 1 (b) shows the distribution of the RP in different Chinese provinces at the close of 2018. The number of RP in eastern and central China is larger than elsewhere. At the end of 2018, the distribution of IS in different Chinese provinces is depicted in Fig. 1(c). The IS is measured in this study using the ratio of the tertiary industry's value to the secondary industry's value. It can be seen that the IS in various provinces presents an obvious regional imbalance.

3.2. Data source and variable determination

In this study, annual data from 2006 to 2018 for 31 Chinese provinces and cities are used as study samples. The regional economic indicator data, traffic data, RP, and IS indicator data are all from the EPS data platform, and the time when cities open HSR is obtained by consulting network news and sorting out. Table 1 describes each variable.

3.3. Methodology

3.3.1. Spatial autocorrelation model

1. Model of global spatial autocorrelation

The basic principle of spatial autocorrelation is the correlation between different locations in geographic space. The most used spatial autocorrelation index is *Moran's I*, which can determine the spatial correlation within a certain range. This method can effectively understand the similarities and differences between different locations, to better explain the spatial heterogeneity of the impact of the HSR opening on the RP and IS in different regions from the spatial dimension. The spatial correlation and degree of aggregation of the RP and the IS before and after the opening of the HSR are analyzed using *Moran's I*. The specific calculation is shown in formula (1):

$$I^P = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j}(P_i - \overline{P})(P_j - \overline{P})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j} \sum_{i=1}^{n} (P_i - \overline{P})^2} \tag{1}$$

Here, the number of study areas is n, that is, the total number of provinces ($n = 31$); P_i, P_j are the value of RP of provinces i and j ($i = 1, 2, \ldots, 31, i \neq j$). w_{ij} is the distance weight between provinces i and j, and $\overline{P}$ is the mean RP value. If $0 < I^P \leq 1$, it will indicate that the regions' RP are positively correlated; if $-1 \leq I^P < 0$, it will indicate a negative correlation; if $I^P = 0$, it will indicate there is no spatial correlation. Similarly, the *Moran's I* index of IS is calculated as formula (2).

Table 1
Variable description of DID model.

Variable	Indicator	Indicator Description	Unit
HSR	High-speed rail	If there is HSR, it is 1; otherwise, it is 0	/
GRP	Gross regional production	Sum of the added value of all industries	100 million yuan
TAX	tax revenue	Value of tax revenue by province	ten thousand yuan
DI	Disposable income	Discretionary income	yuan
RSCG	Total retail sales of consumer goods	The total retail sales of consumer goods by province	100 million yuan
RPV	Railway passenger traffic volume	The number of passengers transported in a given period of time	10 thousand people
RED	Total number of real estate development enterprises	Number of businesses engaged in real estate development in each province	number
RP	Total population at year-end	Resident population	10 thousand people
IS	Industrial structure	Tertiary industry output value/secondary industry output value	/

$$I^S = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(S_i - \overline{S})(S_j - \overline{S})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}\sum_{i=1}^{n}(S_i - \overline{S})^2} \tag{2}$$

in which the value of IS of provinces i and j ($i = 1, 2, \ldots, 31, i \neq j$) are S_i, S_j, severally. $\overline{S}$ is the average value of IS.

The *Moran's I* index is standardized, to examine the significance of spatial autocorrelation between various regions, and formula (3) is as follows:

$$Z^P = \frac{I - E(I^P)}{\sqrt{E\left[(I^P)^2\right] - E(I^P)^2}}, \quad Z^S = \frac{I - E(I^S)}{\sqrt{E\left[(I^S)^2\right] - E(I^S)^2}} \tag{3}$$

Among them, $E(I^P)$ and $E(I^S)$ are the theoretical expectations of I^P and I^S, respectively. If $Z > 0$, then there means positive spatial correlation. And the spatial association is more pronounced the higher the value. If $Z < 0$ represents a negative spatial correlation, and the smaller its value, the greater the spatial difference, otherwise, $Z = 0$ denotes that the value is spatially stochastic.

3.4. Model of local spatial autocorrelation

Local spatial autocorrelation analysis can effectively detect spatial differences caused by spatial correlation and make up for the deficiency of global spatial autocorrelation. Local spatial autocorrelation is a regional component of global spatial autocorrelation, which examines the degree of spatial clustering of similar attribute values in a certain area. To better understand the spatial aggregation state of RP and IS in different regions before and after the opening of HSR, we used the local autocorrelation model to measure the spatial aggregation degree of RP and IS. The distribution of the high and low values of the RP and IS in each province before and after the operation of the HSR is determined by computing the Getis-Ord Gi* index. The calculation is shown in formula (4).

$$G^{P^*}_{\ i} = \frac{\sum_{j=1}^{n} w_{ij}P_j}{\sum_{i=1}^{n} P_i}, \quad G^{S^*}_{\ i} = \frac{\sum_{j=1}^{n} w_{ij}S_j}{\sum_{i=1}^{n} S_i} \tag{4}$$

Similarly, the $G^{P^*}_{\ i}$ and $G^{S^*}_{\ i}$ indices are standardized to obtain the statistics $Z^P(i)^*$ and $Z^S(i)^*$ as shown in formula (5):

$$Z^P(i)^* = \frac{G^{P^*}_{\ i} - E\left[G^{P^*}_{\ i}\right]}{\sqrt{VAR\left[G^{P^*}_{\ i}\right]}}, \quad Z^S(i)^* = \frac{G^{S^*}_{\ i} - E\left[G^{S^*}_{\ i}\right]}{\sqrt{VAR\left[G^{S^*}_{\ i}\right]}} \tag{5}$$

when the value of $Z^P(i)^*$ and $Z^S(i)^*$ is positive and meaningful, shows that there is a high-value spatial agglomeration of RP and IS around region i are the expectation and variance of G_i^*, respectively. When the value of $Z^P(i)^*$ and $Z^S(i)^*$ is significant and negative, shows that there is a low-value spatial agglomeration of RP and IS around region *i*.

3.4.1. Method of pattern evolution analysis

The distribution trend and direction of a set of data are revealed by the standard deviational ellipse approach. The steps mainly include: determining the center of the circle, determining the orientation of the ellipse, and determining the length of the XY axis [43]. The standard deviation ellipse model is a common GIS tool. In this paper, the model describes the spatial distribution characteristics of RP and IS variables and summarizes the spatial dispersion and direction of the two elements. Using this method, we study the effects of the opening of HSR on the spatial evolution trend of RP and IS, and understand the heterogeneity of the effects of HSR on RP and IS. The standard ellipse is calculated as follows:

Step 1. Determine the circle's center. The arithmetic mean center is directly used to calculate the center of the ellipse. formula (6) is as follows:

$$SDE_x = \sqrt{\frac{\sum_{i=1}^{k}(x_i - \overline{X})^2}{k}}, \quad SDE_y = \sqrt{\frac{\sum_{i=1}^{k}(y_i - \overline{Y})^2}{k}} \tag{6}$$

where, x_i, y_i are each element's spatial location coordinates, $\overline{X}, \overline{Y}$ are the arithmetic average center, and k is the number of elements at all points.

Step 2. Determine the orientation of the ellipse. Assuming that the X axis serves as the standard, due north (i.e., the direction indicated by 12 o'clock) is 0°, clockwise rotation, formula (7) of computation is shown below:

$$\tan\theta = \frac{A+B}{C} \tag{7}$$

Here, $A = \left(\sum_{i=1}^{k}\overline{\overline{x}}_i^2 - \sum_{i=1}^{k}\overline{\overline{y}}_i^2\right)$, $B = \sqrt{\left(\sum_{i=1}^{k}\overline{\overline{x}}_i^2 - \sum_{i=1}^{k}\overline{\overline{y}}_i^2\right)^2 + 4\left(\sum_{i=1}^{k}\overline{\overline{x}}_i\overline{\overline{y}}_i\right)^2}$, $C = 2\sum_{i=1}^{k}\overline{\overline{x}}_i\overline{\overline{y}}_i$. And, the differences between x and y coordinates and the average center are represented by $\overline{\overline{x}}_i, \overline{\overline{y}}_i$.

