Exploring the Spatially Heterogeneous Effects of the Built Environment on Bike-Sharing Usage During the COVID-19 Pandemic

Hongtai Yang

Zishuo Guo

Guocong Zhai
Old Dominion University, gzhai001@odu.edu

Linchuan Yang

Jinghai Huo

Original Publication Citation

Yang, H., Guo, Z., Zhai, G., Yang, L., & Huo, J. (2022). Exploring the spatially heterogeneous effects of the built environment on bike-sharing usage during the COVID-19 pandemic. *Journal of Advanced Transportation*, *2022*, 1-15, Article 7772401. https://doi.org/10.1155/2022/7772401

Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 7772401, 15 pages
https://doi.org/10.1155/2022/7772401

Research Article

Exploring the Spatially Heterogeneous Effects of the Built Environment on Bike-Sharing Usage during the COVID-19 Pandemic

Hongtai Yang,[1] **Zishuo Guo**,[1] **Guocong Zhai**,[2] **Linchuan Yang**,[3] **and Jinghai Huo**[1]

[1]*School of Transportation and Logistics, National Engineering Laboratory of Integrated Transportation Big Data Application Technology, National United Engineering Laboratory of Integrated and Intelligent Transportation, Institute of System Science and Engineering, Southwest Jiaotong University, Chengdu 611756, China*
[2]*Department of Civil & Environmental Engineering, Old Dominion University (ODU), Norfolk, VA 23529, USA*
[3]*Department of Urban and Rural Planning, School of Architecture, Southwest Jiaotong University, Chengdu 611756, China*

Correspondence should be addressed to Zishuo Guo; 220213442@seu.edu.cn

Received 12 September 2021; Revised 9 June 2022; Accepted 11 August 2022; Published 28 September 2022

Academic Editor: Ricardo Giesen

Bike-sharing holds promise for available and healthy mobility services during COVID-19 where bike sharing users can make trips with lower health concerns due to social distancing compared to the restricted transportation modes such as public transit and ridesharing services. Leveraging the trip data of the Divvy bike-sharing system in Chicago, this study exploresspatially heterogeneous effects of built environment on bike-sharing usage under the pandemic. Results show that the average weekly ridership declined by 52.04%. To account for the spatially heterogeneous relationship between the built environment and the ridership, the geographically weighted regression (GWR) model and the semiparametric GWR (S-GWR) model are constructed. We find that the S-GWR model outperforms the GWR and the multiple linear regression models. The results of the S-GWR model indicate that education employment density, distance to subway, COVID-19 cases, and ridership before COVID-19 are global variables. The effects between ridership and the built environment factros (i.e., household density, office employment density, and the ridership) vary across space. The results of this study could provide a useful reference to transportation planners and bike-sharing operators to determine the high bike-sharing demand area under the pandemic,thus adjusting station locations, capacity, and rebalancing schemes accordingly.

1. Introduction

The outbreak of COVID-19 has seriously threatened the lives of people around the world. According to Johns Hopkins University in the United States, the cumulative number of deaths due to COVID-19 in the United States has exceeded 650,000, and the cumulative number of confirmed cases has exceeded 40.4 million as of September 8, 2021. During this period, US government agencies have implemented policies to reduce the community spread of the virus, including mandate stay-at-home and social distancing orders [1]. These orders have largely impacted residents' daily travel behaviors and further affected the urban transportation systems [2]. Given that the pandemic may last for a long time, the impactsare expectedto continue.

Consideringthe risk of exposure to COVID-19, people tend to reduce their use of public transportation modes (i.e.,subways and buses)following soical distancing guidances. However, the use of bike sharing has not been severely impacted because users could ride bikes in the open space and keep safe social distances. Studies have shown that when public transportation systemsare considered dangerous during COVID-19 [1], residents usually switch from a high-risk mode to cycling to reduce the risk of infection [3]. As a result, the demand for the use of bike-sharing has changed dramatically compared with the period before the outbreak of COVID-19 [4]. Therefore, understanding how the built environment factors affect the usage of public bicycles under the influence of COVID-19 is necessary because it could provide an important reference to transportation planners and bike-sharing operators to determine the high-demand areas, thusmaking an adjustment to the locations,the capacity, as well as the reblancing schemes of bike-sharing stations.

There has been a rich body of work on the use of public bicycles before the pandemic. Two types of models have been widely used in previous studies, namely global regression models [5, 6] and local regression models [7]. Global models, such as linear regression models and negative binomial regression models, assume that the coefficients of all predictors do not change across space. Despite their wide use, they do not capture the spatial variation in the relationship between predictor and response variables, especially in the case of large study areas [8, 9]. Therefore, the spatial latent class model [10] and the geographically weighted regression (GWR) models [7] are some of the methods adopted to capture this spatial variation. In this study, the GWR model is used. Since GWR models assume that all variables have a spatially varying relationship with the response variable, which may not be true, semiparametric GWR (S-GWR) models have been developed that allow some variables to be global and others to be local.

As a result, this study investigates the spatially varying relationship between the built environment and the bike-sharing ridership during COVID-19 while controlling for the ridership before the pandemic. We intend to answer the following four questions.

(1) Does the bike-sharing ridership increase or decrease due to the outbreak of COVID-19?
(2) If bike-sharing ridership changes, will the change of ridership of each station in proportion to the total ridership change?
(3) If the change of ridership of each station is not in proportion to the total ridership change, what factors, including built environment and demographics, result in this difference?
(4) How do these factors contribute to this difference?

The rest of the paper is structured as follows. The second part is a summary of relevant studies on the usage of bike-sharing. The third part describes the data used in this study. The fourth part presents the model results. The fifth part concludes this study by summarizing the main findings and the limitations.

2. Literature Review

The literature review of related studies is composed of two parts: the influencing factors of public bicycle ridership before the outbreak of COVID-19 and the impact of COVID-19 on bike-sharing usage.

2.1. Factors Influencing Bike-sharing before the Outbreak of COVID-19. Scholars have used different data and models to explore the factors that significantly influence bike-sharing usage. The factors can be divided into two categories: external and internal.

External factors mainly refer to built environment factors, including density, diversity, and design [7, 11–14]. Additionally, demographic factors are also regarded as external factors, including age, private car ownership, and income[15–18]. Other external factors are also included, such as weather conditions (e.g, temperature, humidity, and wind speed), substitution mode, and holidays [6, 17, 19–24]. In addition to this, in the context of the epidemic, some studies have also considered COVID-19-related factors, such as the number of cases and the number of deaths directly related to COVID-19 [25–27]. Internal factors mainly refer to personal preferences and service levels. For example, some studies explored how users' intention to use and fares affect ridership [5, 28–30].

