The Impact of High-Speed Railway on Urban Housing Prices in China: A Network Accessibility Perspective

Xueli Liu[a], Chunxia Jiang[b], Feng Wang[c,*], Shujie Yao[c]

Abstract

Applying the social network analysis (SNA) approach, this paper investigates the impact of high-speed rail (HSR) on urban housing prices from the perspective of network accessibility. Based on a sample of 285 cities in China over 2009-2017, we find a positive effect of HSR network accessibility on urban housing prices. An increase in HSR network accessibility (i.e., degree point or closeness centrality) by one standard deviation causes about a 10.3% increase in average housing prices. Evidence also suggests that this effect varies across regions and housing types. Our results are robust to different model specifications and alternative measures of HSR accessibility. The findings offer insights into the space-time economic laws with important policy implications regarding spatial disparities and regional economic convergence in China.

JEL: O18, R30, R41

Key words: Network accessibility; High-speed railway; Urban housing price; Social network analysis

---

[a] Xueli Liu (xueliliu@tongji.edu.cn), Tongji University, College of Architecture and Urban Planning, Shanghai 200092, China; [b] Chunxia Jiang (chunxia.jiang@abdn.ac.uk), University of Aberdeen, Business School, Aberdeen, AB24 3QY, UK; [c] Feng Wang* (**corresponding author**, wangfeng2008@cqu.edu.cn), and Shujie Yao (yaoshujie@cqu.edu.cn), Chongqing University, School of Economics and Business Administration, Institution of Energy Economics, Chongqing 400030, China.
This research is financially supported by the Natural Science Foundation of China (71973019, 71673033, 71973148), the Social Science Foundation of China (18ZDA005, 19ZDA082), Chongqing Social Science Planning Project (2017YBJJ024), and Fundamental Research Funds for the Central Universities (NO.2020CDJSK02PT26, NO.2019CDSKXYJG0037). The authors are solely responsible for any error or omission herein.

1. Introduction

Transportation infrastructure plays an important role in reshaping economic activities. This happens as the changes in accessibility and proximity to the transportation network induce the location effect and the agglomeration effect. An improved transportation network increases the number of reachable destinations for a given journey time. As results, business operations become more efficient and local and regional economy prosper. Early research was inspired by the introduction of high-speed rail (HSR) in the developed world and the effects of transportation on outcomes (e.g., population density, land rents and output) have been well-documented in the literature.[1] The positive externalities of HSR include increased land value, higher economic growth and productivity, flexible labour market, and reduced inequality (e.g., Duranton and Turner, 2012; Michaels et al., 2012). The recent rapid expansion of HSR in the developing countries has refuelled the research interest on the relationship between transportation infrastructure and economic activities (e.g., Lin, 2017; Shao et al., 2017).

A strand of research addresses the impact of HSR system on property or land prices. Existing studies mainly focus on HSR station area with inconclusive findings. For instance, Hensher et al. (2012) find a positive link between HSR station accessibility and land/property values in 8 out of 15 cities from 8 countries, no corelation in 4 cities, and a negative relationship in 3 cities. Some researchers extend the vision to the HSR corridor and/or the regional network level, suggesting that the impact varies with city size (Chen and Haynes,2015), local economy (Wang et al., 2018), and political considerations (Low and Lee, 2020). The rapid expansion of HSR has resulted in huge variations in the accessibility to cities via direct and indirect connections to the whole network, while the research on the network impact on housing prices is inadequate. This study attempts to fill in the gap by examining the effect of HSR network accessibility on urban housing prices.

The HSR network promotes intercity resources transfer, migration and labors movement. The HSR network induces huge savings in travel time and costs, which makes daily commuting between smaller cities and bigger central cities feasible. The "centripetal force effect" indicates that resource allocation becomes more polarized among large and small cities and housing prices in the larger centrally located cities increase. The "centrifugal force effect" indicates that production resources and population spread from the core metropolises to small cities along

[1] See Redding and Turner (2015) for an excellent literature review.

the corridors of HSR when the network substantially shorten the space-time distance between metropolises and small cities (Hall, 2009), which may pushes up housing prices in small cities. In short, housing prices are likely to increase when cities have improved access to the HSR network. As such, this study tests the hypothesis that *HSR network accessibility has a positive impact on urban housing prices*.

Since 2009, China experienced a rapid booming in HSR construction and the HSR network has gradually taken shape to serve many cities, which provides a unique observation to study this proposition. Figure 1 shows the map of the existing HSR network in China by the end of 2017 when the nationwide operating HSR lines were 25,000 km, with the 5 year-categories colored from light to dark representing the pace of network development. China has built the world's longest HSR network and currently accounts for two-thirds of the world's total HSR networks. This well-connected network has improved cities' HSR network accessibility significantly.[2] Taking advantage of this excellent setting with the aid of robust empirical estimation strategy, this study examines how the rapid development of HSR affect urban housing prices.

[Insert Figure. 1 Here]

This paper extends our knowledge by providing empirical evidence for the impact of HSR on the prosperity of the urban housing markets from a novel perspective of HSR network accessibility. Existing studies mainly examine the impact of HSR presence (i.e., a station), travel time, costs, and travel frequency, using traditional accessibility indices/indicators but ignoring the effect of network accessibility. Martínez and Givoni (2012) point out that any examination of an HSR line must consider a wider geographic area, not just the related cities. Building on the existing literature, we apply a social network analysis (SNA) approach (Wasserman and Faust, 1994) to measure HSR network accessibility with due consideration to the integration between transport networks, especially between conventional rail (CR) and HSR networks. Hence, we examine the spill-over effect of the HSR network on urban housing prices beyond the dimensions of a station, a line or a relatively small region.

The rest of this paper proceeds as follows. Section 2 reviews the literature. Section 3 outlines research methodology and describes the sample and data. Section 4 presents empirical results

[2] The majority of our sample cities have access to CR and all of them have access to highways before our study period (2009-2017). For cities without direct connection to HSR network, we use conventional rail or highway data to calculate the accessibility indices.

and Section 5 discusses and concludes.

2. Literature review

2.1 Accessibility: Definition and measurement

Previous studies define accessibility as the size of each node's opportunity to interact in a traffic network (Hansen, 1959) and the ability to reach desired locations at suitable times using necessary transportation (Moseley, 1979). More recent studies consider accessibility as the ability to reach a designated place at an appropriate time via transportation facilities (Geertman and van Eck, 1995). Extensive research efforts have been devoted to measuring accessibility. The most widely used methods include the distance metric method (Ingeam,1971), the equivalent line method (Breheny, 1978) and the gravity model method (Jensen and Stewart, 1977). Other methods measuring accessibility at the individual level include the space-time method based on time geography (Hagerstrand, 1970), the topological method by logical networks (Wheeler,1999), and the matrix-based topology method at a regional scale (Muraco,1972).

Empirically, researchers employ a variety of indices to proxy accessibility, including travel time (Martínez and Givoni, 2012), travel cost (El-Geneidy et al., 2016), daily frequency of trains (Liu et al., 2020), the number of passengers (Cascetta, 2020) and geographic distance (Shaw et al., 2014). Extant studies mainly measure the HSR accessibility at the station and/or the regional level. Kim et al. (2019) argue that station accessibility by transit is an important factor to understand long-term HSR demand. Martínez and Givoni (2012) use travel time to London as the main benchmark and find limited benefit from the improved accessibility to the proposed new line in terms of geographic spread.

The rapid expansion of HSR network creates huge variations in a city's accessibility via direct connections (station accessibility) and indirect network relationships to the whole transportation system. The extent to which the HSR network benefit can be materialized also depends on the services provided in each city. The HSR accessibility is a network issue and the measurement should include both local (station) and regional accessibility to address the distance attenuation effect (Liu et al., 2020). The literature is inadequate regarding the network accessibility effect.

2.2 HSR, land value and housing prices

The majority of existing studies provide evidence that land values and housing prices are affected by different types of intra-city transport infrastructures, such as metro (Lee et al., 2018; Rohit and Peter, 2018), sub-urban TOD (Mathur and Ferrell, 2013), road (Agarwal, 2015) and BRT systems (Deng and Nelson, 2010). Focus on conventional rail system, studies generally report a positive impact on property and land values, for instance, in Bangkok, Thailand (Chalermpong, 2007), Montreal, Canada (Dubé et al., 2013), Seoul, Korea (Cervero and Kang, 2011), Singapore (Diao, 2015), Wuhan, China (Xu et al., 2016) and Hongkong, China (He, 2020).

HSR is an extremely land-based transport infrastructure with considerable benefits such as travel time savings, productivity improvement, and the prosperity of land value and housing prices. HSR lines may also bring environmental problems like traffic congestion, electromagnetic radiation pollution, noise, and higher crime rates, which also have implication on land value and housing prices. Research on the impact of HSR on land value and housing prices primarily focus on HSR stations (Geng et al., 2015) and empirical evidence is inconclusive. Hensher et al. (2012) review the impact of HSR station accessibility on land and property values in 15 cities from eight countries, reporting a positive impact in eight cities (i.e., London Borough of Camden, Ashford, Lyon, Ciudad Real, Turin, Naples, Kyushu), a negative impact in three cities (i.e., Paris, Milan), and no impact in four cities (i.e., Florence, Berlin, Tainan). Scholars also find that the distance from the properties to HSR stations affect housing prices. Geng et al. (2015) report a positive impact on property prices within the range of 0.891–11.704 km from the HSR station, but a negative impact within the range 0.475– 0.891 km in Beijing (China). Diao et al. (2016) find that inner-city HSR station in Hangzhou shows a positive impact on residential property value while the suburban HSR station in Guangzhou has no significant impact in the short term. With the formation of HSR networks, more recent studies broaden the view to the HSR corridor and/or the region level. Andersson et al. (2010) find no overall increase in housing prices in seven metropolises along an HSR line in southern Taiwan when using intracity geographic distance to an HSR station as the proxy for accessibility. Chen and Haynes (2015) find that the Beijing-Shanghai HSR line has a significant regional impact (both local effects and spill-over effects) on housing prices in medium and small cities, but this effect is negligible in larger provincial capital cities.

