International Journal of
Geo-Information

MDPI

Article

Exploring Spatial Nonstationarity in Determinants of Intercity Commuting Flows: A Case Study of Suzhou–Shanghai, China

Zhipeng Li [1] and Xinyi Niu [1,2,*]

[1] College of Architecture and Urban Planning, Tongji University, Shanghai 200092, China; 2030084@tongji.edu.cn
[2] Key Laboratory of Spatial Intelligent Planning Technology, Ministry of Natural Resources, Shanghai 200092, China
* Correspondence: niuxinyi@tongji.edu.cn

Abstract: The increasing popularity of intercity commuting is affecting regional development and people's lifestyles. A key approach to addressing the challenges brought about by intercity commuting is analyzing its determinants. Although spatial nonstationarity seems inevitable, or at least worth examining in spatial analysis and modeling, the global perspective was commonly employed to explore the determinants of intercity commuting flows in previous studies, which might result in inaccurate estimation. This paper aims to interpret intercity commuting flows from Suzhou to Shanghai in the Yangtze River Delta region. For this purpose, mobile signaling data was used to capture human movement trajectories, and multi-source big data was used to evaluate social-economic determinants. Negative binomial (*NB*) regression and spatially weighted interaction models (SWIM) were applied to select significant determinants and identify their spatial nonstationarity. The results show that the following determinants are significant: (1) commuting time, (2) scale of producer services in workplace, (3) scale of non-producer services in residence, (4) housing supply in residence, (5) year of construction in residence, and (6) housing price in residence. In addition, all six significant determinants exhibit evident spatial nonstationarity in terms of significance scope and coefficient level. Compared with the geographically weighted regression (GWR), SWIM reveals that the determinants of intercity commuting flows may manifest spatial nonstationarity in both residence and workplace areas, which might deepen our understanding of the spatial nonstationarity of OD flows.

Keywords: intercity commuting; determinants; spatial nonstationarity; spatially weighted interaction models; OD flows; mobile signaling data

Citation: Li, Z.; Niu, X. Exploring Spatial Nonstationarity in Determinants of Intercity Commuting Flows: A Case Study of Suzhou–Shanghai, China. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 335. https://doi.org/10.3390/ijgi11060335

Academic Editors: Wolfgang Kainz and Hartwig H. Hochmair

Received: 15 April 2022
Accepted: 2 June 2022
Published: 4 June 2022

1. Introduction

In recent years, high-speed transport infrastructures have reformed the concept of accessibility and mobility. The spread and development of transport technologies, such as high-speed rail (HSR) and freeways, have dramatically transformed the regional functional linkage [1–3]. Additionally, especially with the advent of HSR, which has an excellent performance in reducing travel time, it also significantly stimulates the demand for long-distance trips, provides individuals with more flexible job opportunities, and weakens the locational ties between workplaces and residences [4,5]. High-speed transport infrastructures make commuting flows beyond administrative boundaries and reshapes the megaregional structure [6–8]. These important changes contribute to a trend that the commuting zone of a city has extended beyond the traditional metropolitan area to other cities or regions, and an increasing number of people are commuting across cities. Nowadays, intercity commuting has become more and more common worldwide. In 2020, more than 6% of the labor force in the EU commuted across different NUTS 2 (Nomenclature of Territorial Units for Statistics) subdivisions. The percentage even reached 21% in Belgium, the highest in the EU, followed by the Netherlands, Germany, Austria, and Denmark (all over 10%) [9]. There have been extensive attempts to study the determinants of intercity commuting [10,11] and its impact on regional development [12].

Intercity commuting has exerted a profound influence in many fields. For example, medium and small-sized cities, which used to be geographically isolated from metropolitan centers or other distant metropolises, nowadays can be considered as special subcenters of metropolitan areas due to high-frequency shuttles [13], which results in dramatic local residential development [14]. Increasing labor mobility also contributes to metropolitan integration [15], promotes the growth of regional economic development [16], and reduces regional wage disparities [17]. However, it also has negative effects on energy, ecology, and intercity commuters' physical and mental health [18–20].

China is undergoing a similar process [21]. In the Yangtze River Delta region, one of the most developed regions in China in terms of transport infrastructure and economic development, intercity commuting has become a reality. In Shanghai, the advent of high-speed transport infrastructures such as the 350 km/h Shanghai–Nanjing Intercity Railway, Beijing–Shanghai High-speed Railway, Shanghai's metro line 11 extending to Suzhou, and the highly-developed freeway network have made intercity commuting become a common phenomenon in Shanghai. The number of intercity commuters is rising, and the residential housing development has boomed along some HSR routes or transport corridors. From the perspective of spatial structure, a functional continuous area has gradually taken shape and exceeded the administrative boundary between Suzhou and Shanghai [22]. According to previous studies, long-distance commuting is empirically defined as commuting that takes more than 40 min [23] or covers a distance of over 30 km [24]. In a questionnaire survey conducted in 2018, the interviewed commuters travel from Suzhou to Shanghai by HSR with an average commuting time of 89.56 min and an average distance of over 80 km and, therefore, are typical long-distance commuters [25].

Previous studies tend to posit that uneven regional distributions of social-economic factors, such as economic development, population size, job supply, housing price, etc., are the main significant causes of intercity commuting flows [10,11,26]. However, the aforementioned studies were conducted only from a global perspective. As Fotheringham, et al. [27] noted, it is more appropriate to assume that the relationships might vary over space and coefficient estimates might exhibit significant spatial variations. The global regression model that ignores spatial nonstationarity can only produce "average" or "global" parameter estimates. It is worth noting that compared with the global regression model, the local regression model, able to show spatial variations of parameter estimates, has a better effect on fitting the model and, therefore, provides more accurate information for policy-making.

Moreover, previous studies usually used official statistics as the primary data sources to discuss determinants of intercity commuting flows. However, official statistics are hardly effective in revealing the spatial characteristics of commuting at the inner-city scale due to their low spatial resolution. Moreover, there are no official statistics yet about intercity commuting in China. Mobile signaling data representing the space–time trajectory of commuters can reflect the characteristics of people's activities in a more timely, accurate, and comprehensive manner, which is more suitable for studying the spatial nonstationarity of determinants of intercity commuting flows in China.

Therefore, this study focuses on intercity commuting flows from Suzhou to Shanghai using mobile singling data. The significant determinants of intercity commuting are identified through the global regression model, and the spatial nonstationarity of such determinants is explored through the local regression model.

2. Literature Review

2.1. The Conceptual Model and Determinants of Intercity Commuting Flows

With the development of expanding urbanization, commuting is becoming an increasingly important part of our life, and it has been extensively studied from the perspectives of multiple disciplines. Early studies of its determinants focused on residence or workplace location choices separately, but later on, it became widely accepted that individuals consider these choices dependently and multi-dimensional determinants should be evaluated in a joint model [28]. Spatial interaction is a classical concept in geography that refers to the dynamic flows of elements (e.g., migration, transport, international trade, tourism, etc.) from origins to destinations. According to the spatial interaction model, dynamic flows of elements are co-determined by origin factors, destination factors, and the separation between origins and destinations [29]. Intercity commuting flows can be regarded as a typical spatial interaction because many studies have found that intercity commuting flows are under the combined influence of commuting cost, determinants of residence, and those of workplace [10,30,31].

First of all, the most notable determinant of intercity commuting has significantly reduced commuting costs due to improved transport infrastructure. With the geographical distance staying unchanged, different commuting modes will result in varying commuting costs. For example, compared with freeways, HSR helps to greatly extend the commuting distance while keeping the commuting cost constant, integrating previously remote and isolated areas into the metropolitan region [15]. Therefore, commuting time is a more accurate indicator of intercity commuting cost than geographical distance.

Moreover, the regional difference is another important cause. The previous study found, through economic models, that a significant difference between determinants of the workplace and the place of residence is an essential prerequisite for intercity commuting [32]. Firstly, workplaces with more jobs are more attractive to commuters; thus, intercity commuting may be related to the amount of job supply. Some believe that with the rapid development of communications technologies and transport infrastructure, people can change jobs more flexibly with longer commuting distances for the benefit of their career development. As a result, long-distance commuting has become a substitute for migration [33]. Higher salaries can compensate for higher commuting costs; hence, regional wage disparities can also be a driver. The regional wage level is correlated with industry structure, and the advanced service sector is more competitive [34,35]. It has been proven that industry type is a good indicator of the regional wage level [36].

Housing prices are one of the most notable determinants related to the place of residence [10,11]. Additionally, places with a larger housing supply are more attractive to commuters; therefore, the number of intercity commuters may also be related to housing supply. Previous studies show that people married with children are prone to choose long-distance commuting: it may not be a rational choice in the individual interests of commuters, but it can guarantee better educational resources for their children [37,38]. The year of construction is also a major spatial determinant explored in studies of long-distance commuting. The newer a residential community is, the better its conditions are. Intercity commuters may choose to purchase residential properties in another city for better living conditions. More importantly, it is widely believed that the residents living in older residential communities tend not to choose long-distance commuting because they have a well-established local social network and a stable life, and long-distance commuting can take a toll on family and social networks [39,40].

Therefore, commuting time, job supply, industry type, housing supply, housing price, educational resources, and year of construction are the potential determinants of intercity commuting flows.