In terms of model selections, many studies used multiple linear regression (MLR) models to determine significant influencing variables [6, 12, 31]. Since ordinary least square models cannot account for multicollinearity,capture spatial autocorrelations, and accurately estimate regression coefficients, Hu and Chen used partial least square to deal with the multicollinearity between explanatory variables. They found that the influence of the independent variables like household income on ridership at most stations was spatially different [25]. To further investigate the spatialimpacts, Cox and Hurtubia used spatial regression models to count for the spatial autocorrelation [10, 32–34]. And one of the studies concluded that the usage of dockless bike stations was spatially autocorrelated in commercial areas and road intersections [31]. Singhvi et al. used generalized linear regression models to deal with skewed distribution of the response variable [35–37]. The positive effect of station-CBD distance and the number of entertainment venues on the number of bike-sharing trips was found [33]. Hu et al. adopted a different model, the generalized mixed-effects model, by adding random effects to the generalized linear regression model [26, 38, 39]. Researchers found that areas with more COVID-19 cases, high income, and more educational employment had less human mobility under the impact of COVID-19 [26].

The predicting variables, response variables, and models of the most relevant studies are summarized in Table 1.

2.2. Impact of COVID-19 on Bike-sharing Ridership. The studies related to the impact of COVID-19 on the ridership of bike-sharing are summarized in this section. Bucsky studied the changes in human mobility and travel mode shares in Hungary during COVID-19 [4]. Public bicycles had the smallest decrease in ridership. The studypointed to public bicycles as an alternative to public transportation under COVID-19, giving a stronger rationale for the government to promote cycling. Buehler and Pucher studied the impact of COVID-19 on public bicycle usage through national surveys [41]. They revealed the general trends and changes over time in bicycling in different cities in Europe and the United States from 2019 to 2020. In order to explore the mechanisms of changes in public bicycle use, several

Table 1: Summary of studies on the relationship between the built environment and bike-sharing ridership.

Author	Predicting variable															Response variable	Model
	Density		Diversity		Destination accessibility		Distance to public transit				Design						
	Population	Employment	Mix land use	Percent of different land use types	Commuting distance	Distance to the city center	Number of bus stops	Number of subway stations	Distance to the nearest bus station	Distance to the nearest subway station	Bike-station characteristics	Bike lane characteristics	Road network density	Primary and secondary road density	Intersection density		
Yang et al. [7]	✓	✓			✓	✓					✓		✓	✓		Hourly ridership	MLR; GWR; S-GWR
El-Assi et al. [20]	✓	✓		✓							✓	✓	✓			Hourly ridership	Linear mixed model
Lin et al. [40]																Annual ridership	MLR
Noland et al. [28]	✓	✓	✓	✓							✓	✓				Monthly ridership	Negative binomial regression
Wang and Chen [33]	✓	✓		✓			✓	✓			✓	✓			✓	Monthly ridership	SEM
Hyland et al. [39]	✓	✓		✓		✓	✓	✓		✓	✓	✓				Monthly ridership	Mixed multilayer linear model

scholars have started to study the influencing factors using different methodologies.

Hu and Chen used Bayesian structural time series models and partial least square regression to study the temporal evolution of the impact of COVID-19 on transit ridership in terms of land use, COVID-19 related features, and sociodemographic variables [25]. The number of COVID-19 cases/deaths in the study was positively associated with a decline in ridership of public transportation, opposed toeducational level and income. Hu et al. used the generalized additive mixed model to explore the relationship between trips and influencing factors, including COVID-19-related features, demographic, and employment [26]. The results showed that the number of COVID-19 cases, income level, and educational employment were negatively associated with trips. The nonlinear temporal interactions between various independent variables and bike-sharing usage change were also explored by the same model [27]. The paper illustrated that residential is positively correlated with bike-sharing usage, while car ownership is negatively correlated with it.

In summary, the existing studies addressed the evolution of the effects in the time dimension and handled mixed effects of linearity and nonlinearity. However, the spatial nonstationary relationship has not been considered, which may lead to estimation bias. Therefore, this paper uses GWR and S-GWR models to deal with this issue.

3. Data Description

Chicago is one of the most populous cities in North America, with a large and energetic downtown, which attracts many commuters and visitors. This study uses two data sets, namely, the trip data of the Chicago Divvy bike-sharing system and the Smart Location Database (SLD) developed by the U.S. Environmental Protection Agency. The bike-sharing system of Chicago covers the unrban area of Chicago and two neighboring suburbs with around 600 stations and over 6,000 bikes. The data of the Divvy bike-sharing system include the start/end time of each trip, the start/end station of each trip, the type of membership, and user information. The spatial distribution of Divvy bicycle stations in Chicago is shown in Figure 1. SLD provides the built environment and demographic information aggregated at the level of census block group (CBG).

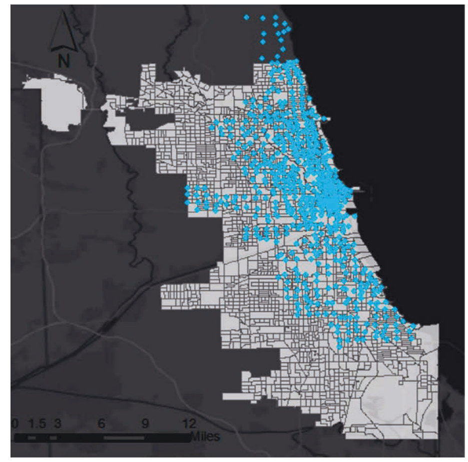


Figure 1: Divvy bicycle stations in the city of Chicago.

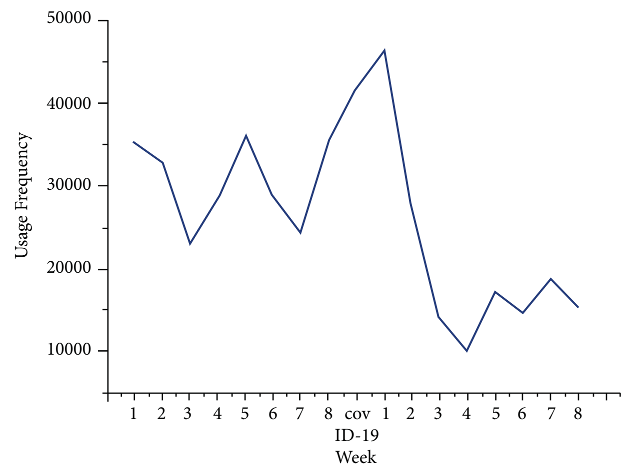


Figure 2: Weekly usage of bike sharing during the period of eight weeks before and after the outbreak of COVID-19.

