International Journal of
Geo-Information

MDPI

Article

Interactions between Bus, Metro, and Taxi Use before and after the Chinese Spring Festival

Jianwei Huang [1], Xintao Liu [1,*], Pengxiang Zhao [2], Junwei Zhang [1] and Mei-Po Kwan [3,4]

1 Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China; jianwei.huang@polyu.edu.hk (J.H.); lsjacob.zhang@connect.polyu.hk (J.Z.)
2 Institute of Cartography and Geoinformation, ETH Zurich, 8093 Zurich, Switzerland; pezhao@ethz.ch
3 Department of Geography and Resource Management, and Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China; mpk654@gmail.com
4 Department of Human Geography and Spatial Planning, Utrecht University, 3584 CB Utrecht, The Netherlands
* Correspondence: xintao.liu@polyu.edu.hk

Received: 28 August 2019; Accepted: 7 October 2019; Published: 10 October 2019

Abstract: Public transport plays an important role in developing sustainable cities. A better understanding of how different public transit modes (bus, metro, and taxi) interact with each other will provide better sustainable strategies to transport and urban planners. However, most existing studies are either limited to small-scale surveys or focused on the identification of general interaction patterns during times of regular traffic. Transient demographic changes in a city (i.e., many people moving out and in) can lead to significant changes in such interaction patterns and provide a useful context for better investigating the changes in these patterns. Despite that, little has been done to explore how such interaction patterns change and how they are linked to the built environment from the perspective of transient demographic changes using urban big data. In this paper, the tap-in-tap-out smart card data of bus/metro and taxi GPS trajectory data before and after the Chinese Spring Festival in Shenzhen, China, are used to explore such interaction patterns. A time-series clustering method and an elasticity change index (ECI) are adopted to detect the changing transit mode patterns and the underlying dynamics. The findings indicate that the interactions between different transit modes vary over space and time and are competitive or complementary in different parts of the city. Both ordinary least-squares (OLS) and geographically weighted regression (GWR) models with built environment variables are used to reveal the impact of changes in different transit modes on ECIs and their linkage with the built environment. The results of this study will contribute to the planning and design of multi-modal transport services.

Keywords: smart card data; taxi GPS trajectories; public transport modes; travel behavior; built environment

1. Introduction

As urbanization accelerates, urban and transportation planners have attempted to improve the sustainability of cities from the perspectives of urban design and urban transportation [1,2]. Especially, multi-modal transportation has already been considered as one of the most effective sustainable strategies for future urban mobility in recent years, as it provides an alternative to the private car through combinations of multiple travel modes such as bus, metro, and taxi [3]. To achieve the goal, one of the critical research topics is to understand the interaction patterns of multi-modal transport modes (i.e., how one type of travel mode competes with or complements other modes in space and time in a city). A better understanding of the interaction patterns of multiple transport modes would shed light on how burgeoning multi-modal transport services are likely to perform, thus potentially assisting urban transportation planning and management departments to better design sustainable multi-modal transportation systems.

A growing body of research regards the study of multi-modal transportation from the perspective of passengers' travel behavior using urban big data [4–6]. Passengers' travel behavior, however, is complex and entails travel modes preference and interactions between different transit modes (taxi, bus, and metro) [7]. Some studies have indicated that the interplay between different transport modes may vary over space and time [8]. It is thus worthwhile to investigate how individuals travel by different transport modes at different times of the day in a city and how this can be empirically demonstrated using urban big data [9]. Most existing studies, however, are either limited to small-scale survey data or focused on the identification of general patterns. Little has been done to explore such interplay patterns from the perspective of transient demographic changes and their linkage with the built environment using big data. Spring Festival, a special event in China each year, causes large-scale inter-urban human movement [10]. A large number of people return to their hometowns for a family reunion before the Spring Festival (a week-long holiday for Chinese New Year) and then go back to the cities where they live after the festival. This phenomenon is particularly evident in Shenzhen, China, which is a young city with a large number of migrant workers. It is still unclear whether and how the large-scale human movement and transient demographic change associated with the Spring Festival influence the interaction patterns of multi-modal transport. In addition, a better understanding of urban multi-modal transportation systems and their relationship with human travel behavior and the urban built environment will potentially facilitate the design and planning of sustainable transportation systems [11].

In this paper, we use smart card data and GPS taxi trajectory data collected in Shenzhen, China, to investigate the patterns of the interplay between bus, metro, and taxi in the context of the transient demographic change around the Spring Festival. The research questions are as follows: (1) How to better evaluate the interaction patterns of multiple transport modes on the basis of the changes in travel patterns caused by transient demographic changes associated with the Spring Festival in a city? (2) How are such patterns related to the urban built environment? To answer these two questions, we first apply a clustering-based method to identify the interaction patterns of multiple transport modes and compare them for two specific periods between which significant demographic change occurred (i.e., before and after the Chinese Spring Festival). Then, we adopt an elasticity change index (ECI) to measure the patterns of transit modes during the Spring Festival and normal periods. Lastly, the relationships between the ECI and the built environment are examined on the basis of both ordinary least-squares (OLS) regression and geographically weighted regression (GWR). The findings reveal that the interaction patterns of the public transit modes are dynamic, and transportation during the Spring Festival affects these patterns and their relationship with the built environment. These results shed light on the planning and design of multi-modal transportation services.

The rest of this paper is structured as follows. In Section 2, we review previous and related work. The methodology of this study is presented in Section 3, including the study area, data preprocessing, as well as methods for inferring changes in the interaction patterns and their linkage to the built environment. In Section 4, the results are presented, including transit behavior changes, the spatial distribution of the ECI, and the relationship between dynamic transit behavior patterns and the built environment. We present our conclusions in the final section and suggest directions for future work.

2. Related Work

2.1. Travel Behaviors in Urban Space

Considerable recent research has indicated that the functionality of distinct urban spaces can be revealed by temporal variations in people's travel behaviors [12–14]. The idea is that urban spaces with similar functions tend to generate similar temporal patterns of travel behaviors [15]. Owing to technological developments, individuals' data can be collected (e.g., GPS trajectories, call details of phone-based communication, smart card data, and social media) and analyzed to investigate the relationship between the spatial distribution of urban functions and travel behaviors. The general analytic framework of this research can be summarized as follows. First, aggregate all the temporal frequency of travel behaviors in a predefined temporal granularity (such as hours of a weekday or a weekend day) into one specific geographic unit (traffic analysis zones [TAZs] or grids). Second, classify urban spaces into several categories using clustering algorithms by measuring the similarity and dissimilarity among different geographic units. This framework has been introduced and developed in much research. For instance, taxi trajectory data and smart card data have been widely used for understanding human mobility patterns and urban dynamics [16–18]. Meanwhile, data sources (e.g., cell phones and social media) with more social attributes are also used to understand the relationship between built environment and travel behavior [19–21]. Despite the obvious relationship between the functionality of urban space and travel behavior, it is generally agreed that one single type of travel data may not be enough to understand such a complex phenomenon.

Urban spaces with different functions normally have different socioeconomic and demographic characteristics, which will also affect travel mode choices to some extent [22]. Meanwhile, the socioeconomic and demographic characteristics in a specific urban space might be influenced by the presence of a large number of migrant workers as a result of the urbanization process [23]. This is so because newcomers with different cultural backgrounds not only would have new work and lifestyles but also may have different travel mode choice patterns [24]. For instance, Blumenberg suggests that migrant workers have contributed to increased travel across all transport modes and a shift in the demographic composition of transit riders [25]. Hu suggests that the travel behavior of migrant workers might be influenced by their birthplace and cultural backgrounds [26]. Moreover, research also points out that the rapid growth of urban migrant workers might increase urbanization and also result in a shift in travel behaviors [27]. How the presence of a large number of migrant workers in a city influences travel behavior patterns is, hence, an important issue. However, existing works about this issue often rely on survey data that have limited spatial and temporal granularity due to the cost and labor intensiveness of collecting these data. Emerging fine-grained spatiotemporal mobility data open an avenue for investigating this problem. Meanwhile, a comparison between special and normal times is required when examining the influence of the presence of a large number of migrant workers on travel behaviors.