In the literature, the HSR network impact on housing prices are under-researched. Against this backdrop, this paper attempts to fill in the gap by applying the SNA approach to measuring HSR accessibility from a wider transportation network perspective. Social network is a sociological concept, referring to the assemblage of social actors and their relationships. The SNA argues that the flow and acquisition of resources are closely related to the connections of the relationship, which determines the flow and allocation of resources (Marcus et al., 2006). The SNA approach can be applied at both the micro and macro levels. The micro SNA approach takes individuals as the subject and examines the relationship between a single actor and other actors, employing various network measures, such as point degree. The macro SNA approach takes the whole network as the subject and analyzes its overall characteristics, using closeness centrality as the accessibility measure (Liu et al., 2020). Following the literature, we treat every city as an actor and CR/HSR as network relationships and employ point degree and closeness centrality as the measures of HSR network accessibility, along with a traditional HSR accessibility measure of average travel time.

2.3 HSR network accessibility and housing prices: Hypothesis

According to the bid rent theory (Alonso, 1964), bid rent is the highest rent that a land user (resident or enterprise) is willing to pay for a certain urban land (a certain location). Competitors with more location-sensitivity and stronger ability to pay rents (i.e., commercial services) are more like to obtain the right to use the land or premise. Cook and Watson (2016) support the ripple effect in the regional spreading of housing price, based on the evidence that housing price changes in London lead to the housing price changes in the surrounding cities. These two theories work simultaneously and the overall impact of the HSR network on the property sector can take place through the primary real estate market (Button, 2012), the secondary real estate market (Chen and Haynes, 2015), and premise-related public services (Dong et al., 2020).

In this study, we focus on the relationship between the inter-city HSR network accessibility and urban housing prices. The cities located along HSR lines gain considerable location advantages. Time saving boosts factor mobility between and within cities, such as population, information and technology. The convenient transportation network allows the production and service resources to flow from small and medium-sized cities along the HSR corridors to core megapolises, characterized as the "centripetal force effect" (Givoni, 2007). Firms are more likely to locate their headquarters or R&D centers in centrally located bigger cities that offer better access to pools of talents (among other factors, such as information). HSR services shorten the space-time distances between different cities, which makes remote sourcing of labour and talents possible. The improved accessibility allows residents to work and live in different cities. Production resources and population can also spread from core megapolises to small cities along the corridors, leading to the "centrifugal force effect" (Garmendia et al., 2008).

Intercity resources transfer, migration and labor movement are likely to push up housing prices in cities with improved access to the HSR network. While the "centripetal force effect" tends to drive up housing prices in larger centrally located cities, people start to move to nearby medium or small-sized cities to reduce living costs as housing prices in those big cities become much more expensive and unaffordable to many households. Also, growing bigger cities are faced with worsening quality of life because of high levels of traffic congestion and pollution (Zheng and Kahn, 2013). HSR allows individuals to enjoy the benefits of urban agglomeration while avoid high housing costs and city's social costs by living in nearby cities. Thus, housing prices also tend to increase in small cities with improved access to HSR network. As such, we test the following hypothesis: *HSR network accessibility has a positive impact on urban housing price*.

3. Research methodology, sample, and data

3.1. Measuring network accessibility

3.1.1. Assumptions

We make certain assumptions when measuring accessibility. First, we consider the passenger flow of CR, HSR and a small part of highways, but excluding intercity passengers using other transport modes such as air, highway, and waterways due to data unavailability.[3] Second, our sample consists of 285 cities in mainland China (excluding Tibet, Hong Kong, Macao, and Taiwan) and the network accessibility is measured between these cities without considering their interactions with international cities. Third, as trains have different designed speeds as well as flexible running speeds (i.e., G, C or D oriented HSR and K, Z or number-oriented CR[4])

[3] For a very small number of cities without HSR or CR, we use the passenger flow of highway. The number of such city in 2009 was 32 while dropped to 11 in 2017.

[4] As our main research interest is the impact of HSR network accessibility on housing prices, we differentiate between HSR

and so do highway lanes have different speed limit, the same model may lead to variations in accessibility measure for the same trip. Following literature (Wang and Duan, 2018), we build a geographic cost raster and set 350km/h for G-oriented HSR train, 250 km/h for C-oriented HSR train, 200 km/h for D-oriented HSR train, 120 km/h for all types of CR train, and 90 km/h for highways based on the standards of the Ministry of Transport of China.[5] Fourth, we prioritize the lowest travel time over the comfort of services and the inconvenience of transfer. More comfortable, faster, but more expensive high-speed trains (i.e., G, C, D grade) are not necessarily preferred to less comfortable, slower but cheaper CR trains or highways. For example, if the travel time of an indirect HSR route between two cities is longer than that of a direct CR route, we assume passengers choose CR. Fifth, trains usually stop for 2 to 3 minutes at small stations between two cities. Such a time is largely neutralized when calculating the travel time of every city to all the other cities and our model ignores the parking time at small stations. Sixth, if there are more than one HSR station in a city, our model recognizes the city as one node and each point represents a city rather than a station.

Finally, and more importantly, among different models, such as B-space, P-space, C-space, we employ the L-space model to construct the HSR infrastructure network (Ferber et al., 2009). The L-space model reflects the direct connection of stations in a network (each station is represented by a node) and has no multiple links between two stations. A station connected to a larger number of peripheral stations is of high importance in the network. Unlike other models that focus on demand side timetable data or passenger flow data, L-space is a supply side model, which fits better the infrastructure characteristics of HSR service. In the B-space model, both routes and stations are represented by nodes and the model is more useful in the analysis of socio-economic connections of locations. When the B-space model sets the projection to station nodes, it is known as the P-space model. The neighbours of a P-space node are all stations that can be reached without changing means of transport. This model is more useful in the analysis of transitivity between stations. When the B-space model sets the complementary

and CR, but do not differentiate among different types of HSR (G, C, CR). CR and HSR have separate infrastructure. HSR trains use newly built or upgraded dedicated lines and HSR stations are associated with large scale development plans, which have significant impact on housing prices. CR travels on low-speed railway lines and there is no large-scale development plan in the surrounding area of CR stations.

[5] We employ the raster cost method for two main reasons. First, our focus is on the variations in housing price across cities and HSR time savings are mainly achieved during the high-speed inter-city journey. Second, data on actual travel time are unavailable for the sample period. The officially published National Rail Timetable of China for HSR only records actual travel time at a certain day rather than annual average (Ren et al., 2020), and the frequency of speed adjustment also cause complexities in measurement of actual travel time. An alternative source is Service Website of China's Railways(https://www.12306.cn) that also reports actual travel time on a daily basis without historical data on actual travel time.

projection to route nodes, it becomes the C-space model, where any two route nodes are neighbours if they share a common station. The C-space model is more appropriate to analyse transitivity across routes.

3.1.2. Measurement: Point degree and closeness centrality and average travel time

This paper employs commonly used point degree and closeness centrality (Wasserman and Faust, 1994) to measure the overall accessibility and network structure. The point degree of a node *i* is the number of other nodes directly connected to node *i* in the whole network, written as in Eq. (1):

$$Acc\_degree_i = \sum_{j=1}^{n} X_{ij} \tag{1}$$

where X_{ij} denotes the attribute of the connections between nodes *i* and *j*, and n is the total number of the nodes.

The closeness centrality of a point is the sum of the geodesic distance between the point and all the other points in the network.[6] It is a commonly used measure to evaluate a city's location advantage in a network. According to Freeman (1978), the formula is written as in Eq. (2):

$$Acc\_closeness_i = \frac{1}{\sum_{j=1;j\neq i}^{n} d_{ij}} \tag{2}$$

where d_{ij} denotes the shortest path distance of railway lines between node *i* and node *j*, and n is the total number of the nodes in a network, $Acc\_closeness_i$ represents the closeness centrality of point *i*, the larger the $Acc\_closeness_i$, the more convenient the route is.

The minimum impedance accessibility analysis is the most popular method in the literature as it quantitatively describes the object and intuitively reflects the space-time characteristics of accessibility. The smaller the value, the better the accessibility, and vice versa. The average of the minimum impedance accessibility from city *i* to all other cities in the network is the so-called the third network accessibility-average travel time. According to Ingram (1971), it can be calculated using Eq. (3):

[6] In an indirect network, the distance between two nodes is the number of connections in the shortest route between the two points, and the average shortest distance in the network is called geodesic distance (Freeman, 1978).

$$Acc\_time_i = \frac{1}{n-1}\sum_{j=1,j\neq i}^{n} a_{ij} \tag{3}$$

Where a_{ij} is defined as the shortest travel time between city i and j, Acc_time_i can be interpreted as a network accessibility between city i and a random city j, and n is the total number of the nodes in a network.

3.2 Empirical model specifications

The empirical estimation strategy to test our hypothesis is as follows. Firstly, the baseline model in Eq. (4) is employed to test the hypothesis directly. Different control variables are included to ensure the stability of the model and the robustness of results. Secondly, the difference-in-difference analysis in Eq. (5) is employed to verify the effect of the HSR network accessibility on housing prices is a causal relationship. Thirdly, to further gauge how the effect of HSR network accessibility on housing prices varies with regions, we introduce a set of dummy variables and their interaction terms with HSR network accessibility in Eq. (6). Finally, the high housing prices, especially in provincial capital cities or municipalities, have raised government concerns, which has motivated our further analysis. Eq. (8) is introduced to gain more detailed information across different housing types using a sub-sample of 35 cities (mostly provincial capital and municipalities).

3.2.1 HSR network accessibility and housing prices: The baseline model

We empirically investigate the relationship between HSR network accessibility and urban housing prices and test the hypothesis. The baseline regression model is shown in Eq. (4):

$$\ln HP_{i,t} = \beta_0 + \beta_1 Acc_{i,t} + \beta_2 X_{i,t} + \mu_i + \varepsilon_{i,t} \tag{4}$$

where $\ln HP_{i,t}$ is the dependent variable, $Acc_{i,t}$ represents the accessibility measures (namely point degree, closeness centrality, and average travel time), $X_{i,t}$ denotes a set of control variables, i denotes cities, t denotes year, μ_i represents the city fixed effect, and ε is an idiosyncratic error term.