2.2. Local Spatial Model Applied in Commuting

Commuting has attracted considerable interest from various disciplines such as economics, geography, and urban planning, which leads to multiple approaches to explain the causes of commuting. Some studies analyze the influence of individual characteristics on commuting [40,41], while others adopt a spatial perspective to analyze the determinants of commuting by economic models or regressions. The advantage of economic models is that they can provide systematic theoretical explanations [42,43] and assess the effects of planning policies combined with empirical data [44]. However, the economic models usually assume that the economic entities such as individuals and firms are homogeneous, making it difficult to reveal the local and heterogeneous spatial patterns. Regression is another approach to analyzing the problem from a spatial perspective, which can be divided into global regression and local regression depending on whether or not spatial variations are considered. The global regression model assumes that samples are independent of each other and the process remains stationary over space, which means that the parameter estimates do not change along with geographical locations. In many real-life situations, spatial data contains locational information and attribute information, therefore, the assumptions of global stationarity may result in inaccurate estimates [45]. To deal with this problem, the local regression considering spatial nonstationarity is a more appropriate way. Spatial nonstationarity, a form of spatial heterogeneity, refers to the common phenomenon that the relationship between variables may vary across the study region. For example, residents in places with underdeveloped transport infrastructure may be more sensitive to distance when making commuting decisions than those in places with good transport infrastructure. Similarly, residents in the downtown area may be less willing to choose long-distance commuting than those in suburban areas.

There are several models that account for spatial nonstationarity, such as moving window regression and spatially adaptive filtering. Geographically weighted regression (GWR) is the most commonly used. The standard GWR conducts local regression based on linear regression while considering the distance decay effect [46]. According to Tobler's First Law of Geography, "everything is related to everything else, but near things are more related to each other". The calibration of GWR is a little complicated but essentially based on the assumption that the data points more closely located to the regression point have greater effects on the model estimation of this regression point. At every regression point, all the weights of samples are recalculated and, hence, a set of local parameter estimates is obtained. As a result, a local parameter surface is constructed to visually show the spatial variation of the relationship described by the coefficient. The main steps of GWR include selecting a spatial weighting function, determining the optimal bandwidth by minimizing the corrected Akaike Information Criterion (AICc) of the model, and then estimating local coefficients [47]. By GWR, spatial nonstationarity has been widely observed in commuting-related studies, such as commuting modes [48], commuting time [49], extreme commuting [50], and sustainable commuting [51].

The aforementioned studies are all based on point or polygon features. However, flow space has become a new perspective in analyzing geographical space. Commuting, in nature, is a functional flow between the workplace and residence; therefore, the flow feature is undoubtedly the best feature type to represent the commuting functional linkages. Currently, the related studies only use the global regression model to explore determinants of intercity commuting flows. One of the reasons for this dilemma is that there were no other appropriate research tools available. There was an early attempt, which needed further improvement, in which Nakaya [52] tried to combine the GWR and spatial interaction model into the origin-specific model to illustrate the spatial nonstationarity of flows. Fortunately, Kordi and Fotheringham [53] subsequently came up with the spatially weighted interaction models (SWIM), an extension of GWR that can better illustrate the spatial nonstationarity of flows. SWIM has been used in some empirical research, such as traffic flows [54] and online travel searches [55]. Compared with the global spatial interaction model, SWIM, based on the more realistic assumption, can be more explanatory of the formation of flows.

3. Study Area and Data Sources

3.1. Study Area

In this study, the workplace is the downtown area of Shanghai and the place of residence is the Suzhou administrative area (Figure 1). As the core city of the Yangtze River Delta region, Shanghai is a regional hub connecting flows of business, capital, population, and information [56]. With higher productivity, Shanghai is a more attractive workplace for commuters from surrounding cities. Its downtown area boasts the highest density of businesses and high-end jobs citywide. Intercity commuters who choose to work in downtown Shanghai, instead of in suburban Shanghai, are more likely for financial benefits and career development. Suzhou is the major city of residence for intercity commuters to Shanghai. Being adjacent to Shanghai, it is an important city in the Yangtze River Delta region. According to the official statistics for 2020, Shanghai's GDP reached RMB 3870 billion yuan and its population 24.9 million, while Suzhou's GDP was RMB 2017 billion yuan and its population 12.75 million. Although the GDP and per capita GDP between Suzhou and Shanghai does not appear very different, Shanghai is the unique center of advanced industrial agglomeration in the Yangtze River Delta region [57]. According to the 2020 Official Statistical Yearbook, the above-scale business revenue of Shanghai and Suzhou by industry type demonstrates a significant gap in industrial development between Shanghai and Suzhou (Figure 2). The development of producer services in Shanghai is much better than in Suzhou. In this paper, spatial units are defined based on the administrative boundaries of the township units in the two cities. The downtown area of Suzhou or Shanghai is the set of township-level units that intersect with the downtown area defined by the official master planning.

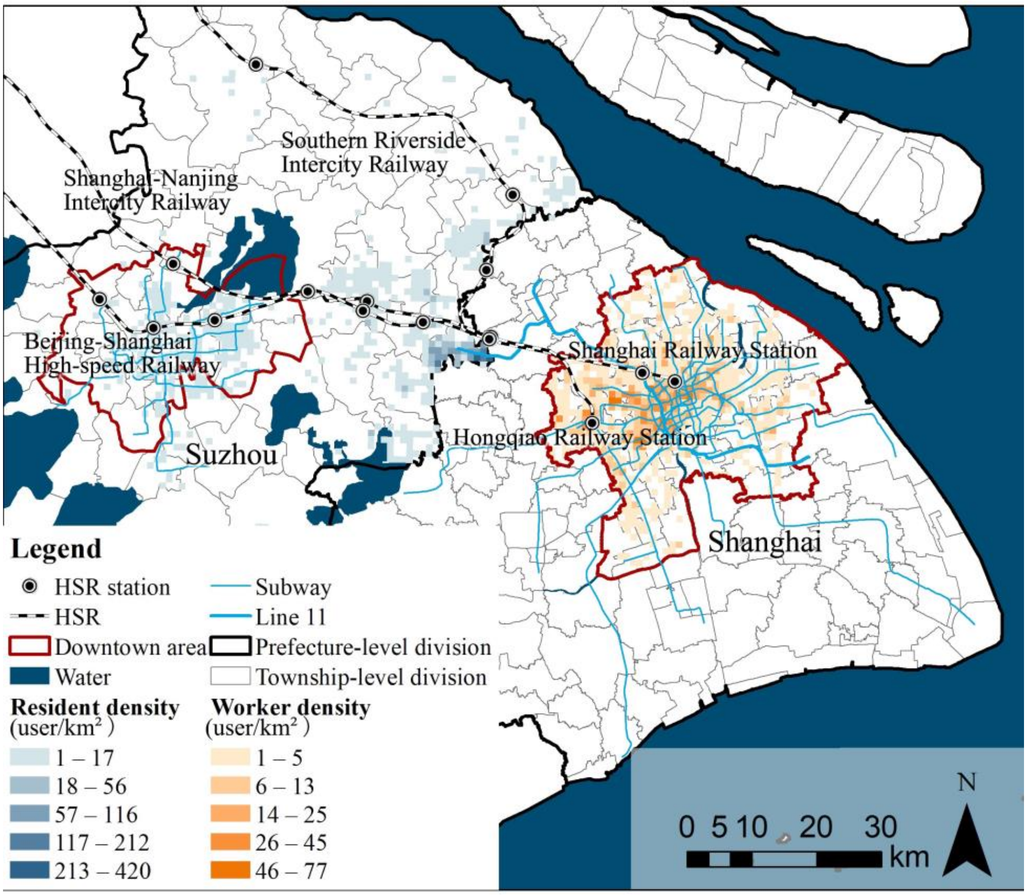

Figure 1. Study area and commuter density in residence and workplace.

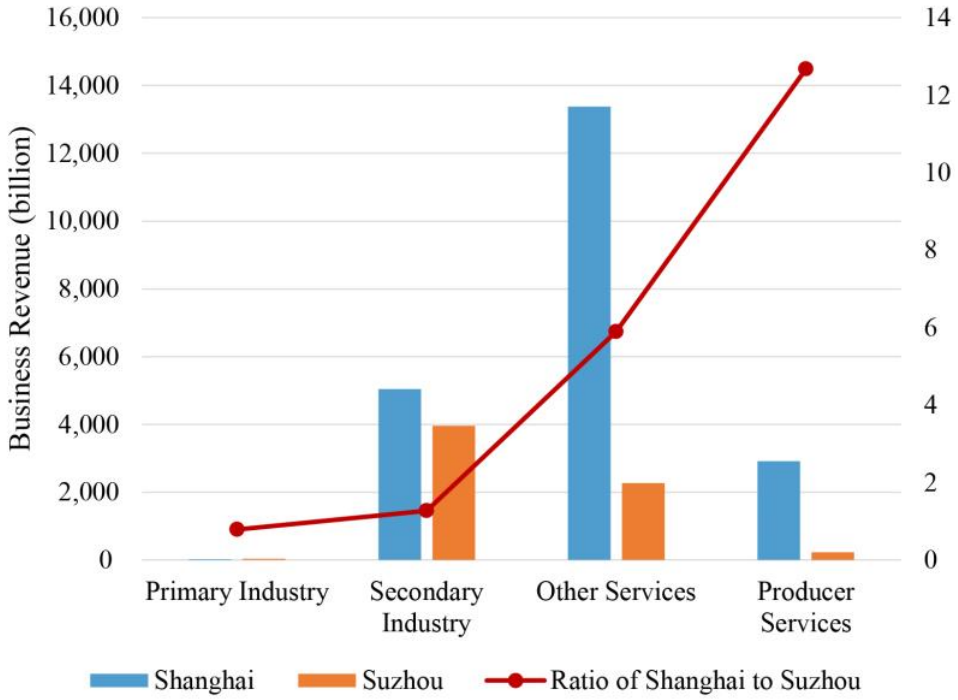

Figure 2. Above-scale business revenue of Shanghai and Suzhou by industry type.