2.2. The Relationship between Transit Demands and the Built Environment

As this introduction reveals, the built environment in past studies was largely summarized in terms of density, diversity, and design (also known as the "three Ds"), and a number of studies have suggested that the three Ds are strongly associated with people's travel behavior at different temporal scales through various mechanisms [28,29]. For example, 'density', which is measured with respect to population, employment, or facilities, not only encourages people to walk and take public transit but also reduces private car dependency [30,31]. 'Diversity' might reduce the need for long-distance travel and increase the need for short-distance travel due to better accessibility to a variety of activity opportunities [32]. 'Design' could entice people to leave their car and enjoy the neighborhood and urban space [33,34]. Thus, to reduce automobile use and encourage public transit use, urban space can be re-designed with a focus on the density of population, employment, and facilities, as well as on the diversity of activity opportunities and land use. Indeed, such an urban design principle is also the basis of compact or transit-oriented cities [35], and reducing private automobile use and encouraging the use of public transit can be an effective means for achieving sustainable urban development. However, redesigning urban space is usually a time- and resource-consuming process, especially when considering a large number of migrant workers who moved into urban areas from rural ones [36]. As Section 2.1 above proposed, the considerable flows of migrant workers can be used to reconstruct travel behavior patterns. Meanwhile, it is still an open question whether the change in transit mode patterns due to the influx of migrant workers would still have a strong association with the built environment.

3. Methodology

3.1. Study Area and Data Preprocessing

As one of the youngest, largest, and most vibrant migrant cities in China [37], Shenzhen is chosen as the study area in this research. It covers a land area of 1997.27 km^2 and had a total permanent population of approximately 12 million in 2017 (Shenzhen Statistics Department, 2017). It was one of China's early special economic zones (SEZs) and is now well known for its high technology and innovative industries. Currently, Shenzhen is one of the most populated and fastest-growing cities in China. As shown in Figure 1, Shenzhen consists of 10 districts (Luohu, Futian, Nanshan, Yantian, Baoan, Longgang, Guangming, Longhua, Pingshan, and Dapeng). The data utilized in this study include smart card data (SCD) of the metro/bus service, GPS trajectories of taxi/bus, and data from TAZs and road network and points of interest (POI) datasets in Shenzhen around the Chinese Spring Festival of 2017 (from 27 January to 11 February 2017). Each item in the SCD file is a tap-in-tap-out record for a metro station or a bus stop, which is associated with a masked user ID, a timestamp, and a metro station ID or a bus line ID. It is straightforward to identify where a metro user gets on or off by matching the station ID with a station map.

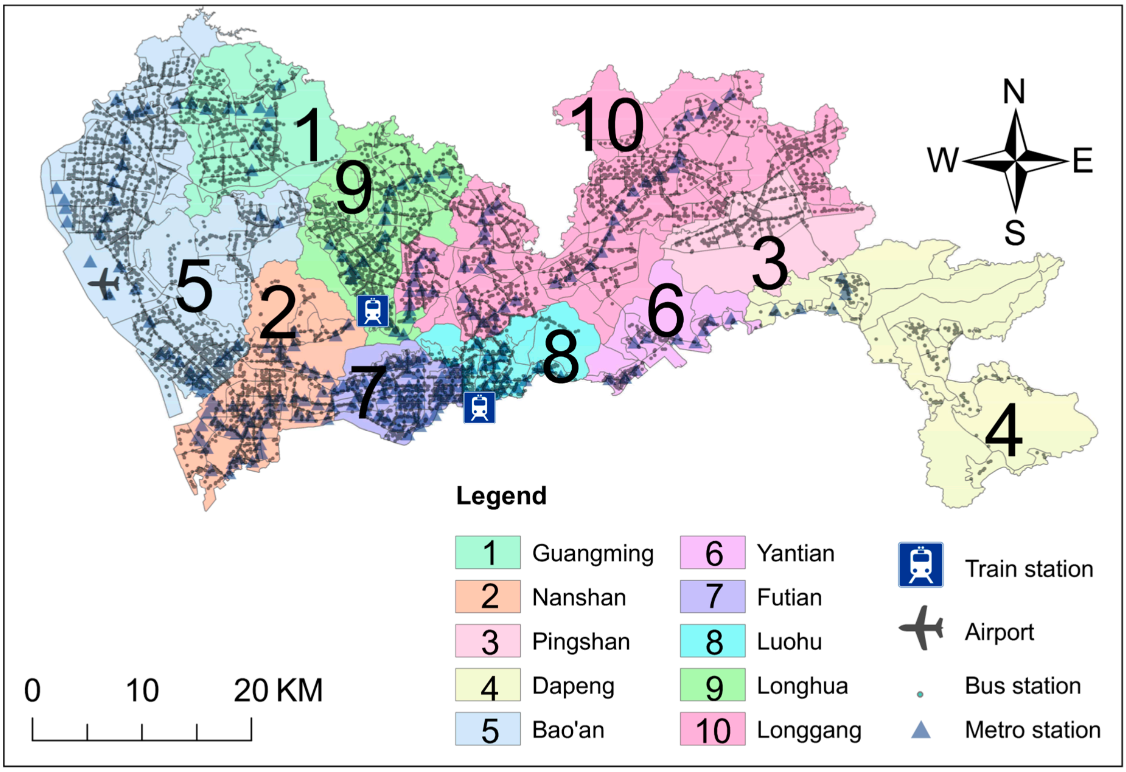

Figure 1. Study area.

However, the SCD data in Shenzhen for bus passengers only record the bus line ID when the user gets on. To solve this problem, we used a time-station matching process, which mainly applied cross-referencing of the bus line ID, timestamp, and GPS timestamp of the bus with the minimum time difference to match the GPS trajectories of buses with the SCD data and infer when and where the bus passenger gets on the bus. The taxi GPS trajectories are in a general format, including a taxi ID, a timestamp, latitude, longitude, occupant status, and so on. Taxi ridership was extracted using the method developed in a previous study [38].

After processing all the SCD and trajectory data, three travel modes (taxi, bus, and metro) were obtained. According to the annual report of the Transport Commission of Shenzhen Municipality (TCSM), for 2017, Shenzhen Spring Festival travel started on 13 January and ended on 22 February 2017. The peaks of the transient migration happened on 21 January and 5 February, which were one week before and one week after the 2017 Spring Festival (27 January). Besides, there is a one-week official holiday for the Chinese Spring Festival, which means that people had to be back to Shenzhen to work on 5 February 2017, and the transport mode pattern would return to normal after this day. To avoid the peaks of the 2017 Spring Festival migration, this study selected the periods 27–30 January as the Spring Festival time and 8–11 February as a normal time for comparative purpose. Figure 2 shows the volume of the three transport modes over different periods in one-hour granularity. It shows that the interactions between bus, metro, and taxi use were quite different between these two periods. For instance, Figure 2 indicates that the volume of ridership in the Spring Festival was around 70% less than that in the normal period, and the travel patterns were also different (e.g., the peak hour appeared at different times).

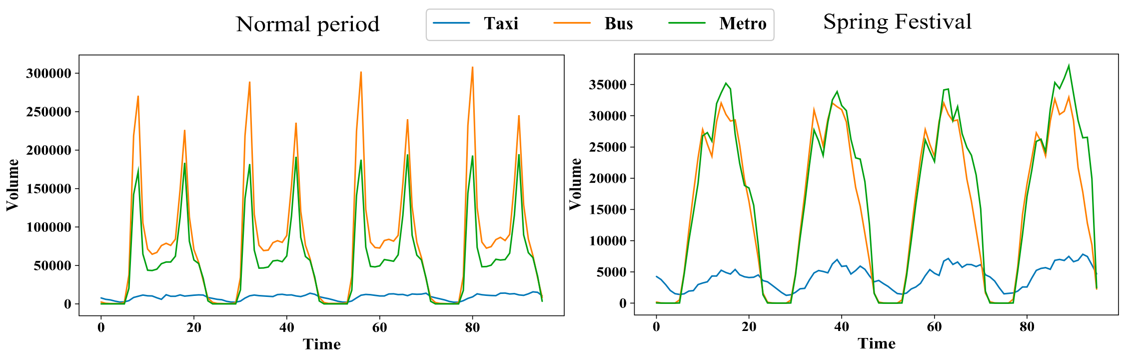

Figure 2. The hourly volume of the three ridership modes over different periods.

Additionally, the basic areal or geographic unit this work used is the TAZ, which is the areal unit most commonly used in conventional transportation planning models. The TAZ dataset was obtained from the Shenzhen transportation committee and consisted of 491 TAZs that cover the entire area of Shenzhen. To characterize the built environment, this work also used a road network dataset and a POI dataset, which were downloaded from OpenStreetMap (OSM) and Amap.

3.2. Deriving the Interaction Patterns of Transit Modes

In this paper, the interaction patterns between different public transit modes (bus, metro, and taxi) were identified on the basis of ridership. A method introduced in the literature was used to measure the spatiotemporal similarities of the patterns of different transit modes usage [9]. Figure 3 shows the workflow of the method. First, we constructed a time series for each TAZ based on the volume of ridership for each transit mode in one-hour granularity and the ratio of different transit ridership, which is the share for each transit mode of the total volume of ridership. Second, we quantified the similarity between different TAZs on the basis of their ratio of different transit ridership. Third, we used a spectral clustering algorithm to categorize the TAZs into different groups before and after the Chinese Spring Festival. Finally, changes in transit mode patterns were derived by comparing the results of the clusters obtained for different periods.

