Transportation Research Part A 199 (2025) 104572

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Fine-grained analysis of transport-demographic relationships: County-level responses to multimodal connectivity across metropolitan and peripheral China

Junxi Qu [a], Xiaoyi Ma [b], Yang Zhou [a], Xianlong Chen [b], Tianren Yang [a,c,d,*]

[a] Department of Urban Planning and Design, The University of Hong Kong, Hong Kong SAR 999077, China
[b] Guangzhou Transport Planning Research Institute Co., Ltd., Guangzhou 510030, China
[c] Urban Systems Institute, The University of Hong Kong, Hong Kong SAR 999077, China
[d] Shenzhen Institute of Research and Innovation, The University of Hong Kong, Shenzhen 518057, China

ARTICLE INFO

Keywords:
Multimodal transport network
Population dynamics
Network centrality
Urban modelling
Urban systems
Computational urban science

ABSTRACT

Despite extensive research on transport-induced demographic change, few studies have systematically investigated how different transport modes shape population distribution across diverse geographical contexts at fine spatial scales. This study explores the relationship between multimodal transport networks and population density using 20 years of county-level panel data, focusing on differential impacts in metropolitan and peripheral counties across China. Employing high-dimensional fixed effects models and centrality measures derived from aviation, high-speed rail and conventional rail networks, our results show that higher degree centrality (i.e., more direct connections) is positively associated with population concentration (+0.018%), while higher inverse closeness centrality (i.e., greater average shortest distance to all other nodes) is negatively associated (−0.005%). These effects are not instantaneous but emerge with significant five-year lags. Metropolitan counties experience approximately 4% greater population gains from improved connectivity than peripheral regions. While connectivity effects are observed across all transport modes, high-speed rail exhibits relatively consistent and positive associations with population growth over longer time lags, although its effects in metropolitan areas are generally weaker than those of aviation and conventional rail. Complementing the regression results, LightGBM-based SHAP analysis reveals substantial spatial heterogeneity: even within the same classification (metropolitan or peripheral), counties with advantageous network positions—such as regional hubs—exhibit markedly stronger demographic responses. The findings offer valuable guidance for urban planning and transport policy, emphasising the need for targeted, mode-sensitive investment strategies that account for regional disparities in transport access and development potential.

1. Introduction

The relationship between transport infrastructure and population distribution has long been a central concern in regional economics and transport geography. As nations invest in extensive transport networks, understanding how these investments shape

* Corresponding author at: Department of Urban Planning and Design, The University of Hong Kong, Hong Kong SAR 999077, China
E-mail address: tianren@hku.hk (T. Yang).

https://doi.org/10.1016/j.tra.2025.104572
Received 19 November 2024; Received in revised form 30 May 2025; Accepted 15 June 2025
Available online 23 June 2025
0965-8564/

demographic patterns becomes increasingly important for policymakers and planners (e.g., Liotta et al., 2023; Ogawa & Fujita, 1980). China presents a particularly compelling case for examining these dynamics, having experienced unprecedented transport infrastructure development alongside dramatic demographic shifts over the past two decades.

Despite the significant benefits of transport investment (Allen & Arkolakis, 2022; Li et al., 2025; Lieske et al., 2021), debate persists regarding which regions benefit most from these networks. Transport systems such as aviation often favour developed regions with advanced economies (Sun et al., 2024), while in less developed areas, transport connectivity remains a critical driver of economic growth, particularly in rural regions (Chen et al., 2023). The varied impacts between metropolitan and peripheral areas, alongside differing effects of various transport modes (e.g., Baum-Snow et al., 2020), challenge our understanding of how transport infrastructure reshapes population patterns.

Previous research has established that transport connectivity influences population distribution through improved accessibility to employment, education and services (Condeço-Melhorado et al., 2014; Guo et al., 2022; Jiao et al., 2014; Lee, 2022; Ling et al., 2024; Wu and Levinson, 2024). However, most studies have focused on single transport mode or limited geographical context, often overlooking the complex, multimodal nature of modern transport systems. Furthermore, the differential impacts on metropolitan versus peripheral regions remain inadequately understood, particularly in rapidly developing economies like China where regional disparities persist despite substantial infrastructure investments.

This research aims to quantify how different aspects of multimodal transport connectivity influence population density across metropolitan and peripheral regions in China. By applying network centrality measures to China's evolving transport systems—including aviation, high-speed rail (HSR) and conventional rail (CR)—we examine the relationship between connectivity and demographic dynamics at the county level, accounting for both immediate and lagged effects.

Network theory provides a valuable framework for conceptualising transport systems as interconnected nodes and links, where connectivity benefits emerge not merely from direct connections but from a relative position within the broader network. Centrality measures derived from network analysis capture these complex relationships, providing nuanced metrics of accessibility and connectivity that extend beyond simple distance or travel time calculations.

This research contributes to the literature in several ways. First, it offers a fine-grained analysis of transport-demographic relationships at the county level, providing greater spatial resolution than previous studies that typically examine city or regional scales. Second, it provides a comprehensive assessment of multiple transport modes (aviation, HSR and CR), revealing how different systems influence demographic patterns with varying magnitudes. Third, it identifies time-dependent relationships between transport development and population density, demonstrating both immediate and lagged demographic responses to connectivity changes. Fourth, it quantifies the differential effects between metropolitan and peripheral regions, providing empirical evidence of how similar transport investments yield uneven population outcomes across China's urban hierarchy. Finally, it reveals important spatial heterogeneity beyond the metropolitan-peripheral divide, showing how specific geographical contexts modify the impact of transport connectivity on population density dynamics.

The remainder of this paper is structured as follows. Section 2 reviews existing studies on transport networks and urban growth. Section 3 presents the methodological framework, including graph development, regression and machine learning models. Section 4 analyses the evolution of network centralities and their impacts on population density. Finally, Section 5 offers conclusions and policy recommendations.

2. Literature review

2.1. Impacts of transport infrastructure and network factors

A large body of literature examines the socioeconomic impacts induced by the existence of transport infrastructure. For instance, transport infrastructure is positively correlated with population, employment, production, mixed land use and high housing price (Deng, 2013; Kasraian et al., 2016; Lakshmanan, 2011; Lieske et al., 2021; Ratner & Goetz, 2013). However, investments in transport infrastructure may not always yield the benefits. An inverted-U impact on urban population and housing price induced by transport investment can be identified in countries such as China, Spain and the United States (Han et al., 2023; Maroto & Zofío, 2016; Seo et al., 2014). Other negative outcomes, such as noise and pollution directly generated from transport infrastructure, can also bring adverse impacts (Friedt & Cohen, 2021; Zhao et al., 2022).

Network factors of transport infrastructure are increasingly introduced to investigate their impacts on urban performance. Service frequency and total passengers are commonly adopted to examine their impacts on population and economic outcomes (Krishnan et al., 2015; Moyano et al., 2018; Wu & Levinson, 2024). Furthermore, market access, employment and productivity are positively correlated with HSR frequency and expanded network (Li et al., 2020; Lin, 2017). Similarly, the frequency of aviation services and port throughput contribute to population and economic growth (Blonigen & Cristea, 2015; Wang et al., 2021). These studies emphasise that the benefits should be primarily attributed to increased accessibility and relative position within the network, rather than the presence of transport infrastructure (Jiao et al., 2020).

Limited studies investigate the impacts of network centrality on urban growth. For instance, in-closeness centrality and degree centrality of railway contribute to economic output (Huang et al., 2020; Tian et al., 2023), whereas out-closeness centrality may reduce it (Huang et al., 2020). In parallel, degree and closeness centrality is positively associated with housing price (Liu et al., 2021). Regarding network expansion, there is a clear trend that the high-ranked centrality of HSR trunks tended to concentrate in cities with dense populations and advanced economics (Jiao et al., 2017). Specific cities may become critical nodes leading to the new transport community in the network. For instance, the aviation and economic networks are largely concentrated in a few primary cities in China, reflecting an aviation community and the unbalance connection (Lao et al., 2016). Some studies display rich-hub effects in developed regions and large cities that govern the network linkages (Ducruet et al., 2018). Typical research examining the impacts of network factors and centralities is summarized in Table 1.

2.2. Varied impacts across socioeconomic backgrounds and transport modes

Moreover, the impacts of transport infrastructure and network factors can be affected by socioeconomic and geographical characteristics (Zhou & Zhang, 2021; Zou & Chen, 2024). An ongoing debate concerns whether the transport infrastructure and networks enhance growth in primal or non-primal areas. On the one hand, transport infrastructure is usually viewed as a key driver of population and economic development in major cities (Baum-Snow et al., 2020), leading to a loss in peripheral areas (Rokicki & Stępniak, 2018; Wang, 2018). Consistent identifications can be seen in the US and China, where central areas characterized by high income and dense economic activities benefit more from transport networks (Allen & Arkolakis, 2022; Baum-Snow et al., 2017). On the other hand, observations in Africa and China also suggest that transport infrastructure may contribute to population and economic growth in peripheral cities (Jedwab & Moradi, 2016; Luo & Zhao, 2021). In addition, some studies indicate that the areas along transport routes benefit more from the expansion of transport networks, while the impacts are marginal in other areas (Cascetta et al., 2020). Regarding network centralities, the degree and closeness centrality of HSR display a greater impact on housing prices in the central regions near the developed eastern areas (Liu et al., 2021). Consistently, centrality indices tend to exert a greater impact in cities with a population of 0.78–1.56 million than in others (Tian et al., 2021).

Diverse impacts also exist among different transport modes. Beyond the positive impacts of highway, railway, aviation and port transport on population and economic growth (e.g., Cascetta et al., 2020; Cristea, 2023; Pratama et al., 2022; Wang et al., 2021), railways may have a greater impact than highways in some cases (Jedwab & Moradi, 2016), while highways can also have a larger impact than railways under different contexts (Jiao et al., 2020). This variation is largely due to different transport modes operating efficiently at different distances and socioeconomic contexts (Liu et al., 2016).

