Travel on the road: does China’s high-speed rail promote local tourism?

Xiaoxiao Zhou¹ · Siyu Chen² · Hua Zhang³

Received: 7 May 2022 / Accepted: 15 July 2022 / Published online: 28 July 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Following a Chinese saying: To be rich, roads first, high-speed rail (HSR) opening and station construction are indispensable for economic developing. Probing the nexus between HSR, as a vital part of modern transportation system, and local tourism development provides a scan for reviving tourism and gaining low-carbon transition after COVID-19 pandemic. Drawing on prefecture-level panel data, this study takes difference-in-difference and instrument variable methods to detect the overall and heterogeneous effects of HSR connection on cities’ tourism development. The results showed that HSR connection had an overall positive effect on cities’ domestic tourist arrivals. The heterogeneity of the effect from HSR to tourism development appears to be that central and western cities, non-resource-based cities, and small cities benefited more from the opening of HSR. From a dynamic perspective, HSR connection promoted local tourism development in the 0 and 1 year of HSR opening but failed to show a positive effect in the long term. Hence, the study proposed some adjustments for evaluating the efficiency of HSR with consideration for the tourism effect, redesigning the system of HSR with consideration for local heterogeneity, and optimizing the HSR environment. These measures can optimize China’s HSR management and the design of HSR systems.

Keywords High-speed rail · Tourism development · Difference-in-differences · Heterogeneity

Introduction

There was an old Chinese saying: To be rich, roads first. As a crucial component of modern transportation systems (Castillo-Manzano et al. 2018; Zhang et al. 2020), high-speed rail (HSR) has developed rapidly and shows a gradually increasing influence on economic developing, especially in China. With almost 20 years of endeavors including preliminary planning, technology improvement, large-scale construction, stagnation, and going global (Li et al. 2021), the largest HSR network has been a buildup in China (China Railway Maps 2021). Compared with traditional modes of travel, HSR has the advantages of saving time by providing comfortable, convenient, and highly efficient transport (Li et al. 2021). HSR development has profoundly changed the mode of transport and led to a suite of impacts on economic, social, and environmental aspects (Castillo-Manzano et al. 2018; Zhang et al. 2020; Li et al. 2021). The tourism industry, particularly, depends heavily on transport (Wang et al. 2018; Chen and Haynes 2015). The mode, scale, structure, and distribution of tourism have been strongly affected by HSR development (Pagliara et al. 2017; Yin et al. 2019; Zhang et al. 2020).

Research on both HSR and tourism has attracted the attention of academics and policymakers (Chen and Haynes 2015; Pagliara et al. 2017; Li et al. 2021). Indeed, from different perspectives, a variety of studies have drawn on diverse data and applied multitudinous methods to detect the nexus between HSR and tourism development. Some economists used case studies to reveal the changes in the arrivals of tourists and peaks of tourism activities (Delaplace et al. 2014; Guirao and Campa 2016). Some others considered the tourism corridor effect and proved that both aggregation and
diffusion effects appear between HSR and tourism development (Dai et al. 2018; Jin et al. 2020). Numerous works have embraced this issue by connecting HSR with tourism development using an econometric method based on panel data, yielding commixed results covering positive effects (Pagliara et al. 2017; Castillo-Manzano et al. 2018; Pagliara and Mauriello 2020), negative effects (Albalate and Fageda 2016; Gao et al. 2019), and non-significant effects (Pagliara et al. 2015; Campa et al. 2016). These studies shape a complicated understanding of the connection between HSR and tourism. The links from HSR to tourism remains controversial and shows heterogeneity (Pagliara and Mauriello 2020). Thus, we focused on the mechanism behind the relationship between HSR connection and tourism development and aimed at a credible and robust empirical test based on China’s city data using difference-in-differences (DID) and instrument variable (IV) methods.

Essentially, few studies have addressed the causality between HSR construction and tourism development in China (Chen and Haynes 2015; Xu et al. 2018). As of 2020, China’s HSR network has a length of over 37,900 km (23,550 mi) and serves 94.7% of the population of more than one million cities. China’s HSR construction supports the new concept of “Green, Intelligent, Safety and Humanity” and endeavors to contribute to sustainable development. Green transformation is trendy in the context of economic recession and sustainability. Coincidentally, tourism is a representative sector of the green industry. In recent years, the role of tourism in the economy has increased. From Fig. 1, both tourism expenditure and its ratio to gross domestic product (GDP) have an upward tendency. In 2019, the added value of tourism and its related industry reached RMB 4.4989 trillion, accounting for 4.56% of GDP (NBSC). HSR shows a direct and strong correlation with regional tourism industry. However, in China, whether the nexus of HSR and tourism development is significant and positive or negative; how HSR connections influence tourism development; and whether the connection between HSR and tourism is heterogeneous among different regions, to address these issues, we focused on the mechanism, level, direction, and heterogeneity of the relation between HSR and tourism in China. Our findings could inform HSR system optimization and tourism policy formulation.

The present study may make some contributions as follows. First, from both macro and microscopic perspectives, we provide a comprehensive mechanism analysis based on a comprehensive literature review. Given that previous studies have mainly attempted to explore the nexus of HSR and tourism from a specific view (Guirao and Campa 2016; Dai et al. 2018), our analysis would identify the pathway from HSR to tourism development by accounting for macro and micro subjects and expecting a comprehensive and deep understanding of them. Second, the prefecture-level city data covering 283 samples used in our study provide detailed and referential cases in this area. Existing studies, especially on China, mainly conduct empirical tests using province-level data or small samples, such as a tourism corridor or a special region or city (Yin et al. 2019; Jin et al. 2020). Third, our methods included DID, IV, and propensity score matching combined with DID (PSM-DID) to treat endogeneity and econometric selection bias, as well as determine robust and believable conclusions.

