


Article

The Influence of the Built Environment on School Children's Metro Ridership: An Exploration Using Geographically Weighted Poisson Regression Models

Yang Liu ¹, Yanjie Ji ^{1,*}, Zhuangbin Shi ²  and Liangpeng Gao ¹

¹ Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, School of Transportation, Southeast University, Dongnandaxue Road 2, Nanjing 211189, China; kmliuyang90@gmail.com (Y.L.); seugaolp@163.com (L.G.)

² Intelligent Transportation System Research Center, Southeast University, Dongnandaxue Road 2, Nanjing 211189, China; shizhuangbin@gmail.com

* Correspondence: jianjie@seu.edu.cn; Tel.: +86-13813996939

Received: 26 October 2018; Accepted: 7 December 2018; Published: 9 December 2018



Abstract: Long-distance school commuting is a key aspect of students' choice of car travel. For cities lacking school buses, the metro and car are the main travel modes used by students who have a long travel distance between home and school. Therefore, encouraging students to commute using the metro can effectively reduce household car use caused by long-distance commuting to school. This paper explores metro ridership at the station level for trips to school and return trips to home in Nanjing, China by using smart card data. In particular, a global Poisson regression model and geographically weighted Poisson regression (GWPR) models were used to examine the effects of the built environment on students' metro ridership. The results indicate that the GWPR models provide superior performance for both trips to school and return trips to home. Spatial variations exist in the relationship between the built environment and students' metro ridership across metro stations. Built environments around metro stations, including commercial-oriented land use; the density of roads, parking lots, and bus stations; the number of docks at bikeshare stations; and the shortest distance between bike stations and metro stations have different impacts on students' metro ridership. The results have important implications for proposing relevant policies to guide students who are being driven to school to travel by metro instead.

Keywords: students; metro ridership; school commuting; spatial variations; geographically weighted regression

1. Introduction

School children's reliance on cars is not only harmful to their daily travel behaviors towards sustainability but also aggravates traffic congestion on urban roads [1–4]; therefore, governments in many developed and developing countries have committed to encouraging school children to go to school using active modes of travel including walking and bicycling [5–8]. However, active modes are most suitable for short-distance travel. When students travel a long distance to school, policies and measures that encourage them to walk or bicycle to school would be ineffective. In China, although some students live near their schools under the *nearby enrollment policy*, there are many students, especially in junior and senior high schools, having to travel a long distance to school due to the uneven distribution of educational resources [9,10]. A similar phenomenon can also be found in many developed countries, such as the United States [11] and New Zealand [12]. Therefore, to reduce the number of car trips needed for escorting caused by spatial inequality, studies focusing on students who have medium and long-distance school commutes are essential.

In contrast to developed countries, the school bus systems in China and some developing countries are not mature. There are two main aspects that limit the widespread usage of school buses in China. First, China lacks favorable conditions for operating bus systems. As mentioned before, increasing numbers of parents would like to send their children to famous schools, which are usually far away from home; thus, many students are scattered across the city. This issue makes it extremely difficult to operate high-efficiency school buses. Second, the supervision and management of school bus operations are poor. On the one hand, there are no laws or regulations that oversees the construction of school buses; at the same time, the safety and quality of vehicles that picking up/dropping off students are not guaranteed, which worries parents [9,13]. On the other hand, schools would be unwilling to bear the costs and safety risk of school bus operation. To summarize, it is difficult to popularize school buses in China in the short term. Therefore, for students whose school travel is long distance, public transportation, especially in the metro area of a metropolis, is their best choice other than using a car. As metro ridership increases, the number of students choosing to travel by car decreases. To guide more students with middle or long commuting distances to travel by metro, it is important to investigate the factors that influence students' metro ridership for school trips.

Although only a few studies have focused on students' metro ridership [14], many studies have pointed out that the built environment has a considerable influence on students' choice of travel mode for school commuting [15–18]. For instance, in areas with high residential density and neighborhood density as well as well-connected and accessible by roads, the likelihood of students travelling by active modes increases [17]. Since students usually walk or bicycle from their homes to metro stations, the characteristics of the built environment around metro stations largely affects whether they choose the metro to travel either to or from school. It is, therefore, essential to examine the impact of the built environment on metro ridership for school trips at the station level. Importantly, the demand distribution of passengers on the metro may vary across space; thus, the effects of the spatial heterogeneity of factors should be considered when conducting an analysis of students' metro ridership. However, most of the existing relevant studies use multiple linear regressions to explore the influencing factors on metro ridership because this type of model can quickly, conveniently and effectively determine the relationship between independent variables and the dependent variable [19,20]. These studies have ignored the possibility of spatial variations. Therefore, to obtain a deeper understanding of students' metro ridership for school trips, it is necessary to find other models to reflect spatial heterogeneity and provide an important basis for transportation planning and management to guide students in using the metro.

Overall, based on the consideration of spatial heterogeneity, this study attempts to use geographically weighted regression (GWR) to investigate the relationship between students' metro ridership and the built environment around metro stations. The overall framework of this paper is as follows: Section 2 reviews existing literature in regard to the built environment factors that affect metro ridership at station level and the application of GWR in the transportation field. Section 3 provides a brief introduction of the study area and data set. Section 4 describes the research methodology. The results of the model and the corresponding discussions are described in Section 5. This is followed by a presentation of the implications for transport policies and specific measures in Section 6. Our conclusions are provided in the final section.

2. Literature Review

A large body of literature exists on school travel mode choice [2,8,9,21,22]; however, few studies have focused on students' metro ridership for school trips using smart card data [23]. The literature review focuses on previous studies that examine the influence of the built environment factors on metro ridership and the application of GWR in the transportation field.

2.1. Influence of the Built Environment on Metro Ridership

Since, as stated previously, there are few studies on the use of metros by primary and secondary school students, this paper uses the built environment factors affecting metro passenger flow in the existing literature as a reference. In the research examining metro ridership at the station level, the characteristics of the built environment are usually divided into land use, intermodal connections and external connectivity with the metro [19]. Previous studies have confirmed that land use is closely related to transit ridership, and population and employment are the most important variables related to land use [24,25]. Zhao et al. [19] introduced six factors related to mixed land use to explore the impacts of land use on metro ridership. They found that the number of education buildings, entertainment venues and shop centers have a significant influence on metro ridership. In terms of intermodal connections with the metro, such as feeder bus lines and park-and-ride (P&R) spaces, also have a significant impact on metro passenger flow [25,26]. Using Nanjing, China as a case study, Zhao et al. [19] explored the influence of the number of feeder bus lines and the bicycle P&R spaces within the metro stations' pedestrian catchment area (PCA) on the passenger flow of the metro. They found that the number of bicycle P&R spaces are closely related to metro ridership, while the number of feeder bus lines are not. Recently, Ji et al. [27] also found that the density of feeder bus stations and bikeshare stations around metro stations have a significant impact on metro-bikeshare transfer trips. In addition, the characteristics of external connectivity also impact transit ridership [19]. Many studies have shown that road density, the number of intersections, and bicycle facilities, etc. can significantly affect students' choices regarding active travel modes [9,15,17].