In summary, previous studies have primarily focused on the effects of transport infrastructure (Wang et al., 2020) and the influence of flow indicators on urban growth (Chakrabarti et al., 2022; Wu & Levinson, 2024). However, uncertainties arise due to varying socioeconomic and geographical characteristics, complicating the impacts of transport networks and leading to inconsistent conclusions (Jiao et al., 2020; Rokicki & Stępniak, 2018). These complexities intensify when multiple transport modes are considered under different backgrounds. To address these research gaps, this study investigates the relationship between centrality indicators and population density across multiple transport modes and urban structures, extending the findings from the existence of transport infrastructure.

3. Data and methodology

This study examines the relationship between multiple modes of transport (aviation, HSR and CR) and population density at the county level across China. These transport modes[1] were selected because they represent significant national infrastructure investments and provide long-distance connectivity that extends beyond regional boundaries.

3.1. Study area and county-level unit definition

All county-level units in China with available data were investigated, including counties (*县*), county-level cities (*县级市*), districts (*市辖区*), banners (*旗*), and autonomous counties (*自治县*). To ensure consistency across the study period (2000–2020), we harmonised administrative boundaries to account for changes in administrative divisions. The resulting dataset comprises 2,627 county-level units,[2] providing comprehensive national coverage. In the context of intercity transportation planning—which is often determined at the national level—three categories of administrative divisions play a particularly critical role in shaping regional development and the political hierarchy: municipalities directly under the central government (*直辖市*), provincial capital cities (*省会城市*), separately planned cities (*计划单列市*, sometimes known as sub-provincial cities). These urban centres typically receive prioritised investment in transportation infrastructure and are designated as regional hubs, benefiting from preferential policy support from the central government. Given their strategic importance, we classify county-level units into two broad categories: metropolitan regions, which include all county-level units located within these key political and regional centres, and peripheral regions, referring to county-level units situated outside these central areas. This classification resulted in 414 metropolitan county-level units as displayed in Fig. 2, which according to the 2020 population census, account for over 26% of China's total population. The peripheral county-level units are more sparsely populated but constitute the majority of China's territory. This metropolitan-peripheral dichotomy

[1] Our research specifically focuses on networked transportation infrastructure that creates differential connectivity advantages between regions. While roads are indeed crucial for local accessibility, they typically provide more uniform coverage across regions compared to the more centralised and hierarchical networks of rail and aviation. Road networks generally develop incrementally and provide more distributed connectivity, whereas rail stations and airports create more concentrated nodes of accessibility that can significantly alter regional development patterns.

[2] Hong Kong, Macau, Taiwan, as well as the Xisha and Nansha Islands, are excluded due to data unavailability. Additionally, some units lacking population statistics in earlier years are also excluded from the regression analysis. Ultimately, 2,627 units out of more than 2,800 remain for analysis.

Table 1
Overview of the studies on transport networks and urban performance.

Literature source	Study area	Transport modes	Network factors	Methods	Key findings
Blonigen and Cristea (2015)	Metropolitan statistical area, the US	Aviation	Passengers	Difference-in-difference	Aviation is positively correlated with population and economic growth.
Ducruet et al. (2018)	Global cities	Port	Centrality indices	Complex network	Larger cities dominate smaller cities Larger cities primarily connect with each other.
Jiao et al. (2020)	Cities, China	HSR Aviation Highway Port	HSR frequency Air throughput Port throughput	Spatial regression	Positive impacts induced by HSR can be attributed to network connectivity rather than the existence of infrastructure. Impacts vary across different locations. HSR's impacts on economic growth are smaller than those of highways but larger than those of aviation and maritime transport.
Huang et al. (2020)	Cities, China	HSR	HSR frequency Centrality	Complex network and spatial regression	HSR centrality and frequency contribute to economic growth. HSR's impacts are affected by the development stage and the position in the network. Higher in-closeness centrality inclines to promote economic growth.
Guo et al. (2020)	Agglomerations, China	HSR	HSR frequency Centrality	Complex network and quadratic assignment procedure	HSR frequency and centrality contribute to economic growth. Developed areas enjoy a more significant impact than others.
Wang et al. (2021)	Entire territory, China	HSR Aviation Highway Port	HSR frequency Air throughput Port throughput	Vector autoregressive and vector error correction model	Transport infrastructure drives economic growth. Maritime transport infrastructure promotes economic growth.
Wu and Levinson (2024)	Counties, China	HSR Aviation Highway	HSR frequency Air throughput	Fixed-effects regression	HSR frequency is positively related to population density. New HSR positively affects population and economic growth shortly after construction. Aviation and port infrastructure contribute to economic growth.

provides a framework to investigate differential responses to transport connectivity between densely populated urban cores and less developed hinterlands.

3.2. *Transport network data*

3.2.1. *Air transport network*

Air transport data were collected from the Civil Aviation Administration of China (CAAC) for 2000, 2005, 2010, 2015, and 2019, including aviation routes between cities, the number of flights and airport throughput of each city. While some source data is catalogued as 2020, we use 2019 operational data to avoid capturing the significant disruptions caused by the COVID-19 pandemic. Intercity aviation routes were assigned directly to airports, as most cities had only one commercial airport. For cities with two airports, such as Shanghai and Beijing, the number of flights was apportioned based on each airport's throughput proportion. Each airport was then assigned to its corresponding county-level unit for analysis.

3.2.2. *High-speed rail network*

HSR data were collected from the China Railway Corporation and railway maps. Due to the rapid development of the HSR network, our analysis focused on 2010, 2015 and 2019. Since the first HSR route in China was constructed in 2008, the start year of the HSR network in this study is set as 2010. Following established definitions, we included lines with operating speeds of at least 250 km/h. For each time point, we constructed a network with nodes representing HSR stations and edges representing direct connections between them.

3.2.3. *Conventional rail network*

CR network data were obtained from the Ministry of Transport and historical railway atlases. Train timetable data for 2010 and 2015 were sourced from the China Passenger Railway Timetable published by the Ministry of Railway, and the data for 2019 was sourced from the China Railway Official Platform. Due to the lack of detailed timetable information for 2000 and 2005, CR stations within 2 km of the railway network in 2000 and 2005 were identified using historical railway maps to represent the network at those time points. We constructed networks for all five time points in our study period (2000–2019). As with the HSR network, nodes represented stations and edges represented direct connections.

Index	Start	Arrive
Rail 1	A	B
Rail 1	B	C
Rail 2	A	C
Air 4	C(D)	D(C)

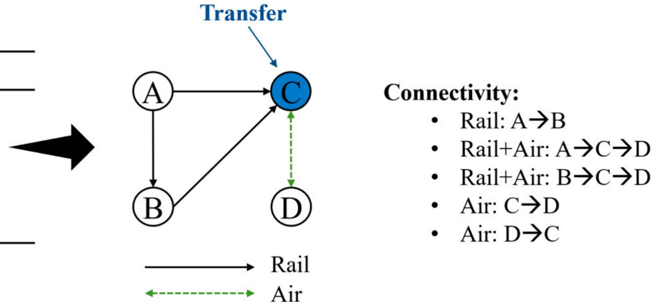

Fig. 1. Topology of multimodal transport network.

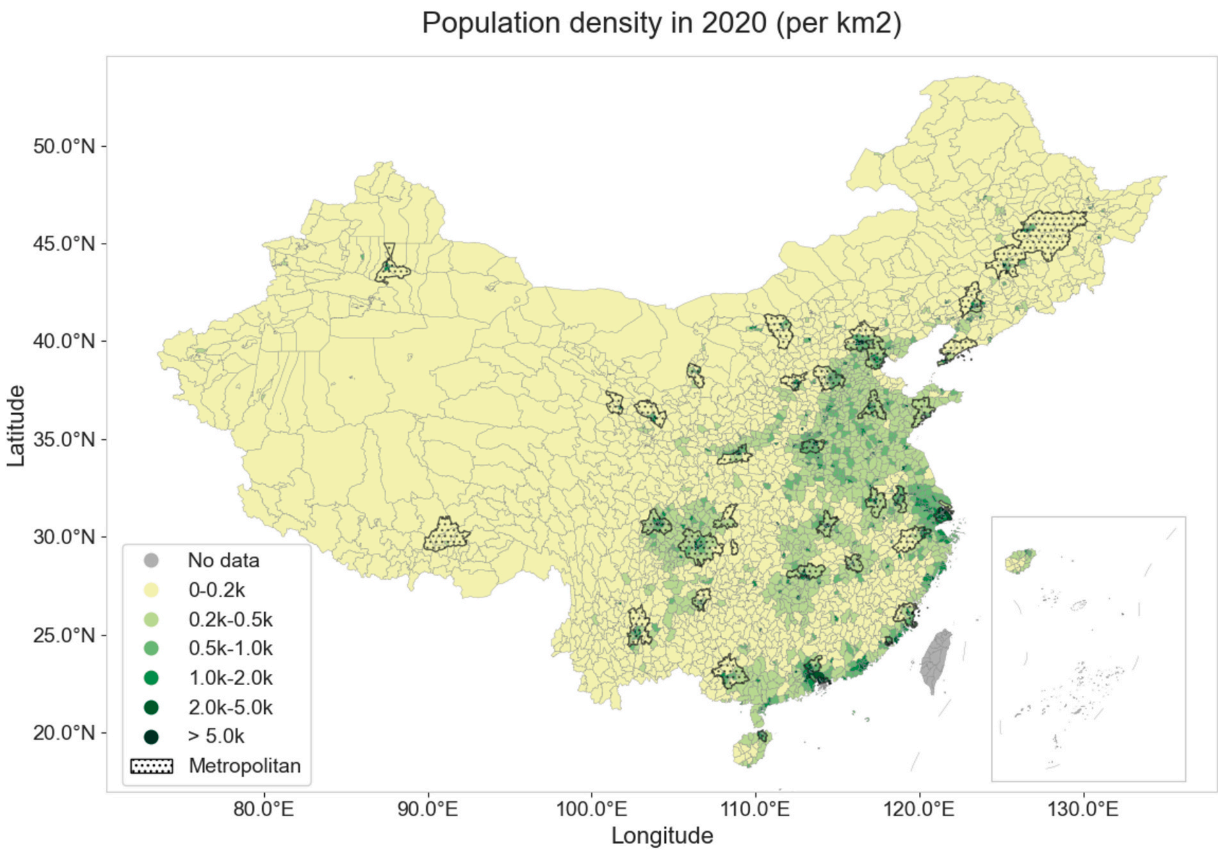

Fig. 2. Population density in 2020. **Notes:** This figure illustrates the spatial distribution of population density in 2020. All maps, including those depicting the railway and aviation networks, are based on the standard map provided by Tianditu. The data has been matched to the map without altering any national, city or county-level boundaries from the standard version (GS(2024)0650). This is available for free download in GeoJSON format at https://cloudcenter.tianditu.gov.cn/administrativeDivision.