The distance between the downtown areas of Suzhou and Shanghai is over 100 km, which would take more than one hour and a half by train before the advent of HSR. Since the Shanghai–Nanjing Intercity Railway and Beijing–Shanghai High-speed Railway were completed in 2010, the traveling time has been reduced to 30 min and railway has become the primary commuting mode for intercity commuters between the two cities. In 2020, the Southern Riverside Intercity Railway went into operation. Although it plays only a limited role in intercity commuting currently, it may manifest more potential in the future. Commuters from Suzhou can get aboard at multiple railway stations in Suzhou and get off at Hongqiao Railway Station or Shanghai Railway Station in downtown Shanghai. Moreover, Shanghai's metro line 11 links Suzhou and Shanghai, meaning residents in the areas neighboring Shanghai can easily reach downtown Shanghai by subway. Intercity commuting from Suzhou to downtown Shanghai boasts ideal transit conditions.

3.2. Data Sources

Previous studies mostly use official statistics at the city level to represent intercity commuting flows. However, currently, there are no official statistics as such in China. Fortunately, mobile singling data have become a widely used data source for studying commuting patterns and spatial structure [58–60]. This paper uses mobile signaling data from China Unicom as the data source. As one of three major telecommunications operators in China, China Unicom renders service to 19% of mobile phone users in the Chinese market. The data from the entire Yangtze River Delta region from 1–30 June 2021 have were collected. Mobile singling data was passively collected each time mobile phone activity occurred (e.g., voice call, SMS, network traffic, automatic contact with a base station). Every record collected consisted of five essential attributes: anonymous user ID, record time, base station, longitude, and latitude (see Table 1 for the data structure).

Table 1. Mobile signal data structure.

Anonymous User ID	Record Time	Base Station	Longitude	Latitude
8415dfadgf6454213d5fdf548	2021/6/5 21:31:35	BS9145370	121.252434	31.458201
8415dfadgf6454213d5fdf548	2021/6/5 22:04:27	BS4800692	121.190352	31.403564
8415dfadgf6454213d5fdf548	2021/6/5 01:29:42	BS2656484	121.174301	31.364021
...	...	...	...	...
1489fe546d45212d5648de84	2021/6/5 10:12:01	BS2445601	120.676156	31.308156
1489fe546d45212d5648de84	2021/6/5 11:25:34	BS8006925	120.807564	31.235461

Other multi-source big data were used to profile the social and economic attributes of spatial units. Corporate data were collected from the National Enterprise Credit Information Publicity System of China, including the number of employees, industry type, and other attributes of all the enterprises in our study area. Residential housing data were collected from "Beike", a well-known real estate trading platform in China, including the number of apartments, housing price, year of building construction, and other attributes of all the residential quarters in our study area. Data about primary and secondary schools, including their total number and locations, were collected from AutoNavi POI.

4. Research Methods

4.1. Constructing Commuting OD Flows with Mobile Signaling Data

To measure the spatial–temporal trajectory of commuters and construct commuting OD flows with mobile signaling data, the most commonly used approach is to identify stop points along a spatial–temporal trajectory and analyze the duration of each stop point throughout the daytime and nighttime. The commuting OD flow construction approach adopted in this study was improved based on Li and Niu [22]. First, the time–space trajectory of every user in a given day was constructed according to time geography, and it can be represented as:

$$T = [r_1, r_2, r_3, \cdots, r_i, \cdots, r_n], r_i = (id_i, t_i, long_i, lat_i) \quad (1)$$

where r_i stands for record point i; $id_i, t_i, long_i, lat_i$ stand for the base station, record time, longitude, and latitude of record point i, respectively; n the total number of record points of the user in that given day. All record points are in chronological order. Figure 3 shows the typical daily spatial–temporal trajectory of a commuter. If two or more contiguous record points share the same base station, they are counted as one stop point. For example, r_3, r_4, r_5 in Figure 3 can be counted as one stop point. In Figure 3, the trajectory of a stop point is indicated as a blue line and the trajectory between two adjunct stop points as an orange line. For each mobile phone user, the overall duration of each stop point from 20:00 at night to 6:00 the next morning is calculated, and the stop point where the user stays for the longest time (and at least more than 1 h) is considered as the user's place of residence on that given day. If no such stop point is found, the user will be identified as one with no place of residence on that day. Most people follow rather steady daily commuting patterns but may occasionally deviate from their usual routes. Therefore, the most frequently identified place of residence (identified as such for more than 50% of days in a month) is regarded as the user's actual permanent place of residence.

Similarly, the overall duration of each stop point between 9:00 and 16:00 is calculated, and the stop point where the user stays for the longest time (and at least more than 1 h) is considered as the user's workplace on that given day. If no such stop point is found, the user will be identified as one with no workplace on that day. The most frequently identified workplace (identified as such for more than 50% of workdays in a month) is regarded as the user's actual permanent workplace.

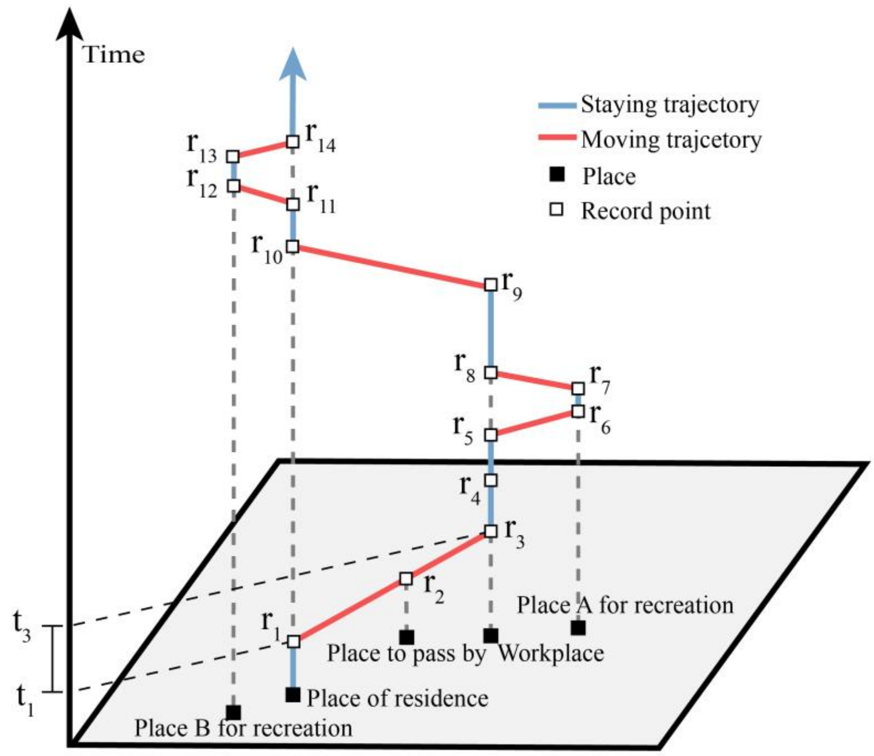

Figure 3. Typical daily spatial–temporal trajectory of a commuter.

In accordance with the above methods, all mobile phone users are classified into four categories: (1) users with the identified place of residence and workplace; (2) users with only identified place of residence; (3) users with only identified workplace; (4) users without either identified place of residence or workplace. The first group of users are believed to be the effective samples of commuters.

In previous studies, estimates of commuting time are usually based on Internet map API [61] or HSR running time [10]. However, intercity commuters from Suzhou to Shanghai have a variety of commuting modes. They can reach the railway station from their residence or workplace on foot, by bike, bus, subway, or car [25]; therefore, the above methods of estimating commuting time may result in much inaccuracy. In this study, mobile signaling data are used to calculate their commuting time, and the principle is that the commuting time of a user is the duration between the earliest record point identified at the permanent workplace and the latest one before that record point at the permanent place of residence ($t_3 - t_1$ in Figure 3).

It is worth noting that the data privacy restrictions prevent us from obtaining original high-resolution datasets of personal trajectories at the longitude and latitude level. It is compulsory to aggregate the latitude and longitude coordinates of workplace and residence into the 1-km grid to obscure the phone users' individual information. Therefore, this study is based on the OD flow data at the 1-km grid level.

The intercity commuters are selected from all effective samples of commuters. Based on estimates made from the mobile singling data in this study, Suzhou–Shanghai intercity commuters account for over 90% of all intercity commuters to Shanghai. In the dataset, 4275 intercity commuters from Suzhou to Shanghai can be identified. The Suzhou–Shanghai intercity commuters identified through mobile signaling data mostly live in the border areas of the two cities, in downtown Suzhou, or along the HSR route and work in the mid-western area of downtown Shanghai (Figure 1).

The modifiable areal unit problem (MAUP) is an inevitable issue in spatial analysis. Township units are used for two reasons. First, the township unit is the lowest unit level in China's socio-economic statistics, and it is also the lowest administrative unit level of local government. The benefit of using the township unit is that the model estimates can be compared with official statistics, causing the conclusions to have more policy implications. Second, larger unit divisions can make the difference within the dependent variable more noticeable. Of course, overly large spatial units tend to make spatial nonstationarity unobserved. Given the above three reasons, it is believed that the township unit is the most suitable choice. Therefore, the intercity commuters from Suzhou to Shanghai are aggregated to construct the intercity commuting flows at the township level. Figure 4 shows flows larger than 2 for clarity, accounting for over 83% of intercity commuters. Table 2 shows the descriptive statistics of intercity commuting flows at the township level.