The dependent variable, ln*HP*, is the logarithm of annual urban average housing prices per square meter of floor. We use three different network accessibility measures in our analysis: *Acc_degree* representing the point degree of a city, which is the number of adjacent cities that can be reached by HSR from a city; *Acc_closeness* denoting the closeness centrality from point *i* to point *j* in the network in year *t*; and *Acc_time* denoting the average travel time from city *i* to all other cities in year *t*. The model also includes a set of control variables with proven effect on housing prices in the literature. To control for the effect of other transportation facilities in the city, we include a dummy variable *airport* that represents whether this city owns an airport and a variable *metro* that takes a value of 1 when the city starts to own an urban metro system and 0 otherwise. ln*pop* represents the logarithm of urban population of total permanent resident (in ten thousand) to control for city size, agglomeration economics and heterogeneous patterns of residential characteristics across cities. An alternative dummy variable for population, *pop_dummy*, is also used to ensure the robustness of our results. It equals 1 if a city's population is greater than the sample average of a particular year, and 0 otherwise. *savings* and *loans* are included as in Zhang et al. (2012), representing financial development of a city, defined as the ratio of total savings and total loans to GDP at the city level, respectively. Environment is an important factor driving housing prices as people increasingly pursue the quality of living. We introduce two ecological environment variables – *greenspace* and *PM2.5* – to the empirical model. *Greenspace* is the proportion of green covered area to the built-up area. *PM2.5* is an indicator for air quality, referring to the annual average atmospheric particulate matter (PM) concentration (in $\mu g/m^3$) that have a diameter of less than 2.5 micrometers. The higher the value the lower the air quality.

We test the robustness of the results using BOX-COX transformations (Box and Cox, 1964) that provides better estimates with abnormal distribution. The transformation is applied to the dependent variable *y* as follow:

$$y^{(\lambda)} = \begin{cases} (y^{\lambda} - 1)/\lambda, & if\ \lambda \neq 0 \\ \log y, & if\ \lambda = 0 \end{cases} \tag{5}$$

Where λ is the transformation parameter indicating different functional forms.

3.2.2 HSR network accessibility and housing prices: The difference-in-difference analysis

In our sample, a large proportion of cities have HSR started to operate since 2012-2013. This allows us to conduct a difference-in-difference analysis to verify the causal relationship between HSR network accessibility and housing prices. It is helpful to resolve the endogeneity encountered by the OLS regression. We construct a sub-sample which includes cities with HSR operating since 2012-2013 as the treatment group, and cities without changes in HSR status (i.e., cities with or without HSR throughout the whole sample period) as the control group. The cities in the control group are matching observations selected based on Mahalanobis distance scoring method (Mahalanobis, 1936).[7] The initial sample of 2012-2013 includes 1687 observations and the final sample after matching contains 1117 observations. We introduce two dummy variables *Post* and *HSR*. *Post* takes a value of 1 from 2012 onwards and 0 otherwise, and *HSR* takes a value of 1 for cities with access to HSR network and 0 otherwise. The point degree and closeness centrality equal to zero for cities without HSR, the difference-in-difference analysis can only apply to average travel time. As our main interest lies on HSR accessibility, we interact network accessibility with *Post*×*HSR* as shown in Eq. (6).

$$lnHP_{i,t} = \beta_0 + \beta_1 Acc_{i,t} + \beta_2 HSR + \beta_3 \, Post + \beta_4 Acc_{i,t} \times HSR + \beta_5 Acc_{i,t} \times Post + \beta_6 Post \times HSR + \beta_7 Acc_{i,t} \times HSR \times Post + \beta_8 X_{i,t} + \mu_i + \varepsilon_{i,t} \qquad (6)$$

where ln$HP_{i,t}$ is the dependent variable, $Acc_{i,t}$ represents the accessibility measure of average travel time, $X_{i,t}$ denotes a set of control variables, *HSR* is a dummy variable for cities with HSR, *Post* is a dummy variable for years after 2012, μ_i represents the city fixed effect, *i* denotes cities, *t* denotes year, and ε is an idiosyncratic error term.

3.2.3 HSR network accessibility and housing prices: Regional heterogeneity

In China, provinces are classified into eastern, middle and western regions, according to the level of economic development and geographical location.[8] We introduce a set of regional dummies and their interaction terms with the network accessibility measures to the baseline model to investigate how the effect of HSR network accessibility on housing prices varies with different regions. *East, Middle*, and *West* takes a value of 1 if a city is in the eastern, middle, and western region, respectively, and zero otherwise, *East* is omitted from the regression for comparison purposes. The regression model with regional dummies is shown in Eq. (7), which can be rewritten for the western, middle, and eastern region, respectively as in Eq. (8):

$$\ln HP_{i,t} = \beta_0 + \beta_1 Acc_{i,t} + \beta_2 Acc_{i,t} \times West_i + \beta_3 Acc_{i,t} \times Middle_i + \beta_4 West +$$

[7] Mahalanobis distance is widely used techniques in cluster analysis and classification. This is a multi-dimensional generalization of measuring how many standard deviations away one data point from the mean.

[8] Eastern provinces covering Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Guangxi and Hainan. Middle provinces covering Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei and Hunan. Western provinces covering Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia and Xinjiang.

$$\beta_5 Middle_i + \beta_6 X_{i,t} + \mu_i + \varepsilon_{i,t} \tag{7}$$

$$\begin{cases} \ln HP_{i,t} = (\beta_0 + \beta_4) + (\beta_1 + \beta_2) Acc_{i,t} + \beta_6 X_{i,t} + \mu_i + \varepsilon_{i,t}, & West_i = 1, Middle_i = 0; \\ \ln HP_{i,t} = (\beta_0 + \beta_5) + (\beta_1 + \beta_3) Acc_{i,t} + \beta_6 X_{i,t} + \mu_i + \varepsilon_{i,t}, & West_i = 0, Middle_i = 1; \\ \ln HP_{i,t} = \beta_0 + \beta_1 Acc_{i,t} + \beta_6 X_{i,t} + \mu_i + \varepsilon_{i,t}, & West_i = 0, Middle_i = 0. \end{cases} \tag{8}$$

where $\ln HP_{i,t}$ is the dependent variable, $Acc_{i,t}$ represents the accessibility measures (point degree, closeness centrality, and average travel time), $West_i$ and $Middle_i$ are dummy variables for western and middle regions, $X_{i,t}$ denotes a set of control variables, μ_i represents the city fixed effect, i denotes cities, t denotes year, and ε is an idiosyncratic error term.

3.2.4 HSR network accessibility and housing prices: Housing type

We further explore how the impact of HSR network accessibility on housing prices varies with different housing types. We introduce a set of housing type dummies and their interaction terms with the HSR network accessibility measure to the baseline model. *Residential, Villa, Office* and *Commercial* take a value of 1 for ordinary-residential, villas and high-grade apartments, office, and commercial building, respectively, and 0 otherwise. *Commercial* is omitted from the regression for comparison purposes. The regression model is shown in Eq. (9):

$$\ln HP_{i,t} = \beta_0 + \beta_1 Acc_{i,t} + \beta_2 Acc_{i,t} \times Residential_i + \beta_3 Acc_{i,t} \times Villa_i + \beta_4 Acc_{i,t} \times Office_i + \beta_5 Residential_i + \beta_6 Villa_i + \beta_7 Office_i + \beta_8 X_{i,t} + \mu_i + \varepsilon_{i,t} \tag{9}$$

where $\ln HP_{i,t}$ is the dependent variable, $Acc_{i,t}$ represents the accessibility measures (point degree, closeness centrality, and average travel time), $Residential_i$, $Villa_i$ and $Office_i$ are dummy variables for different types of housing, $X_{i,t}$ denotes a set of control variables, μ_i represents the city fixed effect, i denotes cities, t denotes year, and ε is an idiosyncratic error term.

3.3. Sample and data

The sample consists of 285 cities over the period 2009-2017, covering all prefecture-level cities, provincial capital cities and municipalities in mainland China except Tibet due to data unavailability. The sample is unbalanced due to missing values, i.e., missing housing price of some small cities in early years. Annual urban housing prices are directedly calculated using data obtained from Statistical Yearbooks of each province (2010-2018), which provide market statistics of cities.[9] The spatial statistical unit is the whole city with the towns and villages under their jurisdiction. Housing price is the average citywide price per square meter of floor, which equals the total sales of marketable housing divided by total floor area of marketable houses sold. We include all types of marketable housing, namely business buildings, office buildings, residential buildings, and others, and exclude land prices. Data on population, GDP, savings, loans and green space are collected from "*China's City Statistical Yearbook* (*2010-2018*)". The data on airport (whether a city possesses an airport) is extracted from "*The official Civil Aviation Industry Development Statistics Bulletin* (*2009-2017*)" published by the Civil Aviation Administration of China (CAAC).[10] The data on metro lines is collected from Journal of Urban Rapid Rail Transit (in Chinese) which publishes annual metro statistics of every city. The data on metro lines are collected from website of metro companies in various cities. Data on *PM2.5* are collected from NASA satellite observation published by CIESIN at Columbia University.[11]

Original data on accessibility – whether a city operates HSR – is collected from "*National Railway Passenger Train Schedules* (*2009-2017*)", obtained from China's railway customer service center[12]. We draw an HSR operation map and construct an HSR complex network by each year over the sample period 2009-2017. We then compute network accessibility measures. We calculate point degree (Eq.1) and closeness centrality (Eq.2) using Gephi8.2 and the average travel time (Eq.3) using ArcGIS10.0. Point degree is defined as the number of adjacent cities reachable by HSR network, closeness centrality is the sum of the geodesic distance between the point and all other points in the network, ranging from 0 to 1, average travel time refers to the average time taken from a city to all the other cities (unit in 100 minutes). We also construct a sub-sample of 35 large and medium-sized, mostly provincial capital cities with detailed data on different types of housing over the sample period 2009-2016 for further analysis. **Table 1** provides descriptive sample statistics.