2.2. Application of GWR in the Transportation Field

GWR is a modeling method that reveals the spatial heterogeneity of influencing factors by allowing regression coefficient estimates to vary with geographic location [28]. This type of model has mostly been used in the fields of medical science, economics, geography and so on [29,30]. In recent years, this model has been applied to research in the field of transportation and has provided many results [31–34]. In the research on public transportation ridership, Chow et al. [35] applied a GWR model to predict bus passenger flow for home-based trips using data from Broward County, Florida. They found that the GWR model generates more accurate predictions than linear regression models. Later, Gardozo et al. [36] used a GWR model to predict the passenger flow of the Madrid metro at the station level. The authors also pointed out that the GWR model is more suitable for the analysis of metro ridership demand than the linear regression model. To improve public transportation usage in Taiwan, Chiou et al. respectively applied global and local regression models to identify the key factors affecting public transportation usage rates in different regions. The authors found that the GWR model has better accommodates of spatial autocorrelation and that most variables have parameters that differ across regions [37]. Similar findings can be found in the study of Tu et al. [38]. Recently, some scholars have tried to improve the GWR model according to the distribution characteristics of the research data. Ma et al. [39] used a geographically and temporally weighted regression model to explore the relationship between the built environment and transit ridership considering the space-time relationship. The authors indicated that the results of geographically and temporally weighted regression model are superior to those obtained using the GWR method and the ordinary least square (OLS) model. In addition, Ji et al. [27] used a geographically weighted Poisson regression model (GWPR) to study the relationship between sociodemographic, travel-related, and built environment variables and transfer volume from a spatial perspective. The result shows that the GWPR model results are better than those of the Poisson regression model.

Generally, previous studies have provided valuable information on factors influencing ridership and GWR modeling. However, these studies focused on the relationship between the passenger flow at the metro station and the surrounding built environment. The relationship between the built environment and students' metro ridership has not been studied using the GWR model. Little is known regarding whether and how the built environment influences on the metro usage rates across

the regions for students who usually live far away from school. Therefore, this paper focuses on the metro ridership at the station level for trips to school and return trips to home in Nanjing, China by using the GWR model; this research can address the gap in the existing literature.

3. Study Context and Data

3.1. Study Area Context

The study area is Nanjing, the capital of Jiangsu province, China, which includes an area of 6587 km² and in 2015 [40], had a population of more than 8.2 million. We choose Nanjing for this case study because (a) Nanjing is a veritable “educational metropolis”. By the end of 2016, there were 232 secondary schools in Nanjing and 349 primary schools (including 54 high schools and 178 junior high schools). In addition, the number of primary and secondary school students was 930,700 [40]. In the face of fierce competition related to college entrance examinations, parents in Nanjing pay more attention to their children’s education issues, and the phenomenon of *choosing schools across districts* is serious. (b) Nanjing is a typical large Chinese city according to the Standards for Categorizing City Sizes [41] and has a relatively well-developed metro network. Six metro lines have been built and are in operation, namely, lines 1, 2, 3, 10, S1 and S8; the total length of the metro network is 225 km. As shown in Figure 1, the metro network in Nanjing includes 113 stations, with an urban region covering 16 stations, a suburban region covering 58 stations and an exurban region covering 39 stations.

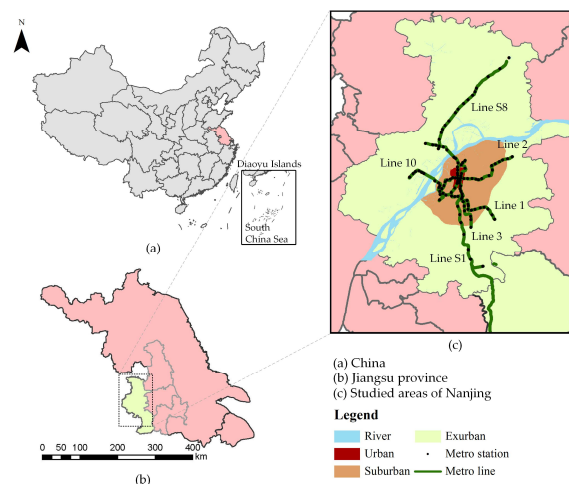


Figure 1. Study areas of Nanjing.

3.2. Data Source

3.2.1. Nanjing Metro Smart Card Data

Smart cards (SCDs) are used for the Nanjing public transit system and the data that can be obtained from the system include the card type, the corresponding station ID and time stamps for each card ID when a passenger enters or exits a metro station. Due to our research aim, this study used data only for student cards. A student SCD is available for students below 18 years old, and primary school students, junior high school students and high school students can use these cards. This study collected SCD data on Nanjing rail transit from 9–31 October 2016, which includes 15 specific weekdays from which samples of trips were extracted. The smart card data set used in this paper was a compilation of approximately 0.36 million transactions made by nearly 29 thousand SCD.

3.2.2. Nanjing Point of Interest (POI) Data

To measure the built environment, the 2016 Nanjing POI data, which were obtained from the Baidu map, were also used in this study. POI data are a kind of point data representing real geographical entities and include spatial information such as latitude and longitude and addresses, and attribute information such as name and category [39]. The data collected for this study include information about the boundaries of the county and district, urban roads, parks, metro stations, bus stations, public bicycle stations, parking lots, financial service areas, commercial buildings, retail industries, hotels, recreation, medical services, research and education, and corporate and residential communities. Built environment indicators were extracted from these data.

4. Methods

4.1. Identifying School Commuters

The research methods used to identify commuters based on SCD data have become mature; however, at present, there are few studies on the identification of school commuters [14]. Many studies have assumed that the place where the cardholder stays for more than several hours (except for the first place, which is considered to be the residence) can be considered to be the place of employment of the cardholder [42–44]. This identification problem occurs because most transit commuters are likely to take buses or the metro regularly with relatively fixed stops at similar times for a long time span [42]. Nevertheless, these methods are not appropriate for identifying school commuting. Because students may be escorted by their parents, quite several students just use the metro for one-way trips. Therefore, an identification method for students and their commuting trips considering escorts was adopted. According to Gu et al. [14], metro commuting students can be identified based on the school schedule and the students’ travel frequency. This study used the following recognition rules: (1) The boarding metro station with the highest frequency for each card ID is the candidate home or school. (2) When multiple stations with the same highest frequency occurs, adjacent stations are merged, and the records of more than two stations with the highest frequency are removed (it is a small probability event). (3) When determining a candidate station, the corresponding station with the highest frequency is identified as another candidate station; (4) Home and school in two candidate stations are determined specifically by the relationship between arrival time/departure time and school time. The flow chart is presented in Figure 2.