Step 3. Determine the XY axis' length. The following is the calculating formula (8):

$$\alpha_x = \sqrt{2}\sqrt{\frac{\sum_{i=1}^{k}\left(\overline{\overline{x}}_i\cos\theta - \overline{\overline{y}}_i\sin\theta\right)^2}{k}},\ \alpha_y = \sqrt{2}\sqrt{\frac{\sum_{i=1}^{k}\left(\overline{\overline{x}}_i\sin\theta + \overline{\overline{y}}_i\cos\theta\right)^2}{k}} \tag{8}$$

Data distribution direction is indicated by the ellipse's long half-axis, while its range is shown by its short half-axis. The direction of the data is clear the bigger the disparity between the values of major axis and minor axis, or the higher flatness. In contrast, when the values of major axis and minor axis go closer, the directionality decreases. If the two halves are the same, it's a circle, and a circle means it doesn't have any directional characteristics. The range of data distribution is shown by the short half axis. The centripetal force is more pronounced the shorter the short half axis is. Conversely, the short half axis is longer, the data will be more discrete. Likewise, the data without any distributional characteristics if the short and long axes are identical. The central position of the data is represented by the central point. Generally speaking, the position of the central point is roughly the same as that of the arithmetic mean as long as the fluctuation of the data is not particularly significant. One standard deviation ellipse surface, which includes the point element set comprising 68 %, is chosen as the output size [44].

3.4.2. PSM-DID model

The DID model (difference-in-differences) is also known as the difference-in-difference model or the multiple difference model. This model was first introduced into mathematical research of economics by Ashenfelter in 1978 [45]. The DID model can avoid the data endogeneity problem to a large extent, and then it is extensively used in various fields to quantitatively evaluate the effect of policy implementation [[11,13,46–49]].

This research views the operation of the HSR as a "natural experiment" and uses the DID method to compare regional economic development before and after the operation of HSR to more accurately assess the influence of HSR on regional economic development. Based on the time nodes of the operation of HSR in various Chinese provinces and the data computation findings in this article, we eventually choose whether the operation of HSR in 2010 to define the samples of HSR cities and non-HSR cities. Among them, the samples of HSR opened in 2010 account for 69.23 % of the total samples. We divided the cities into two groups: those with HSR from 2006 to 2010 and those without HSR from 2011 to 2018.

Step 4. Establish DID Model.

The following is how the DID model of the HSR's level of influence on RP is built:

$$Y_{it}^{P} = \beta_0 + \beta_1 HSR_{it} + \beta_2 X_{it} + \mu_i + \gamma_t + \varepsilon_{it} \tag{9}$$

Y^P is the dependent variable, representing the total population at year-end (RP), where *i* is the city, *t* is the year. The dummy variable HSR_{it} represents whether the province *i* has HSR in the *t* year, that is, the $HSR = 1$ of the year with HSR and the city, and the $HSR = 0$ of the year without HSR. Gross regional product, disposable income, tax revenue, total retail sales of social consumer goods, railway passenger volume, the total number of real estate development enterprises, and so on are all examples of the control variables that make up the term X_{it}. Furthermore, μ_i is the individual fixed effect controlling urban heterogeneity, γ_t is the time fixed effect controlling temporal trend, and ε_{it} is the coefficient vector.

Similarly, the DID model of HSR's impact on IS has the following structure:

$$Y_{it}^{S} = \beta_0 + \beta_1 HSR_{it} + \beta_2 X_{it} + \mu_i + \gamma_t + \varepsilon_{it} \tag{10}$$

Here, Y^S is the dependent variable, and represents the IS. IS is calculated using the value-added ratio of the secondary and tertiary industries [34].

Step 5. Establish the PSM-DID model.

PSM-DID is a statistical evaluation method that combines the difference-in-difference method and the propensity score matching method to investigate the effects of policy. Heckman first put up the idea in 1997 [50,51]. Because the propensity matching score method is combined with the model, compared with the basic differential model, the model takes more into account the sample matching problem, which can effectively eliminate the estimation bias caused by sample bias in DID model.

This paper employs the PSM-DID model to further test the results of the primary regression model after estimating the fundamental DID regression model. The PSM-DID fundamental model is shown in formula (11):

$$ATT = \frac{1}{N}\sum_{i\in(D=1)}\left[\Delta Y_{it} - \sum_{j=(D=0)} w(i,j)\Delta Y_{jt}\right] \tag{11}$$

Among them, ATT represents the treatment group's net effect of a certain policy upon implementation. ΔY_{it} is the difference Y between individuals i in the treatment group before and following the implementation of the policy. ΔY_{jt} is the difference Y between individuals j in the control group before and following the implementation of the policy. $w(i,j)$ stands for the weight of the propensity matching score and N is the number of samples in the treatment group.

4. Results

4.1. Overview

To better comprehend the HSR's effects on regional urbanization, we conducted statistics and visualizations based on the data on urban RP and IS in Chinese provinces from 2006 to 2018. Figs. 2 and 3 show the results. From the perspective of the RP (Fig. 2), we can find that the trend of change in RP in all provinces is consistent from 2006 to 2018. Among them, Guangdong has the highest RP, followed by Shandong. With the continuous opening of HSR lines in 2010, the RP of each province has not changed significantly, this shows that the HSR has little bearing on the expansion of the urban permanent population. The HSR's operation can promote the movement of numerous economic elements, provide more convenient transportation conditions for residents to travel, and relatively reduce the demand for settling in a certain area. This corresponds to the findings of scholars that the cities opened by high-speed rail tend to be concentrated in densely populated areas with developed economies [52].

Fig. 3 shows the change in IS in different provinces from 2006 to 2018. It can be found from Fig. 3 that the changing trend of IS in all provinces is consistent. Among them, Beijing's IS is at the highest level, followed by Hainan. With the continuous introduction of HSR in 2011, the value of IS in each region presents a rising trend, and the growth rate is significantly larger. This shows how the HSR's operation encourages the optimization and upgrading of IS in each area and thus promotes the development of regional urbanization.

4.2. Variables' descriptive statistics

In this paper, the RP and IS at the end of the year are taken as explanatory variables, regional factors, time factors, and HSR factors are taken as explanatory variables, and the total regional GDP, disposable income, tax revenue, total retail sales of social consumer goods, railway passenger volume and the total number of real estate development enterprises are taken as control variables. To eliminate heteroscedasticity, gross regional product, tax revenue, disposable income, total social retail goods, railway passenger volume, the amount of real estate development enterprises, and the year-end total population were all logarithms to make the regression result as accurate as possible. For the chosen variables, descriptive statistics are displayed in Table 2.

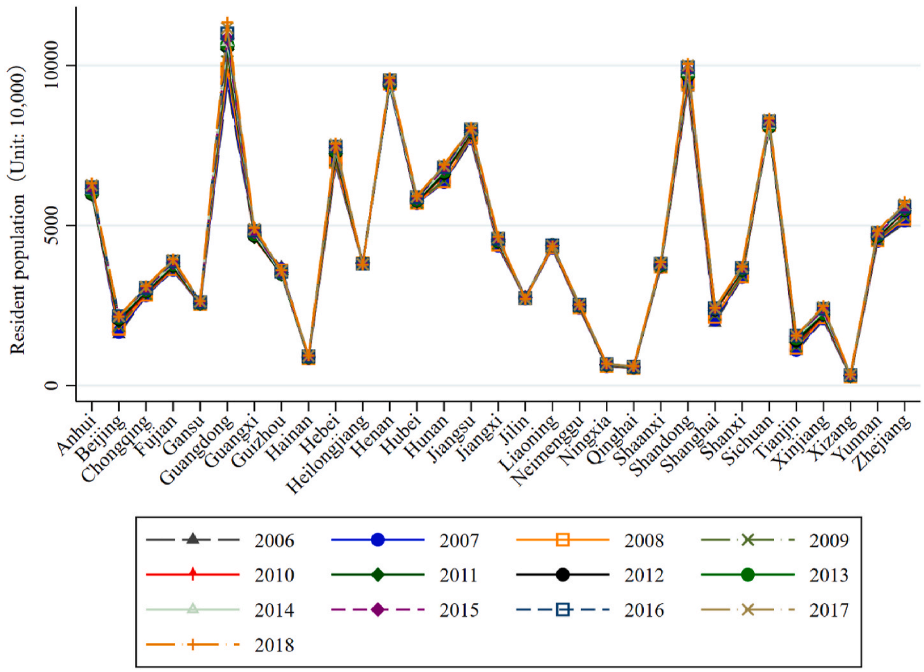

Fig. 2. The RP's spatiotemporal distribution from 2006 to 2018.