3.2.4. County-level assignment of transport infrastructure and multimodal connectivity

We recognise that major transport hubs often serve broader regions beyond the county-level units where they are physically located. For instance, Beijing West Railway Station serves all districts and counties in Beijing (even beyond), not just Fengtai District where it is located. To address this issue while maintaining analytical precision, we adopt a dual approach to infrastructure assignment.

In our primary analysis, we assign connectivity metrics to county-level units based on infrastructure within their administrative boundaries. For multimodal connectivity assessment, we employ a sequential connection approach. Specifically, if County A connects to County B via rail and County B connects to County C via aviation, then we classify County A and County C as indirectly connected in the multimodal network. This methodology captures benefits of intermodal transfers across different transport systems.

It should be noted that this approach does not incorporate radius buffers for transport infrastructure, which represents a methodological limitation. The analysis does not account for potential accessibility benefits derived from proximity to transport nodes in adjacent counties. For example, a county without rail infrastructure but adjacent to a county with a major station receives no connectivity benefit in our primary model.

To partially address potential spatial spillover effects, we conduct supplementary analyses examining the relationship between proximity to transport infrastructure and population dynamics. While this approach acknowledges that transport benefits extend beyond administrative boundaries, we recognise that a comprehensive spatial equilibrium model would better capture competitive and cooperative dynamics between counties (see, e.g., Qu et al., 2024). Such extensions represent promising directions for future research.

3.3. Network centrality measures

3.3.1. Network construction approach

Graph theory pairwise relationships between interconnected entities through vertices (nodes) connected by edges (links) (Bondy & Murty, 1976). Centrality indicators, which stem from graph theory and network analysis, are typically employed to rank the importance of a node within a network, based on its location and connection strength, considering friction like distance (Shanmukhappa et al., 2019; Wang et al., 2011). In this study, we construct a network representation of the multimodal transport system, where nodes represent county-level units with railway stations or airports, and links signify the presence of a connection between two county-level units.

We develop directed networks for both HSR and CR systems, where directional links represent train services from one station to another (Li et al., 2024). This directed approach captures the asymmetrical nature of rail services, where train frequencies, stopping patterns and travel times often differ significantly depending on direction, particularly in China's extensive rail network where service hierarchies create directional disparities. For instance, a smaller station might be served by more trains traveling toward a major hub than in the opposite direction.

In contrast, the aviation network is constructed as an undirected network (Li et al., 2022). This approach reflects the operational reality of commercial air travel, where flights typically operate as round trips with comparable service levels in both directions. While minor scheduling differences exist, they rarely create the significant directional accessibility disparities observed in rail networks.

For all transport modes, beeline (straight-line) distances between connected nodes are used as edge weights in the calculation of distance-based centrality measures such as closeness centrality. While actual travel distances along rail tracks or flight paths would differ from these geodesic distances, beeline distance provides a consistent basis for comparison across all transport modes. County government seats are selected as the vertices for all network construction, as they represent the administrative centres relevant to our analysis of population agglomeration at the county level.

3.3.2. Centrality measures

To measure the importance of these units within the network, two centrality indicators are employed. The first is degree centrality, which is defined as the number of direct connections a county has. It is calculated as: $DC(u) = \sum_{u \le v} a_{v,u}$, where $DC(u)$ represents the degree centrality of county u; and $a_{v,u}$ is a binary variable equal to 1 if counties u and v are connected, and 0 otherwise. Following existing studies (Sun et al., 2024; Wang et al., 2020), the weight is set to 1 in the topology-based centrality. Degree centrality indicates the number of zones that can be reached directly via railway or aviation system. A higher degree centrality signifies greater connectivity of a county-level unit within the multimodal transport networks.

Closeness centrality measures how easily a county-level unit can access other units, and it is calculated as: $CC(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)}$, where $CC(u)$ is the closeness centrality; $d(v,u)$ represents the shortest-path distance between counties v and u; $n-1$ is a normalising constant. Since this analysis includes distance as a factor, we adopt the inverse form of closeness centrality to maintain the consistency: $ICC(u) = \frac{\sum_{v=1}^{n-1} d(v,u)}{n-1}$. This inverse closeness centrality $ICC(u)$ represents the average distance to all other connected zones. A higher value indicates a greater average distance to other county-level zones, suggesting that the county is less centrally located within the transport network.

Fig. 1 presents a simplified network of four county-level units connected by two transport modes. The original railway schedule data is transformed into origin–destination records to align with the aviation route data. In the figure, black lines represent railway connections, while green dashed lines indicate aviation routes. Specifically, the railway topology illustrates directed connections from A to B, A to C, and A to B to C. An undirected aviation connection exists between C and D, making C a transfer point for travel to D via aviation.

When measuring DC, both indegree and outdegree—representing direct connections between nodes—are considered. Therefore, node A, connected to B and C, has a DC value of 2. Node B, connected to A and C, also has a DC value of 2. Node C, connected to A, B, and D, has a DC value of 3, while node D, connected only to C, has a DC value of 1. Regarding ICC, it reflects the average distance from the start node to all other nodes. Assuming all edges have a distance of 1, the ICC of A is 1.33, calculated by the sum of distances 1 (A → B), 1 (A → C), and 2 (A → C → D), divided by the 3 connected nodes (B, C, D). Similarly, the ICC value for B is 1.5, derived from 1 (B → C) plus 2 (B → C → D), divided by 2 nodes (C, D). Since C and D are only connected by aviation, their ICC value is 1.

When constructing the multimodal network, we apply two key principles for calculating centrality measures. First, for DC, when two county-level units are connected by multiple transport modes, we count this as a single direct connection rather than multiple connections. This approach maintains consistency with the fundamental definition of DC as the number of unique directly connected nodes. Second, for ICC, we consider only the shortest path between any pair of counties, regardless of which transport mode provides that shortest connection. This ensures that closeness centrality accurately reflects the minimal network distance between locations. When analysing individual transport modes, these principles are applied separately to each mode-specific network, allowing us to isolate and assess the contribution of each transport mode in our regression models.

3.4. Demographic and socioeconomic data

3.4.1. Data processing

A panel dataset spanning from 2000 to 2020 at five-year intervals was collected to examine the impacts of transport networks on population growth. Population data were derived from the Global Human Settlement Layer (Schiavina et al., 2022), with spatial raster data (1-km resolution) aggregated at the county-level units. To alleviate the effects of area variations, population density was used (using administrative area as denominator). The spatial distribution of population density is displayed in Fig. 2. Higher population densities are concentrated in the central, eastern and southeastern regions, while the western regions have lower population densities despite their larger administrative areas.

Due to data availability, the road network is represented by highway mileage, with data sourced from Gaode Map (pre-2015) and OpenStreetMap (post-2015). Port throughput data were obtained from the Port Statistic Yearbook. However, it is important to highlight that these two transport modes—road networks and ports—were not directly included in constructing the transport network in this study. This decision was made due to our focus on higher-level inter-county connectivity (aviation, HSR and CR) that spans longer distances, while highways primarily serve intra-county and shorter inter-county trips. Additionally, historical highway network data lacks the necessary precision and consistency for centrality analysis across the study period.

3.4.2. Variable definitions and descriptive statistics

Table 2 presents the definitions and descriptive statistics of the variables used in this study. The dependent variable is population density. Control variables include highway density, port throughput and a metropolitan indicator. For network centrality measures, we calculate DC and ICC for the overall transport network as well as separately for aviation, HSR and CR networks.

The descriptive statistics reveal significant variation in both population density and transport network centrality across Chinese counties. For example, while the average population density is 680 people per km^2, the maximum reaches 37,848 people per km^2. Similarly, substantial variation exists in network centrality measures, with some counties having extensive connectivity (maximum *DC* of 101 for the overall network) while others remain entirely unconnected to the formal transport networks measured in this study.

Table 2
Definition and statistics of the variables.

Variables	Descriptions	Mean	S.D.	Min	Max
Dependent Variables					
Population density	Average population density (per km^2)	680	2,072	0	37,848
Control Variables					
Highway	Average highway density (per km)	41.78	66.64	0	1,221
Port	Port throughput (10,000 tons)	258	2,983	0	112,009
Metropolitan dummy	1 for metropolitan county-level units, 0 otherwise	0.15	0.35	0	1.00
Network Centrality					
Overall					
DC	Degree centrality	2.14	4.60	0	101
ICC	Inverse closeness centrality weighted by distance	597	1,331	0	51,990
Aviation					
DC	Degree centrality	0.33	2.92	0	95.00
ICC	Inverse closeness centrality weighted by distance	64.28	376	0	5,278
HSR					
DC	Degree centrality	0.69	2.52	0	38.00
ICC	Inverse closeness centrality weighted by distance	196	974	0	28,681
CR					
DC	Degree centrality	1.82	3.49	0	45.00
ICC	Inverse closeness centrality weighted by distance	567	1,287	0	50,503

Note: S.D. = standard deviation.

3.5. Empirical strategy

3.5.1. High-dimensional fixed effects model

To investigate the relationship between transport network centrality and population growth, we employ a high-dimensional fixed effects (HDFE) model. This approach helps control for various unobserved factors such as development stage, city characteristics and geographical locations that can influence the relationship between transport infrastructure and population dynamics (Guo et al., 2020; Huang et al., 2020). The model is specified as:

$$\ln(popd_{ijt}) = \beta_0 + \beta_1 \ln(C_{ijt}) + \beta_2 \ln(X_{ijt}) + \eta_i + \mu_j + \sigma_t + \varepsilon_{ijt} \tag{1}$$

where $popd_{ijt}$ denotes the population density in county i, city j and year t; C_{ijt} is the set of centrality measures DC and ICC; X_{ijt} represents control variables including port throughput and highway density; η_i and μ_j capture the unobserved fixed effects (e.g., institutional factors) of counties and cities, respectively. σ_t represents time-fixed effects; β_0 is the constant term; and ε_{ijt} denotes the error term. All variables except for the metropolitan/peripheral region dummy are log-transformed to linearise their relationships.