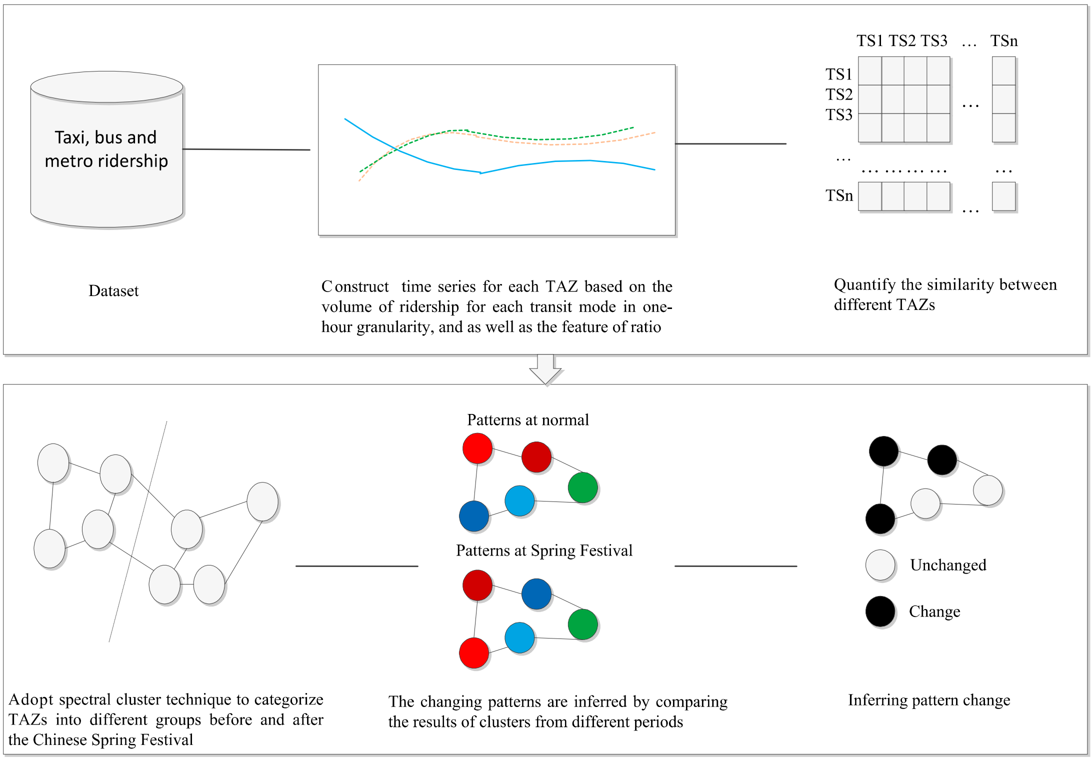

Figure 3. Method to identify multi-modal transit patterns. TAZs: traffic analysis zones.

For each TAZ, we computed three variables to depict it. Specifically, we calculated the volume of ridership for each transit mode in one-hour granularity as a time series during the Spring Festival period and the selected normal period. As introduced above, we took the four-day dataset for each period and denoted the volume of ridership for each transit mode for TAZ i as $\left[V_i^1, V_i^2, \ldots V_i^{96}\right]$ in different periods. For comparing the three transit modes, we determined three ridership volumes of TAZ i $\left\{V_i^{bus}, V_i^{metro}, V_i^{taxi}\right\}$, and the ratios of different transit ridership $\left\{R_i^{bus}, R_i^{metro}, R_i^{taxi}\right\}$:

$$R_i^{mode} = \frac{V_i^{mode}}{V_i^{bus} + V_i^{metro} + V_i^{taxi}} \tag{1}$$

where V_i^m presents the volume of ridership in TAZ i in the time slot m for each transit mode, V_i^{mode} denotes the volume of ridership for bus, metro, or taxi, and R_i^{mode} represents the ratio of ridership volume for bus, metro, or taxi.

Then, we quantified the similarity between different TAZs on the basis of their estimated autocorrelation coefficients. Autocorrelation is a measure of how well a time series matches a time-shifted version of itself, which is also referred to as a lag. The autocorrelation coefficient is widely used to measure the similarity of time series because of its advantage in representing the dependence structure of a time series [39]. Given a time series $T = \{t_1, t_2, t_3 \ldots t_n\}$, its autocorrelation coefficient at lag r is estimated as:

$$p_r = \frac{\sum_{i=r+1}^{n} (t_i - \bar{t})\left(t_{i-r} - \bar{t}\right)}{\sum_{i=r+1}^{n} \left(t_i - \bar{t}\right)^2} \tag{2}$$

where $\bar{t}$ is the mean value of the input feature.

Given a TAZ i and a type of travel mode, the autocorrelation coefficients of the ratio of different transit ridership at r lag is presented as $\{p_{i,r}^{bus}, p_{i,r}^{metro}, p_{i,r}^{taxi}\}$, and the distance between a pair of TAZs is estimated as:

$$Sim_{i,j} = d_{i,j}^2 = \sum_{r=1}^{n} \left(p_{i,r} - p_{j,r}\right)^2 \quad (3)$$

where $d_{i,j}^2$ denotes the squared Euclidean distance, and $Sim_{i,j}$ is the similarity between TAZs i and j. To compare the similarities for distinct transit modes between different TAZs, we obtained the similarity $\{Sim_{i,j}^{bus}, Sim_{i,j}^{metro}, Sim_{i,j}^{taxi}\}$ for each pair of TAZs.

By measuring the similarities for distinct transit modes between different TAZs, the similarity matrices of different transit modes M_{bus}, M_{metro}, and M_{taxi} were obtained. Considering the proportions of each transit modes as a share of total ridership are different between the Spring Festival period and normal period, the integrated similarity matrices were computed as:

$$M_{sim} = W_{bus} \times M_{bus} + W_{metro} \times M_{metro} + W_{taxi} \times M_{taxi} \quad (4)$$

where W_{bus}, W_{metro} and W_{taxi} are proportions of bus, metro, and taxi ridership as shares of total ridership in the Spring Festival and normal periods. By applying this method, we obtained matrices of different periods to measure the similarity of different TAZs regarding the usage of bus, metro, and taxi in different periods.

Lastly, given the similarity matrices M_{sim}, we used spectral clustering to process the similarity matrices [40]. The Calinski–Harabasz Index (CH-index) was used in this work to determine the number of clusters. The CH-index is known as the variance ratio criterion and is applied to evaluate how well a dataset is separated quantitatively, and the maximum CH-index corresponds to the optimal number of clusters [41]. Given a set $X = \{x_1, x_2, x_3, \ldots x_N\}$ for K clusters in the set, C_i is the ith cluster, k_i is the number of elements in the cluster C_i, the CH-index score s is given as the ratio of the between-and within-cluster dispersion; the formula is as follow:

$$s(k) = \frac{Tr(B_k)}{Tr(W_k)} * \frac{N-k}{k-1} \quad (5)$$

where B_k is the between-group dispersion matrix, and W_k is the within-cluster dispersion matrix defined by:

$$W_k = \sum_{i=1}^{N} ||x_i - C_{pi}||^2 \quad (6)$$

$$B_k = \sum_{i=1}^{k} ||C_{pi} - \overline{X}||^2 \quad (7)$$

where x_i represents the ith point, C_{pi} indicates the centroid of the pth cluster, $\overline{X}$ is the center of the whole dataset, n_i is the number of points in the pth cluster. The changes in the patterns are derived by comparing the results of the clusters from different periods.

3.3. Measuring the Patterns of Transit Modes

As discussed in Section 2, travel demand refers to the amount and transportation mode that people would choose under specific conditions. Previous studies demonstrated that travel demand is influenced by various factors, including demographics, economic activity, land use patterns, the built environment, and so on [42]. Meanwhile, existing studies also indicate that the elasticity of demand is a good measure of the changes in demand for a good or service when other factors (e.g., its price) change [43]. The main idea of elasticity is that when a factor (like price) increases or decreases, the demand tends to increases or decrease, and vice versa. Note that the relationship between travel demand change and the change (influx or outflow) in migrant workers that leads to such demand change can be measured on the basis of the 'elasticity' concept. Hence, this work proposes an elasticity change index (ECI) to measure the change in the volume of ridership for a particular transit mode in response to a change in the number of migrant workers during or after the Spring Festival. The index can be expressed by the following formula:

$$ECI_i^{mode} = \sum_{t=1}^{n} \left| \frac{\Delta V_i^{mode}}{\Delta R_i} \right| \quad (8)$$

A change in the transit mode index of one specific transit mode of TAZ i is denoted ECI_i^{mode}. ΔR_i represents the increase or decrease in ridership for all transit modes between the Spring Festival period and normal period, and ΔV_i^{mode} is the increase or decrease in ridership for one specific transit mode between the Spring Festival period and the normal period. Normally, a higher ECI indicates a larger change in the proportion of the change in ridership for a specific transit mode with respect to the changes in ridership for all transit modes.