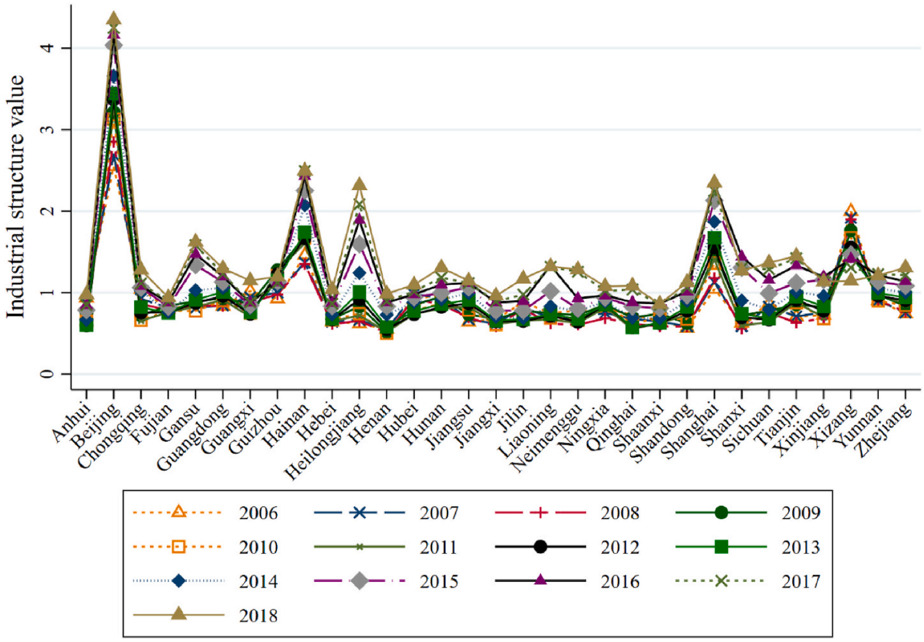

Fig. 3. The IS's spatiotemporal distribution from 2006 to 2018.

Table 2
Variables' descriptive statistics of DID model.

Variable	Description	Average value	Standard deviation	Minimum value	Maximum value	Observation value
HSR	High-speed railway	0.543	0.499	0	1	403
lnGRP	The natural logarithm of gross regional product	9.358	1.078	5.673	11.485	403
lnTAX	The natural logarithm of tax revenue	15.986	1.142	11.667	18.394	403
lnDI	The natural logarithm of disposable income	9.971	0.436	9.091	11.128	403
lnRSCG	The natural logarithm of total retail sales of consumer goods	8.333	1.176	4.496	10.584	403
lnRPV	The natural logarithm of railway ridership	8.398	1.192	3.135	10.438	403
lnRED	The natural logarithm of the total number of real estate development enterprises	7.618	0.995	3.367	9.073	403
lnRP	The natural logarithm of the year-end total population	8.103	0.850	5.652	9.337	403
IS	The industrial structure	1.046	0.576	0.500	4.348	403

4.3. Spatial correlation characteristics of the RP and IS

We utilize *Moran's I* to examine the spatial correlation of the RP and the IS in each province from the viewpoint of spatial correlation. According to formula (3), *Moran's I* of the RP and IS between 2006 and 2018 is obtained, as shown in Fig. 4.

As shown in Fig. 4, *Moran's I* of RP is greater than 0, so we can deduce that RP has a positive spatial correlation. In other words, a place with a large degree of spatial distribution and aggregation has a larger RP. In addition, the RP's *Moran's I* has been in a relatively stable state from 2006 to 2018, indicating that the correlation between the number of the RP and spatial distribution is in a relatively stable state over time. In addition, the IS's *Moran's I* is below 0, indicating that IS is negatively correlated in space. The negative spatial correlation indicates that the correlation becomes significant as the spatial distribution grows more dispersed.

According to formulas (4) and (5), the spatial agglomeration of the RP and IS is obtained, as shown in Figs. 5 and 6. From the perspective of the RP, it is clear from Fig. 5 that the RP's agglomeration status in China remained stable from 2006 to 2018, with a concentration mostly in the eastern and central areas, such as Zhejiang, Hubei, and Fujian. This indicates that both economic factors and traffic factors will have a certain impact on population agglomeration [41]. According to the specific situation analysis of different regions in China. First, economic factors. The eastern and central areas of China are economically developed, with the majority of economically developed cities, strong industrial system support, sound infrastructure and public services, and relatively sufficient employment opportunities. For example, China's GDP totaled 9.90865 trillion yuan in 2019, a 6.1 % increase from the year before. The eastern region contributes half of China's GDP, with a total GDP of 5.1116.12 trillion, or 51.9 % of the national total. Second, traffic. Because it is located in coastal areas and the Yangtze River basin, convenient transportation makes it easier to develop the economy, so more people will gather in the eastern region. Third, climate factors. The Middle East has a temperate climate, which makes it more

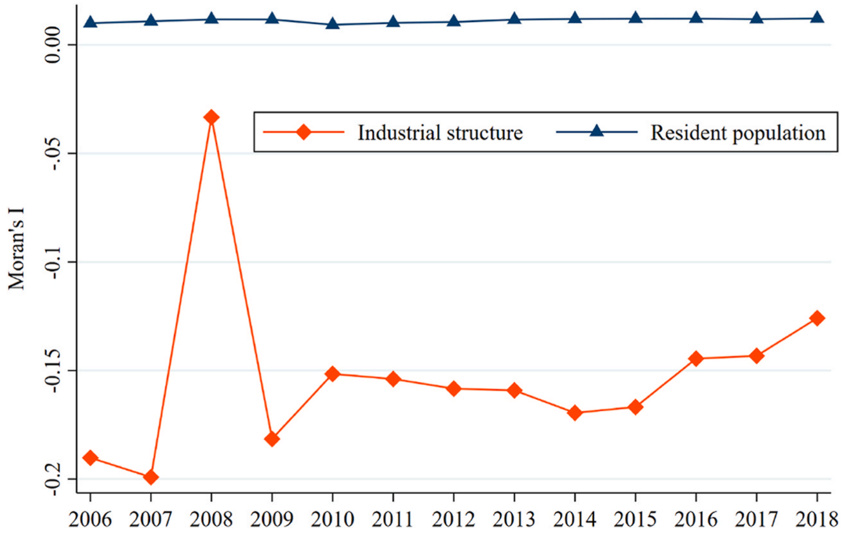

Fig. 4. Global autocorrelation of the RP and IS during the 2006–2018 period.

suitable for residents to live.

The spatial agglomeration of the IS has a clear regional imbalance, as seen in Fig. 6. In the eastern, central, and western zones, there is a significant discrepancy in the aggregation of high and low values. This suggests that economic balance between regions requires improved construction of transportation infrastructure such as high-speed rail, as well as policy guidance and value chain governance [53]. Firstly, before 2010, the IS of central and eastern, and northeast China shows a phenomenon of low-value aggregation and dispersion. In 2010, with the opening of HSR lines, the high-value aggregation areas of IS in central, eastern, and northeast China increase significantly. This is because the eastern and central areas are dominated by the secondary industry, that is, manufacturing and handicraft industries. As a result of the HSR's opening up, many regions' economies have developed, as have the transportation and other service sectors in the eastern and central areas, as well as the proportion of tertiary industry. Xinjiang, Xizang, Qinghai, and other western regions are mainly high-value aggregation areas before 2010, and vice versa after 2010. This is because, with the opening of HSR lines, the expansion of secondary sectors like mining and manufacturing in the western region has been driven by a drop in the ratio of IS, presenting a general area clustering state. Due to the relatively backward economy in western China, the primary industry is the principal industry, which in part demonstrates how the operation of the HSR has helped to optimize and upgrade the region's IS.