[Insert Table. 1 Here]

[9] The rising housing prices form part of inflation and reflect public expectation of real estate appreciation. In this paper, following the mainstream literature on the relationship between housing prices and transportation (Dubé et al., 2013; Agarwal et al., 2015; Chen and Haynes, 2015; He, 2020), we express housing prices in nominal terms.
[10] It will be considered as no airport in following two situations: 1) The airport is licensed for operating permission, but no flights yet; 2) The airport shuts down for a certain year for the reason of long-term maintenance.
[11] https://sedac.ciesin.columbia.edu/
[12] According to the definition of the National Railway Administration of China, the HSR in China includes rail lines served by G, C, and D prefix bullet trains.

4. Empirical results

4.1. The spatial distribution of HSR network accessibility improvement

Table 2 shows the top 20 cities by HSR accessibility improvement from 2009 to 2017, along with HSR accessibility in 2017, based on our accessibility measures. In terms of improvement in point degree, major cities are on the top of the board, such as Nanjing, Chongqing, and Guangzhou, while 8 small cities (out of top 20) are also benefited from the network, such as Tongling and Zhaoqing. In terms of the improvement in closeness centrality, small cities dominate the board with only one exception of Nanchang. As to the average travel time, cities with the most improvement are in the remote areas and normally small, such as Kelamayi, Urumqi, and Beihai. Overall, the results suggest that small cities and cities in the remote areas benefit more from the HSR development since generally there were limited alternative transportation facilities before HSR. HSR is likely to have a stronger spillover effect and substantial economic gains in these areas. Studies (Yao et al., 2019, 2020) find the HSR network helps reduce regional disparity and promote national economic convergence in China in terms of the transport network accessibility. We follow up this issue and examine the regional heterogeneity of the HSR network effect on housing prices in section 4.4.

[Insert Table. 2 Here]

Figure 2 shows the spatial distribution of accessibility improvement in point degree, closeness centrality, and average travel time from 2009 to 2017. **Figure 2(a)** shows that cities with the most significant increase in point degree in the scale 3 to 5 are in the Yangtze River delta, the Pearl River Delta and the Chengdu-Chongqing urban agglomerations. **Figure 2(b)** shows that cities with the most significant improvement in closeness centrality in the scale from 3 to 5 are in Beijing-Tianjin-Hebei urban agglomeration, and inland on-corridor regions. **Figure 2(c) and (d)** shows the spatial distribution of accessibility measured by average travel time in 2009 and 2017, respectively. When there were only two D-trains (HSR) in 2009, almost all the mainland China had poor accessibility, with the average travel time ranging from 47,000 to 184,600 minutes. By 2017, an HSR network of four-vertical and four-horizontal lines took shape. The accessibility to most cities has improved significantly and many cities along the Beijing-Shanghai and the Beijing-Guangzhou HSR lines now have average travel time under 20,000 minutes. Even the most remote cities have a sharp decrease in the average travel time to around 100,000 minutes.

[Insert Figure. 2 Here]

4.2. HSR network accessibility and housing prices

We use the operation of the HSR in China as a natural experiment to test the effect of transport infrastructure on average urban housing prices. We employ three measures of network accessibility – point degree, closeness centrality and average travel time. A city with higher point degree and closeness centrality but lower average travel time represents a high level of accessibility and importance in the transport network. Our accessibility measures are correlated with the year dummies and the inclusion of the year fixed effect in the model results in biased estimates. Following Petersen (2009), we include the city fixed effect in the model with standard errors clustered by year, which applies to all our estimations in this paper.

Table 3 reports the estimation results from the baseline model in Eq. (4), using the OLS estimator in columns (1)-(6). In column (1)-(2), the coefficient on *Acc_degree* is positive and statistically significant at 1% level, suggesting a positive effect of HSR network accessibility on urban housing prices. The results are similar when the network accessibility is proxied by closeness centrality in column (3)-(4). When using the average travel time to measure network accessibility in column (5)-(6), the negative coefficient on *Acc_time* indicates that as network accessibility improves (travel time decreases), housing prices increase. Different measures of network accessibility consistently provide evidence for a positive effect on housing price, which is economically substantial. For a one standard deviation increase in point degree and closeness centrality, the average housing prices will increase by 10.3% (=0.098×1.056) and 11.2% (=7.020×0.016), respectively. For a one-unit reduction in average travel time (100 minutes), the average housing prices will increase by 17.3%. Columns (7)-(9) report results using BOX-COX transformations in Eq. (5). Compared with those from the OLS estimator in columns (1)-(6), the sign and significance level of the coefficient on the network accessibility variable are consistent[13]. The overall evidence supports the hypothesis that *the HSR network accessibility has a positive impact on urban housing prices*.

The baseline model includes a set of control variables and the estimated coefficients also reveal some interesting results. The coefficients on *airport* are positive and significant, that housing prices are higher in cities with an airport than in cities without an airport, e.g., by 14-16% in

[13] The regression coefficients and R-squared are not comparable, due to the dependent variables are different.

columns (1) and (4). The positive and significant coefficient on ln*pop* is consistent with prior expectation and housing prices are higher in more populated cities by about 34-38%. The impact of credit availability is tentative. The coefficient on *loans* is positive, as expected, but mainly significant when the network accessibility is measured by travel time and in the regressions after BOX-COX transformations in columns (7)-(9). The coefficient on *savings* is positive and significant when network accessibility is measured by point degree and closeness centrality columns (1)-(4), but it becomes negative and weakly significant in column (5). When excluding savings from the baseline model, our main results remain unchanged (unreported). The coefficients on *metro*[14] and *PM2.5* are insignificant in most OLS regressions but become significant after BOX-COX transformations. The coefficient on *greenspace* is positive and significant, that the average housing prices will increase by 0.2-0.5% with the 1% increase on green covered area. Overall, the evidence suggests that the housing prices are mainly driven by transportation infrastructure and consumer demand.

[Insert Table. 3 Here]

4.3. HSR network accessibility and housing prices: difference-in-difference analysis

One might question that the effect of HSR network on urban housing prices could be merely co-movements of two variables, rather than a causal relationship. To address this potential concern, we carry out a difference-in-difference analysis to verify that our main result is a causal effect. The difference-in-difference analysis requires a common trend of the treatment and control groups, and the exogeneity and sharpness of the treatment. The HSR construction is part of the national development plan and therefore it is exogeneous to housing prices. When studying the planning and construction of HSR, the sharpness of the treatment may not be strictly adhered. Indeed, in social science research, it is very difficult to carry out experiments under the same strict assumptions like in natural science. We consider that housing price is more responsive to the actual operation of the HSR lines than to the initial announcement of HSR planning. We conduct the analysis in two settings: the full sample and a two-year sample, both being balanced samples.

Since point degree and closeness centrality equal to zero for cities without HSR, we can only

[14] Housing prices can be driven by the improvement of local transportation associated with the development of HSR network. We replace *metro* with *rail mileage* (measured by the length of urban rail transportation, including metro, light rail, tram, and maglev), our main results hold.

employ average travel time as the measure of network accessibility in the analysis. We perform the pre-trend test, which confirms the parallel trends of treated group and control groups. **Table 4** reports the estimation results from Eq. (5). The triple interaction term – *Acc_time×Post×HSR,* is of our particular research interest since it uncovers whether the changes in travel time due to new HSR access in 2012 lead to higher housing prices.[15] Columns (1) and (2) report results from the sample over the period 2009-2017, where alternative measures of population are used for robustness purposes. The coefficient on *Acc_time×Post×HSR* is negative and statistically significant, suggesting that a reduction in average travel time will lead to housing prices to increase more in cities with new HSR network access, compared with housing prices in cities without changes in direct access to the HSR network. When average travel time decreases by one unit (100 minutes), the housing prices will be higher by 8.4-9.9% in cities with new HSR access than in cities without changes in HSR access. We then restrict the sample to two years – 2011 (before) and 2013 (after) and re-estimate Eq. (6). As shown in columns (3)-(4), results mirror those in columns (1)-(2). The coefficient on *Acc_time×Post×HSR* is negative and significant. For a 100 minutes reduction in travel time, housing prices increase faster in 2013 by 10.2-10.8% for treatment group (cities with new HSR access) relative to the control group (cities without changes in HSR access). Overall, the results from this section provide strong evidence for a causal effect of HSR accessibility on urban housing prices.

[Insert Table. 4 Here]

4.4. HSR network accessibility and housing prices: The regional effect

Having established a causal effect of HSR accessibility on urban housing prices, we further explore whether this effect has spatial differentials when cities sharing similar regional environment and location characteristics. **Table 5** reports the estimation results from the regression model with regional dummies in Eq. (7). The coefficients on HSR network accessibility variables (*Acc_degree, Acc_closeness,* and *Acc_time*) are all statistically significant at the 1% significance level with expected signs in all specifications, consistent with our main results from the baseline model (as reported in Table 3). As we include more explanatory variables, the multicollinearity between explanatory variables may be an issue. We

[15] When introducing interaction terms, the coefficients on the main variables (i.e., accessibility) may change and the research focus is on interaction terms. The insignificant coefficient on *time* in column 1 is largely due to the correlation between *population* and *time*.

test the Variance Inflation Factor (VIF) and a mean value of 1.62 with a maximum of 2.45 suggests the degree of multicollinearity shouldn't be a major concern.