3.4. Correlating the ECI with Built Environment Characteristics

As mentioned in Section 2.2, the built environment can be summarized on the basis of the three Ds (density, diversity, and design) [44,45]. Meanwhile, research has also demonstrated that spatial heterogeneity and travel behavior are closely intertwined in the complicated geographic world. The former indicates that different places in a city tend exhibit heterogeneous spatial attributes [46]. In order to uncover the relationship between the ECI and the built environment, we used a GWR model to relate the ECI with built environment measures (i.e., three Ds) at the TAZ level. In this paper, we used seven variables to represent different aspects of the three Ds (see Table 1 for the seven selected variables).

Table 1. Explanations of the abbreviations for the built environment variables. POI: points of interest.

Variables	Abbreviations	Description
Residential density	RoD	The number of residence POI in each TAZ
Workplace density	WoD	The number of workplace POI in each TAZ
Land use diversity	LUD	The characteristics of the land use patterns in each TAZ
Road density	Rod	Street length in each TAZ
Avg. distance between the nearest transit stop and home	DsH	The average street distance between residences and transit stops in each TAZ
Avg. distance between the nearest transit stop and work	DsW	The average street distance between workplaces and transit stops in each TAZ
Density of transit stop	SoD	The number of transit stops in each TAZ

Residential and workplace densities represent the characteristics of density. Accessibility and design are measured on the basis of the average street distance from residence and workplace to transit stops (e.g., metro and bus), the density of transit stops, and road density. In this paper, residence and workplace densities are defined as the number of residential and workplace POIs in each TAZ; road density is defined as street length per unit area contained in each TAZ (kilometer/square kilometer). The average street distance between residence/workplaces and transit stops per TAZ was estimated by matching each POI in the road network dataset and computing the distance between matched points (km). Diversity was measured by land use diversity (LUD), on the basis of the notion and method of entropy. The proportion of eight POI types (i.e., workplace, residences, transport facilities, shopping malls, restaurants, recreational and entertainment places, schools, hospitals, and hotels) was considered in the calculation process. LUD also represents the characteristics of the land use patterns in each TAZ, and its formula is:

$$LUD = -\sum_{i=1}^{n} \frac{p_i * \ln(p_n)}{\ln(n)} \tag{9}$$

In this formula, p_i is the proportion of POIs in each TAZ, n is the number of POI types. On the basis of the built environment variables listed in Table 1, both OLS regression and GWR were applied in this work. We use the corrected Akaike information criterion (AICc) and adjusted R^2 to compare the results of these two models. The AICc is commonly applied to evaluate model quality, and the model with a lower AICc is a better model [47]. We first used an OLS regression model to explain the global relations between the dependent and independent variables. Multicollinearity of the independent variables was assessed by the variance inflation factor (VIF) index: a VIF greater than 10 indicates that multicollinearity exists [48]. The GWR model was proposed for detecting the spatial variations in the relationships between two geographic phenomena by allowing the parameters of a localized regression to be estimated [49]. In this work, the parameters of localized regressions were estimated using the adaptive Gaussian kernel with a fixed bandwidth which was computed using AICc. The model was estimated with GWR4 software [50]. The selected built environment variables were regarded as independent variables, whereas the calculated ECIs for different transit modes were the dependent variables.

4. Results

4.1. Change in the Patterns of Transit Modes between the Two Periods

In this section, we examine the change in the interaction patterns between bus, metro, and taxi in the two periods (Spring Festival and normal period). The TAZs in Shenzhen were divided into subgroups using spectral clustering. The CH-index was utilized to determine the number of clusters and evaluate how well the similarity matrices are separated quantitatively. Figure 4 shows the relations between the number of clusters and the CH-index for different periods. Note that the TAZs for both the Spring Festival and the normal period have a similar optimal number of clusters, which is 5.

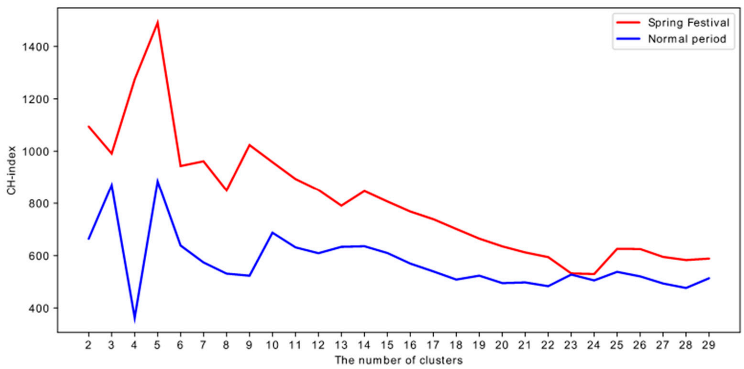

Figure 4. The relation between the number of clusters and the Calinski–Harabasz Index (CH-index).

The results are presented in Figure 5. Figure 5a,b show that there were five clusters of TAZs with distinct patterns of bus, metro, and taxi ridership in normal periods, and Figure 5c,d show such patterns for the 2017 Spring Festival. Under closer scrutiny and by comparing with the results of the literature [9], we identified seven types of distinct ridership or transit mode patterns (denoted as P1–P7) of bus, metro, and taxi between the two periods. All patterns were interpreted as shown in Table 2. In addition, to understand the dynamics of these patterns, we further compared the transit mode patterns during normal periods and the Spring Festival, as shown in Figure 6.

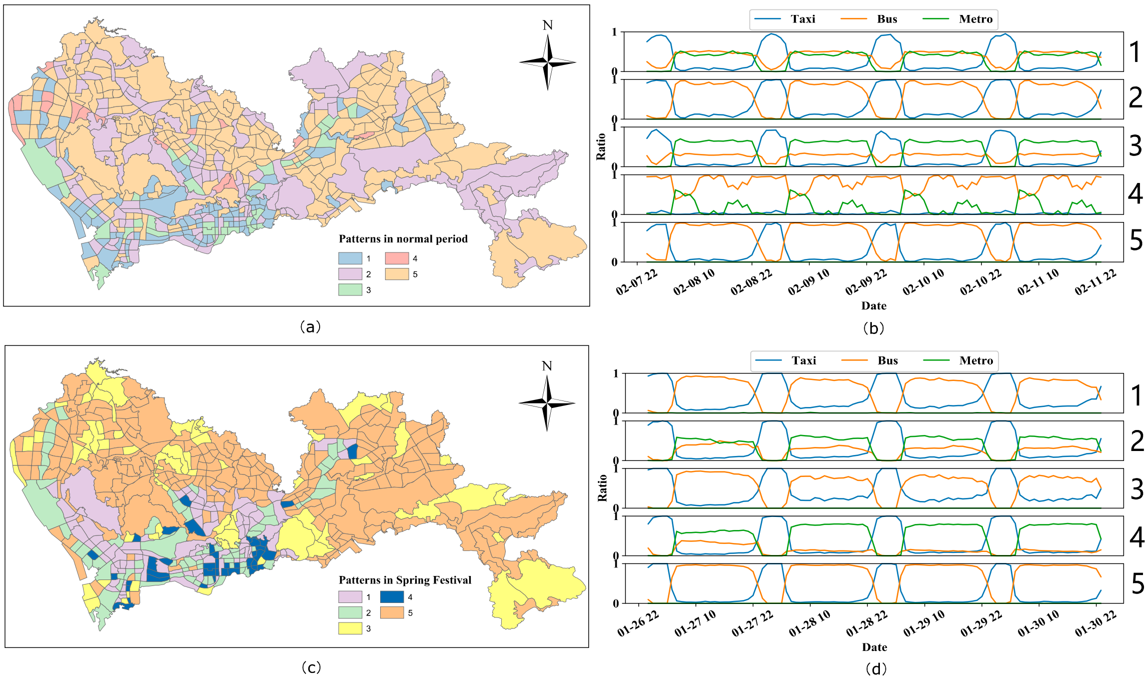

Figure 5. Clusters of TAZs with distinct ridership patterns of public transit modes in normal periods and the Spring Festival: (**a**) spatial distribution of the TAZs in five clusters in the normal period; (**b**) ratio of different transit ridership in the normal period; (**c**) spatial distribution of the TAZs in five clusters in the Spring Festival; (**d**) ratio of different transit ridership in the Spring Festival.

Table 2. Descriptive transit mode patterns on the five clusters in the different periods.