3.1. Changes in the Spatial Distribution of Usage during COVID-19. Considering that the spread of COVID-19 began on 2020/2/26 in the United States, the bike-sharing trip data of the eight weeks before and after the spread week (2/27/2020-3/4/2020) are used in this study to represent the pre-COVID-19 period and peri-COVID-19 period.

The weekly usage of the Divvy system in the eight weeks before and during COVID-19 is plotted in Figure 2. The usage is defined as the sum of the pickup and dropoff trips.

As shown in Figure 2, the usage had fluctuated around 30,000 before COVID-19. During COVID-19, a short rise was observed, followed by a quick decline from nearly 50,000 to 10,000. After that, the usage became stable, being around 18,000. Therefore, the research period after the epidemic in this study is from the next three weeks to the next six weeks (3–6 weeks to the right of the spread week in Figure 2), and the research period before the epidemic is selected symmetrically from the first six weeks to the first three weeks as the research interval (3–6 weeks to the left of the spread week in Figure 2). According to statistics, the average weekly usage before the epidemic was 29,131.5 times, and the average weekly usage during the pandemic was 13,970.25 times. Overall, the usage of bike-sharingduring COVID-19 decreases by 52.04%.

If the ratio of bike-sharing usage during COVID-19 to that before COVID-19 is roughly the same for each station, the change in usage is only caused by the epidemic and has nothing to do with other factors. If this ratio differs greatly from station to station, it shows that the change is not only affected by COVID-19 but also affected by the characteristics of the station and surrounding environment such as built environment and demographics. The histogram of this ratio is shown in Figure 3.

Figure 3 shows that the ridership of most stations decreased during COVID-19 while that of some stations increased and the ratio varies greatly. Figure 4 presents the spatial distribution of the ratio. The two figures indicate that the change of usage is affected not only by the outbreak of COVID-19 but also by other factors such as the built environment.

Figures 5 and 6 present the spatial distributions of ridership. It shows that the stations with high ridership were centered around the CBD before COVID-19 and centered around the stations in the north during COVID-19. From these two figures, we can infer that bike-sharing operators should pay more attention to the possible shortage of bicycles or docks of the stations in the north during COVID-19.

3.2. Response and Predicting Variables. The response variable in this study is the total usage (sum of pickup and drop-off trips) of each station for the four weeks after COVID-19 (3/19/2020–4/15/2020), as shown in Figure 1.

In order to study the factors that affect the change of usage during COVID-19, it is necessary to control the usage before COVID-19. As a result, it is included in the model as a control variable. When selecting built environment factors, this article refers to the "5D" variables (5 types of built environment variables whose names start with D, including density, design, diversity, distance, and destination accessibility) proposed by previous studies as potential variables that may have an impact on the bike-sharing ridership under COVID-19[17]. In the end, a total of 20 predictors, including built environment, demographics, COVID-19-related cases and deaths, and ridership before COVID-19, are selected.

Seventeen variables in the built environment and demographics are derived from SLD. A circular buffer with a radius of 300 meters is drawn around each station. The radius of 300 meters is determined based on the common walking distance between the origin or the destination and the public bicycle station [7]. Values of the predicting variables are extracted based on the buffer.

The number of COVID-19 cases and deaths are obtained from the City of Chicago. Considering that the COVID-19 cases and deaths may have a wider influencing area [27], a circular buffer zone with a radius of 500 meters is drawn around each station to extract the number of COVID-19 cases and deaths.

3.3. Data Processing. The normal distribution of the dependent variable is an assumption of the classical linear regression model that ensures that the parameter regression results are unbiased [42]. The histogram of the response variable follows a skewed distribution, different from the normal distribution. In response to this problem, a logarithmic transformation of the response variable is performed. The transformation has also been adopted by other studies. The transformed results are also shown in Figure 7.

Previous studies have found that demographical and built environment variables would affect the travel patterns of residents during COVID-19 [25, 26]. For example, many people may work remotely or study at home due to the mandated work-from-home order, following social distance guidelines. Most people's home-based trips have a

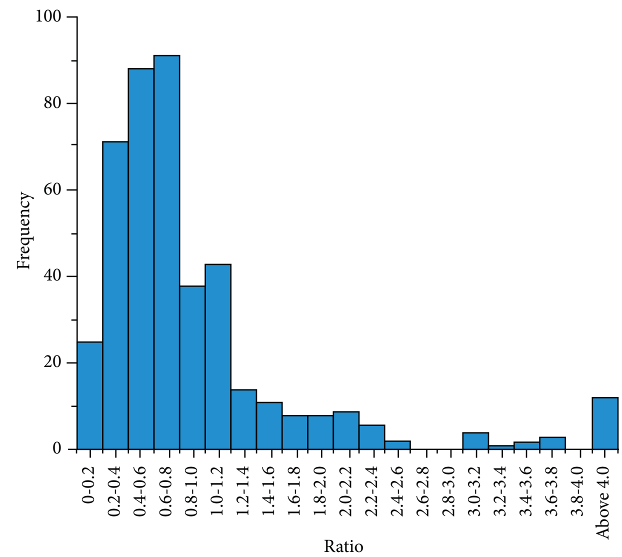


Figure 3: Histogram of the ratio of the ridership during COVID-19 to the ridership before COVID-19.

lower decline than office-based trips. Therefore, controlling for pre-epidemic ridership, household density is assumed to be positively correlated with peri-epidemicusage, while employment density is assumed to be negatively correlated with post-pandemic usage. In addition, variables such as distance to public transportation and proximity to the city center are considered to influence bike sharing usage. So both variables are negatively correlated with peri-pandemic ridership. Thus, this paper includes built environment, socioeconomic, and COVID-19-related variables that may impact peri-epidemic use as the response variable.

In order to eliminate the large difference in the magnitude of the explanatory variables and facilitate result interpretations, the explanatory variables are also logarithmically transformed. The modeling result will represent the elasticity of the response variable to the explanatory variables, which is expressed as the percentage of change in the response variable caused by a 1% change in the explanatory variable. The descriptive statistics of all variables in this study are shown in Table 2.

The formula for employment entropy is as follows [43]:

$$\text{Employment Entropy} = -\frac{\sum_{i=1}^{N}(p_i)\text{In}(p_i)}{\text{In}(N)}, \quad (1)$$

where N represents the number of employment types and p_i is the proportion of employment type i.