As expected, the effect of improved HSR network accessibility on urban housing prices varies with regions. In particular, the effect is much stronger in the middle region, suggested by the statistically significant coefficients on the interaction term between HSR network accessibility and regional dummy – *Acc×Middle*. As shown in column (2), for one standard deviation increase in point degree, the average urban housing prices increase by 12.6% (=(0.035+0.084)×1.056) in the middle region, which is 3.6% higher than that in the eastern region of 8.9% (=0.084×1.056). If average travel time decreases by one unit (100 minutes), the average housing prices increase by 21.5% (=-0.134-0.081) in the middle region but only by 13.4% in the eastern region. When HSR accessibility is measured in terms of closeness centrality, this effect also exists, while the statistical significance is weak. The empirical evidence indicates that HSR benefit the underdeveloped middle region that is geographically adjacent to the more developed east region. This result echoes findings in Zheng and Kahn (2013) that rising housing prices in the secondary cities nearby megacities along HSR lines in China. The insignificant effect of HSR in the remote west region mirrors the findings in Andersson et al. (2010) that the effect of HSR on housing prices is limited perhaps due to high ticket prices and entrenched residential location patterns. Overall, the empirical evidence allows us to conclude that the improvement in HSR network accessibility stimulates regional economic convergence with respect to housing prices and helps to close the gap in urban housing prices between the developed eastern region and underdeveloped adjacent middle region, while this effect remains to be seen in the more remote west region.

[Insert Table. 5 Here]

4.5. HSR network accessibility and housing prices: By housing type

In this section, we use a sub-sample of 35 large and medium-sized cities with detailed data on the prices of different housing types over the period 2009-2016 to further investigate how the impact of HSR accessibility on housing prices varies with housing types. **Table 6** reports the regression results from Eq. (9) with the HSR network accessibility variable only in odd columns and with a full set of interactions terms and control variables in even columns. The estimated coefficients on all HSR accessibility variables are statistically significant with expected signs, consistent with our main results in Table 3. This indicates that sample selection bias is not a major concern for this subsample of 35 cities.

We find evidence that the effect of HSR accessibility on urban housing prices varies with housing types. The commercial building has a lower sensitivity to the improvement in HSR accessibility. For a one-standard deviation increase in point degree, commercial building price, on average, increases by 7.7% (=0.078×0.984), while office building experiences an increase by11.0% (= (0.078+0.034) × 0.984), indicated by the positive and significant (at the 10% level) coefficient on *Acc*×*office*. For a one-standard deviation increase in closeness centrality, the average commercial building price increases by 9.0% (=7.518×0.012), while villa prices increase by 12.5% (=(7.518+2.904)×0.012) and office building price increase by 13.4%(=(7.518+3.648)×0.012). The results in column 6 reveal that housing prices are generally more sensitive to the reduction in travel time, especially for ordinary residential and villa. For a one unit (100 minutes) decrease in travel time, the average price increases by 16.8% for commercial building, 20.8% (=0.168+0.040) for residential, 20.3% (=0.168+0.035) for villa, and 19.3% for office building (=0.168+0.025). This is perhaps because the residents are more correlated with daily commuters, whose decisions on buying residential are more sensitive to travel time reduction.

[Insert Table. 6 Here]

6. Discussions and conclusions

In this paper, we apply the social network analysis approach to measure cities' HSR network accessibility and investigate how the improvement of transportation infrastructure affects urban housing prices in China over the period 2009-2017. We have also applied two heterogeneity tests based on geographical locations and property types to further study the asymmetric effect of HSR network on urban housing price. Using different measures of HSR network accessibility, namely point degree, closeness centrality, and average travel time, our main findings are as follows. *First*, the HSR network accessibility has a significant positive impact on urban housing prices, which is economically substantial. On average, an increase in a city's network accessibility (i.e., point degree and closeness centrality) by one-standard deviation causes a 10-11% increase in housing prices. If average travel time decreases by 100 minutes, housing prices increase by about 17%. *Secondly*, the effect of HSR network accessibility on urban housing prices is stronger in the underdeveloped middle region that is close to the more developed eastern region. *Finally*, the impact of HSR network accessibility on housing prices also varies with housing type. The commercial building has a lower sensitivity to the improvement in HSR accessibility, while HSR accessibility has a stronger effect on the villa and office building market.

There is an extensive literature on the impact of transit lines in urban contexts on the housing market, but very few contributions are proposed on the effect of long-distance rail among cites on the housing market. This paper adds new evidence from the perspective of housing price to the literature on the differential HSR effect. The evidence indicates the externality of infrastructure investment in reducing regional disparity and promoting regional economic convergence as far as housing prices concerned. Chen and Hall (2011) report a differential HSR effect on British economic geography by showing that the renewed economic growth is stronger in towns within a two-hour travel time from London. In Germany, Ahlfeldt and Feddersen (2017) find the economic impact of HSR is about 8.5% in three counties with intermediate stops on the HSR line connecting Cologne and Frankfurt and this spillover declines by 50% for every 30 minutes in additional of travel time, and by 1% beyond 200 minutes. In China, research suggests that HSR tends to reduce regional disparities while enhancing urban economic growth (Qin, 2017, Yao et al., 2019, 2020).

Our findings have important policy implications. First, it advances our understanding of the housing market in China. Housing price appreciation has been phenomenal in major Chinese cities over the past years. Zheng and Saiz (2016) reports an annual housing price appreciation rate of 14.3% for 35 large and medium-sized cities (27.4% for Beijing) during 2006-2013. As real estate price bubbles and housing affordability have become major concerns for the government, we provide timely evidence for policy makers to design future policies. Second, this research provides information relating to the long-term national/regional planning. The HSR network requires huge capital investment. As a quasi-public good, financing HSR projects is often a major issue for all economies. Given its positive externalities on the housing market, governments may consider transferring a certain proportion of housing price premium to finance HSR projects. Indubitably, the key point is that there must be a fair and righteous process for government to collect designated types of taxes from infrastructure spillovers and use it to the public. This topic has attracted widespread discussions lately, many jurisdictions have worked out ways to share taxes among various collectors and recipient entities (Araki and Nakabayashi, 2019). Moreover, the fast-developing Internet-based financial industry is expected to encourage transport infrastructure investments from foreign markets as well as donations from the public (Grant et al., 2020). Hence, this may probably happen in practice in the near future, calling for more research in this direction.

Nevertheless, the results from this study should be interpreted with cautions. Firstly, the average age of the housing stock or the percentage of housing units that are second-hand is an important determinant of house prices; however, it is excluded from our analysis as data are largely unavailable for our sample cities over the sample period. The China's property market in the real sense has a short history of just over 20 years. The housing reform in 1998 marked the comprehensive marketization of China's housing distribution. Together with the fiscal tax sharing system reform in 1994, it brought about China's real estate driven land finance and urban sprawl (Dong, 2018). In this context, most housing buildings in the market are newly built and the second-hand housing transactions account for a relatively small market share. As such, the exclusion of housing age may bias our results, but we consider the effect is not severe. Future research on housing prices should control for this effect, especially as the second-hand housing markets becoming larger and important. One potential source of housing age data is to use python programming to extract data from Lianjia Website (https://lianjia.com/). It is one of China's largest real estate transaction service platforms that provides detailed information on every single house including the year when the housing stock is built. Secondly, housing prices in this study is the average citywide housing price per square meter of floor. As shown in Panel B of Table 2, the prices of different types of housing vary significantly. The average price of luxury villar and high-grade apartment is 67% higher than ordinary residential building, while commercial buildings are slightly more expensive than office buildings. Therefore, the interpretation of the HSR network accessibility effect should consider the housing type. Future research should seek more details data on different types of housing and provide more insightful information. Moreover, without detailed separate housing prices at the HSR station's catchment area, our results reveal the impact of the HSR network on housing prices in the city as whole. Future research should seek more details data on different types of housing and different areas in the city and provide more insightful information. Finally, in this study, we employ the raster cost method to construct HSR network accessibility measures. An alternative measure could be actual travel time. This involves collecting data on actual travel time and can be a direction for future research.