3.5.2. Lagged effects model

Since transport infrastructure impacts may take time to materialise (Wu & Levinson, 2024), we incorporate 5-year lagged terms of centrality measures:

$$\ln(popd_{ijt}) = \beta_0 + \beta_1 \ln(C_{ijt}) + \beta'_1 \ln(C_{ij,t-5}) + \beta_2 \ln(X_{ijt}) + \eta_i + \mu_j + \sigma_t + \varepsilon_{ijt} \tag{2}$$

This specification allows us to distinguish between immediate and longer-term effects of transport connectivity on population density.

3.5.3. Heterogeneity analysis

To examine differential effects between metropolitan and peripheral regions, we interact centrality measures with a metropolitan indicator:

$$\ln(popd_{ijt}) = \beta_0 + \beta_1 \ln(C_{ijt}) + {\beta'}_1 \ln(C_{ij,t-5}) + \beta_2 \ln(X_{ijt}) + \beta_3 \ln(C_{ijt}) \times Metro + \eta_i + \mu_j + \sigma_t + \varepsilon_{ijt} \tag{3}$$

$$\ln(popd_{ijt}) = \beta_0 + \beta_1 \ln(C_{ijt}) + {\beta'}_1 \ln(C_{ij,t-5}) + \beta_2 \ln(X_{ijt}) + \beta'_3 \ln(C_{ij,t-5}) \times Metro + \eta_i + \mu_j + \sigma_t + \varepsilon_{ijt} \tag{4}$$

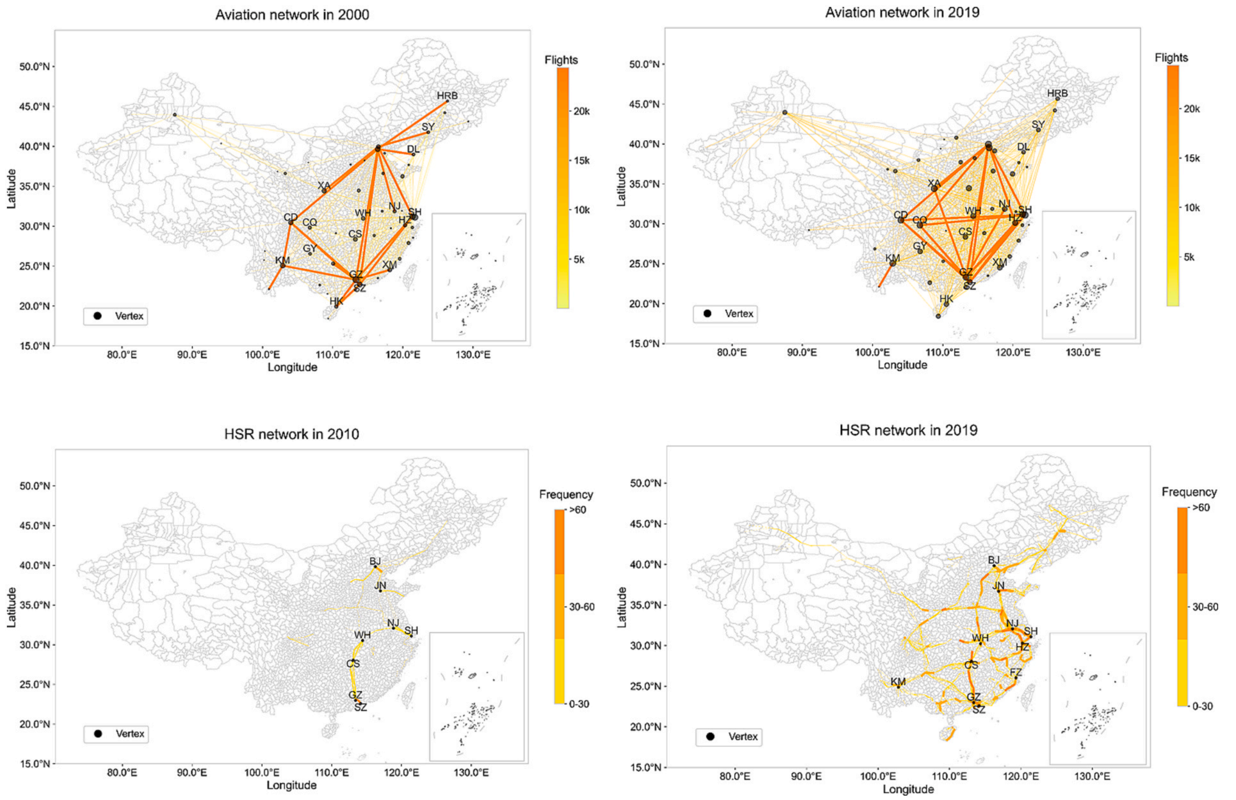

Fig. 3. Spatial expansion of aviation and high-speed rail (HSR) networks in China.

$$\ln(popd_{ijt}) = \beta_0 + \beta_1 \ln(C_{ijt}) + \beta'_1 \ln(C_{ij,t-5}) + \beta_2 \ln(X_{ijt}) + \beta_3 \ln(C_{ijt}) \times Metro + \beta'_3 \ln(C_{ij,t-5}) \times Metro + \eta_i + \mu_j + \sigma_t + \varepsilon_{ijt} \quad (5)$$

These specifications help us address the ongoing debate about whether transport infrastructure primarily benefits metropolitan or peripheral regions.

3.5.4. LightGBM model for local effects

To capture heterogeneous impacts across counties, we employ the LightGBM (Light Gradient-Boosting Machine) model, an advanced machine learning approach effective for high-dimensional and complex tasks (Ke et al., 2017). This model builds decision trees iteratively to improve prediction accuracy by focusing on misclassified instances.

For implementation, we divide the dataset into three parts: 70% for training and 30% for testing. 12.5% of the training dataset is further extracted to serve as a validation set for model evaluation. We incorporate SHAP (Shapley Additive exPlanation) values to enhance model interpretability (Lundberg & Lee, 2017). All variables except the metropolitan dummy are standardised using z-scores. Model performance is evaluated using R-squared and mean-squared error (MSE) metrics.

This machine learning approach complements our econometric models by identifying complex, non-linear relationships between transport network centrality and population patterns that might not be captured in the linear regression framework.

4. Results and discussion

4.1. Spatial-temporal evolution of transport networks and centrality indicators

Transport infrastructure has seen substantial investment over the past two decades. The total built mileage of highways, railways and aviation routes increased from 16,300 km, 68,700 km, and 994,682 km in 2000 to 161,000 km, 146,300 km and 5,597,565 km in 2020, respectively (Statistics, 2024). This rapid expansion has fundamentally reshaped connectivity patterns across county-level administrative units throughout the country.

As shown in Fig. 3, China had 126 airports in 2000, primarily concentrated in southeastern coastal areas such as Shenzhen and Guangzhou, as well as major eastern cities like Shanghai and the national capital, Beijing. By 2019, the number of airports had increased to 239, forming a more interconnected aviation network that now includes key cities such as Beijing, Xi'an, Chengdu, Chongqing, Guangzhou, Shenzhen, Shanghai, Hangzhou and Wuhan.

The railway system underwent even more dramatic transformation. In 2010, the railway network featured only four HSR communities, primarily located in the central, eastern and capital regions. By 2019, however, major corridors running from west to east and north to south had emerged. Due to the substitution of CR by HSR, the number of CR stations decreased. Specifically, there were 3,118 CR stations and 2,604 routes in 2000, compared to 2,012 CR stations (excluding those serving both CR and HSR) and 3,208 CR routes by 2019. Notably, the HSR network had approximately 7,500 routes in operation by 2019.

The spatial pattern of transport development shows clear regional differentiation. HSR stations and airports are predominantly located in or near metropolitan regions, such as densely populated urban agglomerations and provincial capitals. This spatial pattern aligns with existing studies (Guo et al., 2020; Wang, 2018), which found that transport infrastructure primarily concentrates and supports growth in metropolitan areas, leading to a more unbalanced distribution. Eastern and central regions have experienced the most intensive network development, particularly around major metropolitan clusters like Beijing-Tianjin-Hebei, the Yangtze River Delta and the Pearl River Delta. Western regions, while less densely connected, have seen strategic corridor development linking key

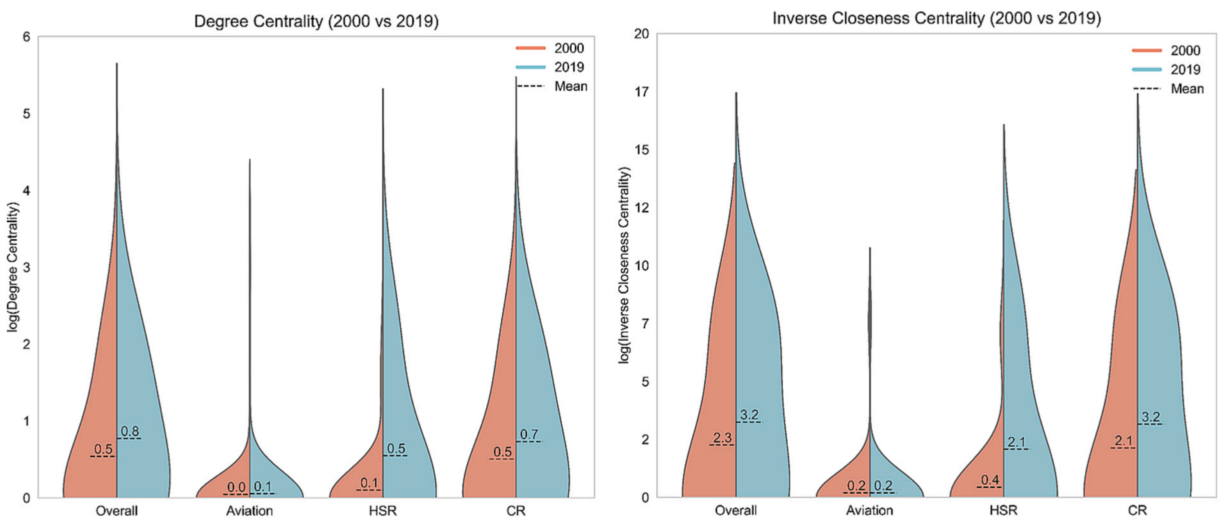

Fig. 4. Centrality evolution during 2000–2019.

urban centres.