4. Methods

4.1. Multiple Linear Regression (MLR). This study establishes the MLR model to analyze factors that influence bike-sharing usage during COVID-19. The model assumes that the relationship between the predictor variables and the response variable is linear and homogeneous across space. Its function is as shown in the following equation.

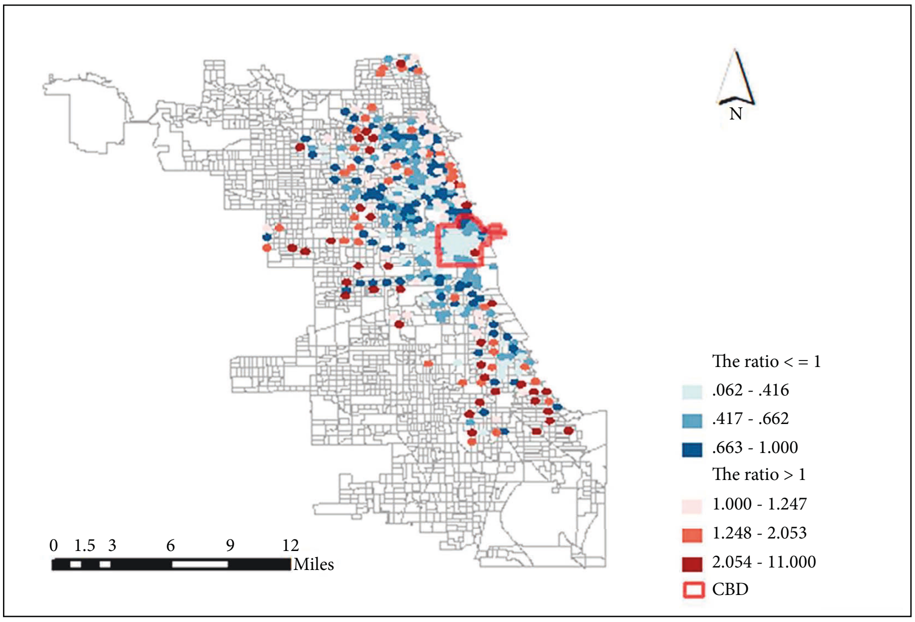


FIGURE 4: Spatial distribution of the ratio of ridership during COVID-19 to that before COVID-19.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon, \quad (2)$$

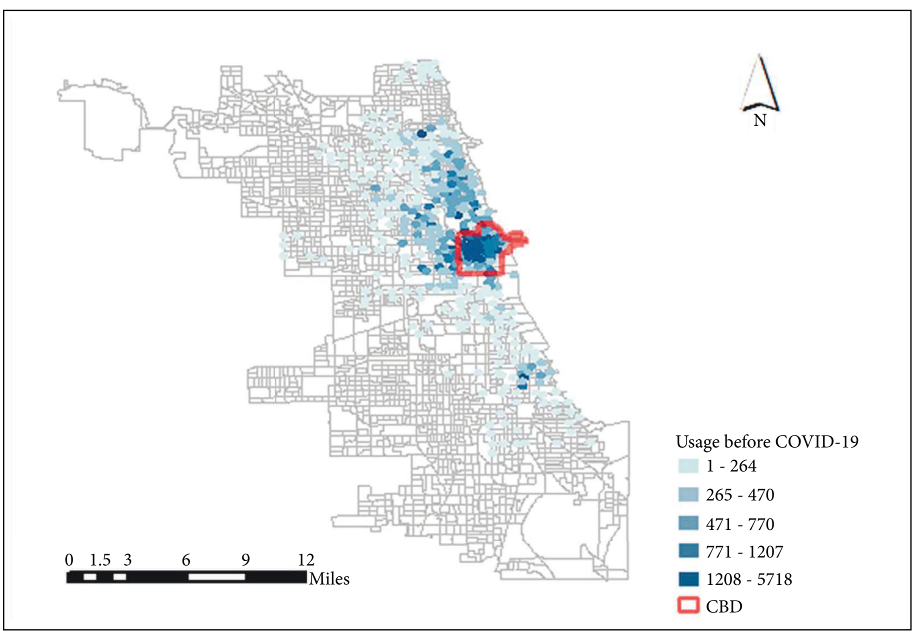


FIGURE 5: Spatial distribution of Divvy usage before COVID-19.

where y is the response variable; $x_1, x_2, \ldots, x_k$ are the predictors; $\beta_0, \beta_1, \ldots, \beta_k$ are the coefficients of the predictors; and ε is the random error, which has an expected value of zero, follows the normal distribution, and is independent of each other [42].

In the MLR model, the parameters are mainly estimated using Ordinary Least Squares (OLS) methods. The objective function is as follows:

$$\min \sum_{i=1}^{n} \left[Y_i - \left(\widehat{\beta}_0 + \widehat{\beta}_1 X_{i1} + \widehat{\beta}_2 X_{i2} \ldots + \widehat{\beta}_k X_{ik} \right) \right]^2, \quad (3)$$

where Y_i is the true value of the i-th response variable; $X_{i1}, \ldots, X_{ik}$ are the k-th predictor of the i-th response variable's predictors; and $\widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_k$ are the estimates of the parameters.