Fig. 1 The tourism and GDP in China (2000–2019)
Literature review and hypothesis

Literature review

From a positive viewpoint, by an extended gravitational model, Wang et al. (2012) found that China’s HSR development can redistribute and transform the tourist market, extend market competition, and reallocate tourism centers. Through a questionnaire analysis of the main Italian cities, Cartenì et al. (2017) found that not only the “Faster services” of HSR, but also “hedonic services” attract more tourists. Pagliara et al. (2017) analyzed 2006–2013 data from 77 Italian municipalities and reported that HSR service has a positive effect on Italian visitors’ arrivals and the number of nights spent. In 28 EU countries for the period 1996–2014, Castillo-Manzano et al. (2018) found that HSR is more conducive to domestic tourism, whereas air travel is more conducive to foreign tourism. Based on 99 Italian provinces, Pagliara and Mauriello (2020) proved that HSR has a satisfying effect on both Italian and foreign tourists and the effect shows spatial heterogeneity.

To the negative results, based on the DID method using Spain’s panel data covering 50 provinces for 1998–2013, Albalate and Fageda (2016) indicated that HSR may have a positive (weak) direct effect on tourism, but it may also lead to a negative indirect effect by impacting air traffic. Drawing on China’s city-level data, Gao et al. (2019) used DID and IV methods and reported that connecting to HSR boosts tourist arrivals but fails on stimulating tourism revenue, resulting in a negative nexus between HSR and tourism revenue per arrival. In addition, the effect was heterogeneous.

Regarding the non-significant effect, Pagliara et al. (2015) analyzed data gathered using a revealed preference survey and found that owing to the high percentage of foreign tourists, HSR has no significant impact on the choice to travel to Madrid. Campa et al. (2016), drawing on 47 provincial panel data in 1999–2015, revealed a positive but lower effect of Spanish HSR construction on foreign arrivals and revenues but not on domestic tourism. Guirao and Campa (2016) examined the cross effect of tourism and HSR with Spain as a case study and showed that the former has positive effects on the latter, but that the HSR cannot bring about additional tourism demand.

Others have also pointed out that the impact from HSR to tourism is the result of multiple interacting forces, which may lead to nonlinearity or conditional results. Albalate et al. (2017) pointed out that, in large cities in Spain, HSR has positive influence on the tourism aspects, but the effects are minimal or even negative to most cities. Xu et al. (2018) constructed an index of the connectivity-accessibility of cities and found that cities in the Yangtze River Delta would suffer the most, whereas central and western cities would gain the most from the HSR network. Gao et al. (2019) acknowledged that the accessibility of HSR can facilitate tourism and stimulate tourism promotion; however, the redistribution from asymmetric accessibility may lead to the agglomeration of tourism resources and hamper the tourism development of surrounding cities.

To explore the link from HSR connection to tourism in China, we reviewed the related literature in Table 1 and

Table 1 High-speed rail and its effect on tourism-related growth in China

| Reference          | Proxy of HSR                                                                 | Sample and methods                                                                 | Conclusions                                                                 |
|--------------------|------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|----------------------------------------------------------------------------|
| Wang et al. 2018   | Urban accessibility and the shortest temporal distance                       | 338 prefecture-level cities in China; principal component analysis and ArcGIS     | Total number associated with tourism-related economic relations ↑; competition in hinterland ↑ |
| Liu and Zhang, 2018| HSR in 2006 and 2014                                                        | 266 prefectural level cities in 2006 and 2014; benchmarking analysis             | Travel times ↓; accessibility ↑; GRP per capita ↑ |
| Gao et al. 2019    | 0, 1 dummy variable*                                                        | China’s city panel data; DID, IV method                                           | Tourism revenue –; tourist arrivals ↑; tourism revenue per arrival ↓     |
| Yin et al. 2019    | The reduction in travel times                                                | The tourism spatial interaction between Beijing and Tianjin; a revised Wilson’s model | Tourism spatial interactions ↑; tourism spatial structure changed         |
| Jin et al. 2020    | HSR in 2008 and 2018                                                        | Three provinces in northeast China; weighted mean travel time and network analysis| One-day and weekend trips ↑; tourism development in nearby cities ↑     |
| Zhang et al. 2020  | 0, 1 dummy variable                                                          | Tourism firms; DID, dynamic DID and placebo test                                 | Tourism firms’ value ↑                                                   |
| Deng et al. 2020   | 0, 1 dummy variable; HSR configurations                                      | China’s city panel data; DID, IV method                                           | Total tourist arrivals ↑                                                 |

*0, 1 dummy variable is set as HSR = 1 if a city is connected to the HSR network, and HSR = 0, otherwise. – means non-significant, ↑ means HSR has a significant positive effect on the explained variable, and ↓ indicates a negative effect.
handled them as comparable to our research. In almost all contexts, HSR shows a positive effect on tourism-related factors in China (Wang et al. 2018; Jin et al. 2020; Zhang et al. 2020). Regarding tourist arrivals and tourism competition, connecting to the HSR system always has a positive influence (Gao et al. 2019; Jin et al. 2020). As for tourism-related revenue and travel times, the impact of HSR appears to be weak and even disappointing (Liu and Zhang 2018; Gao et al. 2019). The divergence and incoherence of the conclusions provide reference to our research but rather prove the necessity and urgency of our work.

**Framework and hypothesis**

To elucidate the connection between HSR and tourism development, we investigated their mechanism from both macro and micro perspectives covering economic development, industrial structure, firm performance, and consumer preference as the intermediary determinants. Figure 2 outlines the framework of our theoretical analysis.

Essentially, for HSR, its promotion of accessibility and mobility is the main reason that it can affect cities’ tourism. In other words, HSR connection brings about acceleration in factor-flow and space–time compression. Connecting to HSR makes factors including capital, labor, information, and other resources flow faster, ultimately reducing the cost of regional resource mobility, expediting resources reallocation and promoting factor productivity. From a macroscopic perspective, these impacts eventually boost the economy and optimize the industrial structure (Cartenì et al. 2017; Li et al. 2020). Economic growth also reversely drives tourism development by providing adequate capital and employment supports and stimulating local consumption, especially tourism consumption (Gao et al. 2019). Meanwhile, the coordinated development of industries simultaneously drives tourism and related industry development. Resource agglomeration brought by HSR connection may benefit key cities but limit the surrounding cities’ tourism, the so-called core-periphery polarization. Gao et al. (2019) noted that redistribution caused by asymmetric accessibility would not be conducive to the tourism development of surrounding cities.