This extensive transport network development is reflected in the evolution of centrality indicators. Fig. 4 displays the evolution of DC during 2000–2019. The expanded blue area in 2019 compared to 2000 signified an improvement in the number of direct connected counties. The larger area above a DC logarithmic value of 0.8 in 2019 indicated that more county-level units achieve values exceeding 0.5 compared to 2000. The distribution revealed that most zones maintain DC values between 0.5 and 2.0.

Similarly, ICC, which reflects the average shortest distance to other connected counties in the network, also changed significantly between 2010 and 2019. An increase in ICC indicates two contrasting phenomena: either decreased efficiency within an existing network (longer travel distances), or more extensive network coverage that now includes more distant counties. In China's case, the observed ICC changes primarily reflect the latter—the expansion of transport networks to connect previously unlinked distant regions. For example, the integration of far western regions like Xinjiang with eastern coastal areas through multimodal connections necessarily increases the average network distance (ICC) while simultaneously improving overall national connectivity.

Among the three transport modes examined, HSR and CR contribute most significantly to growth in DC and ICC. HSR demonstrates the most remarkable increase in connectivity over just a decade, though CR continues to play a key role in the railway system. The aviation network shows modest improvements in both DC and ICC. These patterns align with transport consumption statistics from the National Bureau of Statistics, which indicate rapid growth in railway passengers and freight (from 7% and 10% in 2000 to 23% and 13% in 2019, respectively), whilst aviation exhibits only marginal increases.

4.2. *Impacts of centrality indices on population density*

Prior to quantifying the impacts of the multimodal transport network on population density, we conducted a Wooldridge test to detect serial correlation. The resulting p-value (<0.001) indicated significant serial impact, informing our lagged modelling approach.

As presented in Table 3, DC demonstrates a positive correlation with population density when lagged terms are excluded, with a coefficient of 0.018. Conversely, a 1% increase in ICC correlates with approximately 0.005% decrease in population density. This finding reveals an important distinction between network expansion at the national level and its local effects. While the overall transport network development in China has increased ICC values nationally by connecting more distant regions (as discussed in Section 4.1), our regression results suggest that at the local level, counties with lower ICC values (better proximity to other connected regions) tend to have higher population density.

This relationship highlights a dual nature of transport network development: nationwide expansion increases overall ICC as more distant areas become connected, but areas with advantageous positions in the network (lower ICC) tend to gain population density

Table 3
High-dimensional fixed effects regression results on population density.

Variables	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)
DC	0.018***	0.016***	0.005	0.016***	0.005
	(5.12)	(4.90)	(1.45)	(4.85)	(1.33)
ICC	−0.005***	−0.004***	−0.002***	−0.004***	−0.002***
	(−6.72)	(−5.11)	(−2.96)	(−5.04)	(−2.83)
Lagged DC		0.019***	0.018***	0.005	0.004
		(5.42)	(5.28)	(1.32)	(1.16)
Lagged ICC		−0.004***	−0.004***	−0.002**	−0.002**
		(−5.48)	(−5.28)	(−2.15)	(−2.02)
DC × Metropolitan			0.041***		0.042***
			(5.00)		(5.15)
ICC × Metropolitan			−0.003		−0.004*
			(−1.62)		(−1.72)
Lagged DC × Metropolitan				0.057***	0.056***
				(6.85)	(6.78)
Lagged ICC × Metropolitan				−0.010***	−0.010**
				(−4.83)	(−4.53)
Port	0.003	0.004*	0.004*	0.004*	0.005*
	(1.16)	(1.69)	(1.73)	(1.78)	(1.83)
Highway	0.0002	−0.002***	−0.002***	−0.002***	−0.002***
	(0.47)	(−2.79)	(−2.85)	(−2.84)	(−2.91)
Constant	5.062***	5.070***	5.070***	5.070***	5.070***
	(4128.54)	(3254.60)	(3265.33)	(3262.58)	(3273.71)
County FEs	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
R-squared	0.9984	0.9989	0.9989	0.9989	0.9989
Root MSE	0.0769	0.0662	0.0660	0.0660	0.0657
Observations	13,135	10,508	10,508	10,508	10,580

Notes: Model (1) includes network centrality measures alone. Model (2) incorporates 5-year lagged terms to address serial correlation concerns. Model (3) adds interactions between centrality measures and metropolitan region dummy. Model (4) presents interactions between lagged centrality values and metropolitan region dummy. Model (5) estimates the combined effects of all lagged values and interactions. t-statistics in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

advantages. Counties that function as regional hubs or sit at strategic network positions benefit from their relative proximity to other connected regions.

Lagged centralities display similar impacts: 5-year lagged DC generates positive impacts on population density (0.019), whilst ICC exhibits a negative influence (−0.004). These findings partially align with observations by Wu and Levinson (2024), which confirmed such lagged impacts of HSR and aviation connectivity on population and economic indicators. Similarly, Zhang et al. (2020) found that transport network expansion enhances factor flows and contributes to urban concentration.

According to Model (5), ICC (−0.002) and its lagged value (−0.004) continue to demonstrate a negative correlation with population density. The interaction between the metropolitan county-level dummy and DC (0.042) or lagged DC (0.056) exhibits stronger positive impacts on population density compared to models without interaction terms. Conversely, the interaction between this dummy and ICC (−0.004) and its historical value (−0.010) shows stronger negative effects than models without interactions.

The longstanding debate regarding which areas benefit or experience losses from transport networks is partially addressed by our findings. The results demonstrate that transport network connectivity exhibits greater impacts on population density in metropolitan regions. These observations largely align with studies showing that metropolitan areas derive greater benefits from transport infrastructure (e.g., Allen & Arkolakis, 2022; Baum-Snow et al., 2020), rather than peripheral areas.

In addition, population density positively correlates with port throughput (0.005) but negatively relates with highway mileage density (−0.002). The positive relationship with port throughput likely reflects the historical importance of port cities as commercial centres, while the negative correlation with highway density suggests that extensive road networks may facilitate population dispersion rather than concentration (Qu et al, 2024).

The marginal impacts of the interactions in Model (5) are further predicted and displayed in Fig. 5. For metropolitan counties, the impact on population density increases substantially as DC increases, demonstrating that better-connected metropolitan areas experience significantly greater population density benefits. In contrast, for peripheral counties, this relationship is much weaker—while still positive, the slope is nearly flat, indicating that improved DC has minimal additional impact on population density in these areas.

Regarding ICC, metropolitan counties show a rapidly decreasing population density impact as ICC increases (indicating greater average distances to other network nodes), while peripheral counties exhibit a similar but much more modest negative trend. These patterns suggest that the advantages of strategic network positioning (low ICC) are substantially amplified in metropolitan contexts but have limited effects in peripheral areas. This helps explain why, despite national increases in ICC through network expansion, population density continues to concentrate in well-connected metropolitan regions with locally advantageous network positions. A parallel observation for lagged centrality values is presented in **Supplementary Material**, showing similar but distinct patterns between metropolitan and peripheral counties.

To ensure the robustness of our main findings, we implemented three approaches: First, we incorporated distance to the nearest transport hubs (airports, HSR and CR stations) as described in Section 3.2.4. The results show significant effects of proximity to these transport nodes, demonstrating that transportation networks impact extends beyond directly connected zones. Areas closer to transport infrastructure tend to benefit more, with the greatest advantages accruing to zones at the shortest distances compared to those without direct connections. Second, we replaced population density with real GDP density as the dependent variable, which confirmed our primary findings. Third, we tested alternative zoning systems by classifying areas into three categories: urban districts, county-level cities and other counties, which produced results consistent with our main analysis. Additionally, to address potential endogeneity concerns, we employed 5-year lagged centrality values as instrumental variables (IV). Detailed results from these robustness checks are presented in **Supplementary Material**.

4.3. *Heterogeneous impacts among different transport modes*

Building on our analysis of overall network centrality effects, this section examines how different transport modes uniquely influence population density. As shown in Table 4, the DC measures of aviation, HSR and CR all positively correlate with population density, though with varying significance and magnitude. HSR shows the strongest positive relationship (0.032), followed by aviation

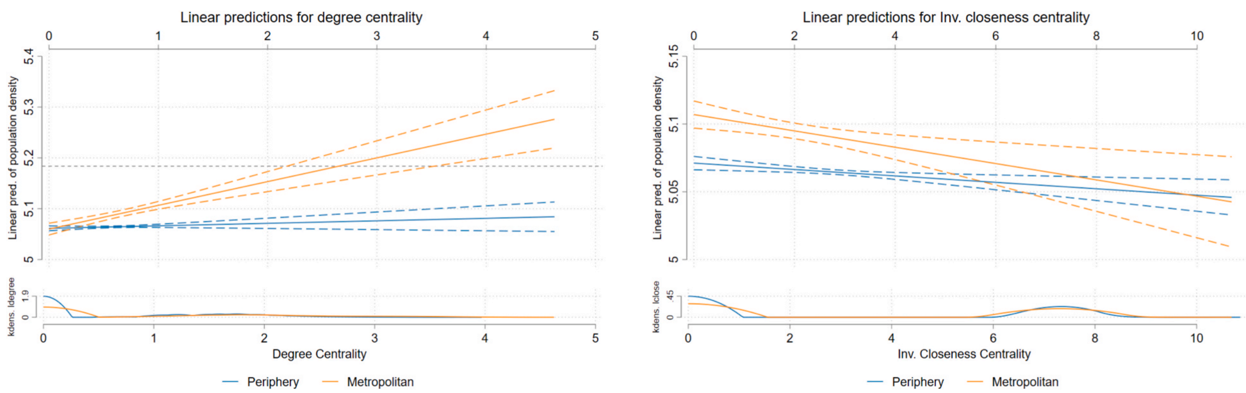

Fig. 5. Marginal impacts of the interaction term.