4.4. Study of the RP and IS's spatial pattern evolution

Based on the standard deviational ellipse model, we obtain the spatial evolution trend of the RP and IS, as shown in Fig. 7. First of all, from 2006 to 2018, the standard deviation ellipse of China's RP and IS is mainly situated in the central and eastern areas of China, showing a "northeastern-southwest" spatial distribution pattern. This indicates that the western and northeast regions will become the main body of the RP and IS upgrading. This is because, firstly, the population of eastern China is concentrated, accounting for 38.6 % of the national population. However, the population of the western area and the central area account for 27.2 % and 26.5 % respectively, which makes the RP present a spatial evolution trend from northeast to southwest, mainly covering the eastern and southern regions of China. Secondly, eastern and southern China have much greater levels of economic development and IS than central and western China. For example, in 2019, retail sales of consumer goods in eastern China accounted for about 50 %, while those in central and western China accounted for 24 % and 20.7 % of the national total. In terms of IS, the tertiary industry accounted for more than 50 % of GDP in the four economic regions. The proportion of the eastern region is the highest, accounting for 56.5 %, and that of the central region is the lowest, accounting for 50.0 %.

Considering the disparity between the spatial pattern distribution of RP and IS. With the opening of the HSR, there are obvious differences in the spatial evolution trend of the RP and IS. As can be shown from Fig. 7, the spatial evolution trend of the RP from 2006 to 2018 has no great difference and the change is small, which manifests that the spatial evolution distribution range and dispersion of the RP are relatively stable. However, from 2006 to 2018, the elliptic flatness of the IS gradually increased, indicating that the directional characteristics of the IS became more and more obvious, showing a trend of spatial divergence in the northeast-to-southwest direction and a trend of centripetal agglomeration in the northwest to southeast direction. The tertiary industry's degree of development in the northeast and west has been advanced by the ongoing opening of HSR lines. Therefore, with the evolution of IS from northeast to southwest, the IS is constantly improved.

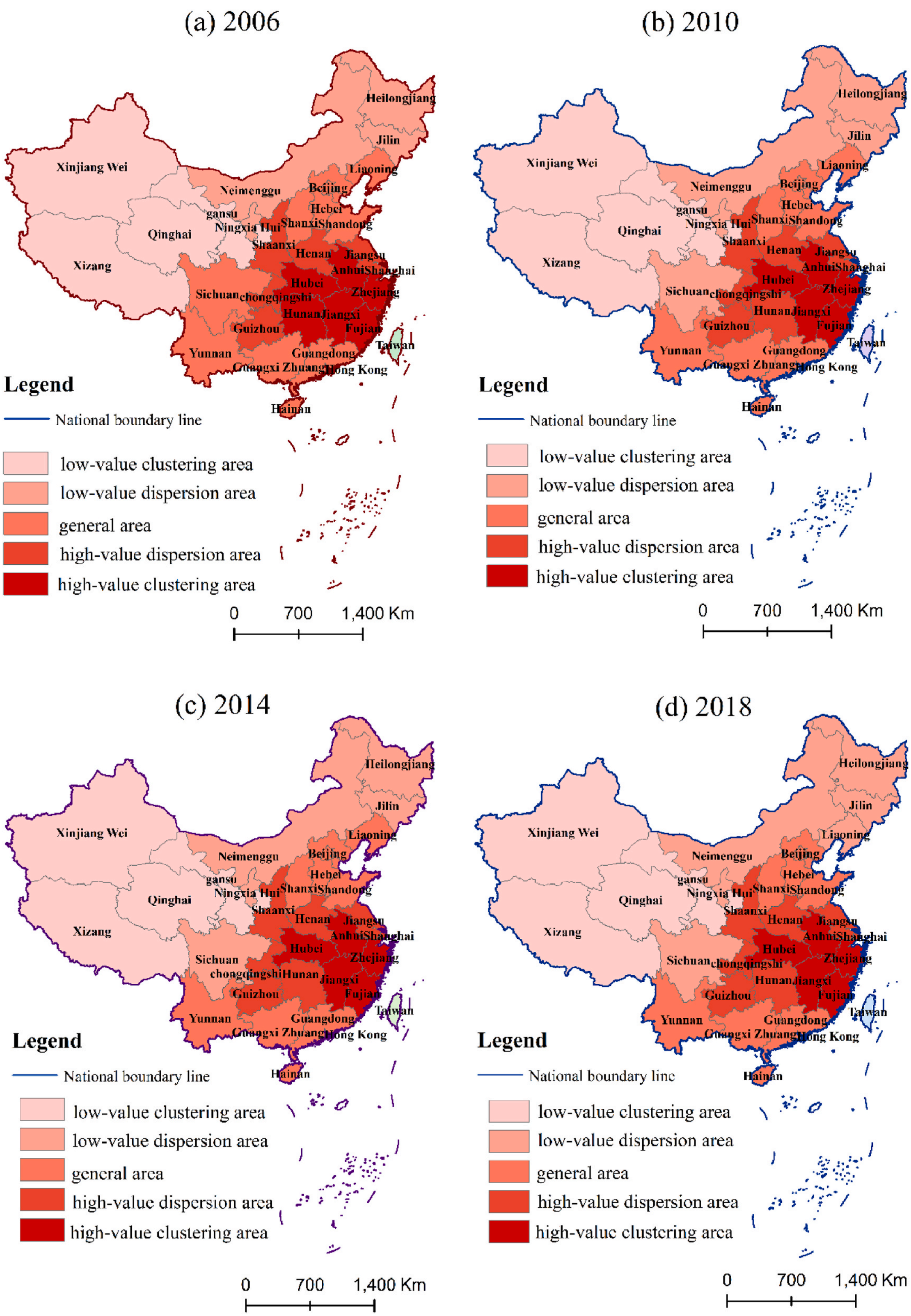

Fig. 5. Distribution of high-low-value clustering sites with the RP in spatial and temporal.

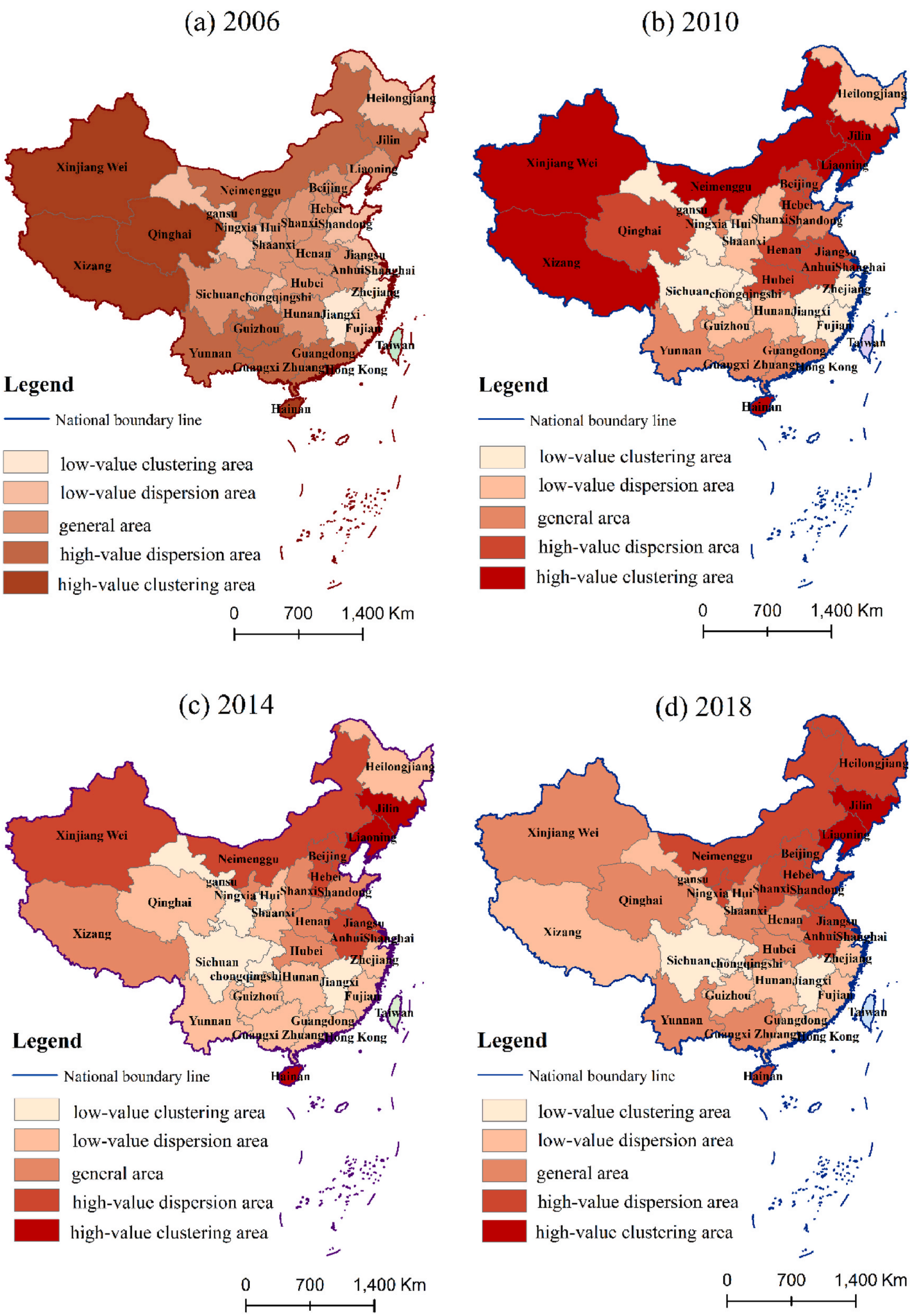

Fig. 6. Distribution of high-low-value clustering sites with the IS in spatial and temporal.