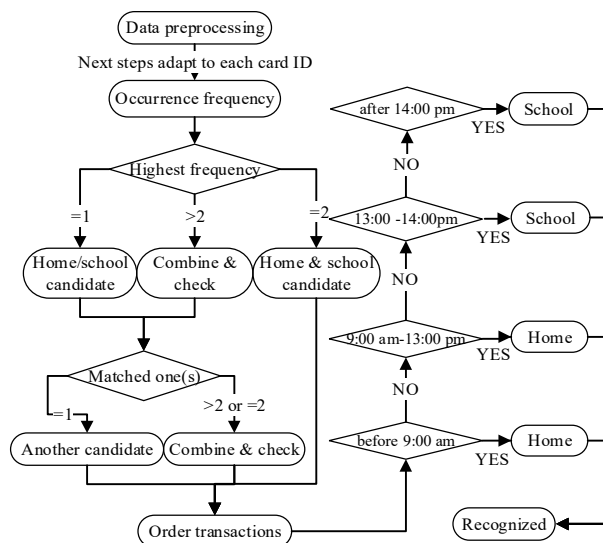


Figure 2. School commuter identification process.

Based on the identification steps, the SCD data indicate that 72,728 students used SCDs, and 28,925 of these students were identified as commuters. The distribution of the number of days that the identified commuters used the metro and commuted by metro during the 15 weekdays under study is shown in Figure 3. Evidently, the regular pattern of trips commuting to school is entirely in accordance with the distribution of the number of travel days for all students. This result verifies the reliability of the above identification method. A total of 351,631 school trips were taken during the 15 weekdays, including 191,055 trips to school (TS) and 160,576 return trips to home (RTH).

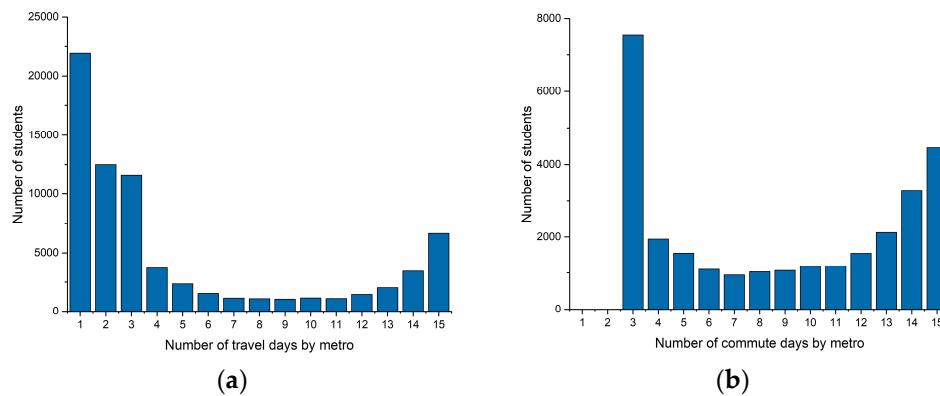


Figure 3. Distribution of the number of days students used the metro and commuted by metro. (a) Number of days the students use the metro; (b) Number of days the students commute.

4.2. Calculating the Dependent Variables

Because students' school commuting trips include TS and RTH, the study needs to consider the students' metro ridership at the station level for both trips. For TS, we measure the ridership of a metro station as the sum of the number of boarding students for which this is the home station and the number of deboarding students for which this is the school station. For RTH, we measure the ridership of a metro station as the sum of the number of boarding students for which this is the school station and the number deboarding students for which this is the home station. Notably, we found that the frequency of metro use by students for TS and for RTH changed little from Monday to Thursday, while the number of students who used the metro to go home from school on Friday increased significantly, as shown in Figure 4. The possible reason for this result is that some of the school students are resident students who usually stay at the school during weekdays and every week, go home on Friday. Therefore, this paper only includes the number of students using the metro from Monday to Thursday in consideration of the unusual characteristics of students using the metro on Fridays. More specifically, we use the average station ridership for TS and for RTH as the dependent variables.

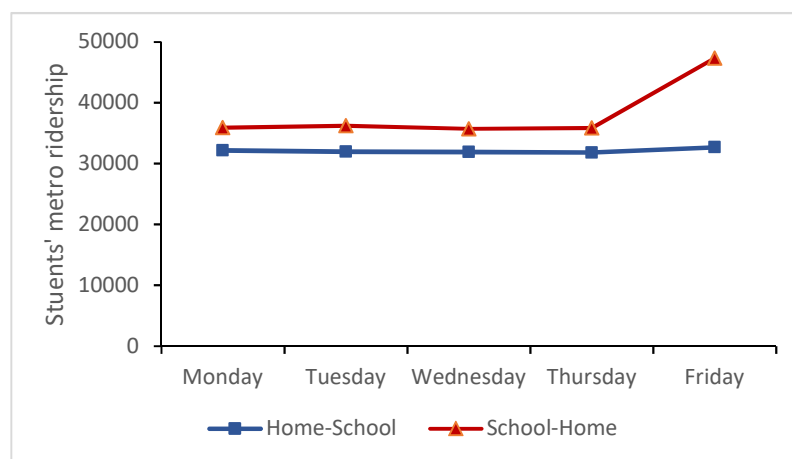


Figure 4. Distribution of the number of students commuted by metro for TS and RTH.

4.3. Selecting the Explanatory Variables

As stated in the literature review section, there are three main aspects of the built environment, namely, land use, connections with the metro and external connectivity, are associated with transit ridership. In this paper, following Xia et al. [45] and Zhao et al. [19], we used circles with a radius of 800 m as a buffer for the metro stations, and extracted the built environment indicators related to the frequency of primary and secondary school students' use of the metro to go to school and return home in the metro stations' PCA, as follows.

First, together with the characteristics of schooling behavior, we used population, employment, and the number of commercial and educational buildings to measure land use. Due to the limitations of our data and following Ma et al. [39], the number of residents and jobs in the PCA have been replaced by the POI of the residential buildings and employment places.

Second, because students who use the metro may need to transfer to either a bus, public bicycles or a car for the remainder of their trip, this study used variables that are related to intermodal connections, following Ji et al. [27], as well as variables related to P&R points.

Third, to measure the external connectivity of metro stations, we included data on road density, intersection density, the closest distance from the bus stop to the metro station, and the closest distance from the bicycle station to the metro station in the PCA.

Finally, a total of 12 variables are selected from three levels: land use, intermodal connection, and external connectivity. All variables are continuous variables, and we had calculated the maximum, minimum, mean, and standard deviation of these variables, as shown in Table 1.

Table 1. Statistical Results of the Various Indicators.