Table 4
Mode-specific impacts of transport network centrality on population density.

Variables	Model (6)	Model (7)	Model (8)	Model (9)	Model (10)
DC (Aviation)	0.022**	0.009	0.015	0.006	0.013
	(2.46)	(0.95)	(1.23)	(0.64)	(1.03)
ICC (Aviation)	−0.002	−0.0003	−0.001	−0.0004	−0.001
	(−1.08)	(−0.17)	(−0.72)	(−0.23)	(−0.64)
DC (HSR)	0.032***	0.010*	−0.002	0.008	0.002
	(6.59)	(1.94)	(−0.38)	(1.55)	(0.32)
ICC (HSR)	−0.004***	−0.001	0.0001	−0.001	−0.001
	(−3.47)	(−0.84)	(0.08)	(−0.43)	(−0.37)
DC (CR)	0.012***	0.012***	0.003	0.011***	0.003
	(3.23)	(3.49)	(0.88)	(2.99)	(0.82)
ICC (CR)	−0.004***	−0.003***	−0.002**	−0.003***	−0.002**
	(−5.18)	(−4.09)	(−2.17)	(−3.71)	(−2.12)
Lagged DC (Aviation)		0.020**	0.014	−0.007	−0.005
		(2.11)	(6.92)	(−0.52)	(−0.38)
Lagged ICC (Aviation)		−0.001	−0.001	0.002	0.002
		(−0.83)	(−0.31)	(1.10)	(1.03)
Lagged DC (HSR)		0.035***	0.034***	0.038***	0.040***
		(5.48)	(5.28)	(4.61)	(4.92)
Lagged ICC (HSR)		−0.005***	−0.005***	−0.008***	−0.008***
		(−3.31)	(−3.05)	(−3.93)	(−3.73)
Lagged DC (CR)		0.014***	0.012***	0.004	0.004
		(3.67)	(3.38)	(0.99)	(0.86)
Lagged ICC (CR)		−0.003***	−0.003***	−0.002*	−0.001
		(−4.10)	(−3.69)	(−1.74)	(−1.53)
DC (Aviation) × Metropolitan			−0.033		−0.021
			(−1.59)		(−1.01)
ICC (Aviation) × Metropolitan			0.011		0.004
			(1.23)		(0.47)
DC (HSR) × Metropolitan			0.018**		0.007
			(1.96)		(0.67)
ICC (HSR) × Metropolitan			0.002		0.004
			(2.27)		(1.36)
DC (CR) × Metropolitan			0.033***		0.031***
			(3.57)		(3.43)
ICC (CR) × Metropolitan			−0.004*		−0.004*
			(−1.76)		(−1.75)
Lagged DC (Aviation) × Metropolitan				0.058***	0.046**
				(3.00)	(2.29)
Lagged ICC (Aviation) × Metropolitan				−0.021***	−0.017**
				(−2.77)	(−2.04)
Lagged DC (HSR) × Metropolitan				−0.017	−0.023*
				(−1.35)	(−1.78)
Lagged ICC (HSR) × Metropolitan				0.012***	0.009***
				(3.61)	(2.74)
Lagged DC (CR) × Metropolitan				0.039***	0.041***
				(4.24)	(4.46)
Lagged ICC (CR) × Metropolitan				−0.008***	−0.008***
				(−3.45)	(−3.59)
Port	0.002	0.003	0.004	0.003	0.004
	(0.72)	(1.09)	(1.46)	(1.35)	(1.47)
Highway	0.0001	−0.002***	−0.002***	−0.002***	−0.002***
	(0.15)	(−3.00)	(−3.05)	(−3.10)	(−3.12)
Constant	5.0595***	5.069***	5.069***	5.070***	5.070***
	(4130.95)	(3274.63)	(3237.41)	(3204.34)	(3106.46)
County FEs	Yes	Yes	Yes	Yes	Yes
City FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
R-squared	0.9984	0.9989	0.9989	0.9988	0.9989
Root MSE	0.0765	0.0658	0.0653	0.0654	0.0651
Observation	13,135	10,508	10,508	10,508	10,508

Notes: Model (6) includes mode-specific centrality measures alone. Model (7) incorporates 5-year lagged terms to address serial correlation concerns. Model (8) adds interactions between centrality measures and metropolitan region dummy. Model (9) presents interactions between lagged centrality values and metropolitan region dummy. Model (10) estimates the combined effects of all lagged values and interactions. t-statistics in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

(0.022) and CR (0.012). Similarly, ICC demonstrates consistently negative relationships with population density across all modes: aviation (−0.002), HSR (−0.004) and CR (−0.004). When lagged effects are included, the pattern persists with positive impacts from lagged DC for aviation (0.020), HSR (0.035) and CR (0.014), as well as negative effects from lagged ICC for HSR (−0.005) and CR (−0.003).

After introducing interaction terms between centrality measures and the metropolitan county-level dummy, we observe how transport networks differentially affect metropolitan and peripheral areas. In Model (10), all three transport modes show distinctive patterns in their metropolitan versus peripheral effects. For aviation, the interaction between lagged DC and metropolitan dummy shows a significant positive impact (0.046), while the interaction with lagged ICC demonstrates negative effects (−0.017). This indicates that aviation connectivity particularly enhances population concentration in metropolitan areas, consistent with the hub-and-spoke structure of air transport networks.

HSR demonstrates unique patterns compared to other modes. The interaction between HSR's DC and metropolitan dummy initially shows positive impacts (0.018), but shifts to a negative impact (−0.023) when all interactions are considered. Notably, lagged HSR's ICC tends to increase population density in metropolitan areas (0.009 in Model (10)), contrary to the general negative relationship between ICC and population density observed elsewhere. This finding suggests that HSR may operate under different spatial dynamics than other modes, potentially enabling metropolitan regions to maintain population growth despite greater distances from economic centres. This distinctive HSR effect may relate to its speed advantages and specialised station placement, which can reshape economic geography at regional scales. These findings partially align with research by Jedwab and Moradi (2016) and Luo and Zhao (2021), which found that peripheral regions can benefit from transport infrastructure, though in our study this effect appears unique to HSR.

For CR, both DC and its lagged values display consistently positive impacts on population density in metropolitan areas (0.031–0.041), while ICC and its historical values show negative impacts (−0.004 to −0.008). This suggests that CR continues to play a fundamental role in shaping population distribution, particularly in metropolitan counties where it has historically served as the backbone of the transport system.

These mode-specific findings complement our spatial heterogeneity analysis in Section 4.2, revealing that transport connectivity effects vary not only by region but also by transport mode. The differential impacts of air, HSR and CR suggest that an integrated multimodal approach to transport planning is essential for balanced regional development, as each mode creates unique patterns of accessibility and connectivity that shape population distribution in different ways.

4.4. Localised and nonlinear effects revealed by the LightGBM model

We employ a machine learning approach named the LightGBM model to explore nonlinear and spatially heterogeneous impacts that might be overlooked in traditional econometric frameworks. For the overall LightGBM model, we obtained an R-squared of 0.7981 and a Mean Squared Error (MSE) of 0.2429, indicating strong predictive power for population density. The mode-specific model performed slightly lower with an R-squared of 0.7764 and an MSE of 0.263. While these performance metrics are somewhat lower than those of the HDFE model, this difference is expected since the HDFE approach incorporates multiple fixed effects (city, county, and year) that account for unobserved institutional, geographic and social factors. The value of the LightGBM model lies not in outperforming HDFE statistically, but in its ability to identify complex, non-linear relationships between transportation networks and population density that linear models might miss.

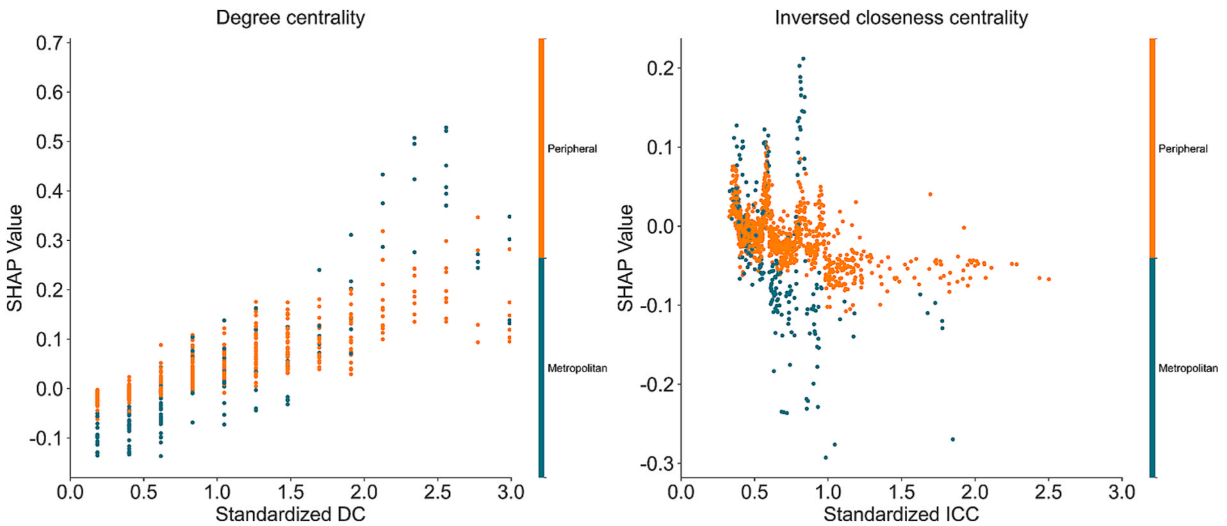

Fig. 6. SHAP values of centrality measures and metropolitan status. **Notes:** The x-axis is limited to the primary range of values for clarity. Positive SHAP values indicate features that increase predicted population density, while negative values indicate features that decrease predicted population density.

Fig. 6 illustrates how the relationship between centrality measures and population density varies across metropolitan and peripheral areas through SHAP values. As DC increases, the SHAP values for both metropolitan and peripheral areas tend to rise, indicating a positive impact on population density. Notably, SHAP values for metropolitan areas are significantly higher than those for peripheral zones when DC exceeds 2, though with greater variance. Conversely, SHAP values tend to decrease as ICC increases, consistent with our HDFE findings. These patterns largely align with the marginal effects shown in Fig. 5, confirming our earlier analysis while providing additional details.