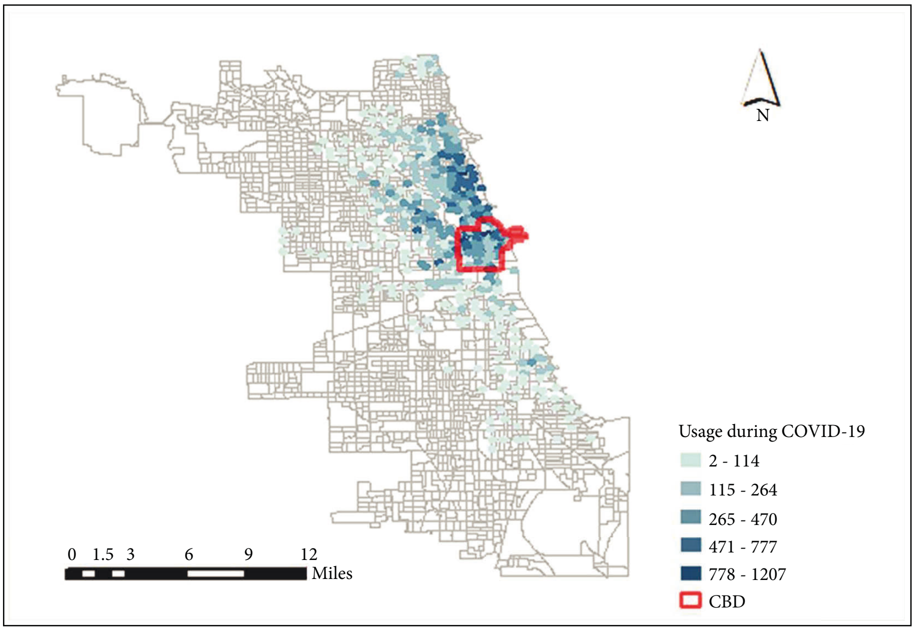
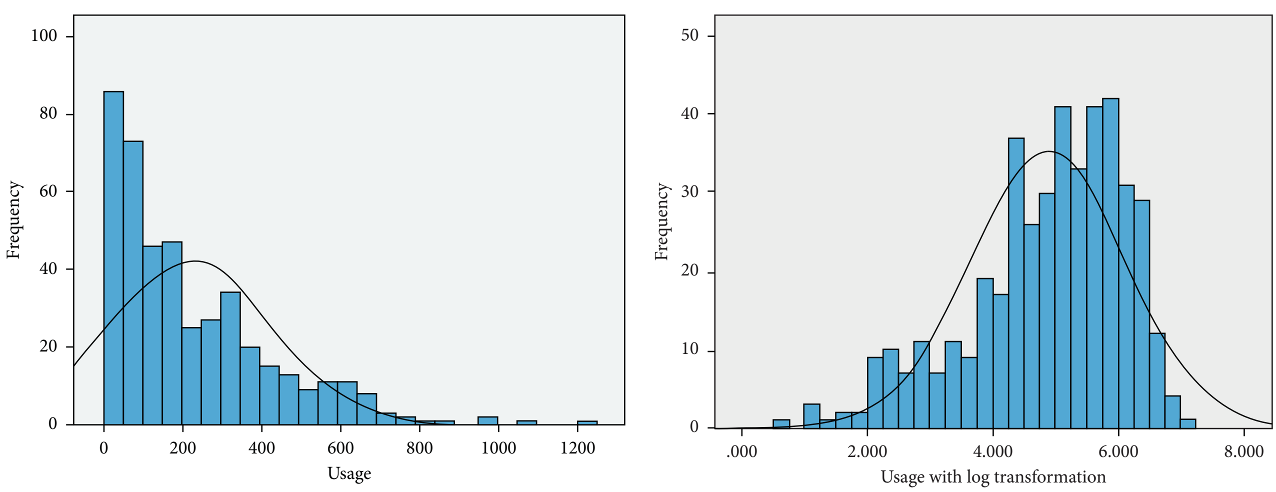


Figure 6: Spatial distribution of Divvy usage during COVID-19.



(a) (b)

Figure 7: Histogram of the bike-sharing usage with and without log transformation: (a) histogram of the bike-sharing usage and (b) histogram of the bike-sharing usage with log transformation.

4.2. Geographically Weighted Regression (GWR). In the case of a large study area, the relationship between the response variable and explanatory variables may vary across space. So the study needs to use a local regression model such as GWR. The GWR model improves the traditional MLR model by allowing the relationship to vary across space. It establishes local regression equations at each station and thus allows the regression parameters to vary with spatial location. Its function is as follows [7]:

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^{n} \beta_{ik}(u_i, v_i) x_{ik} + \varepsilon_i, \tag{4}$$

where y_i represents the bike-sharing usage of station i, β_{ik} is the coefficient of the predictor k of station i, x_{ik} is the predictor k of station i, ε_i is the random error term of station i, and (u_i, v_i) represents the latitude and longitude of station i.

There are a set of coefficients at each public bicycle station for the GWR modeling results. It indicates that the effect of the predicting variables on the response variable varies across space. When estimating the coefficients of each station, weight w_i is assigned based on the distance from other stations to the target station.The coefficients of explanatory variablesare estimated by minimizing the

Table 2: Descriptive statistics of variables.

Category	Variable	Meaning	Source	Mean (log-transformed)	Variance (log-transformed)
Usage	Usage after COVID-19	Usage of each station from 2020/3/19 to 2020/4/15	Divvy's official website	4.872	1.524
Density	Population density	Population density (people/acre)	Smart Location Database	3.336	0.510
	Household density	Household density (households/acre)	Smart Location Database	2.700	0.711
	Entertainment employment density	Entertainment employment density (jobs/acre)	Smart Location Database	0.202	4.540
	Education employment density	Education employment density (jobs/acre)	Smart Location Database	−1.135	8.962
	Retail employment density	Retail employment density (jobs/acre)	Smart Location Database	−0.051	2.962
	Office employment density	Office employment density (jobs/acre)	Smart Location Database	0.396	4.699
	Healthcare employment density	Healthcare employment density (jobs/acre)	Smart Location Database	0.228	3.616
Diversity	Employment entropy	Employment entropy using the formula for eight types of jobs such as retail jobs, factory jobs, and service jobs	Smart Location Database	−0.625	0.342
Design	Auto-oriented links	Length of links only for automobiles per square mile	Smart Location Database	1.316	0.721
	Multimodal links	Length of links for autos and pedestrians per square	Smart Location Database	−3.462	20.677
	Pedestrians-oriented links	Length of links only for pedestrians per square	Smart Location Database	3.027	0.081
Distance from public transit	Distance to subway	Distance from population-weighted centroid to the nearest subway stop	Smart Location Database	5.112	0.432
Destination accessibility	Job accessibility	Jobs within 45 minutes auto travel time	Smart Location Database	12.969	0.189
	Working-age population accessibility	Working age population within 45 minutes auto travel time	Smart Location Database	13.228	0.120
	Distance to CBD	Distance to CBD	Own calculation	1.533	0.825
Demographic	Percentage of HH with no vehicles	Percentage of zero-car households in CBG	Smart Location Database	−1.367	0.377
	Percentage of low-income population	Percentage of workers earning $1250/month or less	Smart Location Database	−0.995	0.898
COVID-19	COVID-19 cases	Cumulative COVID-19 cases from 2020/3/19 to 2020/4/15	Chicago data portal	6.690	0.190
	COVID-19 deaths	Cumulative deaths caused by COVID-19 from 2020/3/19 to 2020/4/15	Chicago data portal	2.419	0.584
Control	Usage before COVID-19	Divvy usage from 2020/1/16 to 2020/2/12	Divvy's official website	5.254	2.751

weighted sum of squares. The objective function for the GWR model is as follows:

$$\min \sum_{j=1}^{n} w_{ij}\left(y_i - \beta_{i0} - \sum_{k=1}^{n} \beta_{ik}(u_i, v_i)x_{ik} \right)^2. \tag{5}$$

The spatial weight reflects the importance of the position. Many ways can be used to calculate the spatial weight. The simplest one is the distance threshold function. The specific function is as follows:

$$w_{ij} = \begin{cases} 1, & d_{ij} \le D, \\ 0, & d_{ij} > D, \end{cases} \tag{6}$$

where D represents the distance threshold and d_{ij} represents the distance between station i and the target station j. To solve the problem of weight discontinuity, the Gaussian function is also often used to express the relationship between weight and distance

$$w_{ij} = e^{-1/2\left(d_{ij}/b\right)^2}, \tag{7}$$

where w_{ij} represents the weight between stations i and the target station j, d_{ij} represents the distance between stations i and the target station j, and b is the bandwidth.