Independent Variables	Unit	Max	Min	Mean	S.D.
Dependent variables					
Trips to school (TS)	Number	32,197	31,826	31,973	139.67
Return trips to home (RTH)	Number	47,324	35,726	35,933	195.47
Land use variables					
Job density (JobD)	Number/km ²	431	0.00	41	72.99
Residential-oriented land use (RLU)	Percentage	1.00	0.00	0.06	0.10
Commercial-oriented land use (CLU)	Percentage	0.83	0.00	0.50	0.19
Educational-oriented land use (ELU)	Percentage	0.70	0.00	0.07	0.10
Transferring-related variables					
Density of P&R (ParkingD)	Number/km ²	40.78	0.00	5.33	6.82
Density of bus stations (BusD)	Number/km ²	40.78	0.00	12.00	10.42
Density of bike stations (BikeD)	Number/km ²	8.95	0.00	1.83	2.37
Docks at bikeshare stations (Docks)	Number/km ²	389.94	0.00	69.80	92.79
External connectivity variables					
Density of intersections (IntersectionD)	Number/km ²	18.54	0.56	5.71	3.86
Density of road networks (RoadD)	km/km ²	27.78	3.98	16.04	5.34
Shortest distance between bus station and metro station (SDtoBus)	m	3695.94	1.98	379.15	660.68
Shortest distance between bike station and metro station (SDtoBike)	m	42,693.46	16.21	5850.36	8792.79

4.4. Modeling GWPR

To investigate the impacts of the built environment on the metro ridership of students from a spatial perspective, a GWR model was used in this paper. As stated before, the GWR model is an extension of the linear regression model, which allows for the assessment of spatial heterogeneity in the relationship between a given independent variable and a set of dependent variables. In the GWR model, the closer the sample point is to a given regression point, the greater the influence on the sample point. For any metro station i , the GWR model uses the following expression [28]:

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^P \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad i = 1, 2, 3 \dots n \quad (1)$$

where (u_i, v_i) is the coordinates of the metro station i in space; and y_i , x_{ik} and ε_i are the dependent variables, the k th independent variable and the error term for location i , respectively. β_{i0} is a constant

variable; and β_k is the coefficient at station i for the independent variable x_{ik} . In addition, P is the number of independent variables. It is worth noting that GWR generally requires that the dependent variable has a normal distribution, and when the dependent variable is the “frequency of metro use by students”, in accordance with the Poisson distribution, the GWPR is more effective than the general GWR model. Based on Poisson regression, the formula for the GWPR can be expressed as follows:

$$y_i = e^{\beta_{i0}(u_i, v_i) + \sum_{k=1}^P \beta_k(u_i, v_i)x_{ik} + \varepsilon_i} \quad i = 1, 2, 3 \dots n \tag{2}$$

The linearization formula can be calculated as follows:

$$\ln y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^P \beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad i = 1, 2, 3 \dots n \tag{3}$$

where $\ln y_i$ is the log rate ratio of y_i . Please note that for each station, the parameter β_k can be different. The closer the observed data is to the station i , the greater the influence on $\beta_k(u_i, v_i)$ is. For the solving principle in detail, please refer to the study of Nakaya et al. [46]. In addition, the corrected Akaike information criterion (AICc) and the Akaike information criterion (AIC) were used to evaluate whether GWR provides a better fit than a global model [28], as a smaller AICc/AIC indicates a better result.

5. Results

5.1. Global Results

Before establishing the global and local models of the GWR, it is essential to perform a collinearity test on each parameter to remove parameters with strong collinearity. For TS and all independent variables, we used STATA software for collinearity testing. The independent variables RLU and BikeD are removed because they have a collinear relationship with the other variables (VIF > 10). Similarly, the independent variables RLU and BikeD were excluded from the collinearity test of the dependent variable RTH and the independent variables. The other tested variables are included in the global Poisson regression models for TS and RTH separately.

After deleting the variable for which the evaluation result was not significant ($p > 0.1$), the results are presented in Table 2. The first model uses TS as the dependent variable, and the second model uses RTH as the dependent variable. It is found that the land use around the metro station has less of an impact on the frequency of metro use by students commuting to school, and only commercial land use (CLU) has a significant impact. In contrast, intermodal connection and external connectivity have a significant impact on the student’s metro ridership for school trips.

Table 2. Global Poisson Modeling Results.

Dependent	Trips to School (TS)			Return Trips to Home (RTH)		
Variable	Coef.	S.E.	Z-Stat.	Coef.	S.E.	Z-Stat.
Intercept	5.377	0.007	746.032 ***	5.474	0.007	782.499 ***
CLU	0.446	0.007	61.512 ***	0.444	0.007	63.789 ***
RoadD	−0.407	0.008	−50.420 ***	−0.368	0.008	−47.939 ***
ParkingD	−0.387	0.011	−36.112 ***	−0.315	0.009	−33.659 ***
BusD	0.404	0.008	48.358 ***	0.393	0.008	50.676 ***
Docks	0.183	0.008	23.052 ***	0.183	0.007	24.677 ***
SDtoBike	−0.575	0.010	−56.195 ***	−0.607	0.010	−58.849 ***
AIC	20,495.239			19,214.216		
AICc	20,496.305			19,215.282		
Percent deviance explained	0.429			0.482		

Note: *** $p < 0.01$.

The table shows that CLU, BusD and Docks have a positive impact on the frequency of students using the metro for commuting. Specifically, in the first model, the coefficient of CLU is 0.446, indicating that after controlling for the other variables, each additional unit of CLU will result in an increase in the expected number of students who use the metro to go to school by 0.562 times ($e^{0.446} - 1$). In the same way, when the other variables remain unchanged, the expected number of students using the metro to go to school will increase by 0.498 times when BusD increases one unit. For each additional unit of Docks, the expected number of students using the metro will increase by 0.201 times. Unlike the above explanatory variables, RoadD, ParkingD and SDtoBike have a negative impact on the frequency of students using the metro for commuting. As the density of such variables within a PCA increases, the students' metro ridership decreases. Similar results can be found in the second model.

The conventional Poisson regression model explains the possible determinants of students' metro ridership for school trips from a global average point of view. However, the regression parameters are fixed across the space and the potential geographical variations of these parameters remains unknown. To explore the impact of spatial heterogeneity on the frequency of students' metro use more deeply, this paper will examine the variables using the GWPR model.

5.2. Local Results

Before establishing the GWPR model, the spatial nonstationarity of the dependent variable is tested. Spatial nonstationarity refers to a change in the relationship or structure of the variables caused by changes in geographic location and can be tested by calculating Moran's I from the residuals of the global regression result [29]. The value of Moran's I ranges from -1 to 1 . When the absolute Moran's I > 0 , a spatial correlation exists. The higher Moran's I is, the more obvious the spatial correlation. It is reasonable to establish a local model only when the regression residual does not have spatial randomness. In this paper, the global Moran's I of both the first model and the second model are 0.167 and 0.273, respectively. This indicates that the relationship between the dependent variable and the independent variable of the model is spatially nonstationary. Therefore, we normalize the significant variables in the global model and include them in the local model. By calculation, each factor has a specific coefficient value for each metro. The results of the factors affecting the frequency of metro use by students for TS and for RTH are shown in Tables 3 and 4.

Table 3. GWPR Modeling Results of TS.