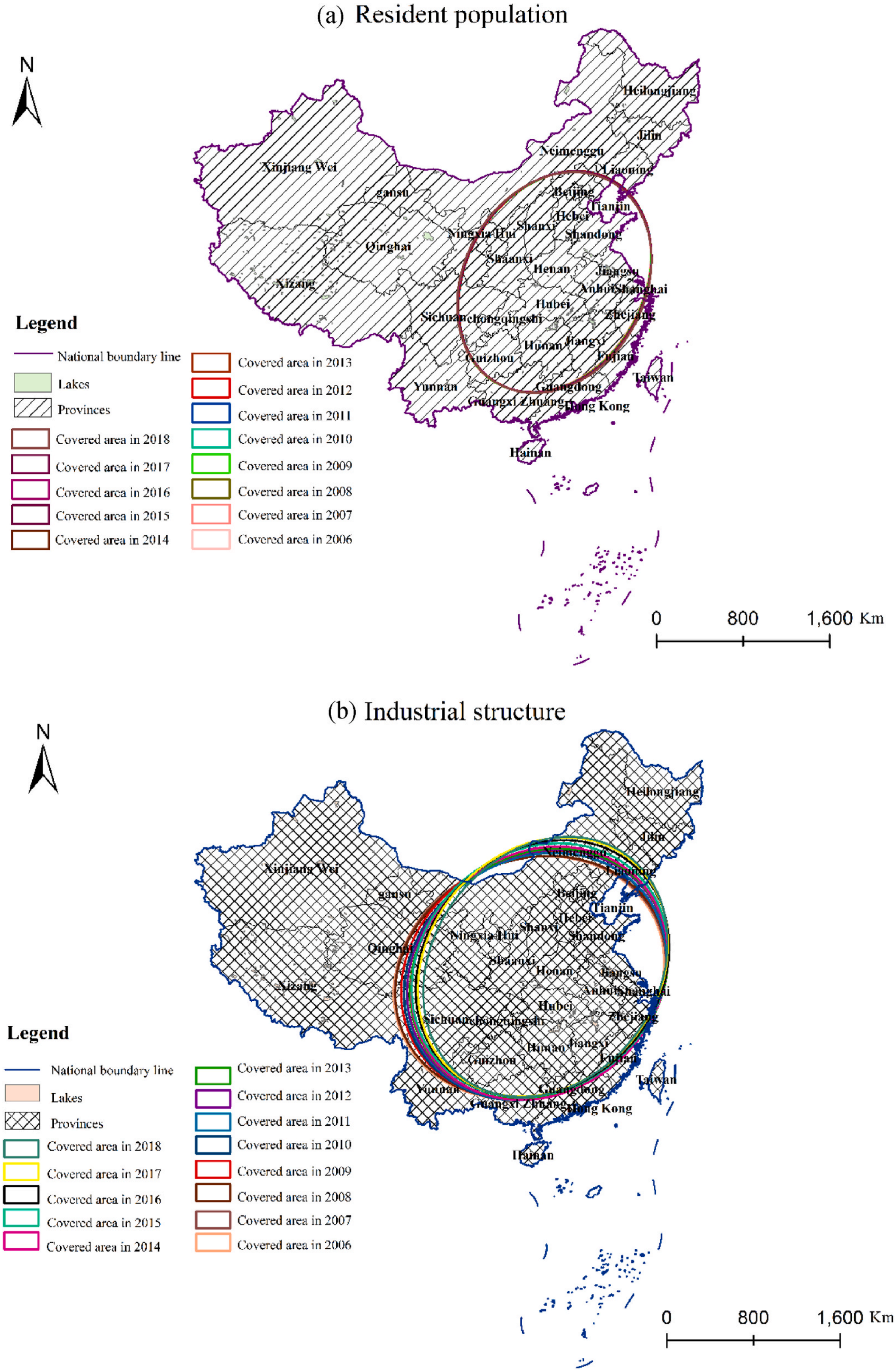

Fig. 7. Spatial distribution center and scope of the RP and IS changes in China between 2006 and 2018.

4.5. Analysis of regression

4.5.1. Research of overall baseline regression results

According to equations (9) and (10), we can get the regression results of DID as shown in Table 3. It can be found that the regression coefficients of HSR to the RP are both positive and statistically significant, showing that there are more permanent urban residents in areas with HSR than those without HSR. It demonstrates that HSR can, to a certain extent, encourage the growth of RP. Because of the dense distribution of towns and cities, the operation of HSR makes it possible for residents to commute between different regions, which facilitates residents' work and other travel activities, and promotes the growth of RP, thus promoting the development process of urbanization.

On the other hand, we examine how the HSR will affect IS. Table 3 shows that since the HSR was opened, the X has continuously improved. The significant and positive regression coefficient shows that HSR fosters the growth of the IS. This is because the HSR's opening can greatly reduce the distance between different regions and encourage the conversion of regional traditional industries into emerging industries. The establishment of HSR lines can make economically backward areas transfer industries from economically developed areas, drive the growth of regional industries, and support the growth of secondary and tertiary sectors, to continuously optimize the urban IS.

4.5.2. Research of local baseline regression results

The effects on RP and IS following the opening of HSR vary significantly between regions. While the operation of the HSR has a minimal influence on various regions' RP, it has a large influence on the IS. From Fig. 8, we can see that the HSR's opening had a positive influence on the RP of Guizhou, Shanghai, Zhejiang, Shanxi, Anhui, Guangxi, and other regions, and promoted the growth of the local population. The operation of the HSR is, however, negatively connected with the RP in Sichuan, Guangdong, Beijing, Hebei, and other regions, which indicates that the operation of HSR accelerates the population migration and reduces the pressure of urban population agglomeration to a certain extent.

Additionally, the operation of HSR has obvious spatial heterogeneity of IS. To evaluate the IS, we look at the proportion of tertiary to secondary industries. The secondary sector is based mostly on manufacturing and processing, whereas the tertiary sector is centered primarily on services and transportation. The HSR's operation has sped up the transportation of economic elements between regions and helped several, including Hainan, Hunan, Zhejiang, and Jilin, aided their tertiary industries. However, the opening of HSR has not done much to advance the IS of Shanghai, Tianjin, and Guangdong. The transportation in these regions is relatively developed and the IS is largely flawless, so the influence of the operation of HSR will be minimal.

4.5.3. Regression test

Based on the PSM-DID model, namely formula (11), the difference-in-differences model and the PSM matching data are used to test the influence of the HSR on the RP and IS. The robustness test's regression outcomes are displayed in Table 4.

According to Table 4, which shows that the PSM-DID model's regression coefficients pass the 5 % significance level test, HSR affects the RP negatively but favors the IS. The DID model developed in this work is stable since the verification result is compatible with the previous DID regression result.

Table 3
Results of the overall baseline regression of DID model.

Variable	(1)	(2)	(3)	(4)
	lnRP	IS	lnRP	IS
HSR	−0.456***	0.267***	−0.025**	0.044**
	(0.084)	(0.053)	(0.033)	(0.049)
lnGRP			0.563***	−2.276***
			(0.105)	(0.253)
lnTAX			−0.260***	0.832***
			(0.049)	(0.120)
lnDI			−0.870***	0.446***
			(0.055)	(0.092)
lnRSCG			0.335***	1.116***
			(0.081)	(0.177)
lnRPV			0.019	0.067
			(0.021)	(0.052)
lnRED			0.230***	0.024
			(0.037)	(0.065)
Individual effect	✓	✓	✓	✓
Time effect	✓	✓	✓	✓
Observations	403	403	403	403
R-squared	0.072	0.053	0.936	0.548

Note: 1. Standard deviation in brackets; 2. ***, **, and * signify significance at the 1 %, 5 %, and 10 % levels of statistics, respectively.