	Normal Period	Abbr.	Spring Festival	Abbr.
Cluster 1	Bus and metro dominate, while taxi complements.	P1	Bus dominates, while taxi complements.	P2
Cluster 2	Bus dominates, while taxi complements.	P2	Bus and metro are critical competitive, while taxi complements.	P3
Cluster 3	Metro dominates, while bus complements.	P3	Bus dominates, while taxi is strongly competitive.	P6
Cluster 4	Bus and metro arecritical competitive.	P4	Metro dominates, while taxi and bus complement.	P7
Cluster 5	Bus dominates.	P5	Bus dominates.	P5

The results revealed that most passengers relied on the bus both in normal periods and during the Spring Festival (e.g., Cluster 5 in both periods). Meanwhile, the TAZs with metro as the dominant travel mode were more likely to be characterized by metro as the dominant travel mode before and after the Spring Festival (e.g., Cluster 3 in normal periods and Cluster 2 in the Spring Festival). Furthermore, the spatial distribution of the TAZs with bus as the dominant travel mode and taxi as complimentary travel mode tended to be dispersed in normal periods and concentrated in the downtown area in the Spring Festival (e.g., Cluster 2 in normal periods and Cluster 1 in the Spring Festival). Besides, the new patterns P6 and P7 (yellow and blue in Figure 5c) appeared in the Spring Festival, while P1 and P4 disappeared.

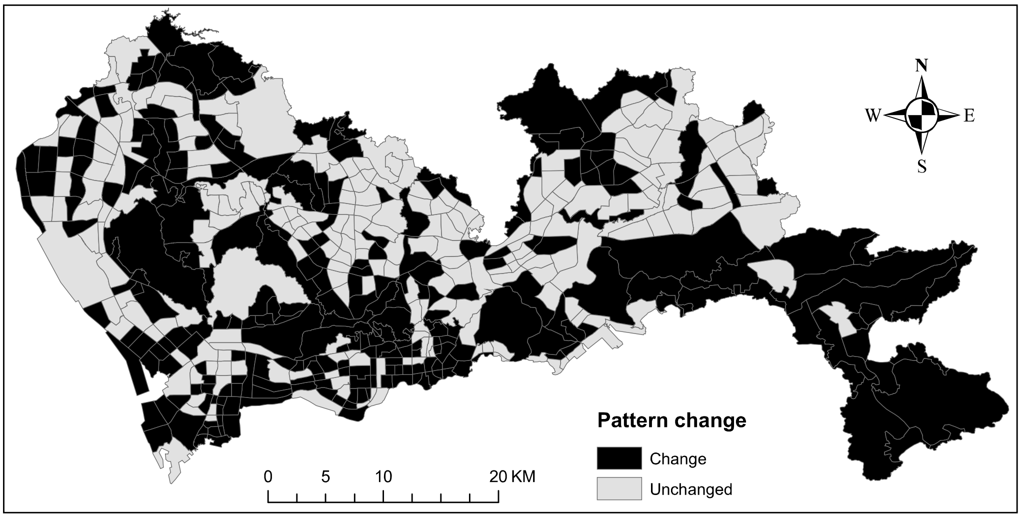

Figure 6. Spatial distribution of pattern changes in different periods.

We identified the following possible reasons for such kind of various interaction patterns in different periods: (a) Commuting travel greatly decreased in the Spring festival, which could bring dramatic change in the patterns (e.g., disappearance of P1 and P4 and appearance of P6 and P7); (b) At the same time, the outflow of migrant workers during the Spring Festival might reduce the congestion of public transport modes, and this would encourage passengers to take public transport (e.g., the increase in the number of TAZs with bus as the dominant travel mode); (c) The increase in the volume of family-based travel in the Spring Festival could stimulate the volume of taxi ridership (e.g., P1 and P6 in the Spring Festival).

4.2. Spatial Characteristics of the ECI

In this section, we used the ECI to measure the change in transit mode patterns associated with the influx/outflow of migrant workers between the Spring Festival and the normal period. Figure 7 shows the distribution of the ECI, the red color representing higher values of the ECI, and the blue color representing lower values. By using the natural breaks classification method, we divided the TAZs into five groups.

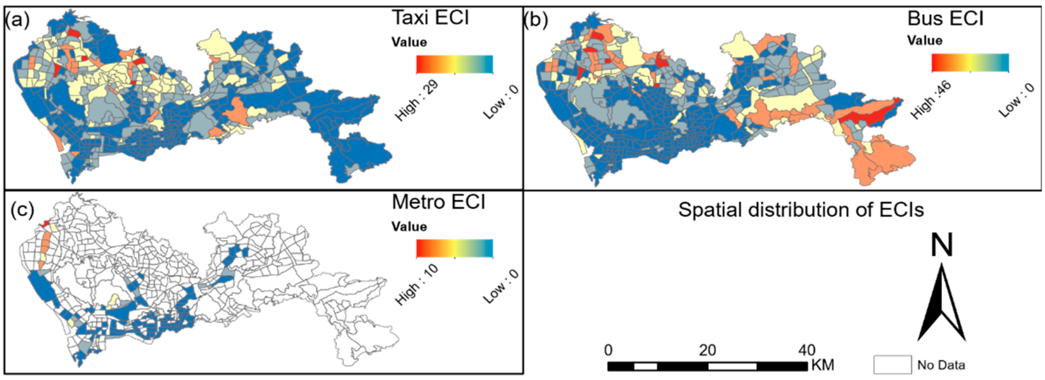

Figure 7. Spatial distribution of elasticity change index (ECI) for three transit modes. (**a**) ECI of taxi; (**b**) ECI of bus; (**c**) ECI of metro.

The results indicated that the ECI of bus and taxi were more sensitive to the influx/outflow of migrant workers between the Spring Festival and the normal period, while the opposite was observed for the ECI of metro. On the one hand, taxi and bus ridership were more scattered in the study area. Specifically, the Guangming, Longhua, Longgang, and Pingshan Districts had higher ECI values for taxi and bus. As Section 3.3 mentioned, the higher the ECI is, the more the transit mode shares change between the Spring Festival and the normal period. Hence, it can be inferred that these areas are suburbs with compact communities and industrial factories which have more complex demographics, because they provide numerous job opportunities for migrant workers who would return to their hometowns to celebrate the Spring Festival with their families. Meanwhile, the Dapeng District had a higher ECI value for bus and a lower ECI value for taxi, because Dapeng is a port area for international trade or a tourist area for vacation. The dominant transit mode in Dapeng in the normal periods was bus, and a dramatic change in transit mode happened during the Spring Festival, since a lot of people who live in Dapeng would return to their hometowns. On the other hand, the Baoan, Nanshan, Futian, and Luohu districts had the lowest ECI values for bus and taxi, because these areas are part of the central city with well-built transportation infrastructure, and travel in these areas is highly dependent on the metro.

4.3. The Relationship between ECI and Built Environment Characteristics

In this section, we discuss the relationship between ECI and built environment characteristics. First, the descriptive statistics for all the built environment variables are reported in Table 3. These built environment indicators were used as independent variables in subsequent regression analysis, and the ECI was used as the dependent variable. Both OLS regression and GWR were used to examine the relationship between the ECI and the built environment variables. Three OLS and GWR models were estimated to relate the ECI at the TAZ level to all the built environment variables. The results are presented in Table 4. The VIF results indicated that OLS estimations were not biased by multicollinearity. However, the OLS regression models explained only 12% with AICc 2784.242 for taxi ECI, 16% with AICc 2807.351 for bus ECI, and 8% with AICc 339.838 for metro ECI. Table 4 also summarizes the GWR results and shows that GWR was more suitable than the OLS models, because the GWR models could explain 35%, 41%, and 93% for taxi ECI, bus ECI, and metro ECI with lower AICc.

Table 3. Descriptive statistics of built environment variables.

Variable	Mean	Standard Deviation
Residential density	0.99	3.70
Workplace density	0.66	0.26
Land use diversity	0.09	0.20
Road density	1.19	1.47
Distance between the nearest transit stop and home	0.66	0.87
Distance between the nearest transit stop and work	0.37	0.26
Stop density	1.1945	4.3706

The OLS results also indicated that road density and transit stop density were negatively associated with the ECIs. It means that the interaction patterns between taxi, bus, and metro would more likely to be consistent in different periods if road density and transit stop density improved. Meanwhile, workplace density was positively associated with the ECIs. This suggests that the interaction patterns between taxi, bus, and metro would more likely change in different periods if workplace density increased. Besides, other built environment variables had a varied impact on the ECIs. For example, residential density and the distance between transit stops and the home had a positive impact on the taxi and bus ECIs, while they had a negative impact on the metro ECI. Land used diversity was negatively associated with the taxi and bus ECIs, while it was positively associated with the metro ECI. The distance between transit stops and workplaces had a negative effect on the taxi ECI, while it had a positive effect on the bus and metro ECIs.