However, the LightGBM model reveals important nonlinearities that were not apparent in the HDFE analysis. For instance, we observe negative SHAP values for the metropolitan dummy in regions with low DC, indicating that metropolitan status can negatively impact population density in poorly connected areas. Taking county-level units within Urumqi as an illustrative case, their connectivity with economically advanced regions, particularly Southeast China, relies heavily on aviation, with limited railway linkage. The degree centrality of HSR and CR is lower than 10. Despite Urumqi's status as the provincial capital of Xinjiang, expansions in transport infrastructure appear to be negatively associated with population growth. A similar pattern is observed in northeastern regions, such as the urban districts of Changchun. These areas share key characteristics: they are geographically remote and lie outside the core economic clusters of central, eastern and southeastern coastal China. As such, they occupy both peripheral positions in the national transport network and marginal roles in the country's economic hierarchy.

Similarly, positive SHAP values for ICC (which typically has a negative relationship with population density) are concentrated in regions with very high ICC values. These zones are predominantly in western regions and parts of northeast China that are far from core economic centres. In these remote areas, long distances become a positive driver of population growth, possibly serving as a form of isolation that fosters local development rather than extraction (Burger et al., 2017; Meijers & Burger, 2017).

The lagged and mode-specific analyses of centrality measures using SHAP values (shown in **Supplementary Material**) reveal similar nonlinear patterns, further confirming that the relationship between transport connectivity and population density involves complex threshold effects and context-dependent dynamics that vary across China's diverse regional landscape.

5. Conclusions

This study provides a fine-grained analysis of the relationship between transport connectivity and population dynamics in China during 2000–2020, revealing significant heterogeneity in how different regions respond to multimodal transport networks. By examining county-level data through the lens of network centrality measures, we have identified several important patterns that advance our understanding of transport-demographic relationships.

Our findings confirm that transport connectivity influences population distribution, but these effects are neither uniform nor immediate. Higher degree centrality (direct connections) positively correlates with population growth, while greater network distances (inverse closeness centrality) demonstrate negative effects. The magnitude of these relationships is modest but statistically significant, with a 1% increase in degree centrality associated with a 0.018% increase in population density. Importantly, these effects exhibit temporal lags, with significant coefficients appearing at 5-year intervals, suggesting that population responses to transport developments unfold gradually over time.

A key contribution of this research is the identification of distinctly different responses between metropolitan and peripheral regions. Metropolitan areas derive approximately 4% greater population benefits from enhanced connectivity (i.e., degree centrality) compared to peripheral regions. This differential effect varies notably across transport modes. Particularly significant is our finding that HSR exhibits unique spatial dynamics compared to other modes, with lagged HSR inverse closeness centrality showing positive effects on population density in metropolitan areas—contrary to the generally negative relationship observed with other modes. This suggests that HSR may reshape regional population patterns differently from conventional rail or aviation networks, potentially due to its distinctive speed–cost advantages and specialised station placement.

Machine learning analysis further reveals that spatial heterogeneity extends beyond the metropolitan–peripheral dichotomy. Even locations with similar administrative designations may experience contrasting outcomes depending on their position within the transport network. For instance, in poorly connected areas, metropolitan status can negatively impact population density, suggesting that administrative designation without adequate connectivity provides little demographic advantage.

These findings yield important implications for transport infrastructure planning and regional development. First, while transport investment has been a key driver of population growth in China, its effectiveness relies not only on service frequency but also on the degree of multimodal integration. Empirical observations reveal a pattern of strong–strong connectivity, whereby regions well-served by one mode of transport tend to exhibit strong performance across other modes as well. Conversely, areas with limited service in one mode often lack adequate coverage in others. This clustering effect underscores a critical consideration for infrastructure planning: if the objective is to promote growth within metropolitan regions, continued investment in enhancing service frequency and improving multimodal connectivity in these areas may be effective. However, if the goal is to mitigate spatial disparities and foster more balanced development between metropolitan and non-metropolitan areas, strategies should focus on strengthening interregional connectivity. Establishing stable transport communities may foster population agglomeration in peripheral regions and stimulate growth beyond already-developed cores. Second, for underdeveloped metropolitan counties and peripheral regions, transport investment alone is insufficient. Infrastructure improvements must be paired with place-based policies aimed at enhancing liveability and economic resilience (Xu et al., 2025). In light of ongoing demographic transitions, including population decline in certain regions, infrastructure planning should be informed by dynamic and long-term cost–benefit assessments that account for projected demographic and economic changes (see, e.g., Yang and Zhang, 2025).

Several limitations warrant consideration in future research. First, due to data availability, our analysis was restricted to five-year lag intervals, which may overlook both shorter-term and longer-term population responses to transport infrastructure developments. Incorporating geographical features and cultural indicators may provide a more comprehensive understanding of the heterogeneous effects of transport interventions across regions. Second, although road transportation is a fundamental factor influencing spatial accessibility and population distribution, its effects are only indirectly accounted for in our current modelling framework. Due to the lack of consistent historical road network data, we have not yet explicitly incorporated road infrastructure into the analysis. Third, our network construction was based on topological structures using beeline distances due to the unavailability of consistent schedule-based and route-level data across the full-time frame. While this approach ensures comparability across modes, it may not fully capture functional connectivity, particularly in regions where travel paths deviate significantly from straight-line distances. Future work may improve the accuracy of connectivity measures by incorporating detailed route and schedule-based information. Fourth, despite our efforts to improve the robustness, concerns about reverse causality remain. We therefore acknowledge this limitation and identify it as an important direction for future research.

Despite the limitations, this fine-grained analysis demonstrates that the relationship between transport connectivity and population change is characterised by significant heterogeneity across different regional contexts and transport modes. Furthermore, intermodal connectivity is increasingly important in understanding the real-world accessibility and mobility impacts of transport network (Zhu et al., 2018, 2019). We acknowledge that the integration of multiple transport modes (e.g., HSR and aviation) and their spatial–temporal coordination can significantly influence population distribution and regional accessibility. The various impacts among metropolitan and periphery areas should be considered. These findings underscore the need for context-sensitive approaches to transport planning and regional development policy, particularly in countries experiencing rapid infrastructure development alongside major demographic transitions. By recognising the differential responses of metropolitan and peripheral regions to various transport modes, policymakers can develop more targeted strategies to promote balanced regional development while addressing the complex challenges of urbanisation and population redistribution.

CRediT authorship contribution statement

Junxi Qu: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Xiaoyi Ma:** Writing – review & editing, Conceptualization. **Yang Zhou:** Writing – review & editing, Formal analysis, Conceptualization. **Xianlong Chen:** Writing – review & editing, Formal analysis. **Tianren Yang:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

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

Acknowledgement

The authors wish to acknowledge funding support from the National Natural Science Foundation of China (52408095), the Hong Kong Research Grants Council (27601324), the Guangdong Basic and Applied Basic Research Foundation (2025A1515010902), the Young Elite Scientists Sponsorship Program by China Association for Science and Technology (2022ONRC001), and Guangzhou ‘‘Lingnan Talent Project’’ Reserve Talent Training Program.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.tra.2025.104572.

Data availability

Data will be made available on request.

References

Allen, T., Arkolakis, C., 2022. The welfare effects of transportation infrastructure improvements. Rev. Econ. Stud. 89 (6), 2911–2957.
Baum-Snow, N., Brandt, L., Henderson, J.V., Turner, M.A., Zhang, Q., 2017. Roads, railroads, and decentralization of Chinese cities. Rev. Econ. Stat. 99 (3), 435–448.
Baum-Snow, N., Henderson, J.V., Turner, M.A., Zhang, Q., Brandt, L., 2020. Does investment in national highways help or hurt hinterland city growth? J. Urban Econ. 115, 103124.
Blonigen, B.A., Cristea, A.D., 2015. Air service and urban growth: Evidence from a quasi-natural policy experiment. J. Urban Econ. 86, 128–146.
Bondy, J. A., & Murty, U. S. R. (1976). *Graph theory with applications* (Vol. 290). Macmillan London.
Burger, M.J., Meijers, E.J., Hoogerbrugge, M.M., Tresserra, J.M., 2017. Borrowed size, agglomeration shadows and cultural amenities in North-West Europe. In *Second Rank Cities in Europe*. Routledge, pp. 60–79.
Cascetta, E., Cartenì, A., Henke, I., Pagliara, F., 2020. Economic growth, transport accessibility and regional equity impacts of high-speed railways in Italy: Ten years ex post evaluation and future perspectives. Transp. Res. A Policy Pract. 139, 412–428.