Additionally, the mobility increase and cost reduction not only result in direct firm performance promotion by reducing costs and removing face-to-face meeting obstacles (Lin 2017) but also strengthen the competition of local firms by expanding the market scope and forcing firms to be more creative and effective, especially tourism and related firms (Matas et al. 2020). Zhang et al. (2020) confirmed that HSR fundamentally influences the business environment by substantially lessening actual travel time and boosting the mobility of production factors. Ma et al. (2021) showed that the opening of HSR can enhance the market potential of a region by improving information flow and face-to-face communication, ultimately manifested in the promotion of regional innovation and entrepreneurship. This effect is more prominent in the large cities of China.

HSR is a viable way to mitigate travel time immediately compared with traditional transport modes (e.g., cars, trains, boats). The space–time compression provides another choice for consumers and further changes the tourists’ preferences. With the availability of “faster train” services, unplanned trips increases, resulting higher tourist arrivals. Jin et al. (2020) proved that HSR connection promoted 1-day and weekend trips and tourism development in nearby cities. Also, HSR can provide “hedonic services” by its high frequency, reliability, and easy access, thus attracting more people to travel with HSR instead of other vehicles. Cartenì et al. (2017) indicated that HSR influences tourists’ choice of travel mode through “faster train” services and “hedonic services.” The travel time reduction by HSR makes long-distance travel reliable, and its high-quality services make long-distance travel acceptable.

---

**Fig. 2** Links between HSR and tourism development

- **Factors free-flow**
  - Cost reduction, factor reallocation and agglomeration, strengthen firms’ competition, deindustrialization; Siphonage, Matthew effect

- **High-speed rail**
  - Space-time compression
    - Replace airplanes, automobile, and other vehicles; Change the preference of consumer behavior

- **Aggregate channels**
  - Economic development
  - Tourist arrivals
  - Industrial structure
  - Travel distance
  - Firm performance
  - Travel modes
  - Consumer preference
  - Tourism structure
  - Substitution effect
However, considering transfer and accessibility, tourists prefer airplanes in international travel. Castillo-Manzano et al. (2018) found that, in EU countries, HSR is beneficial to domestic tourism, whereas airplanes benefit international travel.

In general, considering economic development, industrial structure, and firm performance as intermediary factors, HSR connection ameliorates factor free-flow; leads to mobility cost reduction, resource reallocation, and agglomeration; and finally formats the so-called aggregate channels, via the two competing effects of siphonage and positive effect. Given consumer preferences, HSR provides both “faster train” and “hedonic services” and attracts more tourists and more HSR travels, ultimately enabling the so-called substitution effect, which is more apparent in domestic rather than international tourism. Thus, we proposed the following hypothesis:

Hypothesis 1: HSR has positive but varying effect on cities’ domestic tourism.

The regional disparity of the HSR effects on tourism (Pagliara and Mauriello 2020) has been attributed to two aspects: the asymmetric allocation of the HSR system and the heterogeneity of regional characteristics (Albalate et al. 2017; Xu et al. 2018; Gao et al. 2019). In essence, the asymmetrical distribution of HSRs is produced internally from regional heterogeneity, to some extent. Thus, HSR services may have different effects on different cities with heterogeneous endowments and conditions. A typical classification is the geographic location; in this study, we also paid attention to classifications based on city size and resource dependency.

In China, eastern, central, and western cities have apparent differences in economic, social, environmental, transport, and resource endowments. HSR services also have vital differences in these cities. The divergence of HSR and other endowments explains the variation in the territorial impact of HSR on tourism. In eastern regions, which originally have adequate and advanced transport systems, HSR may have little impact on tourism promotion. To central and western cities with scant traffic channels, HSR construction may lead to a wide promotion of tourism development, which is similar to the so-called catch-up effect (Liu and Zhang 2018; Xu et al. 2018; Gao et al. 2019). Generally, non-resource-based cities always pose abundant tourism resources and have a higher reliance on tourism, but resource-based cities tend to be tourism-lagged. Similarly, owing to the disparity in tourism endowment, HSR’s effects on tourism in resource-based and non-resource-based cities would be heterogeneous; the former benefits little, whereas the latter gains more (Gao et al. 2019). Compared with large cities with more HSR stations, small cities should benefit more from connecting to the HSR system, which satisfies the diminishing marginal utility of HSR construction. Accordingly, on the divergence of the territorial impacts of HSR, we proposed the following hypotheses:

Hypothesis 2a: HSR has more positive effects on tourism development in central and western cities than in eastern cities in China.

Hypothesis 2b: HSR has more positive effects on tourism development in non-resource-based cities than in resource-based cities.

Hypothesis 2c: HSR has more positive effects on tourism development in small cities than in large cities.

Empirical strategy

Empirical framework

Referring to Gao et al. (2019), Zhang et al. (2020), and Deng et al. (2020), we treated HSR connection as a quasi-natural experiment and adopted a DID method. We constructed a benchmark model as follows:

\[
\text{tourism development}_i = \alpha + \beta_1 \text{HSR}_i + \gamma Z + \mu_i + \lambda_t + \epsilon_{it} \tag{1}
\]

The dependent variable \( \text{tourism development} \) can be measured by tourist arrivals or tourism revenue. \( \text{HSR}_i \) is the city-level HSR connection: \( \text{HSR}_i = 1 \) if a city has opened HSR service in year \( t \); otherwise, \( \text{HSR}_i = 0 \). \( Z \) covers all the control variables. \( \mu_i, \lambda_t, \) and \( \epsilon_{it} \) size the time-fixed effect, individual-fixed effect, and the error term, respectively. To avoid potential heteroscedasticity and serial correlation among others, we clustered standard errors at the city level. We defined the “connected” city as a city that has at least one station on the HSR line.