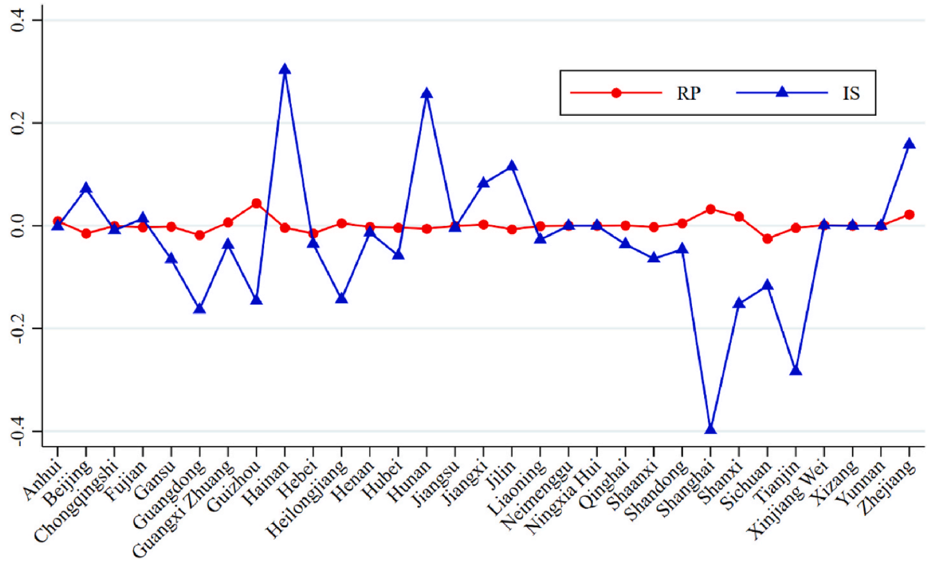

Fig. 8. The distribution of local baseline regression results in different regions.

Table 4
Regression results of PSM-DID.

Variable	(1)	(2)
	lnRP	IS
_treated	−0.053**	0.198***
	(0.092)	(0.061)
_cons	8.067***	0.908***
	(0.076)	(0.051)
N	403	403
r2_a	−0.002	0.023
F	0.333	10.408

Note: 1. Standard deviation in brackets; 2. ***, **, and * signify significance at the 1 %, 5 %, and 10 % levels of statistics, respectively.

4.5.4. Parallelism test

The degree of regional urbanization development is greatly influenced by if and when the HSR is opened in a region. To confirm the influence of the HSR's opening time on the RP and IS of various regions, we perform a parallelism test.

From Fig. 9, we can see that after *t* years, the influence of the HSR on the RP is adversely correlated. Along with the opening of the HSR, the regional permanent population shows a relatively downward trend, which indicates that HSR promotes the population flow between regions and alleviates the pressure of population growth. This is mainly because of the continuous opening of HSR lines, the problems affecting residents' commuting in the past have been solved to a certain extent. The residents' travel is more convenient, the travel time is significantly shortened, and the permanent population of the cities with HSR will be relatively reduced.

Fig. 10 shows the parallelism test of the influence of HSR on IS, and the test outcomes pass the test. This demonstrates the upgrading of IS and the percentage of the tertiary industry's added value that is promoted by the opening of HSR, which is compatible with our previous benchmark regression results and proves the rationality of the previous regression model construction.

The effect of the HSR on IS has a positive correlation, as illustrated in Fig. 10. The added value of the tertiary industry has been continuously increasing, and the IS has been constantly upgrading, ever since the *t* year (after the opening of the HSR). This is because the operation of HSR lines links the social movement of people and logistics in various regions. The accelerated growth of HSR is crucial for enhancing the capacity of railway infrastructure support, accelerating industrial transformation and upgrading, reducing social logistics costs, optimizing IS adjustment, and providing strong support for promoting economic and social development.

5. Conclusion

HSR not only reduces the space distance between various places but also improves residents' travel experience and promotes the upgrading of IS, which is crucial for accelerating the coordinated growth of the regional economy. This research investigates the spatial aggregation pattern and evolutionary distribution of the influence of RP and IS before and after the opening of HSR based on the spatial autocorrelation model and standard deviation ellipse model. Moreover, the spatial heterogeneity of the influence of HSR opening on

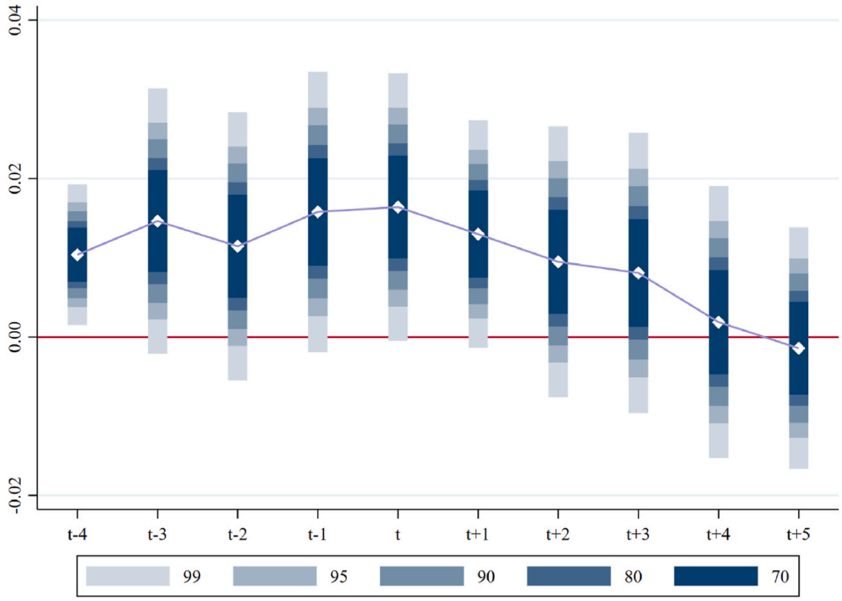

Fig. 9. The parallelism test of the influence of the HSR on RP.

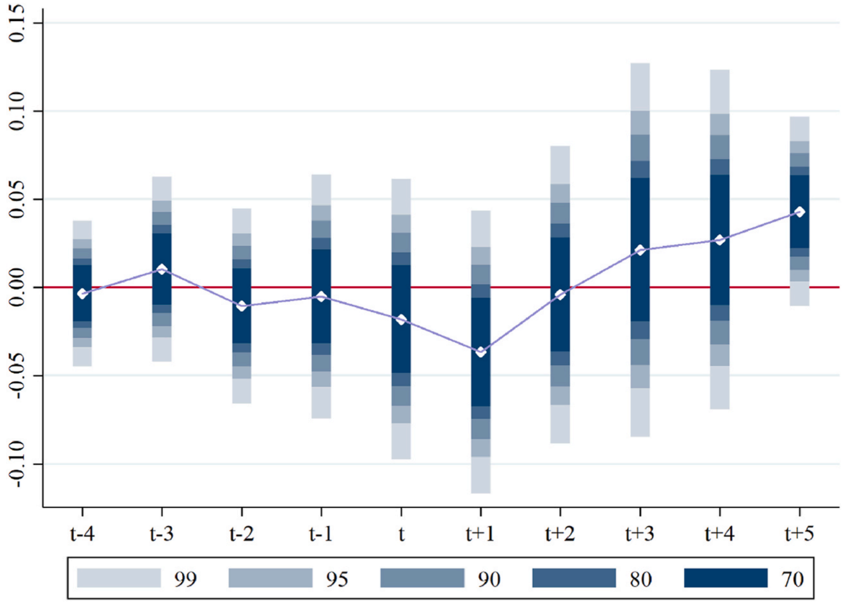

Fig. 10. The parallelism test of the HSR's influence on IS.