Local Terms	Mean	Min	Low Quartile	Median	Upper Quartile	Max	DIFF of Criterion *
Intercept	5.201	2.377	4.94	5.253	5.787	8.069	-648.232
CLU	0.352	-0.145	0.175	0.363	0.562	0.95	-1888.56
RoadD	-0.336	-0.798	-0.493	-0.397	-0.176	0.12	-1244.441
ParkingD	-0.352	-1.078	-0.548	-0.349	-0.228	0.569	-1337.879
BusD	0.329	-0.084	0.196	0.342	0.436	0.84	-879.722
Docks	0.238	-0.181	0.026	0.167	0.448	0.852	-1055.801
SDtoBike	-0.862	-5.192	-1.103	-0.624	-0.382	3.237	-519.315
AIC	10,730.82	AICc	10,748.06	Percent of deviance explained 0.702			

Note: * The positive value of the diff-criterion suggests that nonspatial variability exists in terms of model selection criteria.

Table 4. GWPR Modeling Results of RTH.

Local Terms	Mean	Min	Low Quartile	Median	Upper Quartile	Max	DIFF of Criterion *
Intercept	5.456	3.143	5.188	5.395	5.890	8.451	-695.618
CLU	0.360	-0.090	0.192	0.337	0.546	0.928	-1992.976
RoadD	-0.290	-0.721	-0.433	-0.338	-0.145	0.115	-1065.073
ParkingD	-0.315	-0.889	-0.463	-0.292	-0.215	0.453	-1006.787
BusD	0.328	0.026	0.198	0.329	0.430	0.772	-817.654
Docks	0.246	-0.065	0.090	0.191	0.396	0.745	-865.232
SDtoBike	-0.634	-4.076	-0.961	-0.626	-0.273	3.582	-504.021
AIC	10,724.904	AICc	10,741.48	Percent of deviance explained 0.712			

Note: * The positive value of the diff-criterion suggests that nonspatial variability exists in terms of model selection criteria.

By comparing the AIC/the AICc of the global modal and those of the local model, it is found that the two local models have a lower AIC and AICc than their corresponding global models. This result shows that the regression result of the GWPR model is better than that of the Poisson regression model. In contrast to the fixed coefficients across space in the global model, all selected variables in the local model have spatial variabilities that depend on the diff-criterion values. Tables 2–4 demonstrate how different the coefficients of the variables appear in the global and local models. For instance, the coefficient Docks is positive in the global model but is negative in the local model. This indicates that the influence of the factor Docks on students' metro ridership indeed varies over space and that we need to consider spatial variations in the association between students' metro ridership and the characteristics of the built environment. Since each metro station has a specific coefficient estimate in the GWPR model, the estimation result of this model can reflect the influence of each factor on the frequency of metro use by students for school trips across the regions.

5.3. Analysis and Discussion

To further explore the impact of the spatial heterogeneity of the built environment variables on the frequency of students using the metro for TS and for RTH, we use ArcGIS software to express the coefficient estimates for each metro station as shown in Figures 5 and 6. Figure 5 shows that the effect of all variables on metro ridership for TS varies with the spatial distribution of the metro stations. CLU and BusD have positive impacts on student's metro ridership at the station level. Specifically, the coefficient of CLU is 0.446 in the global model, but in the local model, it increases from urban to exurban areas, with the largest value for exurban areas (Figure 5a). This result indicates that the coefficient in the global model is the average value of the continuous coefficients in the local mode. The global model cannot show that in areas with darker colors, each additional unit of CLU is associated with a higher frequency of students going to school. Perhaps this result occurs because commercial development encourages students to use the metro, but concentrated commercial development tends to attract more passengers at early peak hours, which is not conducive to students using the metro to go to school. In contrast, BusD has a greater positive association with the frequency of metro use by students in the central area, which includes line 2 and line 10 (in Figure 5d). Perhaps this result occurs because some metro commuters in these areas are more willing to transfer to a bus, which helps students use the metro for going to school during the morning peak.

Unlike the above variables, RoadD and ParkingD have negative impacts on the frequency of students using the metro for TS. As RoadD increases, the frequency with which students use the metro decreases. This result is consistent with the study conducted by Broberg and Sarjala [16]. Notably, this negative impact has a greater association with students' metro ridership for TS in south of the Yangtze River (in Figure 5b), which may be due to a higher population density and road density in these areas. Similarly, as ParkingD increases, the frequency with which students use the metro decreases. The main reason may be that the existence of a parking lot or parking space around the school will encourage parents to drive their children to school. Figure 5c shows that this negative impact is more obvious north of the Yangtze River. This result may have occurred because of the relatively low density of public transport facilities in these areas, and the increased density of parking lots encourages students to use a car to go to school.

Unexpectedly, although Docks and SDtoBike have a positive impact on the frequency of students' metro use to go to school in the global model, both positive and negative effects exist throughout the region. This result differs from the results of previous research [27]. In general, the more transfer facilities there are and the closer they are to the metro station, the more positive impact they have on metro passenger flow. However, Figure 5e shows that as the number of docks increase, the frequency of metro use by students who live near line 10 and line 1 decreases. The possible reason for this result is that these two metro lines have a concentrated large number of commuter passengers during the morning peak, and an increase in the number of public bike stations has attracted more commuters to use the metro because they can transfer from a bicycle. Facing travel chaos, primary and secondary

school students, as a vulnerable group, are unwilling to (or their parents will not allow them to) use the metro to go to school. Figure 5f shows that students living in the central area are more willing to use the metro to go to school, as the distance from a public bicycle station to the metro station increases. It may be that some students who travel a short distance prefer to go to school by bicycle rather than using the metro during the morning peak. However, these different spatial distributions cannot be reflected in the global model, which not only reduces the prediction accuracy of the model but is also unfavorable for policy recommendations.

As seen in Figure 6, the impact of the built environment on the frequency with which students use the metro for RTH is basically consistent with that of students using the metro for TS. The difference is that the number of docks has a positive effect only for the frequency with which students use the metro for RTH, as shown in Figure 6e. This result may have occurred because public transport services have been improved due to the difference between school runs and rush hours. This result confirms our previous suspicion that commuting trips during the morning peak have a negative impact on the use of the metro by primary and secondary school students. These results are of great significance for proposing facility planning, and they have policy implications for the use of the metro by students who commute over medium and long distances.