Data and variables

City-level tourism development

As a tertiary industry, tourism is closely related to and difficult to distinguish from related industries, such as the transport and catering industry (Campa et al. 2016). Based on the discussion in the “Literature review and hypothesis” section, tourism development can be measured from the perspectives of tourist arrivals, distance, revenue, and scenic spots (Campa et al. 2016; Gao et al. 2019; Zhang et al. 2020). Moreover, it is reasonable to separately discuss domestic and foreign tourism development. Similar to Chen and Haynes (2015), we mainly considered domestic tourist arrivals, domestic tourism revenue, international tourist arrivals, and international tourism revenue as the measurement of city-level tourism development. Given that international tourists cannot substitute flying with HSR travel, the influence of the
opening of HSR on international tourism is small (Pagliara et al. 2015; Campa et al. 2016). Therefore, we considered the relevant data of domestic tourism as the main research object.

City-level HSR connections

Many indices have been used to proxy HSR (Albalate and Fageda 2016; Gao et al. 2019; Zhang et al. 2020). Given the data limitations and potential endogeneity of indexes, we followed Gao et al. (2019) and set a dummy variable to proxy HSR. Specifically, HSR = 1, if a city is connected to the HSR network; otherwise, HSR = 0.

City-level control variables

Based on previous studies (Campa et al. 2016; Gao et al. 2019; Yin et al. 2019; Deng et al. 2020), we introduced per capita income, industrial structure, population density, FDI, education level, financial expenditure, wage level, and the number of 5A scenic spots into the models as control variables. The detailed definitions of the control variables are presented in Table 2. All of the variables were derived from the China Statistical Yearbook, China City Statistical Yearbook, The Yearbook of China Tourism Statistics, and Statistical Communiqué on National Economic and Social Development of Cities (2004–2017). Variables involving prices were deflated into the year 2000.

Empirical findings

Benchmark results

Table 3 gives the benchmark regression results of the effect of HSR connections on tourism development. Models (1) and (2) are the basic results for all cities; models (3) and (4)
provide the results for the samples excluding municipalities; and models (5) and (6) show the results for the samples excluding municipalities, provincial capitals, and specific plan-oriented cities.

The positive and significant coefficients of HSR in Table 3 implied that HSR had a dramatically positive impact on cities’ tourism development. That is, connecting to HSR and HSR station construction enhanced connectivity and accessibility among cities in China, bringing in more tourism travelers and a boom in local tourism development. Our results affirmed the positive effect of HSR in Hypothesis 1 and agreed with the conclusions in Gao et al. (2019) and Jin et al. (2020). The value of \( p \) in model (2) indicated that, with ceteris paribus, the domestic tourist arrivals in the cities with HSR increased by 4.25% on average compared with cities without HSR. Consequently, after the 2008 HSR opening, connecting to HSR averagely increased domestic tourist arrivals by 0.47% (4.25%/9) annually.

As for the control variables, according to model (2), GDP per capita had a significant and positive coefficient. Therefore, economic development had a positive effect on tourism development, also confirming the results of Gao et al. (2019). The positive coefficients of the ratio of FDI and ratio of financial expenditure indicated that financial support can improve regional tourism facilities and the environment and could drive local tourism development. The ratio of college students showing a negative effect on tourist arrivals may be attributed to the fact that younger people may prefer to travel to other cities. The unexpected coefficient of the number of 5A scenic spots indicated that cities have not made good use of local 5A scenic spots to attract tourists (Gao et al. 2019). The coefficients of the other variables were insignificant, indicating an absence of influence on tourism development.

### Table 3 Effect of HSR on tourism: DID method

| Dependent variable: Log (domestic tourist arrivals) | All cities | Cities excluding municipalities | Cities excluding municipalities, provincial capitals, and specific plan-oriented cities |
|----------------------------------------------------|------------|---------------------------------|-------------------------------------------------|
| (1) HSR                                            | 0.2362***  | 0.2363***                       | 0.2063***                                       |
| (0.0568)                                           | (0.0572)   | (0.0221)                        | (0.0669)                                        |
| GDP per capita                                     | 0.9409***  | 0.9588***                       | 0.6833***                                       |
| (0.3157)                                           | (0.3161)   | (0.1048)                        | (0.3134)                                        |
| Share of tertiary industry                         | 0.0115***  | 0.0122***                      | 0.0099***                                       |
| (0.0042)                                           | (0.0042)   | (0.0006)                        | (0.0041)                                        |
| Population density                                 | 0.5988     | 0.6308                          | 0.6861                                          |
| (0.3946)                                           | (0.3983)   | (0.2690)                        | (0.5293)                                        |
| Ratio of FDI                                       | 0.0228***  | 0.0242***                       | 0.0239***                                       |
| (0.0068)                                           | (0.0069)   | (0.0059)                        | (0.0065)                                        |
| Ratio of college students                          | −0.0082    | −0.0129                         | 0.0033                                          |
| (0.0256)                                           | (0.0262)   | (0.0202)                        | (0.0356)                                        |
| Ratio of financial expenditure                     | 0.0313**   | 0.0310**                       | 0.0200*                                         |
| (0.0122)                                           | (0.0121)   | (0.0038)                        | (0.0115)                                        |
| Wage level                                         | 0.7982***  | 0.7914***                       | 0.9690***                                       |
| (0.2051)                                           | (0.2049)   | (0.0489)                        | (0.2159)                                        |
| Number of 5A scenic spots                           | −0.0307    | −0.0048                         | −0.0155                                         |
| (0.0283)                                           | (0.0267)   | (0.0178)                        | (0.0218)                                        |
| Constant                                           | −14.2171***| −14.5201***                    | −13.9197***                                     |
| (3.0363)                                           | (3.0459)   | (1.9640)                        | (3.4883)                                        |
| City fixed effect                                  | Yes        | Yes                             | Yes                                             |
| Year fixed effect                                  | No         | No                              | Yes                                             |
| Observations                                       | 3546       | 3497                            | 3098                                            |
| R-squared                                          | 0.9282     | 0.9518                          | 0.9197                                          |