4.3. Semiparametric Geographically Weighted Regression (S-GWR). GWR models assume that the coefficients of all predictors vary across space. However, the relationship between some predictors and the response variable may not vary across space. The S-GWR model, as an extension of the GWR model, allows some predictors to be global and others to be local. The expression of the S-GWR model is as follows [7], and the symbols in the equation are the same as the GWR model.

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^{n} \beta_{ik}x_{ik} + \sum_{k=1}^{n} \beta_{ik}(u_i, v_i)x_{ik} + \varepsilon_i. \tag{8}$$

The objective function for the S-GWR model is as follows:

$$\min \sum_{j=1}^{n} w_{ij}\left(y_i - \beta_{i0} - \sum_{k=1}^{p} \beta_{ik}x_{ik} - \sum_{k=1}^{n} \beta_{ik}(u_i, v_i)x_{ik} \right)^2. \tag{9}$$

5. Model Results

This section establishes MLR, GWR, and S-GWR models to explore the relationships between explanatory variables and the response variable.

5.1. Results of MLR Model. There are two ways to construct regression models. One way is to treat the ridership during COVID-19 as the dependent variable and the ridership before COVID-19 as well as other variables as the independent variables. The other way is to treat the ratio of the ridership during COVID-19 to the ridership before COVID-19 as the dependent variable and other variables as the independent variables. Both ways are explored in this study.

The first way to construct the model, which is to treat the ridership during COVID-19 as the dependent variable and the ridership before COVID-19 as well as other variables as the independent variables, is first explored. The backward variable selection method is adopted to select variables. After this method, six explanatory variables are significantly related to the usage of Divvy bike-sharingduring COVID-19 at the5% level. Moreover, the variance inflation factor VIF is less than 5 for these six variables, which indicates there are no multicollinearity issues. The formula for VIF is as follows:

$$\mathrm{VIF} = \frac{1}{\left(1 - R_i^2\right)}, \tag{10}$$

where R_i^2 is the coefficient of determination for the mode using the i-th explanatory variable as the response variable and the rest of the explanatory variables as explanatory variables.

Table 3 presents the model results. The significant variables are household density, education employment density, office employment density, distance to the nearest subway station, COVID-19 cumulative cases, and usage before COVID-19.

Explanations for the relationships between significant predictors and response variables are given below.

During COVID-19, the household density is positively correlated with ridership. This may be due to work-at-home policies that keep residents at home. As a result, more trips are home-based trips. There is also a positive association between the cumulative number of cases and the ridership. This result may be because areas with more bike-sharing trips are those with high travel demand. More people gathering around the area leads to a higher risk of infection. On the other hand, the increase in the number of COVID-19 infections may also make people use bike-sharing to replace other public transportation modes, such as the subway and bus. Moreover, the usage of Divvy bike-sharing before COVID-19 is positively correlated with the usage of Divvy bike-sharing during COVID-19. This shows that stations with a high usage rate before COVID-19 continue to have high usage during COVID-19. The coefficient is 0.626, indicating that if the pre-COVID-19 usage increases by 1%, the peri-COVID-19 usage will increase by 0.626% on average.

The education employment density is negatively correlated. It could be due to that most schools require students and faculty members to stay at home and to take or teach classes online. Similarly, the office employment density is also negatively correlated. This result may also be due to the stay-at-home order and work-from-home policy in Chicago in response to COVID-19. And the number of ridership is negatively correlated with the distance to the nearest subway station. The possible reason is that residents who are close to the nearest subway station after the COVID-19 outbreak have switched from using the subway to using bike-sharing.

However, bike-sharing usage before COVID-19 may weaken the effect of time invariant explanatory variables on bike-sharing usage during COVID-19. To demonstrate the explanatory power of these predicting variables for bike-

Table 3: MLR model results.

Variable	Coefficient	Standard error	P	VIF
Household density	0.288	0.041	0.000	1.953
Education employment density	−0.028	0.011	0.013	1.865
Office employment density	−0.043	0.021	0.047	3.629
Distance to subway	−0.165	0.047	0.000	1.580
COVID-19 cases	0.224	0.069	0.001	1.522
Ridership before COVID-19	0.626	0.022	0.000	2.133

Table 4: MLR model results using the ratio of bike-sharing usage during/before COVID-19 as the response variable.

Variable	Coefficient	Standard error	P	VIF
Household density	0.133	0.045	0.004	1.324
Entertainment employment density	−0.126	0.024	0.000	2.632
Education employment density	−0.072	0.014	0.000	1.930
Working age population accessibility	−0.663	0.221	0.003	1.349
Percentage of HH with no vehicles	0.165	0.066	0.012	1.245

sharing usage during COVID-19, the ratio of bike-sharing usage during COVID-19 to bike-sharing usage before COVID-19 is also developed as the response variable. Considering there are huge differences in the ratio, the log transformation is performed on the dependent variable. The same backward variable selection method is used to screen for significant variables. The results of the model are shown in Table 4.

The model results, shown in Table 4, are different from those of the previous model. The goodness of fit of the model is 0.347, which is not as good as that of the previous model. This is probably because the variable of ridership before COVID-19 accounts for a large proportion of the time invariant components of the response variable, ridership during COVID-19, in the previous model. Household density and education employment density are significant at both mdels. Though there are some differences in the model results in terms of significant variables, it should be noted that the difference in model results is quite common in empirical studies. Sometimes, simply extending the time period of the data could lead to different model results.

Specifically, the model results show a positive correlation between household density and the dependent variable. The work-from-home policy made many people stay at home where office-based trips decreased more dramatically than home-based trips. The proportion of households without a car is also positively correlated with the dependent variable. During the epidemic, the automobile would be travelers' first preference. For households without an automobile, people are less willing to use the metro or bus, which could lead to disease transmission and thus turn to bike-sharing. As a result, bike-sharing riderships in areas with higherhousehold density and lower vehicle ownershipis not affected by the epidemic too much.