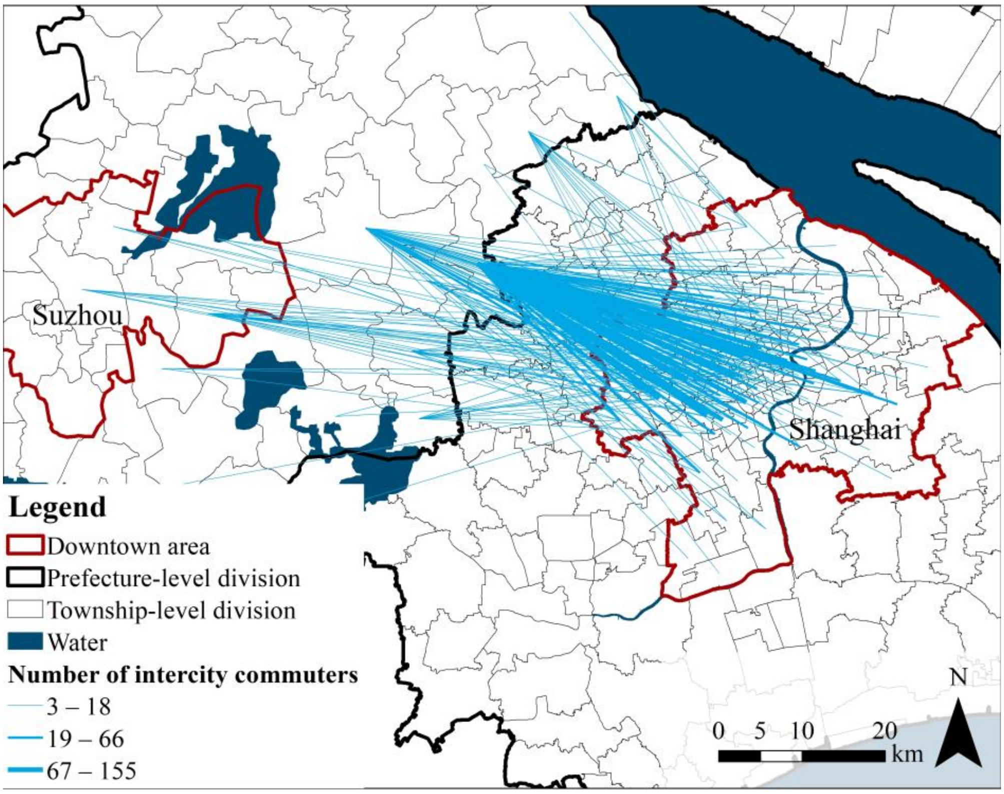

Figure 4. Intercity commuting flows at township level.

Table 2. Descriptive statistics of the intercity commuting flows at township level.

Parameter	Value
Count	835
Sum	4275
Average	5.12
Variance	182.09
Max	155
Min	1
Quantile (25%)	1
Quantile (50%)	1
Quantile (75%)	3
Number of intercity commuters from downtown Suzhou	373
Number of intercity commuters from other areas in Suzhou	3902

4.2. Exploring the Determinants of Intercity Commuting Flows

Building a set of potential determinants is the first step to selecting significant determinants. The descriptive index of job supply is the sum of employees of all enterprises in a spatial unit. Different industry types vary in wage level and their appeal to workers.

As producer services are more attractive, jobs in producer and non-producer services are analyzed separately. The descriptive index of housing supply is the sum of all houses in a spatial unit. The descriptive index of housing price is the average housing price of all house estates in a spatial unit. The descriptive index of the year of construction is the average year of construction for all residential quarters in a spatial unit minus the baseline year (1950). In Suzhou and Shanghai, the major urbanization campaign began in the 1980s and their oldest residential areas were built in the 1950s. The baseline year is substituted into the formula to better reveal the difference in the average year of construction between spatial units. The higher the number is, the newer the residential community is, and the more likely it is for its residents to choose intercity commuting. The descriptive index of educational resources is the ratio of the total number of primary and secondary schools in a spatial unit to the total area of this unit. To fully illustrate the combined influence of workplace and residence on intercity commuting flows, the above determinants both in workplace and residence are tested simultaneously in the global regression model. For descriptive statistics of all independent variables, see Table 3.

Table 3. Descriptive statistics of the independent variables of the intercity commuting flows from Suzhou to Shanghai.

	Variable	Average	Variance	Max	Min
	Commuting time (minutes)	105.562	457.877	149.78	25.02
Workplace	Scale of producer services (ten thousand jobs)	2.992	8.334	15.024	0.023
	Scale of non-producer services (ten thousand jobs)	2.806	3.99	9.617	0.057
	Housing supply (ten thousand houses)	5.048	9.442	14.66	0.316
	Housing price (ten thousand yuan)	7.794	4.061	13.149	3.213
	Year of construction (year)	43.78	114.478	59.815	4
	Density of primary and secondary schools (number of schools per km^2)	2.192	3.844	9.205	0
Residence	Scale of producer services (ten thousand jobs)	5.451	48.741	31.894	0.063
	Scale of non-producer services (ten thousand jobs)	14.386	133.523	40.951	0.591
	Housing supply (ten thousand houses)	15.732	201.696	47.512	0.088
	Housing price (ten thousand yuan)	1.937	0.568	3.699	0.568
	Year of construction (year)	59.838	11.686	69	49.876
	Density of primary and secondary schools (number of schools per km^2)	0.287	0.099	1.398	0.029

The global regression model is the most common method for significance tests of determinants. The most frequently used global regression model is the ordinary least squares (OLS) regression on the assumption that dependent variables are normally distributed. However, Suzhou–Shanghai intercity commuting flows are a non-negative discrete counting variable with features of overdispersion (which means that the variance is much higher than average), consistent with the assumption of negative binomial (*NB*) distribution. Therefore, *NB* regression is the appropriate tool for studying the determinants of Suzhou–Shanghai intercity commuting flows. The formula is as follows:

$$M_{ij} \sim NB[t * \exp\left(\sum_{k=1}^{n} \beta_k \log\left(x_{ijk}\right)\right), \alpha] \tag{2}$$

wherein M_{ij} is the number of intercity commuters from spatial unit i to j; n is the number of determinants; t is the offset parameter; x_{ijk} is the value of determinant k from i to j; β_k is the coefficient of determinant k; α is the dispersion parameter. Adjusted pseudo R^2 and AICc are used to compare the fitting performance of the global model and the local one. Adjusted pseudo R^2 ranges from 0 to 1. Unlike adjusted R^2, it evaluates the relative importance of the model in comparison to the constant-only model from the perspective of log likelihood. The absolute value of AICc is meaningless, while the difference between the AICc values of the two models is a more telling number. As a rule of thumb, if the difference between the

AICc values of the two models is higher than 3, it is reasonable to believe that one is better than the other.

4.3. Exploring the Spatial Nonstationarity of Determinants

The significant determinants selected through the global regression model may be spatially nonstationary. As the global regression model cannot reveal this spatial feature, the local regression model is used to study it. This study adopts SWIM as the local regression model. Although Kordi and Fotheringham [53] only introduced Gaussian model and Poisson model, the subsequent study has helped improve the *NB* model and demonstrated that it is better than the Gaussian model when the dependent variable follows the *NB* distribution [54]. The formula is as follows:

$$M_{ij} \sim NB[t_{ij} * \exp\left(\sum_{k=1}^{n} \beta_{ijk} \log\left(x_{ijk}\right)\right), \alpha] \tag{3}$$

wherein M_{ij} is the number of intercity commuters from spatial unit i to j; n is the number of determinants; t_{ij} is the offset parameter; x_{ijk} is the value determinant k from i to j; β_{ijk} is the coefficient of determinant k from i to j; $\propto$ is the dispersion parameter.

In SWIM, the coordinate of flow ij is (x_i, y_i, x_j, y_j); that of flow $i'j'$ is $\left(x_{i'}, y_{i'}, x_{j'}, y_{\ j'}\right)$; the distance between ij and $i'j'$ is calculated by:

$$d_{(ij)(i'j')} = \sqrt{(x_i - x_{i'})^2 + (y_i - y_{i'})^2 + \left(x_j - x_{j'}\right)^2 + \left(y_j - y_{j'}\right)^2} \tag{4}$$

In GWR, there are two popular methods to calculate the spatial weights matrix: fixed bandwidth and adaptive bandwidth. Fixed bandwidth aims to look for an optimal bandwidth and calculate the spatial weight between flow ij and $i'j'$ by following a Gaussian function. The formula is as follows:

$$w_{(ij)(i'j')} = exp\left[-\frac{1}{2}\left(\frac{d_{(ij)(i'j')}}{b}\right)^2\right] \tag{5}$$

wherein $w_{(ij)(i'j')}$ is the spatial weight between flow ij and $i'j'$; b is the bandwidth; $d_{(ij)(i'j')}$ is the distance between ij and $i'j'$.

Iteratively reweighted least squares is used for the fitting of the coefficients. The bandwidth is chosen through a golden selection process until the AICc value of the model reaches its minimum. For more detailed information on bandwidth identification and coefficient fitting, see Da Silva and Rodrigues [62].

5. Results

5.1. Global Regression Model

Table 4 shows the estimation results from *NB* regression, revealing the "average" performance of determinants of the entire study area. The adjusted pseudo R^2 of the model is 0.143 and the AICc is 3898.5. Commuting time, scale of producer services in workplace, scale of non-producer services in residence, housing supply in residence, year of construction in residence and housing price in residence are significantly correlated with the scale of intercity commuting flows, while other determinants fail to pass the significance test. When it comes to the plus and minus of the six significant determinants, the model estimates are in line with the previous studies [10,39,40].

Table 4. Estimated coefficients from *NB* regression.