***, ***, and * denote significance at 1%, 5%, and 10% levels, respectively. Standard errors are clustered at the city level and are reported in parentheses. The following tables are the same. The regressions are controlled for city fixed effects in columns 1, 3, and 5, and for city and year fixed effects in columns 2, 4, and 6.
To provide insight into the main influence of HSR on tourism, we introduced other indexes of tourism development into the baseline model as shown in Table 4. The HSR had a positive influence on local tourism development, which can be attributed to the main effect of HSR on domestic tourist arrivals. In the regression with other tourism indicators, the HSR shows no significant impact on cities’ tourism revenue and international tourist arrivals. In other words, the effect of HSR connection was mainly observed for domestic tourist arrivals rather than tourist revenue and foreign arrivals, partly confirming the conclusions of Gao et al. (2019). The reason may be similar to the claim of Guirao and Campa (2016): a constrained indicator of tourists may be the main reason for the lack of significant effects of HSR on tourism in Spain.

### Heterogeneity

The heterogeneity of HSR’s effect is a critical issue (Pagliara and Mauriello 2020). According to the analysis in the “Framework and hypothesis” section, we identified the heterogeneity of the HSR connection effect from the perspectives of geographical location, city size, and resource endowment (shown in Table 5).

Table 5 shows that the coefficients of HSR are strikingly positive in models (2) and (3), but not in model (1), implying that connecting to HSR has a positive effect for the central and western cities but not for the eastern cities in China. As less developed regions, central and western China gain more profits by connecting to HSR. Liu and Zhang (2018) pointed out that the opening of HSR can reduce travel time, and given that the western region itself has a large travel time base, it could benefit greatly from HSR. Although the opening of HSR could reduce accessibility gaps between regions, the differences persist. Meanwhile, the eastern regions already have mature transportation systems, high connectivity and accessibility of cities, and flourishing tourism, thereby gaining less from HSR connection. Hypothesis 2a was thus confirmed.

The results of models (5) and (6) showed a significant positive coefficient in the resource-based city regression and a non-significant coefficient in the non-resource-based city regression. Therefore, non-resource-based cities can enjoy the benefits of HSR, whereas resource-based cities do not. This regional disparity may be because non-resource-based cities have abundant tourism resources and a higher reliance on tourism, whereas resource-based cities tend to be tourism-lagged. Owning to the difference in tourism endowments, HSR’s effects on tourism in resource- and non-resource-based cities were heterogeneous: the former benefited little, whereas the latter gained more. The results were similar to those reported by Gao et al. (2019). Hypothesis 2b was therefore supported.

In models (6), (7), and (8), the HSR had a significant coefficient in the regression with small cities but a non-significant coefficient with large and medium-sized cities. That is, in China, only small cities could benefit from the connectivity and accessibility brought by connecting to HSR. These results were contrary to the findings in Spain by Albalate et al. (2017). The potential explanation behind these phenomena is that for large and medium cities, it may be difficult to promote tourism by connecting to HSR as a convenient transportation network. Thus, we observed a marginal positive utility of HSR in small cities featuring primary transport networks and lower connectivity. Hypothesis 2c was proved to some degree. In general, not all cities had their tourism boosted by HSR. In the process of HSR development, attention should be paid to avoiding the mismatch of demand and deliberate choices of overinvestment, overdesign, and over-quality (Beria et al. 2016). Ongoing endeavor is needed to optimize the composition of different sized cities and towns in the HSR plan of China (Liu and Zhang 2018).
Robustness checks

To prove the validity of our results, we performed robustness checks according to Gao et al. (2019) and Zhang et al. (2020). The corresponding results are listed in Table 6.

In Table 6, model (1) gives the regression results with the alternative measures of tourism development, using domestic tourist arrivals per capita, calculated as the rate of the ratio of domestic tourist arrivals to year-end population. Model (2) shows the results for the sample excluding outliers. Model (3) provides the results based on the samples after 2005, and model (4) shows the results with an alternative measures of industrial structure using a comprehensive index of industrial structure (CIIS), where

\[ CIIS = \frac{\text{Output value of primary industry}}{\text{GDP}} + 2 \times \frac{\text{Output value of secondary industry}}{\text{GDP}} + 3 \times \frac{\text{Output value of tertiary industry}}{\text{GDP}}. \]

Model (5) displays the results of adding highway passengers to the existing control variables, and model (6) documents the results of considering the impact of airports by introducing a dummy variable for whether the city has a civil aviation airport. All the coefficients of the main explanatory variable HSR were significantly positive in the six models. Regardless of some adjustments to the measurement, regression method, sample selection, and variable controlling, the HSR connection makes positive promotion on local tourism development in China. Hence, the core conclusion of the above analysis has been confirmed and is robust and reliable.

Estimation with PSM-DID

To enhance the cogency of the main conclusion, we used the PSM-DID model to solve the problem of sample selection bias. We obtained the propensity score by a logit regression with the control variables in the benchmark model as the explanatory variables and HSR as the explained variable. HSR and non-HSR cities were matched by their propensity score. Table 7 gives the results. The advantage of the PSM-DID is that it can ameliorate the potential systematic disparities in HSR and non-HSR cities.

In models (1) to (7), the matching data are the cross-sectional data of the control variables in 2003, 2004, 2005, 2006, and 2007; the average value in 2003–2007; and the average value before the connection and opening of the HSR, respectively. The coefficients of HSR were significant at the 5% level and located in the interval (0.0469, 0.1182) in the seven models. Thus, HSR connection and station construction can stimulate domestic tourist arrivals and further enhance local tourism development. These results bolstered the robustness of the core conclusion in the above analysis.
To avoid the probable bias from the pre-existing effect, referring to the decision on HSR connection and station construction being closely based on economic and social conditions (Faber 2014), we conducted a placebo test (Ma et al. 2021) with false HSR opening times by assuming that the opening year of HSR was 2004, 2005, 2006, and 2007, respectively. The corresponding outcomes in Table 8 show insignificant coefficients of false HSR and demonstrate the lack of a preexisting confounding factor that disturbs our core conclusion.