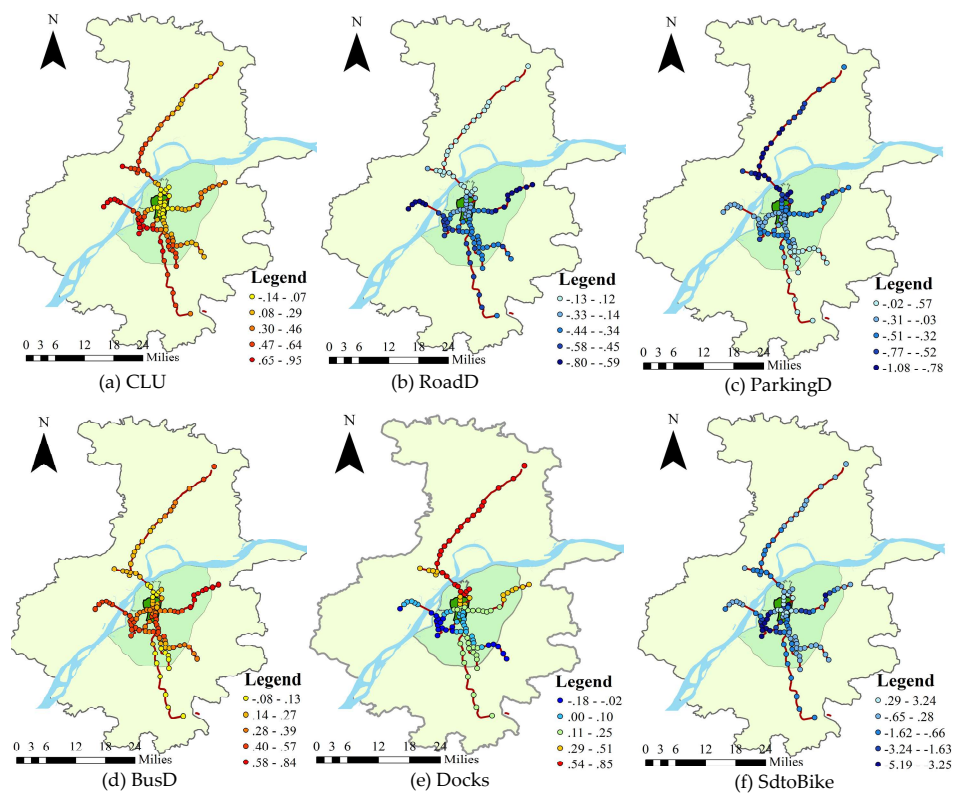


Figure 5. Spatial distribution of estimated coefficients of variables for TS.

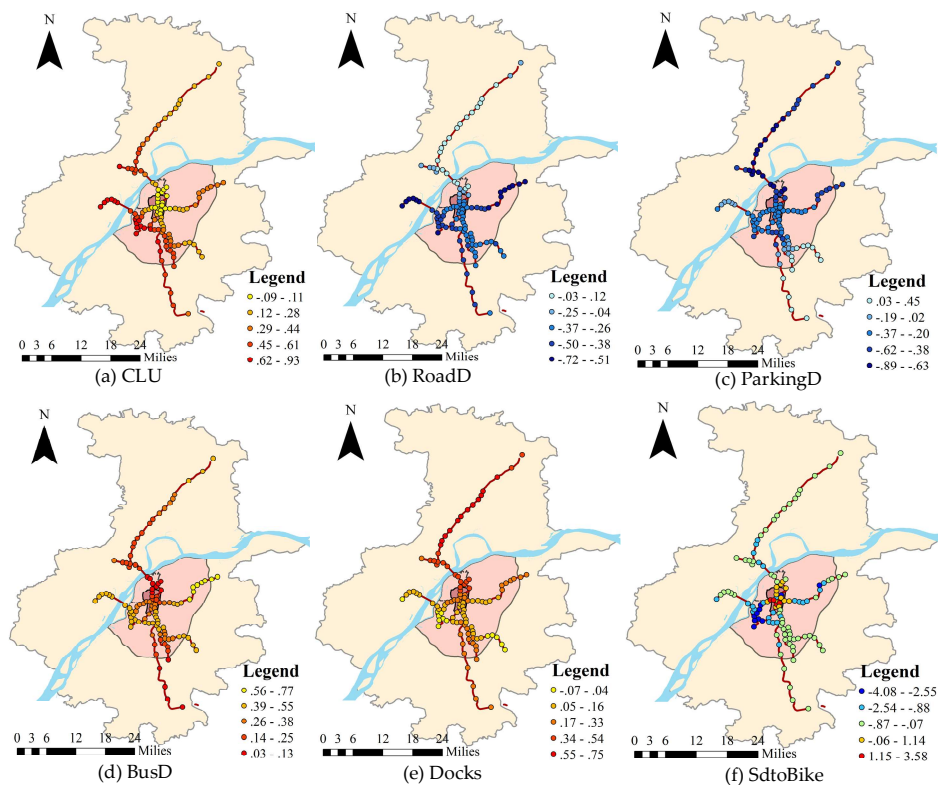


Figure 6. Spatial distribution of estimated coefficients of variables for RTH.

6. Implications and Limitations

6.1. Research Implications

In Nanjing and most other Chinese cities, for most parents, a major factor when choosing a school for their children is the quality of schools. Only after choosing a school will they decide whether, who and how to escort their children according to the school's location and other factors. Thus, students' choice of school travel mode, especially for students with independent travel ability, largely depends on the built environment around their home [47]. The results from this study reveal that land use and transport facilities have different impacts on the frequency of metro use by students across space, and these results have valuable implications for guiding students who travel long distances in metropolitan areas to use the metro transport system more.

First, although the commercial buildings that exist along the metro lines have a significant positive influence on students' metro ridership in suburban and exurban areas, this effect is not significant in the central area. This result suggests that commercial development around metro stations in the suburbs and exurbs should be encouraged. Notably, since an increase in the density of commercial areas attracts more people and traffic, setting a long time for walk signal at intersections within 800 m of the metro station may help encourage students to use the metro, especially in the area south of the Yangtze River in Nanjing.

Second, an increase in the density of bus stations also has a significant positive impact on students' metro ridership, especially around line 2 and line 10 in Nanjing. This result means that for students who go to school using the metro, the bus is more a feeder mode than an alternative mode. Therefore, enhancing the layout of bus stops around line 2 and line 10 will contribute to increasing students' metro ridership by improving the accessibility of these metro stations and shifting some of the commuting traffic off the metro. In addition, to attract more students who travel a medium or long distance to use public transit the government should encourage bus companies to add feeder buses to the metro

stations or to the longer distance local bus routes for primary and secondary school students during school runs.

Third, our results indicate that public bicycles have an important influence on students' metro ridership for TS. Since public bicycles first began to be available in the urban areas of Nanjing in 2015, the city realized the full coverage of the public bicycle network in all the districts at the end of 2017. However, according to our results stated above, the number of docks at bikeshare stations has a negative impact on students use of the metro to go to school. This result suggests that appropriate traffic management measures should be implemented after planning and designing transport infrastructure. Primary and middle school students, as a vulnerable group, prefer not to travel during the morning peak when there is a crowded and noisy environment in public transport systems. Our descriptive analyses also indicate that the average station ridership for TS is less than that for RTH. To solve this problem, staggered shifts should be implemented during the morning peak to accommodate both employees and students in some suburban and exurban areas (such as Qixia District, and Jianye District).

Moreover, the density of parking lots is closely related to the frequency of metro school commuting. As the density of the parking lots is increased, students become less willing to use the metro. Nevertheless, the number of parking facilities around the metro stations cannot be decreased because, on the one hand, the parking lots around metro stations are provided not only for parents but also for commuters or other social vehicles. On the other hand, some students, especially younger children with a long travel distance, must be escorted by their parents. If the number of parking facilities around metro stations are reduced, then the surplus parking demand will transfer to urban roads, which will inevitably cause traffic chaos. Establishing a mature school bus system is the only way to reduce the number of car trips used for these students in the future.