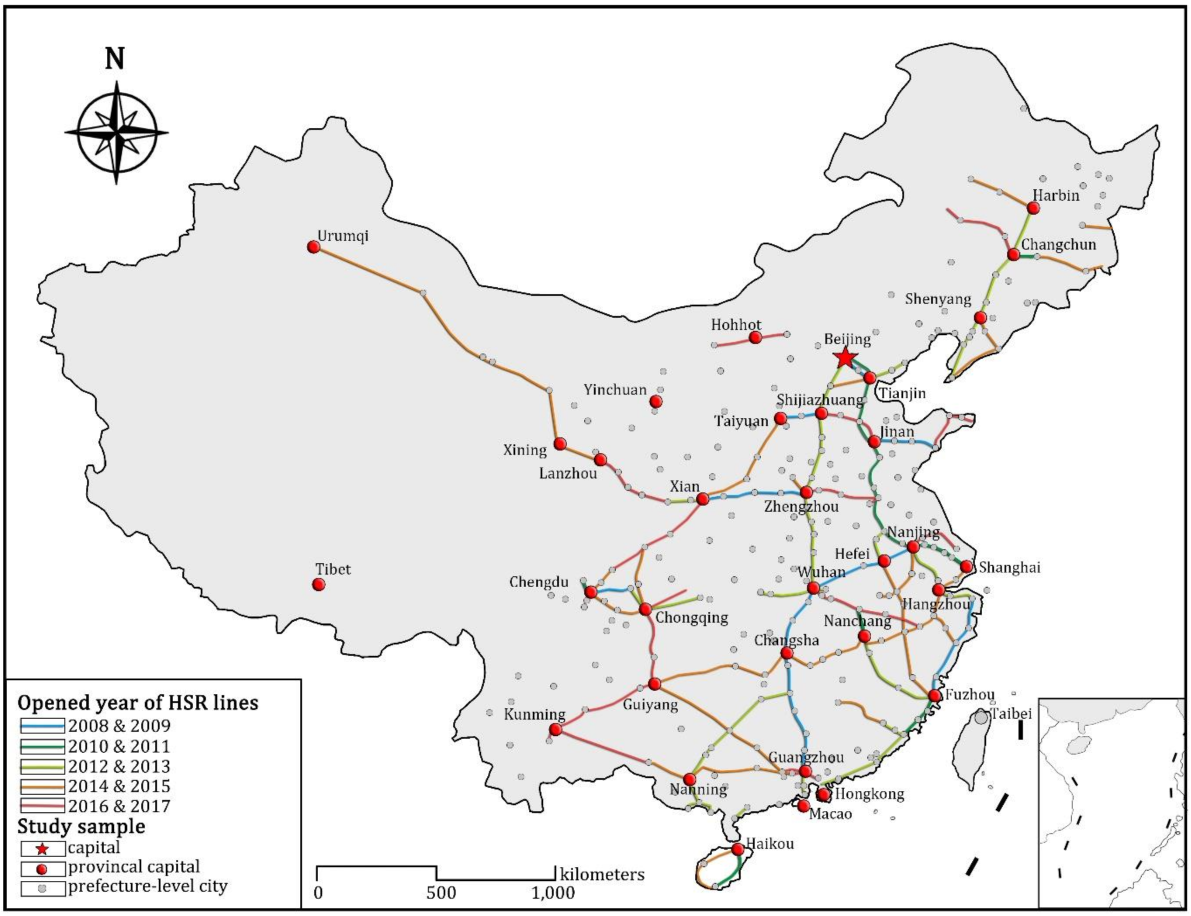
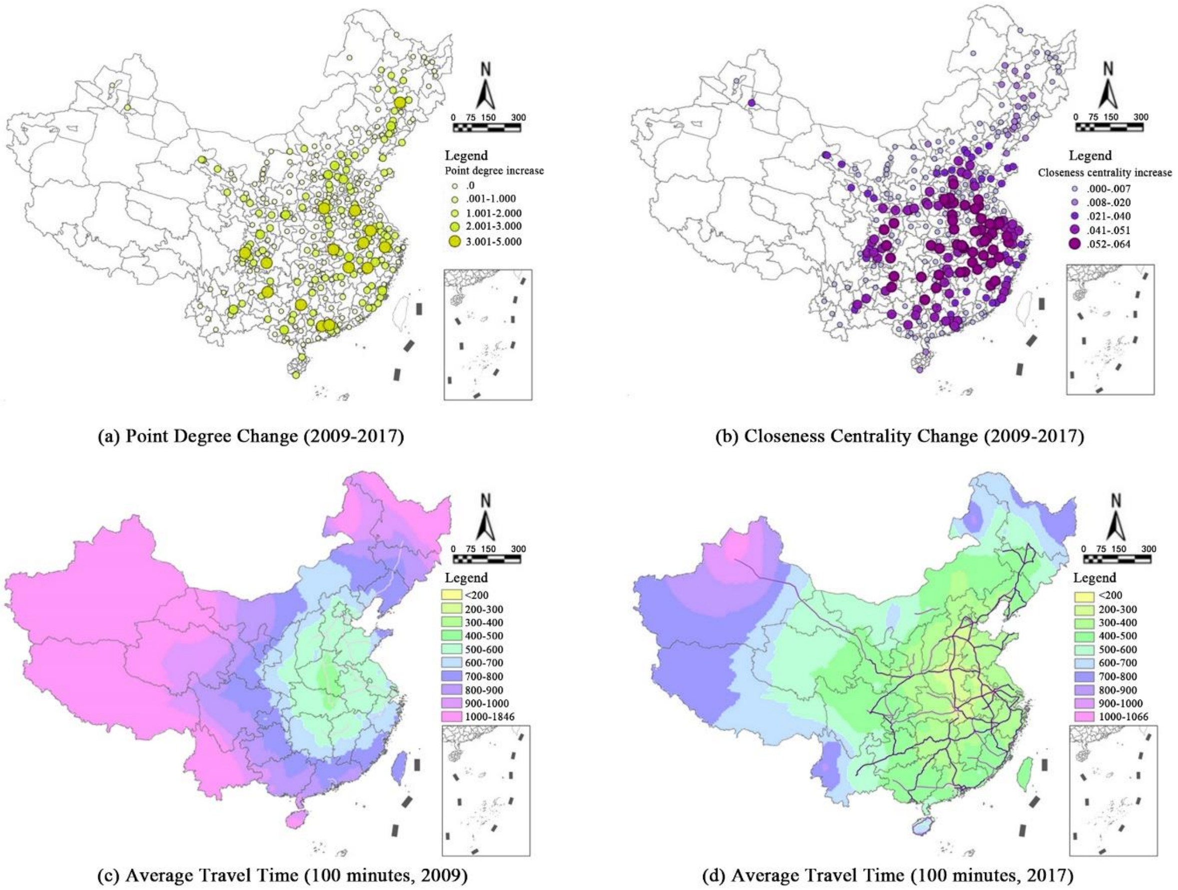


Figure 1. China's high-speed rail lines opened in every two-year from 2008 to 2017 with the 5-year categories coloured from blue to red representing the pace of network development.

Source: By authors based on data collected from the National Railway Administration of China.



(a) Point Degree Change (2009-2017)

(b) Closeness Centrality Change (2009-2017)

(c) Average Travel Time (100 minutes, 2009)

(d) Average Travel Time (100 minutes, 2017)

Figure 2. Spatial distribution of accessibility improvement from 2009 to 2017.

Table 1. Variable definition and summary statistics

Variable	Definition	Panel A: The full sample of 285 cities				Panel B: A subsample of 35 cities			
		Mean	S. D.	Min	Max	Mean	S. D.	Min	Max
HP	Annual average housing price per square meter of floor	4703.6	3015.6	1295.8	47935.8				
Commercial	Commercial building price per square meter of floor					13109.7	5424.4	5010	35437
Residential	Residential price per square meter of floor					8216.4	5281.8	2811	45498
Villa	Villa price per square meter of floor					13729.2	9770.0	2538	71306
Office	Office building price per square meter of floor					11916.7	6628.6	1948	45313
airport	A dummy variable: 1 for a city has an airport, 0 otherwise	0.449	0.497	0	1	0.567	0.496	0	1
pop	Population: total permanent resident in ten thousand	142.06	181.86	15.1	2449	168.16	244.46	24.83	1345.2
savings	The ratio of savings to GDP	0.844	0.395	0.043	4.136	1.130	0.562	0.314	3.193
loans	The ratio of loans to GDP	1.569	2.326	0.062	84.661	2.288	6.118	0.188	84.661
metro	A dummy variable: 1 for a city has metro, 0 otherwise	0.068	0.251	0	1	0.446	0.498	0	1
greenspace	The percentage of green covered area to built-up area	38.80	7.66	0.36	95.25	40.564	0.305	25.05	62.46
PM2.5	Annual average PM2.5 concentration in $\mu g/m^3$	36.51	16.43	4.68	86.48	38.192	0.968	10.987	81.79
Authors' computed HSR network accessibility measures									
Acc_degree	Point degree: the number of adjacent cities reachable	0.666	1.056	0	6	0.640	0.984	0	3
Acc_closeness	Closeness centrality	0.009	0.016	0	0.066	0.007	0.012	0	0.041
Acc_time	Average travel time from a city to all other cities in 100 minutes.	5.509	1.964	0	18.467	5.133	1.478	2.965	11.885

Notes: The subsample of 35 cities include 4 types of housing/building each year for 8 year over the sample period 2009-2016, with the total number of observations of 1120.

Table 2. Top 20 cities by HSR accessibility improvement from 2009 to 2017

This table shows the top 20 cities by HSR accessibility improvement from 2009 to 2017, along with their HSR accessibility in 2017, based on our calculated accessibility measures of point degree, closeness centrality, and average travel time (in 100 minutes). The cities in italics and underlined are small cities.

Rank	city	Point degree		city	Closeness centrality		city	Average travel time	
		2017	2009-17		2017	2009-2017		2017	2009-2017
1	Nanjing	6	5	*Xinyang*	0.0634	0.0636	*Kelamayi*	264.56	790.32
2	Guangzhou	6	5	*Zhumadian*	0.0621	0.0623	Urumqi	266.84	742.45
3	Chongqing	5	5	*Luohe*	0.0612	0.0614	*Yuxi*	269.75	551.89
4	Wuhan	6	4	*Ezhou*	0.0606	0.0609	*Puer*	269.94	550.37
5	Chengdu	5	4	*Chuzhou*	0.059	0.0592	*Lincang*	270.38	547.06
6	Zhengzhou	5	4	Nanchang	0.0589	0.0592	*Jiayuguan*	273.17	530.50
7	Nanchang	4	4	*Loudi*	0.0587	0.0590	*Jiuquan*	273.76	526.98
8	Guilin	4	4	*Jinzhou*	0.0586	0.0589	Kunming	274.14	513.34
9	Guiyang	4	4	*Jiujiang*	0.0585	0.0588	*Jiaozuo*	274.61	500.61
10	Changchun	4	4	*Kaifeng*	0.0585	0.0588	*Zhangye*	275.24	487.92
11	Hangzhou	4	4	*Huanggang*	0.0585	0.0588	*Qujing*	276.67	484.26
12	*Tongling*	4	4	*Huangshi*	0.0584	0.0587	*Jinchang*	277.67	457.36
13	*Zhaoqing*	4	4	*Bengbu*	0.0582	0.0583	*Baise*	278.33	455.20
14	*Xuzhou*	4	4	*Yichun*	0.0579	0.0583	*Beihai*	279.67	453.92
15	*Shangrao*	4	4	*Xinyu*	0.0577	0.0581	*Baoshan*	279.91	451.66
16	Xian	4	3	*Tongling*	0.0575	0.0577	*Lijiang*	280.20	451.66
17	*Quzhou*	3	3	*Xuzhou*	0.0574	0.0575	*Wuwei*	281.82	442.62
18	*Guangyuan*	3	3	*Huaihua*	0.057	0.0572	*Fangchenggang*	283.54	433.25
19	*Jiujiang*	3	3	*Huangshan*	0.0569	0.0571	*Weihai*	283.90	432.41
20	*Dezhou*	3	3	*Shangrao*	0.0568	0.0568	*Shanwei*	284.18	428.98