### Table 6 Robustness checks

|                   | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  |
|-------------------|------|------|------|------|------|------|
|                   | HSR  | Controls | City fixed effect | Year fixed effect | Observations | R-squared |
|                   | 0.0370* (0.0215) | 0.0473** (0.0215) | Yes | Yes | 3546 | 0.9430 |
|                   | 0.0336* (0.0196) | 0.0423* (0.0221) | Yes | Yes | 3546 | 0.9507 |
|                   | 0.0377* (0.0219) | 0.0401* (0.0222) | Yes | Yes | 3546 | 0.9663 |
|                   | 0.0377* (0.0219) | 0.0401* (0.0222) | Yes | Yes | 3546 | 0.9518 |
|                   | 0.0377* (0.0219) | 0.0401* (0.0222) | Yes | Yes | 3546 | 0.9521 |
|                   | 0.0377* (0.0219) | 0.0401* (0.0222) | Yes | Yes | 3546 | 0.9519 |

Without special statement, outcome variable and control variables are the same with Table 5.

### Table 7 Effects of HSR on tourism: PSM-DID method

|                   | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  |
|-------------------|------|------|------|------|------|------|------|
|                   | 2003 | 2004 | 2005 | 2006 | 2007 | 2003–2007 | 2003–2007 |
|                   | HSR  | Controls | City fixed effect | Year fixed effect | Observations | R-squared | R-squared |
|                   | 0.0500** (0.0233) | Yes | Yes | Yes | 3294 | 0.9482 | 0.9416 |
|                   | 0.0538** (0.0236) | Yes | Yes | Yes | 2843 | 0.9436 | 0.9416 |
|                   | 0.0613** (0.0238) | Yes | Yes | Yes | 2960 | 0.9430 | 0.9416 |
|                   | 0.0469** (0.0235) | Yes | Yes | Yes | 3147 | 0.9479 | 0.9416 |
|                   | 0.0580** (0.0239) | Yes | Yes | Yes | 3043 | 0.9397 | 0.9416 |
|                   | 0.0616** (0.0240) | Yes | Yes | Yes | 3058 | 0.9417 | 0.9416 |
|                   | 0.1182*** (0.0324) | Yes | Yes | Yes | 2033 | 0.9416 |

Outcome variable and control variables are the same with Table 5. The matching variables are the control variables in the baseline model. We use the second-order nearest neighbor matching in the caliper. Columns 1–7 show the matching data for the control variables in 2003, 2004, 2005, 2006, and 2007; average value in 2003–2007; and average value before HSR connection, respectively.

### Table 8 Estimation results for false opening years of HSR

|                   | Sample period: 2003–2007 |
|-------------------|--------------------------|
|                   | 2004 | 2005 | 2006 | 2007 |
|                   | (1)  | (2)  | (3)  | (4)  |
|                   | False HSR | Controls | City fixed effect | Year fixed effect |
|                   | −0.0483 (0.0793) | Yes | Yes | Yes |
|                   | −0.0204 (0.0483) | Yes | Yes | Yes |
|                   | −0.0038 (0.0420) | Yes | Yes | Yes |
|                   | 0.0111 (0.0428) | Yes | Yes | Yes |
|                   | Observations | 1235 | 1235 | 1235 | 1235 |
|                   | R-squared | 0.9391 | 0.9390 | 0.9390 | 0.9390 |

The sample period in this table is 2003–2007. Outcome variable and control variables are the same with Table 5. In columns 1–4, the key independent variables are artificial pseudo variables, assuming that the opening year of HSR is 2004, 2005, 2006, and 2007, respectively. The treated cities remained unchanged.

### Placebo test

#### Setting the false opening year of HSR

To avoid the probable bias from the pre-existing effect, referring to the decision on HSR connection and station construction being closely based on economic and social conditions (Faber 2014), we conducted a placebo test (Ma et al. 2021) with false HSR opening times by assuming that the opening year of HSR was 2004, 2005, 2006, and 2007, respectively. The corresponding outcomes in Table 8 show insignificant coefficients of false HSR and demonstrate the lack of a preexisting confounding factor that disturbs our core conclusion.

#### Randomly generated HSR status

To separate the benchmark regression results from “change findings caused by the influences of missing variables,” we referred to Li et al. (2016) and Jia...
et al. (2021) to execute the placebo test by randomly selecting HSR opening cities. First, the same number of cities in the treatment group is randomly selected according to the number of cities with HSR opening each year. Second, with the samples, a dummy treatment variable \( HSR_{it}^{false} \) is set artificially. Finally, aligned with the benchmark model, regressions are 500 times and 1000 times based on the dummy variable \( HSR_{it}^{false} \). Figure 3 shows the distribution of the estimated coefficients and their \( p \)-values.

The estimated coefficients of \( HSR_{it}^{false} \) were distributed around 0. The means of the coefficients in the 500 and 1000 simulations were \(-0.000573\) and \(0.000148\), respectively. Comparatively, the coefficient in the benchmark regression was \(0.0425\), larger than most coefficients in the simulations. Regarding the distribution of the \( p \)-value, in the 500 simulations, only 23 coefficients of \( HSR_{it}^{false} \) were located in the interval larger than 0.0425, and the corresponding \( p \)-value was equal to or less than 0.1. That is, in the 500 simulations, the coefficient of benchmark regression was true 95.4\% (1–23/500) of the time. Similarly, in the 1000 simulations, only 53 coefficients of \( HSR_{it}^{false} \) were in the interval larger than 0.0425, and the corresponding \( p \)-value was equal to or less than 0.1, indicating that the coefficient of benchmark regression was true 94.7\% (1–53/1000) of the time. The above results implied that the positive link from HSR connection to tourism development avoided the omitted variable bias at the 90\% level at least.