6.2. Research Limitations

This study has some limitations. First, this paper focuses on built environment factors affecting students' metro ridership for TS and RTH at the station level; however, factors at the individual level including students' socioeconomic attributes, information about their family, and information about the parents, and personal attributes are not included in the modeling due to the limitations of the SCD data. However, the student's family, parents and personal attributes are important factors that influence whether students use the metro to travel [9,10,48]. Second, in this paper, land use is measured by POI because data on building size cannot be obtained. Third, since this study only focuses only on students who commute in the metro system, students not using a metro smart card have not been considered. In future studies, we could extend this study to analyze more factors that may affect students' school commuting, such as the socioeconomic attributes of card holders and their parents as well as the socioeconomic variables associated with the metro stations' PCA, by combining a questionnaire and smart card data. Furthermore, we could establish a semi-parametric GWPR model in addition to the local model and compare their performance, thus providing more detailed insight into the significance of different travel groups.

7. Conclusions

This paper focuses on the primary and secondary school students' commuting behavior in the metro system and explores metro ridership at the station level for TS and RTH by using the smart card data (SCD) of Nanjing. A global Poisson regression model and a GWPR model were used to examine the spatial distributions of the built environment factors associated with students' metro ridership. The results show that GWPR models provide superior performance for both TS and RTH because the global model has misspecification by adopting a spatially averaged associations between correlates. However, spatial variations exist in the relationship between the built environment and students' metro ridership across metro stations. Specifically, built environments, including commercial-oriented land use; the density of roads, parking lots, and bus stations; the number of docks at bikeshare stations;

and the shortest distance between bike stations and metro stations around different metro stations have different impacts on students' metro ridership. According to these findings, recommendations for urban planners and policymakers are further proposed to guide students who are currently being driven to school to instead, travel instead by metro.

This research makes two main contributions. First, most of the existing research on student behaviors focuses on the choice of active travel modes (bicycling and walking), while neglecting the fact that long-distance school commuting is the key to the parental choice of car travel for the students. This paper explores the metro ridership of commuting students by using three weeks of SCD data and addresses the gap in the existing literature. A GWPR model was first used to explore the relationship between the built environment and students' metro ridership for TS and RTH. Using this approach not only improves the accuracy of the model but also provides a theoretical basis for developing differentiated policies and measures for guiding students who live in different regions to use the metro.

Author Contributions: Conceptualization, Y.L. and Y.J.; Data curation, Y.L. and Z.S.; Formal analysis, Y.L. and L.G.; Methodology, Y.L. and Z.S.; Writing—original draft, Y.L.; Revised the paper, Y.L. and Y.J.

Funding: This research was funded by the Projects of International Cooperation and Exchange of the National Natural Science Foundation of China [NO. 51561135003], the Key Project of National Natural Science Foundation of China [NO.51338003], the Scientific Research Foundation of Graduated School of Southeast University [NO. YBJ1842] and the China Scholarship Council (CSC) (No. 201806090205).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lu, M.; Sun, C.; Zheng, S. Congestion and pollution consequences of driving-to-school trips: A case study in Beijing. *Transp. Res. Part D* **2017**, *50*, 280–291. [[CrossRef](#)]
2. Li, S.; Zhao, P. The determinants of commuting mode choice among school children in Beijing. *J. Transp. Geogr.* **2015**, *46*, 112–121. [[CrossRef](#)]
3. Mitra, R. Independent Mobility and Mode Choice for School Transportation: A Review and Framework for Future Research. *Transp. Rev.* **2013**, *33*, 21–43. [[CrossRef](#)]
4. He, B.-J.; Zhao, D.-X.; Zhu, J.; Darko, A.; Gou, Z.-H. Promoting and implementing urban sustainability in China: An integration of sustainable initiatives at different urban scales. *Habitat Int.* **2018**, *82*, 83–93. [[CrossRef](#)]
5. McDonald, N.C. Children's mode choice for the school trip: The role of distance and school location in walking to school. *Transportation* **2008**, *35*, 23–35. [[CrossRef](#)]
6. Sidharthan, R.; Bhat, C.R.; Pendyala, R.M.; Goulias, K.G. Model for Children's School Travel Mode Choice. *Transp. Res. Rec. J. Transp. Res. Board* **2011**, *2213*, 78–86. [[CrossRef](#)]
7. Ermagun, A.; Samimi, A. Promoting active transportation modes in school trips. *Transp. Policy* **2015**, *37*, 203–211. [[CrossRef](#)]
8. Sener, I.N.; Lee, R.J.; Sidharthan, R. An examination of children's school travel: A focus on active travel and parental effects. *Transp. Res. Part A Policy Pract.* **2018**. [[CrossRef](#)]
9. Zhang, R.; Yao, E.; Liu, Z. School travel mode choice in Beijing, China. *J. Transp. Geogr.* **2017**, *62*, 98–110. [[CrossRef](#)]
10. Liu, Y.; Ji, Y.; Shi, Z.; He, B.; Liu, Q. Investigating the effect of the spatial relationship between home, workplace and school on parental chauffeurs' daily travel mode choice. *Transp. Policy* **2018**, *69*, 78–87. [[CrossRef](#)]
11. He, S.Y.; Giuliano, G. School choice: Understanding the trade-off between travel distance and school quality. *Transportation* **2017**. [[CrossRef](#)]
12. Mandic, S.; Sandretto, S.; Bengoechea, E.G.; Hopkins, D.; Moore, A.; Rodda, J.; Wilson, G. Enrolling in the Closest School or Not? Implications of school choice decisions for active transport to school. *J. Transp. Health* **2017**, *5*, S16–S17. [[CrossRef](#)]
13. Li, J.; Zhang, K.; Guo, J.; Jiang, K. Reasons Analyzing of School Bus Accidents in China. *Procedia Eng.* **2012**, *45*, 841–846. [[CrossRef](#)]