RP and IS is explored by the PSM-DID model from both a global and local perspective. The following is a summary of the key research findings:

Firstly, the RP and IS demonstrate a regional spatial imbalance. The spatial correlation of the RP is positive while that of the IS is negative. Eastern and central China are where the RP is chiefly concentrated. The agglomeration of IS in space, which congregates in eastern, central, and western China, displays an obvious regional imbalance. In addition, the spatial evolution pattern of the RP and IS shows a general evolution pattern from northeast to southwest.

Secondly, there is no doubt about the differences between the consequences of the HSR opening on the RP and IS. On the whole, HSR affects the IS favorably while adversely impacting the RP. Locally, the influence of the HSR opening on RP and IS is spatially heterogeneous. For example, whereas Sichuan, Guangdong, Beijing, and Hebei are negatively impacted by the operation of the HSR, Guizhou, Shanghai, Zhejiang, Shanxi, Anhui, and Guangxi are positively affected. On the other hand, the opening of HSR has an adverse correlation with the IS of Shanghai, Tianjin, and Guangdong. But it has promoted the IS of Hainan, Hunan, Zhejiang, and Jilin provinces.

HSR has a multifaceted and intricate effect on regional economic growth. A comparative research of the effects of the operation of HSR on various regional economies, as well as the spatial relevance and evolution process of regional economic development has certain guiding significance for optimizing HSR stations and connecting urban transportation networks. The building and operation of HSR may effectively quicken the movement of the population and stimulate the upgrading of IS between different regions, which helps enhance urbanization. The relevant transport sector will continue to strengthen the construction of regional integration and establish a more perfect transport network according to policies and plans, to foster the healthy growth of urbanization.

This paper uses a spatial model and econometric method to study the impact of HSR opening on RP and IS, as well as the degree of aggregation and correlation in time and space. However, there are some limitations in the breadth and depth of the research methods and analysis in this paper. Firstly, the selection of indicators in the model mainly depends on the platform data and the selection of variables in the previous literature. Therefore, the selection range of indicators in the control variables and more detailed multi-source data can be added in future research. Secondly, the research methods and models in this paper are not comprehensive and specific. More diverse spatial econometric models and quantitative analysis methods should be added. Thirdly, due to the limitation of data acquisition in the time dimension, longer time series data can be used in the future to explore the impact of HSR on urban economic elements. In addition, the differences in prosperity and IS in different regions are not only related to the opening of HSR, but other potential factors that may affect the level of urbanization development are also worthy of further discussion.

Funding

This research was supported by the Social Science Key Fostering project "Research on the Realization Mechanism and Key Policies for Carbon Peak and Neutralization of Transportation in China" (grant numbers 2022JBW1001)

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Xuanxuan Xia: Conceptualization, Formal analysis, Writing – original draft. **Hongchang Li:** Conceptualization, Funding acquisition, Writing – review & editing. **Kun Wang:** Methodology, Software, Writing – original draft. **Yixian Liu:** Methodology, Software, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are very grateful to the editors and anonymous reviewers for their comments on the manuscript.

References

[1] T.T. Deng, D.D. Wang, S.Y. Cheng, The impact of high speed railway on urban service industry agglomeration, J. Finance Econ. 43 (7) (2017) 119–132, https://doi.org/10.16538/j.cnki.jfe.2017.07.010.
[2] Y. Huang, Spatial and temporal heterogeneity of the impact of high-speed railway on urban economy: empirical study of Chinese cities, J. Transport Geogr. 91 (2021), 102972, https://doi.org/10.1016/j.jtrangeo.2021.102972.
[3] M. Yin, L. Bertolini, J. Duan, The effects of the high-speed railway on urban development: international experience and potential implications for China, Prog. Plann. 98 (2015) 1–52, https://doi.org/10.1016/j.progress.2013.11.001.
[4] C. Wang, W. Meng, X. Hou, The impact of high-speed rails on urban economy: an investigation using night lighting data of Chinese cities, Res. Transport. Econ. 80 (2020), 100819, https://doi.org/10.1016/j.retrec.2020.100819.
[5] T. Hiramatsu, Job and population location choices and economic scale as effects of high speed rail: simulation analysis of Shinkansen in Kyushu, Japan, Res. Transport. Econ. 72 (2018) 15–26, https://doi.org/10.1016/j.retrec.2018.06.007.
[6] S. Liu, Y. Wan, H.K. Ha, Y. Yoshida, A. Zhang, Impact of high-speed rail network development on airport traffic and traffic distribution: evidence from China and Japan, Transport. Res. Part A Policy and Practice 127 (2019) 115–135, https://doi.org/10.1016/j.tra.2019.07.015.
[7] C. Wang, H. Yang, H. Yuan, The impact of railway reform on corporate export: the case of China, Transport. Res. Pol. Pract. 118 (2018) 627–647, https://doi.org/10.1016/j.tra.2018.10.013.
[8] Z. Wang, L. Liang, Z. Sun, X. Wang, Spatiotemporal differentiation and the factors influencing urbanization and ecological environment synergistic effects within the Beijing-Tianjin-Hebei urban agglomeration, J. Environ. Manag. 243 (2019) 227–239, https://doi.org/10.1016/j.jenvman.2019.04.088.
[9] T.B. Dong, Influence and countermeasures of high-speed railway economy on regional coordinated development of Pearl River Delta, Industrial & Science Tribune 16 (12) (2017) 23–26, https://doi.org/10.3969/j.issn.1673-5641.2017.12.010.
[10] C.M. Zhang, W.Q. Wang, W.T. Li, W.L. Xiang, Analysis on regional economic characters of Lan-Xin High Speed Railway based on panel data, J. Railw. Sci. Eng. 14 (1) (2017) 12–18, https://doi.org/10.3969/j.issn.1672-7029.2017.01.003.
[11] D.C. Fang, M.Y. Sun, S.O. Business, An evaluation on the impact of high-speed rail on economic development of city clusters in the Yangtze River delta——an empirical study based on difference in difference model, East China Economic Management 30 (2) (2016) 42–47, https://doi.org/10.3969/j.issn.1007-5097.2016.02.007.