Valid and endogeneity of HSR

Validity of parallel trend assumption and dynamic DID

Following Li et al. (2016) and Dong (2018), we probed the validity of the parallel trend assumption by event study. The econometric model was constructed as follows:

\[
tourism_{it} = \alpha + \sum_{k=5}^{5} \beta_k D_{it}^k + \gamma Z_i + \mu_t + \lambda_i + \epsilon_{it}
\]

(2)
where \( i \) and \( t \) denote city and year, respectively. The explained variable \( \text{tourism development} \) it was consistent with the “Empirical framework” section. \( D_k \) it is a dummy for the “event,” HSR opening. Considering the data interval (2003–2016) and the original year of HSR opening (2008), we documented the opening year of HSR of a city with \( s_i \); if \( t - s_i \leq -5 \), \( D_{it}^5 = 1 \); otherwise, \( D_{it}^5 = 0 \). Generally, if \( t - s_i = k \), \( D_{it}^k = 1 \); otherwise, \( D_{it}^k = 0(k \in [-5, 5]) \). With \( t - s_i \geq 5 \), then \( D_{it}^5 = 1 \); otherwise, \( D_{it}^5 = 0 \). The coefficient \( \beta_k \) captures the effect that HSR opening has on local tourism development. Other variables in Eq. (2) are the same as in Eq. (1). The regression of Eq. (2) documents the validity of the parallel trend assumption and the effect of dynamic DID.

From Fig. 4, the results did not reject the null hypothesis that cities had no systematic difference before HSR opening, thereby verifying the validity of the parallel trend assumption. In the 0 and 1 year of HSR opening, the estimations of \( \beta_k \) were positive, with their 95% confidence intervals at (0.0003, 0.1637) and (0.0009, 0.1773), respectively. Thus, tourism development increased by 0.082 and 0.0891 in the original year and the first year of HSR opening, respectively. From the year 2, the coefficients of \( D_k \) became positive but non-significant, meaning that the positive effect of HSR weakened and disappeared after 2 years of HSR opening. Firstly, the opening of HSR shows the benefits of substitution and space–time compression on local tourism development. Over time, the substitution effect weakened and the Matthew effect (Siphonage) led to more little cities failing to benefit from HSR opening. As the result, the average effect of HSR on cities’ tourism ended up being negative and undesired. Overall, the conclusion of the dynamic DID was reasonable and acceptable.

### Estimation with instrumental variables

After determining the causality between HSR and tourism development, we aimed to discuss the bias from the underlying endogeneity of HSR opening. Referring to Faber (2014)
and Wang et al. (2018), we constructed an IV based on cities’ slope values. The superiority of slope as an IV is that, as an exogenous geographic variable, slope has no direct nexus with economic variables and considering the cost and difficulty of HSR construction, slope shows a direct relationship with the HSR connection and station construction. Given the invariance of cities’ slopes, we constructed a dummy variable with the product of slope and year as the IV of HSR. The results are presented in Table 9.

Ten coefficients of the dummy variables were apparently negative in the first stage of the type I models. Only one coefficient was negative and significant in the first stage of the type II models. The first-stage $F$ values were 208.854 and 92.912, respectively, indicating that the slope of cities increased the cost of HSR and had a negative effect on HSR connection. In the second stage of the IV regressions, the coefficient was dramatically positive in the type I model, indicating that HSR connection had a positive effect on cities’ tourism development, consistent with the core conclusion of our study. However, the coefficient in the type II model was positive but non-significant. Considering the $F$ test, it is reasonable to claim that in the IV regressions, HSR promotes local tourism development and adheres to the above results.

**Conclusion**

Our study focused on the effects from HSR to tourism development and analyzed the city-level data covering 2013–2016. We found that, first, HSR connections had an overall positive effect on cities’ tourism development. Second, there is distinct heterogeneity in the HSR’s effects from multidimensional sights. Specifically, central and western cities, non-resource-based cities, and small cities benefit more from the HSR opening. Finally, cities with and without HSR had no systematic disparity, and HSR connection apparently promoted local tourism development in the 0 and 1 year of HSR opening but tended to fail to promote local tourism in the long term.

Based on the above conclusion, some adjustments are proposed to optimize China’s HSR management and the design of HSR systems. The positive effect of HSR on tourism development should be adequately considered in the strategy and evaluation of HSR connections and the construction of a city to ensure a comprehensive HSR management. In the design of an optimized HSR system, the construction should consider the heterogeneity among the effects of HSR to match local conditions and encourage a catch-up effect. From a long-term perspective, the HSR environment and travel attributes should be improved by connecting the structure of the HSR system with local tourism features to better drive the recovery of tourism in the post-epidemic era.

**Author contribution** All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Hua Zhang and Siyu Chen. The first draft of the manuscript was written by Xiaoxiao Zhang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

**Funding** This work was supported by the National Natural Science Foundation of China (Grant No. 71804001, 71703075).

**Data availability** The datasets used or analyzed during the current study are available from the yearbooks or the corresponding author on reasonable request.

**Code availability** Not applicable.

**Declarations**

**Ethics approval and consent to participate** Not applicable.

**Consent for publication** Not applicable.

**Competing interests** The authors declare no competing interests.

**References**

Albalate D, Fageda X (2016) High speed rail and tourism: empirical evidence from Spain. Transp Res Part A-Policy Pract 85:174–185

Albalate D, Campos J, Jiménez J (2017) Tourism and high speed rail in Spain: does the AVE increase local visitors? Ann Tour Res 65:71–82. https://doi.org/10.1016/j.antrs.2017.05.004

Bera P, Grimaldi R, Albalate D, Bel G (2016) Delusions of success: costs and demand of high-speed rail in Italy and Spain. Transp Policy 68:63–79. https://doi.org/10.1016/j.tranpol.2016.08.011

Campana JL, López-Lambas M, Guirao B (2016) High speed rail effects on tourism: Spanish empirical evidence derived from China’s modelling experience. J Transp Geogr 57:44–54. https://doi.org/10.1016/j.jtrangeo.2016.09.012