Table 3. HSR network accessibility and housing price

Variables	Dependent variable: Urban average housing prices (in logs)								
	OLS	OLS	OLS	OLS	OLS	OLS	BOX-COX	BOX-COX	BOX-COX
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Accessibility									
Acc_degree	0.109***	0.098***					0.0002***		
Acc_closeness			7.283***	7.020***				0.011***	
Acc_time					-0.171***	-0.173***			-0.0002***
Control variables									
airport	0.145**	0.163**	0.143**	0.156**	0.056**	0.068**	0.0003***	0.0003***	0.00008***
ln*pop*	0.377***	0.372**	0.335***	0.339**	0.070**	0.052	0.0006***	0.0005***	0.00003
savings	0.162**	0.163**	0.157**	0.160**	-0.030*	-0.031	0.0003***	0.0003***	-0.00002
loans	0.011	0.011	0.012	0.011	0.007**	0.006**	0.00002***	0.00002***	0.000006***
metro		0.058		0.032		-0.025	0.00005	0.00001	-0.00006**
PM2.5		-0.006		-0.006		-0.003	-0.00001***	-0.000009***	-0.000002***
greenspace		0.005**		0.004**		0.002**	0.000006***	0.000008***	0.000002***
Constant	6.399***	6.392***	6.813***	6.768***	9.860***	10.038***	0.482***	0.476***	0.430***
City fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. Observations	2361	2095	2361	2095	2361	2095	2095	2095	2095
BOX-COX λ							-2.034	-2.063	-2.310
adj. R^2	0.803	0.804	0.811	0.809	0.897	0.896	0.227	0.248	0.574

Notes: All regressions control for the city fixed effects with standard errors clustered by year. The significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table 4. HSR network accessibility and housing prices: Difference-in-difference analysis

variables	Dependent variable: Urban average housing prices (in logs)			
	2009-2017	2009-2017	2011 & 2013	2011&2013
	(1)	(2)	(3)	(4)
Accessibility				
Acc_time	-0.020	-0.026**	-0.030**	-0.037***
Acc_time×Post×HSR	-0.084**	-0.099***	-0.102**	-0.108**
Post	0.475***	0.500***	0.111**	0.120**
HSR	-0.640***	-0.699***	-0.591**	-0.641**
Acc_time×Post	-0.003	-0.004	0.003	0.003
Acc_time×HSR	0.119***	0.131***	0.109**	0.120**
Post×HSR	0.510***	0.592***	0.562**	0.589**
Control variables				
airport	0.132**	0.147***	0.139**	0.161***
pop	0.180***		0.186***	
pop_dummy		0.277***		0.281***
savings	-0.019	-0.053	0.029	-0.005
loans	0.007	0.005	0.004	0.002
metro	0.199**	0.288***	0.173	0.279**
PM2.5	-0.002	-0.002	-0.004*	-0.004*
greenspace	0.007**	0.008***	0.007*	0.008**
Constant	7.003***	7.768***	7.369***	8.159***
City fixed effect	Yes	Yes	Yes	Yes
Observations	976	976	254	254
adj. R^2	0.493	0.506	0.352	0.377

Notes: The sample include cities with HSR operating since 2012/2013 as the treatment group and cities without changes in HSR status (with or without HSR throughout the sample period) as the control group. *HSR* is a dummy for cities with high speed railway and *Post* is a dummy for the year 2012. The significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table 5. HSR network accessibility and housing prices: The regional effect

variables	Dependent variable: Urban average housing prices (in logs)					
	Acc_degree		*Acc_closeness*		*Acc_time*	
	(1)	(2)	(3)	(4)	(5)	(6)
Accessibility						
Acc	0.130***	0.084***	8.813***	5.998***	-0.151**	-0.134***
Acc×West	-0.002	0.007	0.620	6.315*	-0.024	-0.068**
Acc×Middle	0.025*	0.035**	0.123	1.319*	-0.065***	-0.081***
West	-1.478***	-0.204	-1.604***	-0.394	0.642	1.375*
Middle	-1.586***	-0.128	-1.756***	-0.355	-1.405***	-1.034***
Control variables						
airport		0.159**		0.155**		0.057**
ln*pop*		0.375**		0.348**		0.091*
savings		0.166**		0.159**		-0.051**
loans		0.011		0.012		0.006**
metro		0.061*		0.033		-0.019
PM2.5		-0.006		-0.005		-0.003
greenspace		0.005**		0.004**		0.002**
constant	9.386***	6.388***	9.513***	6.691***	10.302***	9.618***
City fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2407	2095	2407	2095	2407	2095
adj. R^2	0.777	0.831	0.789	0.810	0.898	0.903

Notes: All regressions control for the city fixed effects with standard errors clustered by year. The significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

Table 6. HSR network accessibility and housing prices: Housing types

Variables	Dependent variable: urban average housing prices (in logs)					
	Acc_degree		*Acc_closeness*		*Acc_time*	
	(1)	(2)	(3)	(4)	(5)	(6)
Accessibility						
Acc	0.125**	0.078*	11.050***	7.518**	-0.206***	-0.168***
Acc×Residential		0.013		1.902		-0.040**
Acc×Villa		0.028		2.904*		-0.035**
Acc×Office		0.034*		3.648**		-0.025***
Residential		-0.540***		-0.545***		-0.325**
Villa		-0.058**		-0.060**		0.140*
Office		-0.169***		-0.172***		-0.021
Control variables						
airport		0.210**		0.214**		0.038
pop_dummy		-0.079*		-0.055		0.005
savings		0.147*		0.135*		0.054
loans		0.004		0.004		0.001
metro		0.164		0.246***		0.086
PM2.5		-0.005		-0.006		-0.001
greenspace		1.045*		0.911*		0.152
constant	9.791***	9.230***	9.860***	9.285***	10.914***	10.688***
City fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1120	1052	1120	1052	1120	1052
adj. R^2	0.569	0.756	0.578	0.764	0.637	0.809

Notes: Residential: ordinary residential housing; villa: villa and high-grade apartments; commercial: housing for business use. Commercial is omitted from the regression for comparison purposes. All regressions control for the city fixed effects and standard error clustered by year. The significance levels at 10%, 5%, and 1% are indicated by *, **, and ***, respectively.

References

Agarwal, S., Koo, K. M., Sing, T. F., 2015. Impact of electronic road pricing on real estate prices in Singapore. *Journal of Urban Economics*, 90, 50-59. https://doi.org/10.1016/j.jue.2015.09.004

Ahlfeldt G M, Feddersen A., 2017. From periphery to core: measuring agglomeration effects using high-speed rail [J]. *Journal of Economic Geography* 18(2), 355-390. https://doi.org/10.1093/jeg/lbx005

Alonso, W., 1964. Location and land use: toward a general theory of land rent. Harvard University Press. https://doi.org/10.4159/harvard.9780674730854

Andersson, D.E., Shyr, O.F., Fu, J., 2010. Does high-speed rail accessibility influence residential property prices? Hedonic estimates from southern Taiwan. *Journal of Transport Geography* 18(1), 166-174. https://doi.org/10.1016/j.jtrangeo.2008.10.012

Araki, S., and S. Nakabayashi, eds. 2019. Tax and Development Challenges in Asia and the Pacific. *Handbook on high speed rail and quality of life.* Tokyo: Asian Development Bank Institute.

Box, G.E.P. and Cox, D.R., 1964. An Analysis of Transformations. *Journal Royal Statistical Society Series B*, 26, 211- 252.

Breheny, M.J., 1978. The measurement of spatial opportunity in strategic planning. *Regional Studies* 12(4), 463-479. https://doi.org/10.1080/09595237800185401

Button, K., 2012. Is there any economic justification for high-speed railways in the United States? *Journal of Transport Geography* 22, 300-302. https://doi.org/10.1016/j.jtrangeo.2012.01.025

Cascetta E., CartenìA., Henke I., Pagliara F., 2020. Economic growth, transport accessibility and regional equity impacts of high speed railway in Italy: ten years ex post evaluation and future perspectives. *Transportation research part A: policy and practice* 139, 412-428. https://doi.org/10.1016/j.tra.2020.07.008

Cervero, R., Kang, C.D., 2011. Bus rapid transit impacts on land uses and land values in Seoul, Korea. *Transport Policy* 18(1), 102-116. https://doi.org/10.1016/j.tranpol.2010.06.005

Chalermpong, S., 2007. Rail transit and residential land use in developing countries: hedonic study of residential property prices in Bangkok, Thailand. *Transportation Research Record* 2038(1), 111-119. https://doi.org/10.3141/2038-15

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

704. https://doi.org/10.1016/j.jtrangeo.2010.08.010

Chen, Z.H., Haynes, K.E., 2015. Impact of high speed rail on housing values: an observation from the Beijing–Shanghai line. *Journal of Transport Geography* 43, 91-100. https://doi.org/10.1016/j.jtrangeo.2015.01.012

Cook, S., Watson, D., 2015. A new perspective on the ripple effect in the UK housing market: Comovement, cyclical subsamples and alternative indices. *Urban Studies* 53(14), 3048-3062. https://doi.org/10.1177/0042098015610482

Deng, T., Nelson, J.D. 2010. The Impact of Bus Rapid Transit on Land Development: A Case Study of Beijing, China, *World Academy of Science, Engineering and Technology* 66, 1196-1206.