Furthermore, the GWR results indicated that the impacts of all built environment variables on the ECIs were geographically heterogeneous. The estimated local coefficients of the three GWR models are presented in Figures 8–10. The local estimates of selected independent variables were mapped.

The red color represents large estimated coefficients, while the blue color represents low coefficients. Meanwhile, the white areas indicate TAZs with no related data. These figures show various patterns of spatial heterogeneity in the relationships between travel modes and the built environment. As the results in Figure 5 show, travel mode in TAZs with metro stations was dominated by the metro in the Spring Festival and the normal period, thus the metro continued to be the dominant travel mode. The results of the GWR models presented a similar picture, as Figures 7–9 show. The effects of the built environment on the ECIs of taxi and bus were non-significant in TAZs with metro stations. In addition, residential density, workplace density, land use density, and the distance between transit stations and home/work had different but significant impacts on all ECIs in different locations in Shenzhen.

Table 4. Results of ordinary least-squares (OLS)-based and geographically weighted regression (GWR)-based models.

OLS					GWR	
Parameter	**Estimated Value**	**Standard Error**	***p*-Value**	**AIF**	**Mean Estimated**	**Standard Error**
Taxi ECI						
Rod	0.19	0.06	0.001	1.9	0.50	0.56
Wod	0.03	0.01	0.004	1.9	0.71	0.44
LUD	−1.40	0.92	0.013	1.2	−0.49	0.81
Rod	−0.14	0.06	0.012	1.1	−0.36	0.93
DsH	0.18	0.41	0.165	1.1	0.10	0.36
DsW	−0.55	0.64	0.389	1.1	0.01	0.27
SoD	−0.17	0.04	0.000	1.4	−1.39	1.46
Adjusted R^2	0.12				0.35	
AICc	2784.242				2646.423	
Bus ECI						
Rod	0.41	0.07	0.000	1.9	0.18	1.18
Wod	0.03	0.02	0.011	1.9	0.04	0.60
LUD	−0.56	1.14	0.017	1.1	1.98	0.80
Rod	−0.36	0.07	0.000	1.1	−0.45	2.78
DsH	0.58	0.51	0.263	1.1	0.74	0.91
DsW	0.79	0.81	0.320	1.1	0.19	0.53
SoD	−0.29	1.14	0.000	1.1	−0.33	2.28
Adjusted R^2	0.16				0.41	
AICc	2807.351				2694.210	
Metro ECI						
Rod	−0.11	0.14	0.450	1.4	−0.39	2.57
Wod	0.01	0.02	0.628	1.4	0.21	1.02
LUD	0.65	0.63	0.307	1.0	−1.67	6.96
Rod	−0.03	0.03	0.270	1.1	−0.10	0.25
DsH	−0.92	0.51	0.077	1.1	2.93	16.51
DsW	1.08	0.41	0.010	1.1	−0.59	4.34
SoD	−0.21	0.09	0.024	1.1	−1.29	5.16
Adjusted R^2	0.08				0.93	
AICc	339.838				106.694	

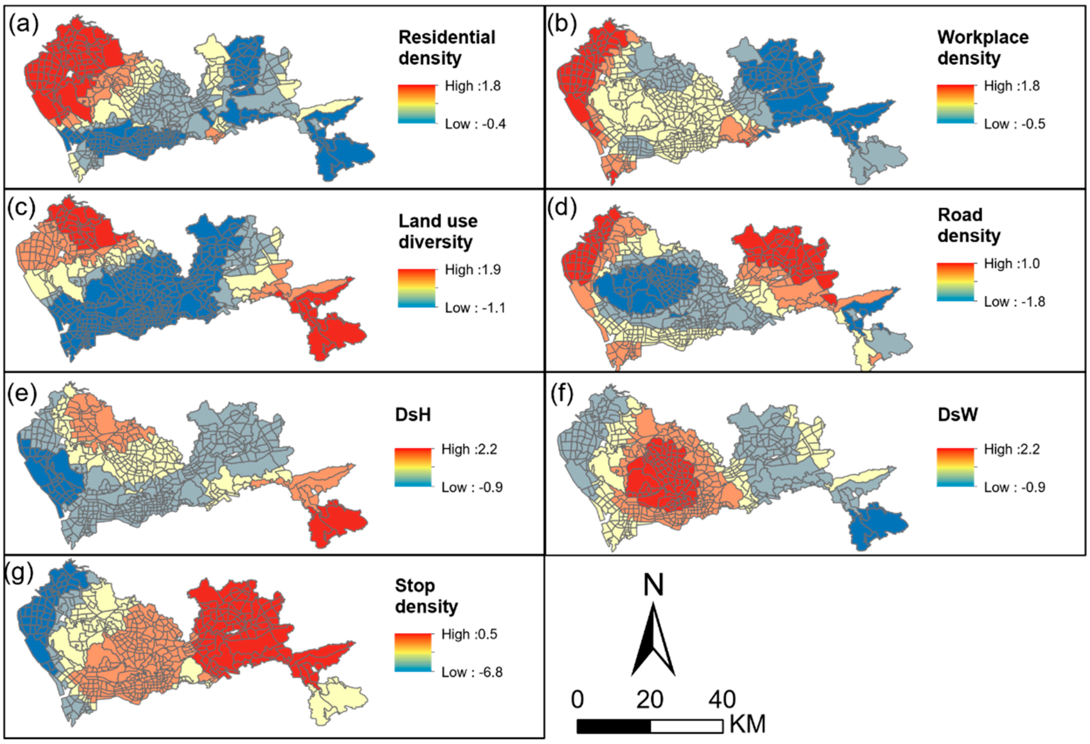

Figure 8. Local coefficients for taxi ECI. (**a**) Residential density; (**b**) workplace density; (**c**) land use diversity; (**d**) road density; (**e**) distance between the nearest transit stop and home (DsH); (**f**) distance between the nearest transit stop and work (DsW); (**g**) stop density.

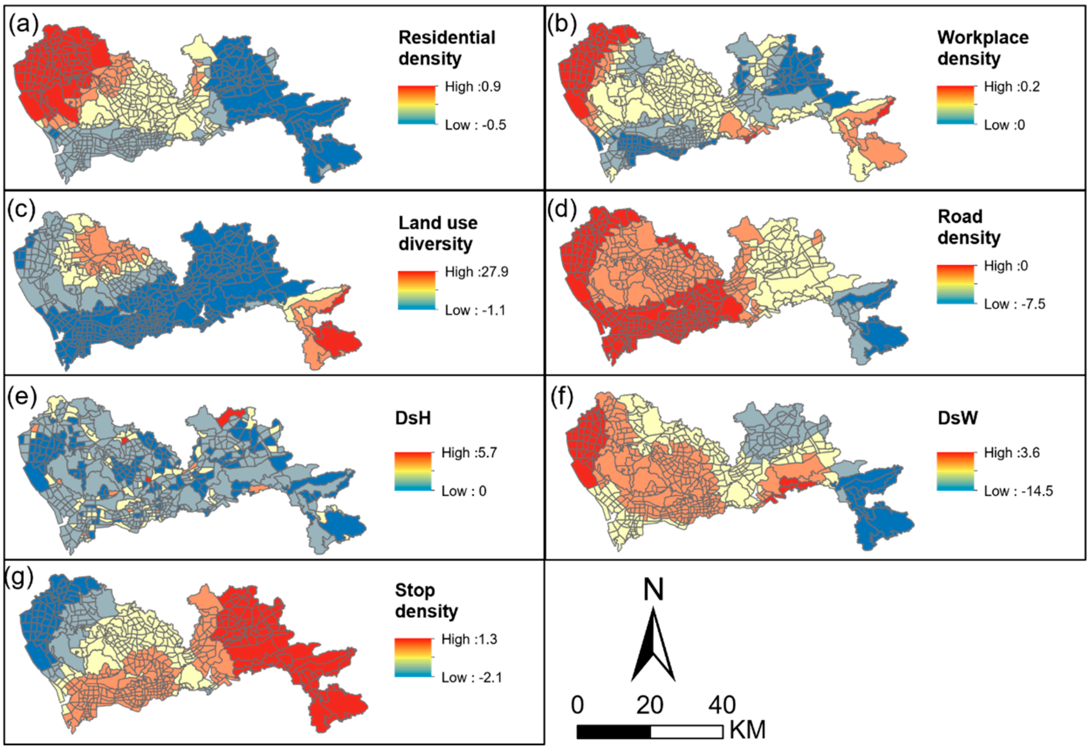

Figure 9. Local coefficients for bus ECI. (**a**) Residential density; (**b**) workplace density; (**c**) land use diversity; (**d**) road density; (**e**) distance between the nearest transit stop and home (DsH); (**f**) distance between the nearest transit stop and work (DsW); (**g**) stop density.