Carteri A, Pariota L, Henke I (2017) Hedonic value of high-speed rail services: quantitative analysis of the students’ domestic tourist attractiveness of the main Italian cities. Transp Res Part A-Policy Pract 100:348–365. https://doi.org/10.1016/j.tra.2017.04.018

Castillo-Manzano JI, Castro-Nuño M, López-Valpuesta L, Pedregal-Tercero DJ, Garrido-Michó JM (2018) High speed rail: fast tracking tourism in the EU? Ann Tour Res 71:64–66. https://doi.org/10.1016/j.antrs.2018.02.005

Chen Z, Haynes K (2015) Impact of high-speed rail on international tourism demand in China. Appl Econ Lett 22:57–60. https://doi.org/10.1080/13504851.2014.925043

China Railway Maps (2021) By 2020, China keeps the world’s largest high speed rail (HSR) network with a length totaling over 37,900 km (23,550 mi). https://www.travelchinaguide.com/china-trains/high-speed/
Dai X, Xu M, Wang N (2018) The industrial impact of the Beijing-Shanghai high-speed rail. Travel Behav Sci 12:23–29. https://doi.org/10.1016/j.tbs.2018.03.002
Delaplace M, Pagliara F, Perrin J, Mermet S (2014) Can high speed rail foster the choice of destination for tourism purpose. Procedia Soc Behav Sci 111:166–175. https://doi.org/10.1016/j.prosocbs.2014.01.049
Deng T, Gan C, Du H, Hu Y, Wang D (2020) Do high speed rail configurations matter to tourist arrivals? Empirical evidence from China’s prefecture-level cities. Res Transp Econ 100952. https://doi.org/10.1016/j.retrec.2020.100952
Dong XF (2018) High-speed railway and urban sectoral employment in China. Transportation Research Part A: Policy and Practice 116:603–621. https://doi.org/10.1016/j.tra.2018.07.010
Faber B (2014) Trade integration, market size, and industrialization: evidence from China’s national trunk highway system. Rev Econ Stud 81:1046–1070. https://doi.org/10.1093/restud/rdu010
Gao Y, Su W, Wang K (2019) Does high-speed rail boost tourism growth? New evidence from China. Tour Manage 72:220–231. https://doi.org/10.1016/j.tourman.2019.05.091
Guirao B, Campa JL (2016) Cross effects between high speed rail lines and tourism: looking for empirical evidence using the Spanish case study. Transp Res Procedia 14:392–401. https://doi.org/10.1016/j.trpro.2016.05.091
Jia R, Shao S, Yang L (2021) High-speed rail and CO2 emissions in urban China: a spatial difference-in-differences approach. Energy Econ 99:105271. https://doi.org/10.1016/j.eneco.2021.105271
Jin S, Yang J, Wang E, Liu J (2020) The influence of high-speed rail on ice–snow tourism in northeastern China. Tour Manage 78:104070. https://doi.org/10.1016/j.tourman.2019.104070
Li N, Chen J, Tsai I, He Q, Chi S, Lin Y, Fu T (2016) Potential impacts of electric vehicles on air quality in Taiwan. Sci Total Environ 566–567:919–928. https://doi.org/10.1016/j.scitotenv.2016.05.015
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
Li H, Dong X, Jiang Q et al. (2021) Policy analysis for high-speed rail in China: Evolution, evaluation, and expectation. Transport Policy 106:37–53
Lin Y (2017) Travel costs and urban specialization patterns: evidence from China’s high speed railway system. J Urban Econ 98:98–123. https://doi.org/10.1016/j.jue.2016.11.002
Liu L, Zhang M (2018) High-speed rail impacts on travel times, accessibility, and economic productivity: a benchmarking analysis in city-cluster regions of China. J Transp Geogr 73:25–40
Ma L, Niu D, Sun W (2021) Transportation infrastructure and entrepreneurship: evidence from high-speed railway in China. Econ Rev 65:101577. https://doi.org/10.1016/j.econrev.2020.101577
Matas A, Raymond J, Roig J (2020) Evaluating the impacts of HSR stations on the creation of firms. Transp Policy 99:396–404. https://doi.org/10.1016/j.tranpol.2020.09.010
National Bureau of Statistics of China
Pagliara F, Mauriello F (2020) Modelling the impact of high speed rail on tourists with geographically weighted Poisson regression. Transp Res Part A-Policy Pract 132:780–790. https://doi.org/10.1016/j.tranpol.2019.12.025
Pagliara F, Pietra A, Gómez JJ, Vassallo JM (2015) High speed rail and the tourism market: evidence from the Madrid case study. Transp Policy 37:187–194. https://doi.org/10.1016/j.tranpol.2014.10.015
Pagliara F, Mauriello F, Garofalo A (2017) Exploring the interdependencies 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.tranpol.2017.09.022
Wang X, Huang S, Zou T, Yan H (2012) Effects of the high speed rail network on China’s regional tourism development. Tour Manag Perspect 1:34–38. https://doi.org/10.1016/j.tmp.2011.10.001
Wang D, Niu Y, Qian J (2018) Evolution and optimization of China’s urban tourism spatial structure: a high speed rail perspective. Tour Manage 64:218–232. https://doi.org/10.1016/j.tourman.2017.08.010
Xu W, Zhou J, Yang L, Li L (2018) The implications of high-speed rail for Chinese cities: connectivity and accessibility. Transp Res Part A-Policy Pract 116:308–326. https://doi.org/10.1016/j.tranpol.2018.06.023
Yin P, Lin Z, Prideaux B (2019) The impact of high-speed railway on tourism spatial structures between two adjoining metropolitan cities in China: Beijing and Tianjin. J Transp Geogr 80:102495–102495
Zhang A, Liu L, Liu G (2020) High-speed rail, tourist mobility, and firm value. Econ Model 90:108–116. https://doi.org/10.1016/j.econmod.2020.05.004

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.