In contrast, entertainment employment density and education employment density are negatively correlated with the dependent variable. This is probably because the epidemic severely impacted the entertainment business, leading to employment loss. Regarding the effect of education employment density, both employees and students are instructed to give lectures or study at home, which reduces trips to and from schools. The working-age population accessibility is the working-age population that can be accessed by driving an automobile in 45 minutes. It is also negatively related to the dependent variable. It may be because, by dividing the population into a working population and a nonworking population, the trips performed by the working population have a higher percentage of decrease compared to those made by the nonworking population due to the work-from-home policy. Thus, by controlling for the household density, the working-age population accessibility is negatively related to the dependent variable.

Each of the two models has its own advantage. In this paper, we adopt the first model to explore the spatially varying relationship between explanatory variables and the dependent variable.

5.2. GWR Model Results. The modeling results of the GWR model are shown in Table 5.

The overall R-squared of the GWR model is 0.865. The R-squared of each station is visualized in Figure 8.

The R-squared for each station is above 0.5, with a considerable proportion being above 0.8, which indicates that the GWR model fits the data well. The generally lower goodness-of-fit for downtown areas may be due to the fact that ridership at these stations is more variable than at other stations. In addition, the variables selected do not accurately explain the dynamic changes in ridership because of the work-at-home policy.

5.3. S-GWR Model Results. The S-GWR model is also constructed.Table 6 presents the modeling results of the S-GWR model. To determine whether each predicting variable is global or local, all variables are first assumed to be global variables. The difference (DIFF) of criterion of each variable is calculated. The DIFF indicates the difference in comparison metrics (AICc) between the GWR and S-GWR models. In general, a positive value for the DIFF of criterion

Table 5: Coefficients of each variable in the GWR model.

Variable	Min	Lower quartile	Median	Upper quartile	Max
Intercept	4.650583	4.796953	5.013346	5.089241	5.200212
Household density	−0.086092	0.193541	0.277111	0.305725	0.373394
Education employment density	−0.174001	−0.086815	−0.050423	−0.017703	0.182649
Office employment density	−0.245413	−0.150852	−0.040037	0.046748	0.177302
Distance to subway	−0.286487	−0.115407	−0.097054	−0.072675	0.084666
COVID-19 cumulative cases	−0.136410	0.006409	0.035964	0.059660	0.180760
Ridership before COVID-19	0.818073	0.935148	0.982256	1.030879	1.144220

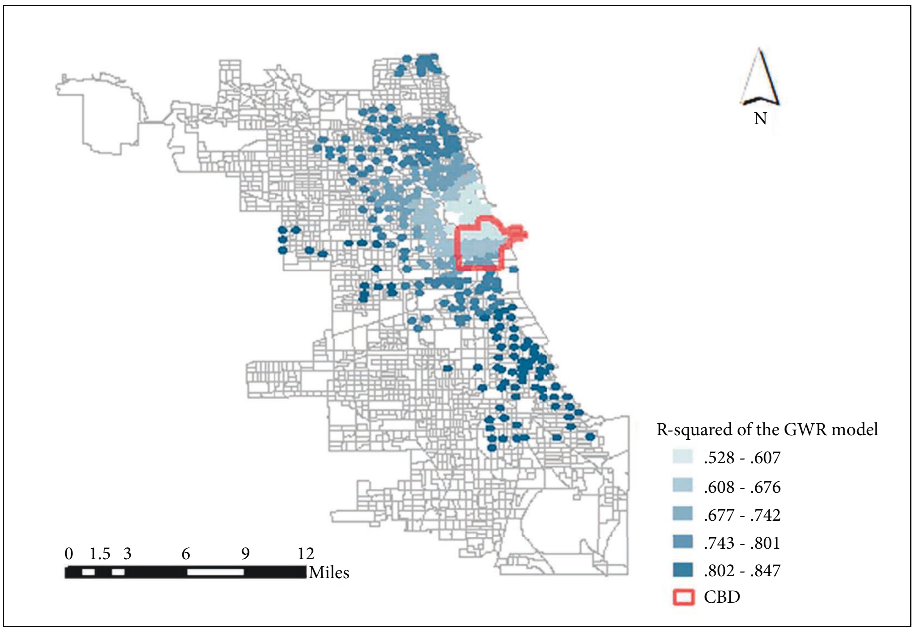


Figure 8: Spatial distribution of the R-squared of the GWR model.

Table 6: Variable type test.

Variable	F	DOF for F	DIFF
Intercept	8.902237	2.268	−15.953818
Household density	5.598979	3.301	−11.813178
Education employment density	1.722706	4.106	1.953052
Office employment density	2.630903	3.698	−1.785874
Distance to subway	2.141936	3.503	0.113963
COVID-19 cases	1.213646	2.911	2.971311
Ridership before COVID-19	1.358139	3.954	3.411965

indicates that there is no spatial variability [7]. In other terms, if DIFF > 0, the predicting variable is regarded as a global variable; otherwise, it is a local variable.

Compared with the MLR modeling results, the intercept for GWR or S-GWR is a local variable that varies with the geographical location, which could be observed in previous studies [7, 41]. As summarized in Table 6, the household density and the office employment density are local variables in the model, while the others are global variables.

The coefficients of the two local variables will be presented below. The interpretations of the results will also be given.

The spatial distribution of the coefficients of household density is shown in Figure 9. When household density is a significant factor, it usually has a positive correlation with the usage of public bicycles during the epidemic, which is consistent with the conclusions of the previous study [27]. The coefficients of household density in the center and eastern regions were relatively small, ranging from 0.161 to 0.286. This could be attributed to many high-income residents who are more likely to commute by private cars instead of public bikes in these two regions.