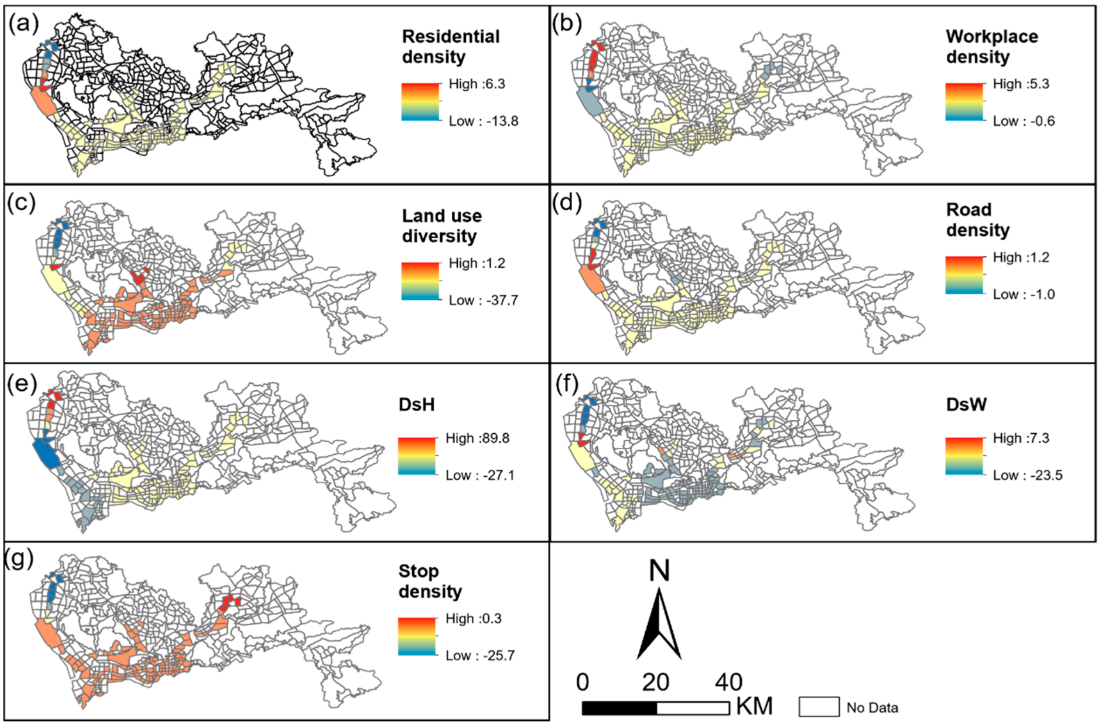

Figure 10. Local coefficients for metro ECI. (**a**) residential density; (**b**) workplace density; (**c**) land use diversity; (**d**) road density; (**e**) distance between the nearest transit stop and home (DsH); (**f**) distance between the nearest transit stop and work (DsW); (**g**) stop density.

5. Conclusions and Discussion

Understanding the interactions between different travel modes can shed new light on how to better design multi-modal transportation systems for building sustainable cities. This paper addresses two questions: (1) How does the interaction patterns of multiple transport modes vary over space and time on the basis of ridership? (2) How are these patterns related to the built environment? To answer the first question, we used a clustering-based method to identify the interaction patterns of multiple transport modes and compare them before and after the Chinese Spring Festival. For the second question, a change in transit mode index (ECI) was used to measure the change in travel mode patterns between the Spring Festival and the normal period. Meanwhile, both OLS and GWR models were used to examine the relationships between the ECI and the built environment. The contributions of this paper are mainly in the following aspects.

First, the spatiotemporal interaction patterns among bus, metro, and taxi were presented for a normal period and the Spring festival via a cluster-based method. Meanwhile, the geographic distribution of these changing patterns was also examined by comparing the results. This analysis can be helpful for multi-modal traffic management around the Spring Festival. For example, the TAZs in which the interaction patterns between transport modes changed dramatically between the two periods probably need to make suitable arrangements to cope with the different ratios of various transport modes, whose details can be determined according to the patterns in normal periods and the Spring Festival.

Second, we used a transit mode index (ECI) to measure the variations of the interaction patterns between the Spring Festival and the normal period. Normally, a higher ECI for a transit mode means a larger change between the Spring Festival and normal periods. The spatiotemporal characteristics of the ECIs were also explored, which reveal that the ECI of bus and taxi were more sensitive than the ECI of the metro. Besides, suburban areas with compact communities, industrial factories, international trade, and tourist areas also had higher ECI values. This is because these areas include complex demographics and provide numerous job opportunities for migrant workers who would return to their hometowns during the Spring Festival. Meanwhile, the area with the well-built environment located in the central city had lower ECI values, since people's travel mode in these areas is highly dependent on metro before and after the Spring Festival.

Third, OLS and GWR models were used to examine the associations between the ECIs and the built environment variables. The results showed that the GWR models had an adjusted R squared of 0.93 for the metro, 0.41 for bus, and 0.35 for taxi, and their lower AICc indicated that these models are more suitable than the OLS models. Moreover, the detailed effects of the built environment on the interaction patterns were discussed. These results contribute mainly to future urban planning rather than to current traffic management. For example, improving road density and transit stop density may reduce the pattern change in a TAZ, while increasing workplace density may have the opposite effect. Furthermore, the impacts of residential density and the distance between transit stops and home suggest that it is necessary to balance the effects of built environment factors in multi-modal transport planning.

Overall, this work reveals many interesting discoveries about the interaction patterns between bus, metro, and taxi use before and after the Chinese Spring Festival and their relationships with the built environment. The results enrich the existing literature concerning the interplay patterns between different public transport modes. Nevertheless, there are also several limitations in this work. Considering that the temporal coverage of existing datasets is limited, it is difficult to apply this analytic framework to other special periods with different cultural backgrounds (e.g., The National Day). It is also not possible to integrate broader transport modes (e.g., ride-sharing, walk, private car) into the analysis because of data limitations. Hence, future work should focus on investigating the interactions between various transport modes (e.g., all public transport modes, bicycling, ride-sharing, and private cars) in different periods and exploring their relationships with the built environment in the context of specific cultural backgrounds.

Author Contributions: Conceptualization, Jianwei Huang and Xintao Liu; Data processing and analysis, Jianwei Huang and Junwei Zhang; Method development and implementation: Jianwei Huang, Xintao Liu, Pengxiang Zhao; Writing—original draft, Jianwei Huang, Xintao Liu, Pengxiang Zhao, Junwei Zhang and Mei-Po Kwan; Writing—revising and editing, Jianwei Huang, Xintao Liu, Pengxiang Zhao, Junwei Zhang and Mei-Po Kwan.

Funding: The research was supported by grants from the Hong Kong Polytechnic University Start-up Research Fund Program (1-ZE6P). In addition, Mei-Po Kwan was supported by a grant from the National Natural Science Foundation of China (#41529101).

Acknowledgments: The authors thank the Shenzhen Institutes of Advanced Technology of the Chinese Academy of Sciences for the Smart Card data and taxi GPS trajectory data used in this study, and the anonymous reviewers, who helped to improve this work.

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

References

1. Bettencourt, L.M.; Lobo, J.; Helbing, D.; Kühnert, C.; West, G.B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 7301–7306. [CrossRef] [PubMed]
2. Kammen, D.M.; Sunter, D.A. City-integrated renewable energy for urban sustainability. *Science* **2016**, *352*, 922–928. [CrossRef] [PubMed]
3. Klinger, T. Moving from monomodality to multimodality? Changes in mode choice of new residents. *Transp. Res. Part A Policy Pract.* **2017**, *104*, 221–237. [CrossRef]
4. Kim, C.; Parent, O.; vom Hofe, R. The role of peer effects and the built environment on individual travel behavior. *Environ. Plan. B Urban Anal. City Sci.* **2018**, *45*, 452–469. [CrossRef]
5. Gao, Q.L.; Li, Q.Q.; Yue, Y.; Zhuang, Y.; Chen, Z.P.; Kong, H. Exploring changes in the spatial distribution of the low-to-moderate income group using transit smart card data. *Comput. Environ. Urban Syst.* **2018**, *72*, 68–77. [CrossRef]
6. Welch, T.F.; Gehrke, S.R.; Widita, A. Shared-use mobility competition: A trip-level analysis of taxi, bikeshare, and transit mode choice in Washington, DC. *Transp. A Transp. Sci.* **2018**, 1–13. [CrossRef]