Diao, M., Qin, Y., Sing, T. F. 2015. Negative externalities of rail noise and housing values: evidence from the cessation of railway operations in Singapore. *Real Estate Economics,* 878-917. https://doi.org/10.1111/1540-6229.12123

Diao, M., Zhu, Y., Zhu, J. 2017. Intra-city access to inter-city transport nodes: the implications of high-speed-rail station locations for the urban development of Chinese cities. *Urban Studies*, 54(10), 2249-2267. https://doi.org/ 10.1177/0042098016646686

Dong, X., Zheng, S., Kahn, M.E.2020. The role of transportation speed in facilitating high skilled teamwork across cities. *Journal of Urban Economics*, 115. https://doi.org/10.1016/j.jue.2019.103212

Dong, X., 2018. Reform of China's housing and land systems: the development process and outlook of the real estate industry in China. *Chinese Journal of Urban & Environmental Studies*, 05(04). https://doi.org/10.1142/S2345748117500270

Dubé J., Thériault M., Rosiers F. D., 2013. Commuter rail accessibility and house values: The case of the Montreal South Shore, Canada, 1992-2009. *Transportation research part A: policy and practice*, 54, 49-66. http://dx.doi.org/10.1016/j.tra.2013.07.015

Duranton, G., Turner, M.A. 2012. Urban growth and transportation. *Review of Economic Studies*, 79(4), 1407–1440. https://doi.org/10.1093/restud/rds010

El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., Loong, C., 2016.The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice* 91, 302-316. https://doi.org/10.1016/j.tra.2016.07.003

Ferber, C., Holovatch, T., Holovatch, Y., Palchykov, V., 2009. Public transport networks: empirical analysis and modeling. *European Physical Journal B*, 68(2), 261-275.

https://doi.org/10.1140/epjb/e2009-00090-x

Freeman, L.C., 1979. Centrality in social networks' conceptual clarification. *Social Networks* 1(3), 215-239. https://doi.org/10.1016/0378-8733(78)90021-7

Garmendia, M., Urena J.M., Ribalaygua C., Leal J., Coronado J.M., 2008. Urban residential development in isolated small cities that are partially integrated in metropolitan areas by high speed train. *European Urban and Regional Studies* 15(3), 249-264. https://doi.org/10.1177/0969776408090415

Geertman, S.C.M., Ritsema van Eck, J.R., 1995. GIS and models of accessibility potential: an application in planning. *International Journal of Geographical Information Systems* 9(1), 67-80. https://doi.org/10.1080/02693799508902025

Geng, B., Bao, H., Liang, Y., 2015. A study of the effect of a high-speed rail station on spatial variations in housing price based on the hedonic model. *Habitat International*, 49, 333-339. https://doi.org/ 10.1016/j.habitatint.2015.06.005

Givoni M., 2007. Development and impact of the modern high-speed train: A review. *Transport Reviews* 26(5), 593-611. https://doi.org/10.1080/01441640600589319

Grant B.S., and Shreyas B., 2020. Governance Institutions: Key Elements for the Integrated Planning and Equitable Deliverability of High-Quality Transport Infrastructure. *Handbook on High-Speed Rail and Quality of Life.* Tokyo: Asian Development Bank Institute.

Hall P., 2009. Magic carpets and seamless webs: opportunities and constraints for high-speed trains in Europe. *Built Environment* 35(1), 59-69. https://doi.org/10.2148/benv.35.1.59

Hansen, W.G., 1959. How accessibility shapes land use. *Journal of the American Institute of Planners* 25(2), 73-76. https://doi.org/10.1080/01944365908978307

He, S. Y., 2020. Regional impact of rail network accessibility on residential property price: Modelling spatial heterogeneous capitalization effects in Hongkong. *Transportation Research Part A: Policy and Practice* 135, 244-263. https://doi.org/10.1016/j.tra.2020.01.025

Hensher, D., Li, Z., Mulley, C., 2012. The impact of high speed rail on land and property values: A review of market monitoring evidence from eight countries. *Road & Transport Research 21*, 3-14. https://doi.org/ 10.1109/MVT.2012.2218144

Ingram, D.R., 1971. The concept of accessibility: A search for an operational form. *Regional Studies* 5(2), 101-107. https://doi.org/10.1080/09595237100185131

Jensen, F., Stewart N.F., 1977. A sensitivity analysis of the gravity model. *Information Systems and Operational Research* 15(3), 308-321.

https://doi.org/10.1080/03155986.1977.11731678

Kim, J., Li, Y.T., Schmöcker, J.D., 2019. Regional heterogeneity in Taiwan HSR demand developments: station accessibility and its effect on usage adoption. *European Planning Studies* 27(3):555-573. https://doi.org/10.1080/09654313.2018.1562654

Lee, C., Ryu, K., Choi, K., Kim J.Y., 2018. The dynamic effects of subway network expansion on housing rental prices using a repeat sales model. *International Journal of Urban Sciences* 22(4), 529-545. https://doi.org/10.1080/12265934.2018.1487331

Li, T., Rong, L., 2020. A comprehensive method for the robustness assessment of high-speed rail network with operation data: a case in china. *Transportation Research Part A: Policy and Practice*, 132, 666-681. https://doi.org/ 10.1016/j.tra.2019.12.019

Li, X., Huang, B., Li, R., Zhang, Y., 2016. Exploring the impact of high speed railways on the spatial redistribution of economic activities - Yangtze River Delta urban agglomeration as a case study. *Journal of Transport Geography* 57, 194-206. https://doi.org/10.1016/j.jtrangeo.2016.10.011

Lin, T., Xia, J., Robinson, T.P., Goulias, K.G., Church, R.L., Olaru, D., Tapin, J., Han, R., 2014. Spatial analysis of access to and accessibility surrounding train stations: a case study of accessibility for the elderly in Perth, Western Australia. *Journal of Transport Geography* 39, 111-120. https://doi.org/10.1016/j.jtrangeo.2014.06.022

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

Liu, S., Wan Y., Zhang A., 2020. Does China's high-speed rail development lead to regional disparities? A network perspective. *Transportation Research Part A: Policy and Practice* 138, 299-321. https://doi.org/10.1016/j.tra.2020.06.010

Low, J., Lee, B. K., 2020. A data-driven analysis on the impact of high-speed rails on land prices in Taiwan. *Applied Sciences*, 10(10), 3357. https://doi.org/ 10.3390/app10103357

Mahalanobis, P. C., 1936. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*. 2 (1): 49–55.

Marcus, S.E., Moy, M., Coffman, T., 2006. Social network analysis. In: Cook, D.J., Holder L.B., Mining graph data. John Wiley & Sons Inc., New Jersey. 443-468. https://doi.org/10.1002/9780470073049.ch17

Martínez Sánchez-Mateos, H.S., Givoni, M., 2012. The accessibility impact of a new High-Speed Rail line in the UK – a preliminary analysis of winners and losers. *Journal of*

Transport Geography 25, 105-114. https://doi.org/10.1016/j.jtrangeo.2011.09.004

Mathur S., Ferrell C., 2013. Measuring the impact of sub-urban transit-oriented developments on single-family home values. *Transportation research part A: policy and practice* 47, 42-55. https://doi.org/10.1016/j.tra.2012.10.014

Michaels, G., Rauch, F., Redding, S. J. 2012. Urbanization and structural transformation. *The Quarterly Journal of Economics*, 127(2), 535–586. https://doi.org/10.1093/qje/qjs003

Moseley, M. J. 1979. Accessibility: The Rural Challenge, London, Methuen.

Muraco, W. A., 1972. Intra-urban accessibility. *Economic Geography* 48(4), 388-405. https://doi.org/10.2307/142890

Petersen, M., 2009. Estimating standard errors in finance panel data sets: comparing approaches. *The Review of Financial Studies* 22(1), 435–480. https://doi.org/10.1093/rfs/hhn053.

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

Redding, S.J., Turner, M.A., 2015. Transportation costs and the spatial organization of economic activity. *Handbook of Regional & Urban Economics*, 5(8), 1339-1398. https://doi.org/10.1016/B978-0-444-59531-7.00020-X

Ren, X., Chen, Z., Wang, F., Dan, T., Liu, C., 2020. Impact of high-speed rail on social equity in china: evidence from a mode choice survey. *Transportation Research Part A Policy and Practice*, 138, 422-441. https://doi.org/10.1016/j.tra.2020.05.018

Rohit, S., Peter, N. 2018. Does urban rail increase land value in emerging cities? value uplift from Bangalore metro. *Transportation Research Part A: Policy and Practice* 117, 70-86. https://doi.org/ 10.1016/j.tra.2018.08.020

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

Shaw, S., Fang Z., Lu, S., Tao, R., 2014. Impacts of high speed rail on railroad network accessibility in China. *Journal of Transport Geography* 40, 112-122. https://doi.org/10.1016/j.jtrangeo.2014.03.010

Wang, L., Duan, X., 2018. High-speed rail network development and winner and loser cities in megaregions: the case study of Yangtze river delta, china. *Cities*, 83(DEC.), 71-82. https://doi.org/ 10.1016/j.cities.2018.06.010

Wang, L., Yuan, F., Duan, X., 2018. How high-speed rail service development influenced commercial land market dynamics: a case study of Jiangsu province, china. *Journal of Transport Geography*, 72, 248-257.https://doi.org/ 10.1016/j.jtrangeo.2018.09.010

Wasserman, S., Faust, K., 1994. Social network analysis: Methods and applications. Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511815478

Wheeler, D.C., O'Kelly M.E., 1999. Network topology and city accessibility of the commercial internet. *Professional Geographer* 51(3), 327-339. https://doi.org/10.1111/0033-0124.00169

Xu, T., Zhang M., Aditjandra P. T., 2016. The impact of urban rail transit on commercial property value: New evidence from Wuhan, China. *Transportation Research Part A: Policy and Practice* 91, 223-235. https://doi.org/10.1016/j.tra.2016.06.026

Yao, S., Zhang, F., Wang, F., Ou, J., 2019. Regional Economic Growth and the Role of High-speed Rail in China. *Applied Economics* 51(32), 3465-3479. https://doi.org/10.1080/00036846.2019.1581910

Yao, S., Fang, J. He, H., 2020. Can time-space compression promote urban economic growth? Evidence from China's high-speed rail projects. *China & World Economy*, 28(5), 90-117. https://doi.org/10.1111/cwe.12339

Zhang, J., Wang, L., Wang, S., 2012. Financial development and economic growth: Recent evidence from China. *Journal of Comparative Economics* 40(3), 393-421. https://doi.org/10.1016/j.jce.2012.01.001

Zheng, S., Kahn, M.E., 2013. China's bullet trains facilitate market integration and mitigate the cost of mega city growth. *Proceeding of National Academy of Sciences of the United States of America* 110 (14), E1248-E1253. https://doi.org/10.1073/pnas.1209247110

Zheng, S.Q., Saiz, A., 2016. Introduction to the special issue "China's urbanization and housing market". *Journal of Housing Economics* 33, 1-3. https://doi.org/10.1016/j.jhe.2016.07.001