Chakrabarti, S., Kushari, T., Mazumder, T., 2022. Does transportation network centrality determine housing price? J. Transp. Geogr. 103, 103397.
Chen, L., Lu, Y., Nanayakkara, A., 2023. Rural road connectivity and local economic Activity: Evidence from Sri Lanka's iRoad program. Transp. Policy 144, 49–64.
Condeço-Melhorado, A., Tillema, T., de Jong, T., Koopal, R., 2014. Distributive effects of new highway infrastructure in the Netherlands: the role of network effects and spatial spillovers. J. Transp. Geogr. 34, 96–105.
Cristea, A.D., 2023. The role of aviation networks for urban development. J. Reg. Sci. 63 (4), 947–980.
Deng, T., 2013. Impacts of transport infrastructure on productivity and economic growth: Recent advances and research challenges. Transp. Rev. 33 (6), 686–699.
Ducruet, C., Cuyala, S., El Hosni, A., 2018. Maritime networks as systems of cities: The long-term interdependencies between global shipping flows and urban development (1890–2010). J. Transp. Geogr. 66, 340–355.
Friedt, F.L., Cohen, J.P., 2021. Perception vs. reality: the aviation noise complaint effect on home prices. Transp. Res. Part D: Transp. Environ. 100, 103011.
Guo, Y., Cao, L., Song, Y., Wang, Y., Li, Y., 2022. Understanding the formation of City-HSR network: A case study of Yangtze River Delta, China. Transp. Policy 116, 315–326.
Guo, Y., Li, B., Han, Y., 2020. Dynamic network coupling between high-speed rail development and urban growth in emerging economies: Evidence from China. Cities 105, 102845.
Han, D., Attipoe, S.G., Han, D., Cao, J., 2023. Does transportation infrastructure construction promote population agglomeration? Evidence from 1838 Chinese county-level administrative units. Cities 140, 104409.
Huang, Y., Hong, T., Ma, T., 2020. Urban network externalities, agglomeration economies and urban economic growth. Cities 107, 102882.
Jedwab, R., Moradi, A., 2016. The permanent effects of transportation revolutions in poor countries: evidence from Africa. Rev. Econ. Stat. 98 (2), 268–284.
Jiao, J., Wang, J., Jin, F., 2017. Impacts of high-speed rail lines on the city network in China. J. Transp. Geogr. 60, 257–266.
Jiao, J., Wang, J., Jin, F., Dunford, M., 2014. Impacts on accessibility of China's present and future HSR network. J. Transp. Geogr. 40, 123–132.
Jiao, J., Wang, J., Zhang, F., Jin, F., Liu, W., 2020. Roles of accessibility, connectivity and spatial interdependence in realizing the economic impact of high-speed rail: Evidence from China. Transp. Policy 91, 1–15.
Kasraian, D., Maat, K., Stead, D., Van Wee, B., 2016. Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies. Transp. Rev. 36 (6), 772–792.
Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. LightGBM: A highly efficient gradient boosting decision tree. Adv. Neural Inf. Process. Syst. 30.
Krishnan, V., Kastrouni, E., Pyrialakou, V.D., Gkritza, K., McCalley, J.D., 2015. An optimization model of energy and transportation systems: Assessing the high-speed rail impacts in the United States. Transp. Res. Part C Emerging Technol. 54, 131–156.
Lakshmanan, T.R., 2011. The broader economic consequences of transport infrastructure investments. J. Transp. Geogr. 19 (1), 1–12.
Lao, X., Zhang, X., Shen, T., Skitmore, M., 2016. Comparing China's city transportation and economic networks. Cities 53, 43–50.
Lee, J.K., 2022. New rail transit projects and land values: The difference in the impact of rail transit investment on different land types, values and locations. Land Use Policy 112, 105807.
Li, H., Li, J., Zhao, X., Kuang, X., 2022. The morphological structure and influence factors analysis of China's domestic civil aviation freight transport network. Transp. Policy 125, 207–217.
Li, J., Yu, Q., Ma, D., 2024. Does China's high-speed rail network promote inter-city technology transfer?——A multilevel network analysis based on the electronic information industry. Transp. Policy 145, 11–24.
Li, Y., Chen, Z., Wang, P., 2020. Impact of high-speed rail on urban economic efficiency in China. Transp. Policy 97, 220–231.
Li, Y., Xia, X., Huang, Q., 2025. Port shipping connectivity as a new driver of urban exports in the context of dual circulation: Evidence from China. Transp. Policy 163, 73–90.
Lieske, S.N., van den Nouwelant, R., Han, J.H., Pettit, C., 2021. A novel hedonic price modelling approach for estimating the impact of transportation infrastructure on property prices. Urban Stud. 58 (1), 182–202.
Lin, Y., 2017. Travel costs and urban specialization patterns: Evidence from China's high speed railway system. J. Urban Econ. 98, 98–123.
Ling, C., Niu, X., Yang, J., Zhou, J., Yang, T., 2024. Unravelling heterogeneity and dynamics of commuting efficiency: Industry-level insights into evolving efficiency gaps based on a disaggregated excess-commuting framework. J. Transp. Geogr. 115, 103820.
Liotta, C., Viguié, V., Creutzig, F., 2023. Environmental and welfare gains via urban transport policy portfolios across 120 cities. Nat. Sustainability 6 (9), 1067–1076.
Liu, H., Stockwell, N., Rodgers, M.O., Guensler, R., 2016. A comparative life-cycle energy and emissions analysis for intercity passenger transportation in the US by aviation, intercity bus, and automobile. Transp. Res. Part D: Transp. Environ. 48, 267–283.
Liu, X., Jiang, C., Wang, F., Yao, S., 2021. The impact of high-speed railway on urban housing prices in China: A network accessibility perspective. Transp. Res. A Policy Pract. 152, 84–99.
Lundberg, S.M., Lee, S.-I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 30.
Luo, H., Zhao, S., 2021. Impacts of high-speed rail on the inequality of intercity accessibility: A case study of Liaoning Province. China. Journal of Transport Geography 90, 102920.
Maroto, A., Zofío, J.L., 2016. Accessibility gains and road transport infrastructure in Spain: A productivity approach based on the Malmquist index. J. Transp. Geogr. 52, 143–152.
Meijers, E.J., Burger, M.J., 2017. Stretching the concept of 'borrowed size'. Urban Stud. 54 (1), 269–291.
Moyano, A., Martínez, H.S., Coronado, J.M., 2018. From network to services: A comparative accessibility analysis of the Spanish high-speed rail system. Transp. Policy 63, 51–60.
Ogawa, H., Fujita, M., 1980. Equilibrium land use patterns in a nonmonocentric city. J. Reg. Sci. 20 (4), 455–475.
Pratama, A.P., Yudhistira, M.H., Koomen, E., 2022. Highway expansion and urban sprawl in the Jakarta Metropolitan Area. Land Use Policy 112, 105856.
Qu, J., Yang, T., Nam, K.-M., Kim, E., Chen, Y., Liu, X., 2024. Transport network changes and varying socioeconomic effects across China's Yangtze River Delta. J. Transp. Geogr. 121, 104051.
Ratner, K.A., Goetz, A.R., 2013. The reshaping of land use and urban form in Denver through transit-oriented development. Cities 30, 31–46.
Rokicki, B., Stępniak, M., 2018. Major transport infrastructure investment and regional economic development–An accessibility-based approach. J. Transp. Geogr. 72, 36–49.
Schiavina, M., Freire, S., MacManus, K., 2022. GHS-pop r2022a—GHS population grid multitemporal (1975–2030). Joint Research Centre Data Catalogue.
Seo, K., Golub, A., Kuby, M., 2014. Combined impacts of highways and light rail transit on residential property values: A spatial hedonic price model for Phoenix, Arizona. J. Transp. Geogr. 41, 53–62.
Shanmukhappa, T., Ho, I.-W.-H., Chi, K.T., Leung, K.K., 2019. Recent development in public transport network analysis from the complex network perspective. IEEE Circuits Syst. Mag. 19 (4), 39–65.
Statistics, N. B. o. (2024). *National Data*. Retrieved from https://www.stats.gov.cn/.
Sun, X., Wandelt, S., Zhang, A., 2024. On the relationship between air connectivity and economic development: A comparative analysis of inequality evolution for 2000–2019. Transport Economics and Management 2, 310–321.
Tian, M., Li, T., Ye, X., Zhao, H., Meng, X., 2021. The impact of high-speed rail on service industry agglomeration in peripheral cities. Transp. Res. Part D: Transp. Environ. 93, 102745.
Tian, M., Wang, Y., Wang, Y., 2023. High-speed rail network and urban agglomeration economies: Research from the perspective of urban network externalities. Socioecon. Plann. Sci. 85, 101442.
Wang, C., Kim, Y.-S., Kim, C.Y., 2021. Causality between logistics infrastructure and economic development in China. Transp. Policy 100, 49–58.
Wang, C., Lim, M.K., Zhang, X., Zhao, L., Lee, P.-T.-W., 2020a. Railway and road infrastructure in the Belt and Road Initiative countries: Estimating the impact of transport infrastructure on economic growth. Transp. Res. A Policy Pract. 134, 288–307.

Wang, J., Mo, H., Wang, F., Jin, F., 2011. Exploring the network structure and nodal centrality of China's air transport network: A complex network approach. J. Transp. Geogr. 19 (4), 712–721.
Wang, L., 2018. High-speed rail services development and regional accessibility restructuring in megaregions: A case of the Yangtze River Delta, China. Transp. Policy 72, 34–44.
Wang, W., Cai, K., Du, W., Wu, X., Tong, L.C., Zhu, X., Cao, X., 2020b. Analysis of the Chinese railway system as a complex network. Chaos Solitons Fractals 130, 109408.
Wu, B., Levinson, D.M., 2024. A multi-modal analysis of the effect of transport on population and productivity in China. J. Transp. Geogr. 116, 103856.
Xu, Y., Chen, C., Deng, W., Dai, L., Yang, T., 2025. Spatial eco-socio-economic trade-offs inform differentiated management strategies in mega-urban agglomerations. npj Urban Sustainability. DOI: 10.1038/s42949-025-00231-x.
Yang, T., Zhang, Y., 2025. Urban systems science and cross-scale dynamics: A conceptual framework to advance integrated planning and governance. Transactions in Planning and Urban Res. 4 (1), 3–13.
Zhang, P., Zhao, Y., Zhu, X., Cai, Z., Xu, J., Shi, S., 2020. Spatial structure of urban agglomeration under the impact of high-speed railway construction: Based on the social network analysis. Sustain. Cities Soc. 62, 102404.
Zhao, J., Yan, J., Ran, Q., Yang, X., Su, X., Shen, J., 2022. Does the opening of high-speed railways improve urban livability? Evidence from a quasi-natural experiment in China. Socioecon. Plann. Sci. 82, 101275.
Zhou, Z., Zhang, A., 2021. High-speed rail and industrial developments: Evidence from house prices and city-level GDP in China. Transp. Res. A Policy Pract. 149, 98–113.
Zhu, Z., Zhang, A., Zhang, Y., 2018. Connectivity of intercity passenger transportation in China: A multi-modal and network approach. J. Transp. Geogr. 71, 263–276.
Zhu, Z., Zhang, A., Zhang, Y., 2019. Measuring multi-modal connections and connectivity radiations of transport infrastructure in China. Transportmetrica A: Transport Science 15 (2), 1762–1790.
Zou, W., Chen, L., 2024. The impact of high-speed railway on firms' productivity. Int. Rev. Econ. Financ. 92, 1374–1394.