	Determinant	Coefficient	Standard Deviation	Significance
	(Intercept)	−18.037	3.8598	0.000
	Commuting time	−1.840 ***	0.2345	0.000
Workplace	Scale of producer services	0.273 ***	0.0751	0.000
	Scale of non−producer services	0.049	0.0834	0.557
	Housing supply	0.065	0.0715	0.365
	Year of construction	0.077	0.1854	0.677
	Housing price	0.397	0.2246	0.077
	Density of primary and secondary schools	−0.142	0.1187	0.230
Residence	Scale of producer services	−0.133	0.0996	0.181
	Scale of non−producer services	0.713 ***	0.1333	0.000
	Housing supply	0.497 ***	0.1262	0.000
	Year of construction	3.906 ***	0.5226	0.000
	Housing price	−0.538 **	0.1811	0.003
	Density of primary and secondary schools	0.273	0.6045	0.652

Note: *** means a significance level of 0.001, and ** that of 0.01.

To be more specific, the coefficient of the scale of producer services in workplace is 0.273 and that of the non-producer services in residence is 0.713, which means that the presence of higher-level sectors in workplace and the presence of lower-level sectors in residence are positively correlated with the scale of intercity commuting flows. Scale of non-producer services in workplace does not pass the significance test, which means that intercity commuters prefer jobs in producer services and are not willing to choose jobs in non-producer services. The coefficient of housing supply in residence is 0.497, which means that a place with a larger housing supply is more likely to be chosen by intercity commuters as the place of residence. The year of construction in residence has a coefficient of 3.906, which means that the presence of newly built residential communities is positively correlated with intercity commuting as the social network is less developed in such places. The coefficient of housing price in residence is −0.538, which means that higher housing price is negatively correlated with intercity commuting flows. The coefficient of commuting time is −1.840, which means that higher commuting time is negatively correlated with intercity commuting flows.

5.2. Local Regression Model

The local regression model can demonstrate the spatial nonstationarity of determinants. The six determinants that pass the significance test in the global regression model are substituted into SWIM. Through a golden selection process, it is revealed that AICc can reach its minimum when 12.06 km is chosen as the bandwidth. In SWIM, the adjusted pseudo R^2 is 0.871, and the AICc is 3559.5. By comparing them with their counterparts in *NB* regression, it can be seen that SWIM shows a much better fitting performance and can better explain the determinants of intercity commuting flows.

See Table 5 for the descriptive statistics of the SWIM coefficient estimation results. The averages from the local regression model and global regression model are similar, both showing that the commuting time and housing price in residence are negatively correlated with intercity commuting flows and that the other four determinants are positively correlated. However, the six significant determinants vary greatly in the number of significant flows. A large number of flows pass the significance test for the year of construction in residence and the scale of producer services in workplace, which means that most commuting flows are under the influence of the two determinants. The scale of non-producer services in residence has the least wide influence as the number of significant flows is only 107. Although housing price has attracted wide discussion in previous studies of intercity commuting, it is significant only in 17.13% of all intercity commuting flows, which means that it does not have much influence on Suzhou–Shanghai intercity commuting flows.

Table 5. Descriptive statistics of the SWIM coefficient estimation results.

Determinant	Average	Variance	Max Value	Min Value	The Number of Significant Flows	Significant Ratio
Intercept	−35.122	30.751	−21.018	−44.449	466	55.81%
Commuting time	−1.584	0.035	−1.089	−2.059	342	40.96%
Scale of producer services in workplace	0.459	0.005	0.566	0.282	554	66.35%
Scale of non-producer services in residence	0.974	0.001	1.044	0.878	107	12.81%
Housing supply in residence	1.094	0.066	1.581	0.816	223	26.71%
Year of construction in residence	8.734	8.001	10.894	1.458	618	74.01%
Housing price in residence	−2.248	0.643	−1.033	−2.997	143	17.13%

5.3. Spatial Nonstationarity

In regression models, there are usually two fundamental descriptive indexes for the parameter estimates: the *p*-value that tests the statistical significance of the relationship between the dependent and the independent variable, and the coefficient value that refers to the response of the dependent variable to a one-unit change in one independent variable while all other independent variables keep constant. Spatial nonstationarity can also be interpreted from those two perspectives. The spatial variation of whether the OD flows pass the significance test can be considered as the spatial nonstationarity in terms of significance scope; the spatial variation of coefficient values that pass the significance test can be considered as the spatial nonstationarity in terms of coefficient level. As the spatial location of a flow is jointly determined by its origin and destination, the standard to determine whether spatial nonstationarity exists in terms of significance scope and coefficient level is whether we can observe spatial variation at the origin or destination.

5.3.1. Spatial Nonstationarity of Industry Determinants

Industry determinants include the scale of producer services in workplace (Figure 5a) and scale of non-producer services in residence (Figure 5b). For most OD flows, the scale of producer services in workplace passed the significance test, and its influence is spatially extensive. We can see that the residence units of those OD flows are mostly near the boundary between Suzhou and Shanghai. The greater the distance between a residence unit and the boundary, the fewer OD flows pass the significance test in this unit. However, along the HSR line, even in the residence units far from the boundary, there is a considerable amount of OD flows whose scale of producer services in workplace passed the significance test. This means that HSR can make advanced jobs available and attractive even if the places of residence are very far from the workplace. In most spatial units in downtown Shanghai, there are OD flows whose scale of producer services in workplace passed the significance test, which means in terms of significance scope, the spatial nonstationarity of this determinant can hardly be observed in the workplace area. Although the scale of non-producer services in residence passed the significance test in the global regression model, its influence was rather limited: only a few OD flows starting from the residence units near the boundary to the west and central areas of downtown Shanghai passed the significance test.

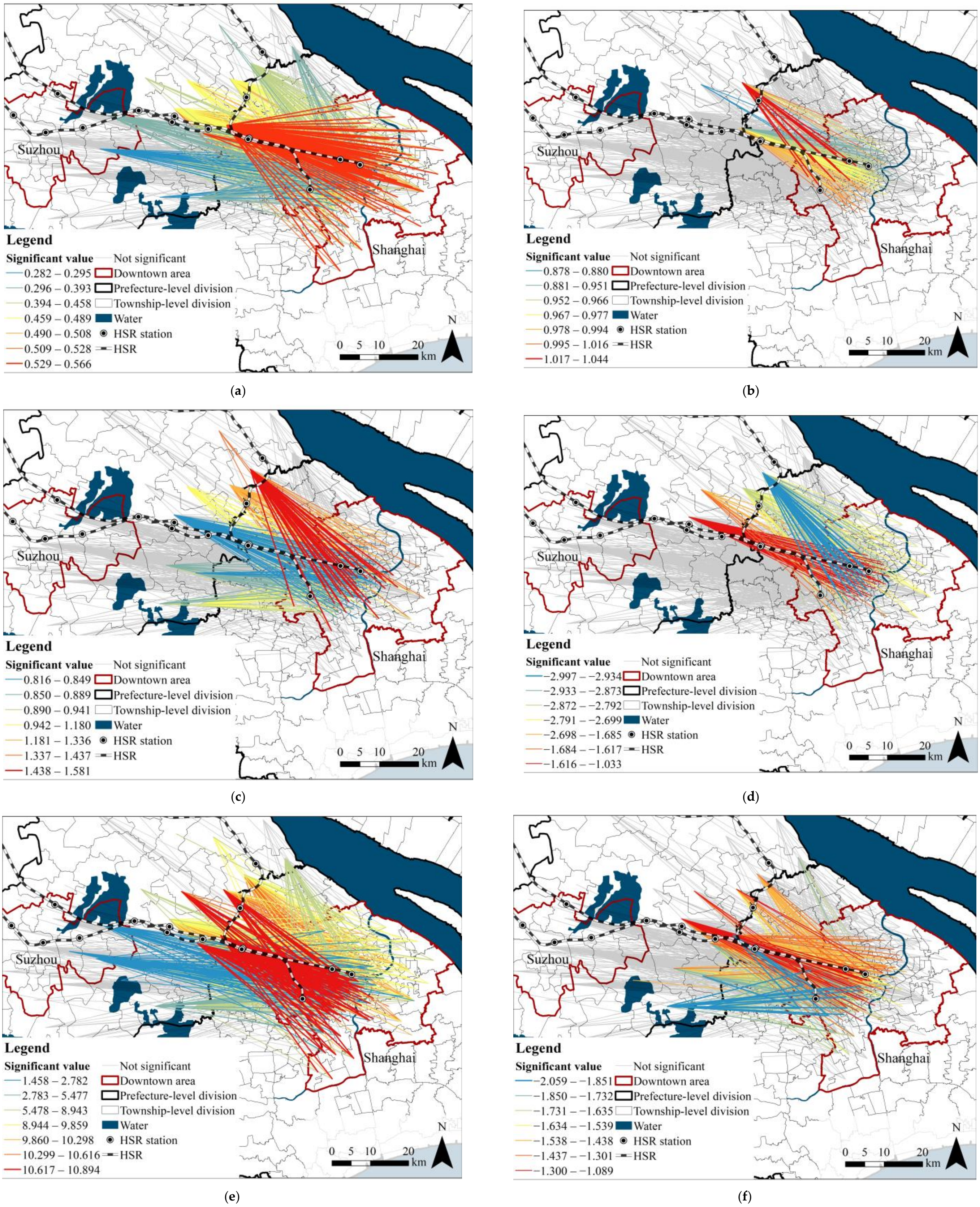

Figure 5. Spatial variation in coefficient estimates. (**a**) Scale of producer services in workplace; (**b**) Scale of non-producer services in residence; (**c**) Housing supply in residence; (**d**) Housing price in residence; (**e**) Year of construction in residence; (**f**) Commuting time.