14. Gu, Y.; Liu, Y.; Ji, Y.; Ma, X. School Commuting Pattern in Metro System Across Different Loyalty Groups. In Proceedings of the Transportation Research Board 97th Annual Meeting, Washington, DC, USA, 7–11 January 2018.
15. Mitra, R.; Buliung, R.N.; Roorda, M.J. Built Environment and School Travel Mode Choice in Toronto, Canada. *Transp. Res. Rec. J. Transp. Res. Board* **2010**, *2156*, 10–1443. [[CrossRef](#)]
16. Broberg, A.; Sarjala, S. School travel mode choice and the characteristics of the urban built environment: The case of Helsinki, Finland. *Transp. Policy* **2015**, *37*, 1–10. [[CrossRef](#)]
17. Rothman, L.; Buliung, R.; To, T.; Macarthur, C.; Macpherson, A.; Howard, A. Associations between parents' perception of traffic danger, the built environment and walking to school. *J. Transp. Health* **2015**, *2*, 327–335. [[CrossRef](#)]
18. Ikeda, E.; Stewart, T.; Garrett, N.; Egli, V.; Mandic, S.; Hosking, J.; Witten, K.; Hawley, G.; Tautolo, E.S.; Rodda, J. Built environment associates of active school travel in New Zealand children and youth: A systematic meta-analysis using individual participant data. *J. Transp. Health* **2018**, *9*, 117–131. [[CrossRef](#)]
19. Zhao, J.; Deng, W.; Song, Y.; Zhu, Y. What influences Metro station ridership in China? Insights from Nanjing. *Cities* **2013**, *35*, 114–124. [[CrossRef](#)]
20. Gan, Z.; Yang, M.; Feng, T.; Timmermans, H. Understanding urban mobility patterns from a spatiotemporal perspective: Daily ridership profiles of metro stations. *Transportation* **2018**. [[CrossRef](#)]
21. Singh, N.; Vasudevan, V. Understanding school trip mode choice—The case of Kanpur (India). *J. Transp. Geogr.* **2018**, *66*, 283–290. [[CrossRef](#)]
22. Ross, T.; Buliung, R. A systematic review of disability's treatment in the active school travel and children's independent mobility literatures. *Transp. Rev.* **2018**, *38*, 349–371. [[CrossRef](#)]
23. Huang, X.; Tan, J. Understanding spatio-temporal mobility patterns for seniors, child/student and adult using smart card data. *Int. Arch. Photogramm. Remote. Sens. Spat. Inf. Sci.* **2014**, *XL-1*, 167–172. [[CrossRef](#)]
24. Gutiérrez, J.; Cardozo, O.D.; García-Palomares, J.C. Transit ridership forecasting at station level: An approach based on distance-decay weighted regression. *J. Transp. Geogr.* **2011**, *19*, 1081–1092. [[CrossRef](#)]
25. Kubly, M.; Barranda, A.; Upchurch, C. Factors influencing light-rail station boardings in the United States. *Transp. Res. Part A* **2004**, *38*, 223–247. [[CrossRef](#)]
26. Sohn, K.; Shim, H. Factors generating boardings at Metro stations in the Seoul metropolitan area. *Cities* **2010**, *27*, 358–368. [[CrossRef](#)]
27. Ji, Y.; Ma, X.; Yang, M.; Jin, Y.; Gao, L. Exploring Spatially Varying Influences on Metro-Bikeshare Transfer: A Geographically Weighted Poisson Regression Approach. *Sustainability* **2018**, *10*, 1526. [[CrossRef](#)]
28. Mcmillen, D.P. Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. *Am. J. Agric. Econ.* **2004**, *86*, 554–556. [[CrossRef](#)]
29. Huang, B.; Wu, B.; Barry, M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 383–401. [[CrossRef](#)]
30. Yang, J.; Bao, Y.; Zhang, Y.; Li, X.; Ge, Q. Impact of Accessibility on Housing Prices in Dalian City of China Based on a Geographically Weighted Regression Model. *Chin. Geogr. Sci.* **2018**, *28*, 505–515. [[CrossRef](#)]
31. Huang, Y.; Wang, X.; Patton, D. Examining spatial relationships between crashes and the built environment: A geographically weighted regression approach. *J. Transp. Geogr.* **2018**, *69*, 221–233. [[CrossRef](#)]
32. Mitra, R.; Buliung, R.N. The influence of neighborhood environment and household travel interactions on school travel behavior: An exploration using geographically-weighted models. *J. Transp. Geogr.* **2014**, *36*, 69–78. [[CrossRef](#)]
33. AsadKhattak, X. Role of travel information in supporting travel decision adaption: Exploring spatial patterns. *Transportmetrica* **2013**, *9*, 316–334.
34. Feuillet, T.; Commenges, H.; Menai, M.; Salze, P.; Perchoux, C.; Reuillon, R.; Kesse-Guyot, E.; Enaud, C.; Nazare, J.A.; Herberg, S. A massive geographically weighted regression model of walking-environment relationships. *J. Transp. Geogr.* **2018**, *68*, 118–129. [[CrossRef](#)]
35. Chow, L.F.; Zhao, F.; Liu, X.; Li, M.T.; Ubaka, I. Transit Ridership Model Based on Geographically Weighted Regression. *Transp. Res. Rec. J. Transp. Res. Board* **2006**, *1972*, 105–114. [[CrossRef](#)]
36. Cardozo, O.D.; García-Palomares, J.C.; Gutiérrez, J. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Appl. Geogr.* **2012**, *34*, 548–558. [[CrossRef](#)]
37. Chiou, Y.C.; Jou, R.C.; Yang, C.H. Factors affecting public transportation usage rate: Geographically weighted regression. *Transp. Res. Part A Policy Pract.* **2015**, *78*, 161–177. [[CrossRef](#)]

38. Tu, W.; Cao, R.; Yue, Y.; Zhou, B.; Li, Q.; Li, Q. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. *J. Transp. Geogr.* **2018**, *69*, 45–57. [[CrossRef](#)]
39. Ma, X.; Zhang, J.; Ding, C.; Wang, Y. A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership. *Comput. Environ. Urban Syst.* **2018**, *70*, 113–124. [[CrossRef](#)]
40. Bureau, N.S. *Nanjing National Economic and Social Development Statistics Bulletin*; Nanjing Statistics Bureau: Nanjing, China, 2015.
41. Government, T.C. Notice on Adjusting the Standards for Categorizing City Sizes. Available online: http://english.gov.cn/policies/latest_releases/2014/11/25/content_281475015213546.htm. (accessed on 20 November 2014).
42. Long, Y.; Thill, J.C. Combining smart card data and household travel survey to analyze jobs–housing relationships in Beijing. *Comput. Environ. Urban Syst.* **2013**, *53*, 19–35. [[CrossRef](#)]
43. Le, M.K.; Bhaskar, A.; Chung, E. Passenger Segmentation Using Smart Card Data. *IEEE Trans. Intell. Transp. Syst.* **2015**, *16*, 1537–1548.
44. Medina, S.A.O. Inferring weekly primary activity patterns using public transport smart card data and a household travel survey. *Travel Behav. Soc.* **2016**, *12*, 93–101. [[CrossRef](#)]
45. Xia, J.; Nesbitt, J.; Daley, R.; Najnin, A.; Litman, T.; Tiwari, S.P. A multi-dimensional view of transport-related social exclusion: A comparative study of Greater Perth and Sydney. *Transp. Res. Part A Policy Pract.* **2016**, *94*, 205–221. [[CrossRef](#)]
46. Nakaya, T.; Fotheringham, A.C.; Charlton, M. Geographically weighted Poisson regression for disease association mapping. *Stat. Med.* **2010**, *24*, 2695–2717. [[CrossRef](#)] [[PubMed](#)]
47. Wang, Y.; Chau, C.K.; Ng, W.Y.; Leung, T.M. A review on the effects of physical built environment attributes on enhancing walking and cycling activity levels within residential neighborhoods. *Cities* **2016**, *50*, 1–15. [[CrossRef](#)]
48. McDonald, N.C. Household interactions and children’s school travel: The effect of parental work patterns on walking and biking to school. *J. Transp. Geogr.* **2008**, *16*, 324–331. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).