[12] M.R. Jiang, Y. Fan, Y. Xia, Q.R. Chen, Y. Ye, Environmental impact and economic assessment of Chinese high speed rail investments, China Population, Resources and Environment 27 (2) (2017) 75–83, https://doi.org/10.3969/j.issn.1002-2104.2017.02.012.
[13] X.X. Liu, H. Zhang, Y. Cheng, The impact of high-speed railway on urban real estate prices: a study based on differential mode, Inquiry into Economic Issue (8) (2018) 28–38. CNKI:SUN:JJWS.0.2018-08-005.
[14] Y. Wang, M. Nian, Did high-speed trains promote the development of regional economy, Shanghai Economic Review (2) (2014) 84–93, https://doi.org/10.19626/j.cnki.cn31-1163/f.2019.11.004.
[15] S.M. Jia, C.L. Qin, The influence of high-speed railway to the equilibrium of China's regional economic pattern, Areal Research and Development 34 (2) (2015) 13–20, https://doi.org/10.3969/j.issn.1003-2363.2015.02.003.
[16] L.H. Wang, W. Xia, Y.M. Ning, Study on disequilibrium of distribution of high-speed railway cities in China: based on the comparison with the ordinary railway, Urban Development Studies 24 (7) (2017) 68–78, https://doi.org/10.3969/j.issn.1006-3862.2017.07.010.
[17] Y.F. Wang, P.F. Ni, Economic growth spillover and spatial optimization of high-speed railway, China Industrial Economics (2) (2016) 21–36, https://doi.org/10.19581/j.cnki.ciejournal.2016.02.003.
[18] K.Z. Zhang, D.J. Tao, The economic distribution effect of transportation infrastructure: evidence from the opening of high-speed rail, Econ. Perspect. (6) (2016) 62–73. CNKI:SUN:JJXD.0.2016-06-006.
[19] X.Y. Lin, Z.H. Shi, S. Luo, D. Wu, M.T. Shi, An analysis of the attractiveness of high-speed railway to talents, J. Beijing Jiaot. Univ. 14 (3) (2015) 7–16, https://doi.org/10.16797/j.cnki.11-5224/c.2015.03.002.
[20] C.F. Huang, Q. Han, The evolution of the high speed rail operation and the distribution pattern of population flow in the Yangtze River delta, Shanghai Economic Review (7) (2021) 39–54, https://doi.org/10.19626/j.cnki.cn31-1163/f.2021.07.005.
[21] J. Zhang, H.S. Chen, A study of agglomeration shadow during urbanization in the new era——the case for high-speed rai, China Economic Quarterly 22 (1) (2022) 281–302, https://doi.org/10.13821/j.cnki.ceq.2022.01.14.
[22] C.R. Bhat, J. Guo, A mixed spatially correlated logit model: formulation and application to residential choice modeling, Transp. Res. Part B Methodol. 38 (2) (2004) 147–168, https://doi.org/10.1016/s0191-2615(03)00005-5.
[23] Y. Shen, A.J.D. Silva e, L.M. Martínez, Assessing High-Speed Rail's impacts on land cover change in large urban areas based on spatial mixed logit methods: a case study of Madrid Atocha railway station from 1990 to 2006, J. Transport Geogr. 41 (2014) 184–196, https://doi.org/10.1016/j.jtrangeo.2014.09.007.
[24] O.A. Smirnov, Modeling spatial discrete choice, Reg. Sci. Urban Econ. 40 (5) (2010) 292–298, https://doi.org/10.1016/j.regsciurbeco.2009.09.004.
[25] O.A. Smirnov, K.J. Egan, Spatial random utility model with an application to recreation demand, Econ. Modell. 29 (1) (2012) 72–78, https://doi.org/10.1016/j.econmod.2010.09.026.
[26] I.N. Sener, R.M. Pendyala, C.R. Bhat, Accommodating spatial correlation across choice alternatives in discrete choice models: an application to modeling residential location choice behavior, J. Transport Geogr. 19 (2) (2011) 294–303, https://doi.org/10.1016/j.jtrangeo.2010.03.013.
[27] X.C. Wang, K.M. Kockelman, J.D. Lemp, The dynamic spatial multinomial probit model: analysis of land use change using parcel-level data, J. Transport Geogr. 24 (2012) 77–88, https://doi.org/10.1016/j.jtrangeo.2012.06.011.
[28] W. Cao, X. Fang, H. Zhang, The structural and spatial properties of the highspeed railway network in China: a complex network perspective, Journal of Rail Transport Planning & Management 9 (2019) 46–56, https://doi.org/10.1016/j.jrtpm.2018.10.001.
[29] L. Zheng, F. Long, Z. Chang, J. Ye, Ghost town or city of hope? The spatial spillover effects of high-speed railway stations in China, Transport Pol. 81 (2019) 230–241, https://doi.org/10.1016/j.tranpol.2019.07.005.
[30] D. Fang, Y. Sun, The reconstruction of the spatial structure of the Yangtze River Delta city Group in the high-speed rail era-Based on the social network analysis, Econ. Geogr. 35 (10) (2015) 50–56, https://doi.org/10.15957/j.cnki.jjdl.2015.10.008.
[31] F.A. Paulo de, G.S. Porto, Solar energy technologies and open innovation: a study based on bibliometric and social network analysis, Energy Pol. 108 (2017) 228–238, https://doi.org/10.1016/j.enpol.2017.06.007.
[32] C.W. Xiao, Y. Liu, P. Li, L.G. Jiang, Evaluation of the high-speed rail new town based on coordination of urban spatial expansion and population growth, J. Nat. Resour. 31 (2016) 1440–1451, https://doi.org/10.11849/zrzyxb.20151208.
[33] W. Xu, J. Zhou, L. Yang, L. Li, The implications of high-speed rail for Chinese cities: connectivity and accessibility, Transport. Res. Pol. Pract. 116 (2018) 308–326, https://doi.org/10.1016/j.tra.2018.06.023.
[34] F. Wang, X. Wei, J. Liu, L. He, M. Gao, Impact of high-speed rail on population mobility and urbanisation: a case study on Yangtze River Delta urban agglomeration, China, Transport. Res. Pol. Pract. 127 (2019) 99–114, https://doi.org/10.1016/j.tra.2019.06.018.
[35] S. Wang, J. Wang, X. Liu, How do urban spatial structures evolution in the high-speed rail era? Case study of Yangtze River Delta, China, Habitat Int. 93 (2019), https://doi.org/10.1016/j.habitatint.2019.102051. Article ID 102051.
[36] C. Liu, K.A. Mamonov, A geographic information system for territorial development of the marine regions of China 1 (13) (2020) 10, https://doi.org/10.15589/smi2020.
[37] P. Zhang, Y. Zhao, X. Zhu, Z. Cai, J. Xu, S. Shi, Spatial structure of urban agglomeration under the impact of high-speed railway construction: based on the social network analysis, Sustain. Cities Soc. 62 (2020), 102404, https://doi.org/10.1016/j.scs.2020.102404.
[38] A. Monzón, E. Ortega, E. López, Efficiency and spatial equity impacts of high-speed rail extensions in urban areas, Cities 30 (2013) 18–30, https://doi.org/10.1016/j.cities.2011.11.002.
[39] Z.W. Xu, J. Song, Spread or backwash: reshaping effect of economic spatial distribution from high-speed rail in China, East China Economic Management 33 (2) (2019) 71–78, https://doi.org/10.19629/j.cnki.34-1014/f.180515024.
[40] X.I. Qiangmin, J.I. Peng, Does the development zone promote population urbanization? Evidence from China, Land Use Pol. 131 (2023), 106736, https://doi.org/10.1016/j.landusepol.2023.106736.
[41] C. Lin, J. Liu, W. Li, Influence of the high-speed railway (HSR) construction on industrial structure transformation, Enterprise Inf. Syst. 17 (2) (2023), 1942998, https://doi.org/10.1080/17517575.2021.1942998.
[42] Z. Tang, H. Bi, J. Sun, Y. Xu, Z. Wang, Research on the coopetition relationship and the development of city-industry integration of urban agglomeration under the effect of HSR, Front. Environ. Sci. 11 (2023), 1082664, https://doi.org/10.3389/fenvs.2023.1082664.
[43] D.W. Lefever, Measuring geographic concentration by means of the standard deviational ellipse, Am. J. Sociol. 32 (1) (1926) 88–94, https://doi.org/10.1086/214027.
[44] A. Mitchel, The ESRI Guide to GIS Analysis, in: Spartial Measurements and Statistics, ume 2, ESRI Guide to GIS analysis, 2005.
[45] O.C. Ashenfelter, D. Card, Using the Longitudinal Structure of Earnings to Estimate the Effect of Training Programs, 1984, https://doi.org/10.3386/w1489.
[46] N.N. Zhu, J.J. Zhu, S.F. Liu, et al., Evaluation of affection on architecture energy-saving policy in China, Chinese Journal of Management Science 16 (S1) (2008) 576–580, https://doi.org/10.16381/j.cnki.issn1003-207x.2008.s1.035.
[47] Y.F. Yan, Z.X. Zhao, R. Wang, The effect of export rebate on CO2 emission: an analysis based on iron and steel industry, Word Economy Studies (4) (2012) 46, 50+88. CNKI:SUN:JING.0.2012-04-007.
[48] L. Chen, J. Xia, Policy effect of China's college enrolment expansion on the innovation efficiency: based on the framework of quasi experiment and difference-in-difference method, Chin. J. Popul. Sci. (5) (2015) 45–57+127. CNKI:SUN:ZKRK.0.2015-05-007.
[49] C. Li, Y.X. Zhang, The policy effect of China's reform of 'Replacing Business Tax with Value-added Tax': a test based on difference-in-difference method, Public Finance Research (2) (2015) 44–49, https://doi.org/10.19477/j.cnki.11-1077/f.2015.02.012.
[50] J.J. Heckman, H. Ichimura, P.E. Todd, Matching as an econometric evaluation estimator: evidence from evaluating a job training programme, Rev. Econ. Stud. 64 (4) (1997) 605–654, https://doi.org/10.2307/2971733.

[51] J.J. Heckman, H. Ichimura, P. Todd, Matching as an econometric evaluation estimator, Rev. Econ. Stud. 65 (2) (1998) 261–294, https://doi.org/10.1111/1467-937x.00044.
[52] J. Jiao, J. Wang, F. Jin, Impacts of high-speed rail lines on the city network in China, J. Transport Geogr. 60 (2017) 257–266, https://doi.org/10.1016/j.jtrangeo.2017.03.010.
[53] X. Lin, Z. Zhang, M. Wang, Value and governance of high-speed railway, Frontiers of Engineering Management 4 (4) (2017) 463, https://doi.org/10.15302/j-fem2017054.