OD flows with the residence units farther from the boundary of the two cities tend to have lower coefficients for the scale of producer services in workplace, and the OD flows starting from the same residence unit tend to have highly similar coefficients. OD flows whose scale of producer services in workplace has higher significant influence with coefficients from 0.529~0.566 all start from the specific residence unit that is next to the boundary and closest to downtown Shanghai, while workplace units are mostly located in the eastern part of downtown Shanghai. Moreover, these OD flows have longer distance than generally necessary for intercity commuters in this residence unit, which also implies that attractive jobs can compensate for long commuting distance. The OD flows arriving at the east end of downtown Shanghai generally have a higher coefficient for the scale of producer services in workplace. Different from the performance of the scale of producer services in workplace, OD flows from the same residence unit to different workplace units have various coefficients of the scale of non-producer services in residence. However, OD flows from different residence units to the same workplace unit have similar coefficients, which means that this determinant is more sensitive to the location change of workplace and less so to that of residence. It is evident that OD flows arriving at the southwest end of downtown Shanghai tend to have high coefficients for the scale of non-producer services in residence, whereas those arriving at the northeast end of downtown Shanghai tend to have low coefficients.

In summary, both in terms of significance scope and coefficient level, the scale of producer services in workplace and scale of non-producer services in residence are spatially nonstationary.

5.3.2. Spatial Nonstationarity of Residence Determinants

Residence determinants include housing supply in residence (Figure 5c), housing price in residence (Figure 5d), and year of construction in residence (Figure 5e). OD flows whose housing supply in residence and housing price in residence passed the significance test mostly start from the residence units near the boundary of Suzhou and Shanghai and arrive at the central and western areas of downtown Shanghai. For most OD flows, the year of construction in residence passed the significance test, and its influence is spatially extensive. We can see that the residence units of those OD flows are mostly near the boundary between Suzhou and Shanghai. The greater the distance between a residence unit and the boundary, the fewer OD flows pass the significance test in this unit. However, along the HSR line, even in the residence units far from the boundary, there is a considerable amount of OD flows whose year of construction in residence passed the significance test. This means that HSR can significantly reduce the travel time and make the loss of social network acceptable even if the places of residence are very far from the workplace. In most spatial units in downtown Shanghai, there are OD flows whose year of construction in residence passed the significance test, which means in terms of significance scope, the spatial nonstationarity of this determinant can hardly be observed in the workplace area.

OD flows that have higher coefficients of housing supply in residence are all from the specific residence unit near the boundary of Suzhou and Shanghai to the central area of downtown Shanghai, and OD flows that have higher coefficients of housing price in residence have similar patterns. For the two determinants, OD flows starting from the same residence unit tend to have similar coefficients while those arriving at the same workplace unit have greatly varying coefficients. This means that the two determinants are more sensitive to the location change of residence and less so to that of workplace. For OD flows whose year of construction in residence has higher significant influence with coefficients from 10.617~10.894, their residence units are mostly located near the boundary of Suzhou and Shanghai, and their workplace units are mostly located in the southwest end of downtown Shanghai. For OD flows starting from the residence units farther from the boundary of the two cities, the coefficient of year of construction in residence is lower. This determinant performs in a highly similar manner in OD flows starting from the same residence unit while no such similarity is observed in the workplace area. This means that this determinant is more sensitive to the location change of residence and less so to that of workplace.

In summary, housing supply in residence, housing price in residence, and year of construction in residence are spatially nonstationary both in terms of significance scope and coefficient level.

5.3.3. Spatial Nonstationarity of Commuting Time

OD flows whose commuting time passed the significance test mostly start from the residence units near the boundary of Suzhou and Shanghai and arrive at the central and western area of downtown Shanghai (Figure 5f). Interestingly, although HSR, for its role in reducing commuting time, is one of the most important causes for why intercity commuting has boomed, OD flows starting from the residence units along the HSR line are the least affected by commuting time and HSR does not extend the spatial influence of commuting time in residence areas. On the contrary, OD flows starting from residence units far from HSR are more affected by commuting time. OD flows that are more subject to the influence of commuting time with coefficients from $-2.059\sim-1.851$ mostly start from the residence units near the boundary of Suzhou and Shanghai and arrive at the workplace units near the transport hub (Hongqiao Railway Station) in downtown Shanghai. As the northeast end of downtown Shanghai is far from the transport hub, OD flows arriving there are less subject to the influence of commuting time. Commuting time is sensitive to spatial change in both residence and workplace, and spatially nonstationary both in terms of significance scope and coefficient level.

6. Conclusions

Determinants of intercity commuting flows are usually studied from a global perspective, which can easily result in inaccurate estimation. In many real-life situations, spatial nonstationarity is a more reasonable assumption. *NB* regression can select the significant determinants of intercity commuting flows from Suzhou to Shanghai, and SWIM can study their spatial nonstationarity. The conclusions are as follows:

1. Intercity commuting flows are positively correlated with the scale of producer services in workplace, scale of non-producer services in residence, housing supply in residence, and year of construction in residence; while they are negatively correlated with housing price in residence and commuting time. On the other hand, the scale of non-producer services in workplace, scale of producer services in residence, housing supply in workplace, year of construction in workplace, housing price in workplace, and density of primary and secondary schools both in workplace and residence are not significant determinants. It is evident that intercity commuting flows are a joint result of multiple determinants regarding both workplace and residence;
2. The six determinants that pass the significance test exhibit evident spatial nonstationarity in terms of significance scope and coefficient level. In comparison to the global regression model, the local regression model is more fitting for intercity commuting flows. In terms of significance scope, the scale of non-producer services in residence, housing supply in residence, housing price in residence, and commuting time are spatially nonstationary in both residence and workplace areas while the other two determinants are spatially nonstationary mainly in the residence area. In terms of coefficient level, the scale of producer services in workplace and commuting time are spatially nonstationary in both residence and workplace areas; the scale of non-producer services in residence is spatially nonstationary mainly in the workplace area; and the other determinants are spatially nonstationary mainly in the residence area. It is worth noting that for the scale of producer services in workplace and year of construction in residence, some residence units farther from the boundary still have significant OD flows because HSR can greatly reduce commuting time to turn commuting restrictions into non-restrictions. However, HSR does not play a similar role in other significant determinants;

3. The determinants of intercity commuting flows may manifest spatial nonstationarity in both residence and workplace areas. The spatial nonstationarity of flows is more complicated than that of points or polygons. Spatial nonstationarity usually means that the farther two spatial features are from each other, the more probable that the coefficients of determinants may vary. The coordinates of a flow are co-determined by its origin and destination; hence, the spatial nonstationarity of its coefficients is not restricted to only workplace or residence. For example, commuting time is spatially nonstationary in both workplace and residence areas. Compared with GWR models, SWIM demonstrates the spatial nonstationarity of intercity commuting flows in a more comprehensive way: the determinants of residence and workplace are all included in the joint model, and more importantly, it also allows us to study their spatial nonstationarity in residence and workplace areas simultaneously.

7. Discussion

7.1. The Understanding of Spatial Nonstationarity

Fotheringham, et al. [63] believe that spatial variation in people's attitudes and preferences is one of the major causes of spatial nonstationarity; it has been proven in previous empirical studies [50,64]. The spatial nonstationarity of the Suzhou–Shanghai intercity commuting flows can also be tentatively explained from this perspective.

Generally speaking, within the administrative area of Suzhou, the downtown area is better-developed socially and economically than the rest (e.g., the areas bordering Shanghai). Compared with the intercity commuters living in downtown Suzhou, those living in the border areas of Suzhou and Shanghai pay less rent and live in less ideal conditions. When making commuting decisions, they attach more importance to high salaries, but at the same time, they also value cost-effectiveness of residence. Therefore, the six determinants all exert significant influence on the intercity commuters in the border areas of Suzhou and Shanghai. Within such areas, spatial nonstationarity is also very obvious. For example, OD flows whose scale of producer services in workplace has higher significant influences with coefficients from 0.529~0.566 all start from the specific residence unit next to the boundary and closest to downtown Shanghai. The reason might be that the intercity commuters living there are more eager for jobs with high salaries.

Although HSR is the major commuting mode of intercity commuters living in downtown Suzhou, they care neither about commuting time nor living costs. The only significant determinants influencing their decisions are the scale of producer services in workplace and year of construction in residence. This may mean that they choose intercity commuting because those high-end jobs are only available in downtown Shanghai. The biggest positive effect of HSR on them is not that it has reduced their commuting cost, but that it has made intercity commuting a reality.

Spatial nonstationarity is not a new topic concerning studies of commuting determinants. However, as required by the research questions or due to the lack of research tools, previous studies have to choose the spatial interaction model or gravity model to evaluate the joint influences of the determinants of origins and destinations. Meanwhile, previous studies have made great efforts to optimize the gravity model, such as considering more independent variables [65] and exploring the influence of choosing inappropriate distance–decay function forms [66] or coefficients [67]. However, there has been no consensus so far. SWIM may deal with the above problems because it has a more flexible model structure to allow more variables. Moreover, GWR can reduce the estimation bias caused by incorrect function forms, which is also one of the reasons for spatial nonstationarity [63].

Another way to compromise is to adopt the GWR models based on the polygon or point feature to reveal spatial nonstationarity, though flow feature is the most fundamental descriptive feature type for commuting. The common method is converting commuting flows into a certain spatial unit ratio and studying the spatial nonstationarity of the relationship between this ratio and the various independent variables. However, this can only reveal the spatial nonstationarity in residence or workplace areas. For example, a study of extreme commuting discussed the spatial nonstationarity of socio-economic and land-use variables in residence areas [50]. There is no doubt that compared with the GWR model, SWIM is less developed and flawed in many aspects. The spatial nonstationarity of flows is more complicated and sometimes even difficult to interpret in an intuitive way. There is not yet a more feasible visualization technique for the cases with too many flows. Nevertheless, it is worth noting that, despite its flaws, SWIM can better reveal the spatial nonstationarity of OD flows and enhance our understanding of commuting.