7. Wang, F.; Ross, C.L. New potential for multimodal connection: Exploring the relationship between taxi and transit in New York City (NYC). *Transportation* **2017**, *46*, 1051–1072. [CrossRef]
8. Chen, C.; Gong, H.; Paaswell, R. Role of the built environment on mode choice decisions: Additional evidence on the impact of density. *Transportation* **2008**, *35*, 285–299. [CrossRef]
9. Yue, M.; Kang, C.; Andris, C.; Qin, K.; Liu, Y.; Meng, Q. Understanding the interplay between bus, metro, and cab ridership dynamics in Shenzhen, China. *Trans. Gis* **2018**, *22*, 855–871. [CrossRef]
10. Xu, J.; Li, A.; Li, D.; Liu, Y.; Du, Y.; Pei, T.; Ma, T.; Zhou, C. Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Appl. Geogr.* **2017**, *87*, 85–96. [CrossRef]
11. Limtanakool, N.; Dijst, M.; Schwanen, T. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. *J. Transp. Geogr.* **2006**, *14*, 327–341. [CrossRef]
12. Liu, Y.; Kang, C.; Gao, S.; Xiao, Y.; Tian, Y. Understanding intra-urban trip patterns from taxi trajectory data. *J. Geogr. Syst.* **2012**, *14*, 463–483. [CrossRef]
13. Gao, S. Spatio-temporal analytics for exploring human mobility patterns and urban dynamics in the mobile age. *Spat. Cogn. Comput.* **2015**, *15*, 86–114. [CrossRef]
14. Liu, X.; Kang, C.; Gong, L.; Liu, Y. Incorporating spatial interaction patterns in classifying and understanding urban land use. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 334–350. [CrossRef]
15. Grauwin, S.; Sobolevsky, S.; Moritz, S.; Gódor, I.; Ratti, C. Towards a comparative science of cities: Using mobile traffic records in New York, London, and Hong Kong. In *Computational Approaches for Urban Environments*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 363–387.
16. Goulet-Langlois, G.; Koutsopoulos, H.N.; Zhao, J. Inferring patterns in the multi-week activity sequences of public transport users. *Transp. Res. Part C Emerg. Technol.* **2016**, *64*, 1–16. [CrossRef]
17. Zhong, C.; Batty, M.; Manley, E.; Wang, J.; Wang, Z.; Chen, F.; Schmitt, G. Variability in regularity: Mining temporal mobility patterns in London, Singapore and Beijing using smart-card data. *PLoS ONE* **2016**, *11*, e0149222. [CrossRef] [PubMed]
18. Zhao, P.; Qin, K.; Ye, X.; Wang, Y.; Chen, Y. A trajectory clustering approach based on decision graph and data field for detecting hotspots. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1101–1127. [CrossRef]
19. Yuan, Y.; Raubal, M. Extracting dynamic urban mobility patterns from mobile phone data. In *International Conference on Geographic Information Science*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 354–367.
20. Xu, Y.; Belyi, A.; Bojic, I.; Ratti, C. Human mobility and socioeconomic status: Analysis of Singapore and Boston. *Comput. Environ. Urban Syst.* **2018**, *72*, 51–67. [CrossRef]
21. Yoshimura, Y.; Sobolevsky, S.; Bautista Hobin, J.N.; Ratti, C.; Blat, J. Urban association rules: Uncovering linked trips for shopping behavior. *Environ. Plan. B Urban Anal. City Sci.* **2018**, 367–385. [CrossRef]
22. Nasri, A.; Zhang, L. The analysis of transit-oriented development (TOD) in Washington, DC and Baltimore metropolitan areas. *Transp. Policy* **2014**, *32*, 172–179. [CrossRef]
23. Price, M.; Benton-Short, L. *Migrants to the Metropolis: The Rise of Immigrant Gateway Cities*; Syracuse University Press: New York, NY, USA, 2008.
24. Chatman, D.G.; Klein, N. Immigrants and travel demand in the United States: Implications for transportation policy and future research. *Public Work. Manag. Policy* **2009**, *13*, 312–327. [CrossRef]
25. Blumenberg, E. Moving in and moving around: Immigrants, travel behavior, and implications for transport policy. *Transp. Lett.* **2009**, *1*, 169–180. [CrossRef]
26. Hu, L. Changing travel behavior of Asian immigrants in the US. *Transp. Res. Part A Policy Pract.* **2017**, *106*, 248–260. [CrossRef]
27. Ewing, R.; Cervero, R. Travel and the built environment: A meta-analysis. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294. [CrossRef]
28. Zhang, L.; Hong, J.; Nasri, A.; Shen, Q. How built environment affects travel behavior: A comparative analysis of the connections between land use and vehicle miles traveled in US cities. *J. Transp. Land Use* **2012**, *5*, 40–52. [CrossRef]
29. Etminani-Ghasrodashti, R.; Ardeshiri, M. Modeling travel behavior by the structural relationships between lifestyle, built environment and non-working trips. *Transp. Res. Part A Policy Pract.* **2015**, *78*, 506–518. [CrossRef]
30. Stevens, M.R. Does compact development make people drive less? *J. Am. Plan. Assoc.* **2017**, *83*, 7–18. [CrossRef]

31. Sun, B.; Ermagun, A.; Dan, B. Built environmental impacts on commuting mode choice and distance: Evidence from Shanghai. *Transp. Res. Part D Transp. Environ.* **2017**, *52*, 441–453. [CrossRef]
32. Chen, H.; Jia, B.; Lau, S. Sustainable urban form for Chinese compact cities: Challenges of a rapid urbanized economy. *Habitat Int.* **2008**, *32*, 28–40. [CrossRef]
33. Reilly, M.; Landis, J. The influence of built-form and land use on mode choice: Evidence from the 1996 Bay Area Travel Survey. *IURD WP* **2002**, *4*.
34. Larrañaga, A.M.; Rizzi, L.I.; Arellana, J.; Strambi, O.; Cybis, H.B.B. The Influence of built environment and travel attitudes on walking: A case study of Porto Alegre, Brazil. *Int. J. Sustain. Transp.* **2016**, *10*, 332–342. [CrossRef]
35. Shum, K.; Watanabe, C. From Compact City to Smart City: A Sustainability Science & Synergy Perspective. *J. Environ. Sci. Eng. A* **2017**, *4*, 200–208.
36. Saunders, D. *Arrival City: How the Largest Migration in History Is Reshaping Our World*; Vintage: New York, NY, USA, 2011.
37. Tao, L.; Hui, E.C.; Wong, F.K.; Chen, T. Housing choices of migrant workers in China: Beyond the Hukou perspective. *Habitat Int.* **2015**, *49*, 474–483. [CrossRef]
38. Zhao, P.; Liu, X.; Shen, J.; Chen, M. A network distance and graph-partitioning-based clustering method for improving the accuracy of urban hotspot detection. *Geocarto Int.* **2017**, *34*, 293–315. [CrossRef]
39. D'Urso, P.; Maharaj, E.A. Autocorrelation-based fuzzy clustering of time series. *Fuzzy Sets Syst.* **2009**, *160*, 3565–3589. [CrossRef]
40. Von Luxburg, U. A tutorial on spectral clustering. *Stat. Comput.* **2007**, *17*, 395–416. [CrossRef]
41. Caliński, T.; Harabasz, J. A dendrite method for cluster analysis. *Commun. Stat. Theory Methods* **1974**, *3*, 1–27. [CrossRef]
42. Litman, T. *Understanding Transport Demands and Elasticities*; Victoria Transport Policy Institute: Victoria, BC, Canada, 2017.
43. Smith, B.W.; Taplin, J.H. Derivation of demand elasticities from travel choice elasticities. *J. Transp. Econ. Policy (JTEP)* **2015**, *49*, 79–97.
44. Cervero, R.; Kockelman, K. Travel demand and the 3Ds: Density, diversity, and design. *Transp. Res. Part D Transp. Environ.* **1997**, *2*, 199–219. [CrossRef]
45. Zhao, P.; Kwan, M.P.; Zhou, S. The uncertain geographic context problem in the analysis of the relationships between obesity and the built environment in Guangzhou. *Int. J. Environ. Res. Public Health* **2018**, *15*, 308. [CrossRef]
46. Goodchild, M.F. The validity and usefulness of laws in geographic information science and geography. *Ann. Assoc. Am. Geogr.* **2004**, *94*, 300–303. [CrossRef]
47. Tu, J.; Xia, Z.G. Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. *Sci. Total Environ.* **2008**, *407*, 358–378. [CrossRef] [PubMed]
48. Wheeler, D.C. Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environ. Plan. A* **2007**, *39*, 2464–2481. [CrossRef]
49. Fotheringham, A.S.; Brunsdon, C.; Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*; John Wiley & Sons: Hoboken, NJ, USA, 2003.
50. Nakaya, T. GWR4 User Manual: Windows application for geographically weighted regression modelling. *GWR4 Dev. Team* **2014**, *23*, 14.