7.2. The Representativeness of the Case and the Planning Strategies

Intercity commuting usually occurs between megacities and surrounding small- to medium-sized cities because the downtown megacities can provide high-level occupations, and small- to medium-sized cities have cost-effective living conditions [13,14,68]. It is worth noting that both Shanghai and Suzhou are megacities with populations of more than 10 million people. Suzhou can offer relatively high-level occupations, and the living cost of downtown is not low. However, compared with Suzhou, Shanghai is the unique economic center of advanced industrial agglomeration in the Yangtze River Delta region, which is attractive enough for intercity commuters from downtown Suzhou despite the long-distance commuting. The intercity commuters living near the boundary of Suzhou and Shanghai are more in line with the job and residential preferences of those commuters from surrounding cities to megacities mentioned in previous studies. The unique spatial distribution of socio-economic factors in workplace and residence makes the spatial nonstationarity evidently manifested. The case study of Suzhou–Shanghai can echo the typical pattern of intercity commuting and reveal a novel pattern of intercity commuting.

Local models considering spatial nonstationarity can more accurately analyze the real world, allowing for more targeted policy-making. For intercity commuters living near the boundary of Suzhou and Shanghai, they are forced to choose intercity commuting because of the inability to pay for high living costs in downtown Shanghai or even in the suburbs of Shanghai. Therefore, increasing housing supply and reducing housing prices in the suburbs of Shanghai, or enhancing the polycentric structure of employment in Shanghai, are effective measures to shorten the commuting distance for these intercity commuters. However, intercity commuters living in downtown Suzhou are not sensitive to the living cost but more eager for career development and personal pursuits. Perhaps optimizing transfer service efficiency and reducing travel time are effective measures to improve their commuting satisfaction.

7.3. Data Limitations and Future Prospects

Revealing the comprehensive commuting patterns is the essential precondition for analyzing spatial nonstationarity, which is why mobile singling data was chosen in this study. However, there are still some questions for future research that stem mainly from data limitations. First, the lack of census data at the township level makes it difficult to accurately control residents' average preferences. Fortunately, the census data at the county level imply that the socio-economic attributes of Suzhou residents do not vary over space significantly. Meanwhile, intercity commuters make up a low percentage of residents, which implies that the average preferences of residents may not affect the model estimates. Second, this study is based on the OD flow data at the 1-km grid level rather than the original trajectories of intercity commuters due to the privacy restrictions, which makes it hard to speculate on commuting mode and substitute it into the model. Commuting cost generally includes monetary cost and time cost. Since commuting cost is the indicator of the separation between workplace and residence in the spatial interaction model, ignoring monetary cost may bias the model estimates, especially for commuting time. However, a questionnaire survey conducted in 2020 demonstrated that nearly half of the intercity commuters from Suzhou to Shanghai have a monthly income of more than RMB 20,000 yuan [20]. The monetary cost of commuting per month is between RMB 800 and 2000 yuan, which seems to be less than 10% of monthly income for most intercity commuters. The previous study shows that commuters prefer reducing commuting time despite the higher monetary costs when incomes rise [69]. Therefore, the negative influence of missing monetary cost is probably insignificant.

In the research of intercity or long-distance commuting, official surveys and censuses may be more appropriate data sources. They usually contain data collected through stratified random sampling or national census. Not only can they reveal the overall commuting features in the study area holistically, but they can also show individual features and attributes that can help us directly and accurately evaluate the commuting preferences of its residents. However, such data sources usually cost too much and are hardly accessible. More importantly, they are not updated in a timely enough manner. Another common data source that can reveal individual features and attributes is self-launched questionnaire surveys, but they are not accurate enough in revealing the regional comprehensive commuting characteristics. Mobile signaling data are low-cost, easily accessible, updated, and of a high resolution and high sampling rate, which makes it an increasingly common data source in commuting studies. The combination of surveys and big data may become a paradigm in commuting studies in the future. Big data can be used to show spatial features, study spatial determinants, and build spatial models, and surveys can be used to interpret and test the findings from the perspective of individual features.

Author Contributions: Conceptualization, Zhipeng Li and Xinyi Niu; methodology, Zhipeng Li and Xinyi Niu; software, Zhipeng Li; formal analysis, Zhipeng Li and Xinyi Niu; resources, Xinyi Niu; data curation, Xinyi Niu and Zhipeng Li; writing—original draft preparation, Zhipeng Li; writing—review and editing, Xinyi Niu; visualization, Zhipeng Li; supervision, Xinyi Niu; project administration, Xinyi Niu; funding acquisition, Xinyi Niu. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (51878457).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw mobile singling data are not publicly available to preserve privacy. Social-economic data was obtained from public websites. Township-level data for exploring social-economic characteristics can be requested from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, C.L.; Hall, P. The impacts of high-speed trains on British economic geography: A study of the UK's InterCity 125/225 and its effects. *J. Transp. Geogr.* **2011**, *19*, 689–704. [CrossRef]
2. Chen, C.L.; Hall, P. The wider spatial-economic impacts of high-speed trains: A comparative case study of Manchester and Lille sub-regions. *J. Transp. Geogr.* **2012**, *24*, 89–110. [CrossRef]
3. Baum-Snow, N.; Brandt, L.; Henderson, J.V.; Turner, M.A.; Zhang, Q. Roads, railroads, and decentralization of Chinese cities. *Rev. Econ. Stat.* **2017**, *99*, 435–448. [CrossRef]
4. Ren, X.; Wang, F.; Wang, C.; Du, Z.; Chen, Z.; Wang, J.; Dan, T. Impact of high-speed rail on intercity travel behavior change. *J. Transp. Land Use* **2019**, *12*, 265–285.
5. Heuermann, D.F.; Schmieder, J.F. The effect of infrastructure on worker mobility: Evidence from high-speed rail expansion in Germany. *J. Econ. Geogr.* **2019**, *19*, 335–372. [CrossRef]
6. Zhang, W.; Fang, C.; Zhou, L.; Zhu, J. Measuring megaregional structure in the Pearl River Delta by mobile phone signaling data: A complex network approach. *Cities* **2020**, *104*, 102809. [CrossRef]
7. Wang, F.; Wei, X.; Liu, J.; He, L.; Gao, M. Impact of high-speed rail on population mobility and urbanisation: A case study on Yangtze River Delta urban agglomeration, China. *Transp. Res. Part A Policy Pract.* **2019**, *127*, 99–114. [CrossRef]
8. Mohino, I.; Loukaitou-Sideris, A.; Urena, J.M. Impacts of high-speed rail on metropolitan integration: An examination of London, Madrid and Paris. *Int. Plan. Stud.* **2014**, *19*, 306–334. [CrossRef]
9. Commuting between Regions. Available online: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210610--1 (accessed on 3 March 2022).
10. Guirao, B.; Campa, J.L.; Casado-Sanz, N. Labour mobility between cities and metropolitan integration: The role of high speed rail commuting in Spain. *Cities* **2018**, *78*, 140–154. [CrossRef]

11. Parenti, A.; Tealdi, C. The role of job uncertainty in inter-regional commuting: The case of Italy. *Growth Chang.* **2019**, *50*, 634–671. [CrossRef]
12. Martinus, K.; Suzuki, J.; Bossaghzadeh, S. Agglomeration economies, interregional commuting and innovation in the peripheries. *Reg. Stud.* **2020**, *54*, 776–788. [CrossRef]
13. Garmendia, M.; Romero, V.; Ureña, J.M.D.; Coronado, J.M.; Vickerman, R. High-speed rail opportunities around metropolitan regions: Madrid and London. *J. Infrastruct. Syst.* **2012**, *18*, 305–313. [CrossRef]
14. Garmendia, M.; de Urena, J.M.; Ribalaygua, C.; Leal, J.; Coronado, J.M. Urban residential development in isolated small cities that are partially integrated in metropolitan areas by high speed train. *Eur. Urban Reg. Stud.* **2008**, *15*, 249–264. [CrossRef]
15. Garmendia, M.; Ureña, J.D.; Coronado, J. Long-distance trips in a sparsely populated region: The impact of high-speed infrastructures. *J. Transp. Geogr.* **2011**, *19*, 537–551. [CrossRef]
16. Guirao, B.; Casado-Sanz, N.; Campa, J.L. Labour opportunities provided by Spanish high-speed rail (HSR) commuting services in a period of financial crisis: An approach based on regional wage disparities and housing rental prices. *Reg. Stud.* **2020**, *54*, 539–549. [CrossRef]
17. Hazans, M. Does commuting reduce wage disparities? *Growth Chang.* **2004**, *35*, 360–390. [CrossRef]
18. Ferreira, J.P.; Barata, E.; Ramos, P.N.; Cruz, L. Economic, social, energy and environmental assessment of inter-municipality commuting: The case of Portugal. *Energy Policy* **2014**, *66*, 411–418. [CrossRef]
19. Kissinger, M.; Reznik, A. Detailed urban analysis of commute-related GHG emissions to guide urban mitigation measures. *Environ. Impact Assess. Rev.* **2019**, *76*, 26–35. [CrossRef]
20. Wang, L.; Zhang, S.; Sun, W.; Chen, C.L. Exploring the physical and mental health of high-speed rail commuters: Suzhou-Shanghai inter-city commuting. *J. Transp. Health* **2020**, *18*, 100902. [CrossRef]
21. Wu, Q.; Perl, A.; Sun, J. Bigger and different: Beginning to understand the role of high-speed rail in developing China's future supercities. *Transp. Res. Rec.* **2016**, *2546*, 78–87. [CrossRef]
22. Li, K.; Niu, X. Delineation of the Shanghai Megacity Region of China from a Commuting Perspective: Study Based on Cell Phone Network Data in the Yangtze River Delta. *J. Urban Plan. Dev.* **2021**, *147*, 04021022. [CrossRef]
23. Cassel, S.H.; Macuchova, Z.; Rudholm, N.; Rydell, A. Willingness to commute long distance among job seekers in Dalarna, Sweden. *J. Transp. Geogr.* **2013**, *28*, 49–55. [CrossRef]
24. Sandow, E.; Westin, K. The persevering commuter–Duration of long-distance commuting. *Transp. Res. Part A Policy Pract.* **2010**, *44*, 433–445. [CrossRef]
25. Chung, H.; Yang, Y.; Chen, C.I.; Vickerman, R. Exploring the effects of built environment, location and accessibility on travel time of long-distance commuters in Suzhou and Shanghai, China. *Built Environ.* **2020**, *46*, 342–361. [CrossRef]
26. Guirao, B.; Lara-Galera, A.; Campa, J.L. High Speed Rail commuting impacts on labour migration: The case of the concentration of metropolis in the Madrid functional area. *Land Use Policy* **2017**, *66*, 131–140. [CrossRef]
27. Fotheringham, A.S.; Charlton, M.; Brunsdon, C. The geography of parameter space: An investigation of spatial non-stationarity. *Int. J. Geogr. Inf. Syst.* **1996**, *10*, 605–627. [CrossRef]
28. Guo, J.; Feng, T.; Timmermans, H.J. Co-dependent workplace, residence and commuting mode choice: Results of a multi-dimensional mixed logit model with panel effects. *Cities* **2020**, *96*, 102448. [CrossRef]
29. Fotheringham, A.S.; O'Kelly, M.E. *Spatial Interaction Models: Formulations and Applications*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1989; Volume 1.
30. Ogura, L.M. Effects of urban growth controls on intercity commuting. *Urban Stud.* **2010**, *47*, 2173–2193. [CrossRef] [PubMed]
31. Chen, H.; Voigt, S.; Fu, X. Data-driven analysis on inter-city commuting decisions in Germany. *Sustainability* **2021**, *13*, 6320. [CrossRef]
32. Suh, S.H. The possibility and impossibility of intercity commuting. *J. Urban Econ.* **1988**, *23*, 86–100. [CrossRef]
33. Green, A.E.; Hogarth, T.; Shackleton, R.E. Longer distance commuting as a substitute for migration in Britain: A review of trends, issues and implications. *Int. J. Popul. Geogr.* **1999**, *5*, 49–67. [PubMed]
34. Wessel, T. Economic change and rising income inequality in the Oslo region: The importance of knowledge-intensive business services. *Reg. Stud.* **2013**, *47*, 1082–1094. [CrossRef]
35. Hsu, P.F. Inter-industry wage premiums and industry-specific productivity in Taiwan. *Appl. Econ.* **2005**, *37*, 1523–1533. [CrossRef]
36. Cervero, R. Jobs-housing balancing and regional mobility. *J. Am. Plan. Assoc.* **1989**, *55*, 136–150. [CrossRef]
37. LaMondia, J.J.; Aultman-Hall, L.; Greene, E. Long-distance work and leisure travel frequencies: Ordered probit analysis across non–distance-based definitions. *Transp. Res. Rec.* **2014**, *2413*, 1–12. [CrossRef]
38. Andersson, M.; Lavesson, N.; Niedomysl, T. Rural to urban long-distance commuting in Sweden: Trends, characteristics and pathways. *J. Rural. Stud.* **2018**, *59*, 67–77. [CrossRef]
39. Dargay, J.M.; Clark, S. The determinants of long distance travel in Great Britain. *Transp. Res. Part A Policy Pract.* **2012**, *46*, 576–587. [CrossRef]
40. Mitra, S.K.; Saphores, J.D.M. Why do they live so far from work? Determinants of long-distance commuting in California. *J. Transp. Geogr.* **2019**, *80*, 102489. [CrossRef]
41. Cheng, J.; Yan, R.; Gao, Y. Exploring spatial heterogeneity in accessibility and transit mode choice. *Transp. Res. Part D Transp. Environ.* **2020**, *87*, 102521.
42. Ogura, L.M. Urban growth controls and intercity commuting. *J. Urban Econ.* **2005**, *57*, 371–390.

43. Li, Z.C.; Ma, J.C. Investing in inter-city and/or intra-city rail lines? A general equilibrium analysis for a two-city system. *Transp. Policy* **2021**, *108*, 59–82. [CrossRef]
44. Yang, T. Understanding commuting patterns and changes: Counterfactual analysis in a planning support framework. *Environ. Plan. B Urban Anal. City Sci.* **2020**, *47*, 1440–1455. [CrossRef]
45. Boots, B.; Okabe, A. Local statistical spatial analysis: Inventory and prospect. *Int. J. Geogr. Inf. Sci.* **2007**, *21*, 355–375. [CrossRef]
46. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [CrossRef]
47. Charlton, M.; Fotheringham, S.; Brunsdon, C. *Geographically Weighted Regression*; White Paper; National University of Ireland: Maynooth, Ireland, 2009.
48. Tao, X.; Fu, Z.; Comber, A.J. An analysis of modes of commuting in urban and rural areas. *Appl. Spat. Anal. Policy* **2019**, *12*, 831–845. [CrossRef]
49. Lloyd, C.; Shuttleworth, I. Analysing commuting using local regression techniques: Scale, sensitivity, and geographical patterning. *Environ. Plan. A* **2005**, *37*, 81–103. [CrossRef]
50. Bai, X.; Zhai, W.; Steiner, R.L.; He, Z. Exploring extreme commuting and its relationship to land use and socioeconomics in the central Puget Sound. *Transp. Res. Part D Transp. Environ.* **2020**, *88*, 102574. [CrossRef]
51. Jang, W.; Yuan, F.; Lopez, J.J. Investigating Sustainable Commuting Patterns by Socio-Economic Factors. *Sustainability* **2021**, *13*, 2180. [CrossRef]
52. Nakaya, T. Local spatial interaction modelling based on the geographically weighted regression approach. *Geol. J.* **2001**, *53*, 347–358.
53. Kordi, M.; Fotheringham, A.S. Spatially weighted interaction models (SWIM). *Ann. Am. Assoc. Geogr.* **2016**, *106*, 990–1012. [CrossRef]
54. Zhang, L.; Cheng, J.; Jin, C. Spatial interaction modeling of OD flow data: Comparing geographically weighted negative binomial regression (GWNBR) and OLS (GWOLSR). *ISPRS Int. J. Geo. Inf.* **2019**, *8*, 220. [CrossRef]
55. Xu, J.; Jin, C. Exploring spatiotemporal heterogeneity in online travel searches: A local spatial model approach. *Geogr. Tidsskr. Dan. J. Geogr.* **2019**, *119*, 146–162. [CrossRef]
56. Cheng, Y.; LeGates, R. China's hybrid global city region pathway: Evidence from the Yangtze River Delta. *Cities* **2018**, *77*, 81–91. [CrossRef]
57. Hu, S.; Song, W.; Li, C.; Zhang, C.H. The evolution of industrial agglomerations and specialization in the Yangtze River Delta from 1990–2018: An analysis based on firm-level big data. *Sustainability* **2019**, *11*, 5811. [CrossRef]
58. Yang, X.; Fang, Z.; Yin, L.; Li, J.; Zhou, Y.; Lu, S. Understanding the spatial structure of urban commuting using mobile phone location data: A case study of Shenzhen, China. *Sustainability* **2018**, *10*, 1435. [CrossRef]
59. Zhao, P.; Liu, D.; Yu, Z.; Hu, H. Long commutes and transport inequity in China's growing megacity: New evidence from Beijing using mobile phone data. *Travel Behav. Soc.* **2020**, *20*, 248–263. [CrossRef]
60. Hadachi, A.; Pourmoradnasseri, M.; Khoshkhah, K. Unveiling large-scale commuting patterns based on mobile phone cellular network data. *J. Transp. Geogr.* **2020**, *89*, 102871. [CrossRef]
61. Peng, Z.; Bai, G.; Wu, H.; Liu, L.; Yu, Y. Travel mode recognition of urban residents using mobile phone data and MapAPI. *Environ. Plan. B Urban Anal. City Sci.* **2021**, *23*, 9981. [CrossRef]
62. Da Silva, A.R.; Rodrigues, T.C.V. Geographically weighted negative binomial regression—incorporating overdispersion. *Stat. Comput.* **2014**, *24*, 769–783. [CrossRef]
63. Fotheringham, A.S.; Charlton, M.E.; Brunsdon, C. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. *Environ. Plan. A* **1998**, *30*, 1905–1927. [CrossRef]
64. Kar, A.; Le, H.T.; Miller, H.J. What Is Essential Travel? Socioeconomic Differences in Travel Demand in Columbus, Ohio, during the COVID-19 Lockdown. *Annu. Am. Assoc. Geogr.* **2021**, 1–24. [CrossRef]
65. Simini, F.; Barlacchi, G.; Luca, M.; Pappalardo, L. A Deep Gravity model for mobility flows generation. *Nat. Commun.* **2021**, *12*, 1–13.
66. De Vries, J.J.; Nijkamp, P.; Rietveld, P. Exponential or power distance-decay for commuting? An alternative specification. *Environ. Plan. A* **2009**, *41*, 461–480. [CrossRef]
67. McArthur, D.P.; Kleppe, G.; Thorsen, I.; Ubøe, J. The spatial transferability of parameters in a gravity model of commuting flows. *J. Transp. Geogr.* **2011**, *19*, 596–605.
68. Ureña, J.M.; Menerault, P.; Garmendia, M. The high-speed rail challenge for big intermediate cities: A national, regional and local perspective. *Cities* **2009**, *26*, 266–279. [CrossRef]
69. Van Ommeren, J.; Dargay, J. The optimal choice of commuting speed: Consequences for commuting time, distance and costs. *J. Transp. Econ. Policy JTEP* **2006**, *40*, 279–296.