The spatial distribution of the coefficients of office employment density is shown in Figure 10. When office employment density is a significant factor, it is usually negatively correlated with ridership. It is probably because the stay-at-home order during COVID-19 reduced the number of people working in the office and reduced the usage of bike-sharing in areas with high office employment density. Therefore, this negative impact is obvious in the city center. Also, there is an area in the south where office employment density is significant because of the proximity of this area to the location of the University of Chicago. The office employment in this area is mainly the faculty members and staff of the university. During the epidemic, most of the faculty members and staff of the university are required to work from home. The rate of faculty members and staff who work from home is higher than that of other types of jobs. Thus, the office employment density is negative at a

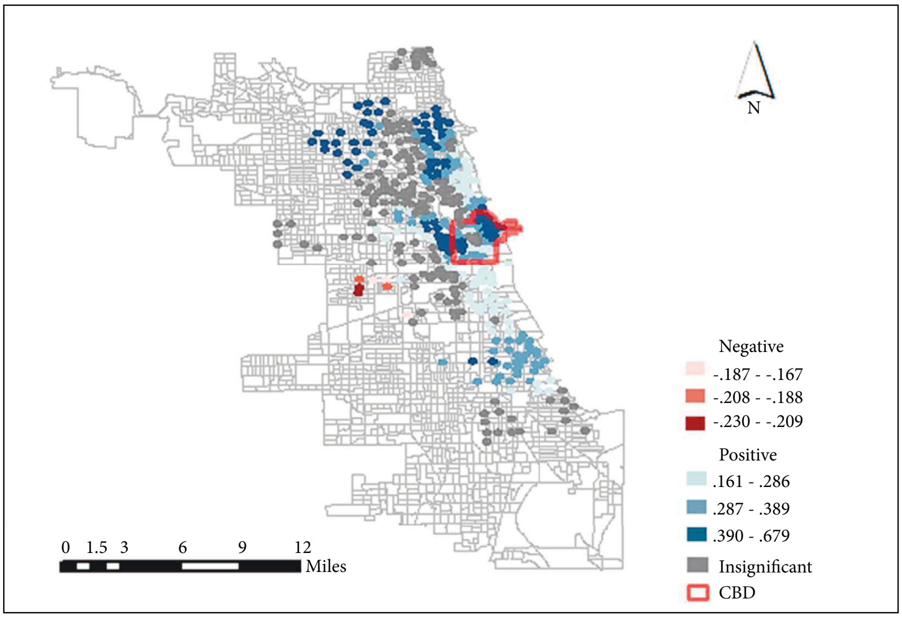


Figure 9: Distribution of household density regression coefficients.

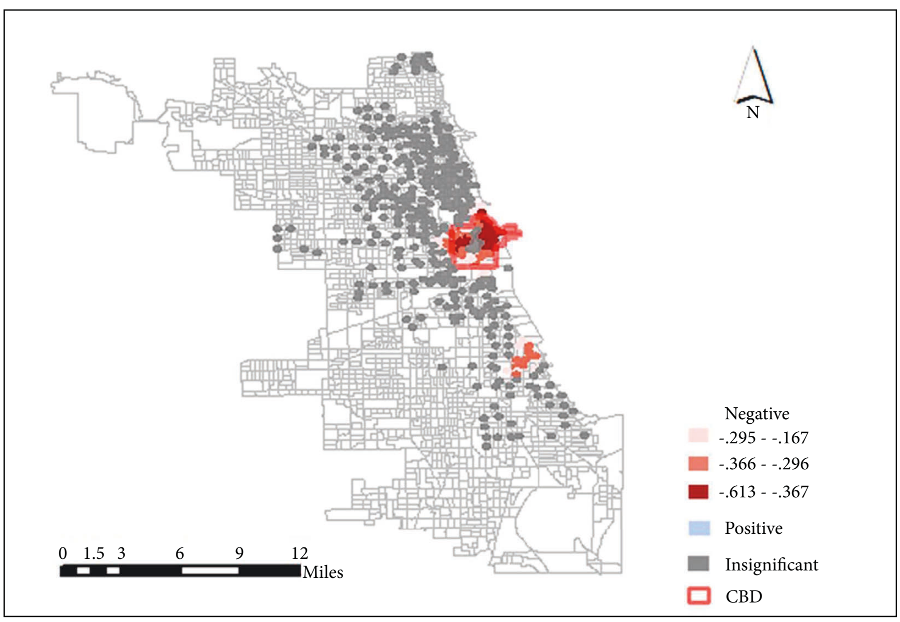


Figure 10: Distribution of office employment density regression coefficients.

significant level. Overall, the magnitude of the negative coefficients of household density and office employment density is larger in downtown areas than in the southern regions. Such relationships may be due to the higher percentage of downtown shutdowns than on the south side.

The overall R-squared of the S-GWR model is 0.886, which is higher than that of the GWR model. The spatial distribution of the R-squared of the S-GWR model is shown in Figure 11. The goodness-of-fit of the stations in the city center is still not high, which shows that the dramatic change in ridership in these areas caused by COVID-19 is difficult to capture.

5.4. Model Comparison. The results of the MLR, GWR, and S-GWR models are compared and shown in Table 7.

Since the models with a smaller sum of squares of residuals, −2 log-likelihood, AIC, AICc, and higher R-squared are regarded as better models, the S-GWR model is the best among the three models.

6. Conclusion

This study investigates the built environment factors that influence the bike-sharing ridership of the Chicago Divvy system during COVID-19 while controlling for the ridership

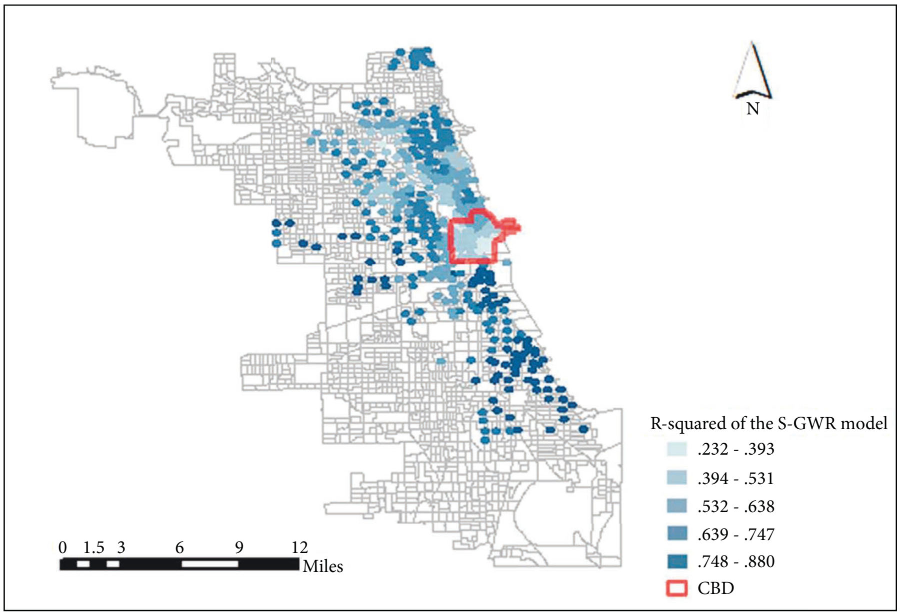


Figure 11: Spatial distribution of the R-squared of the S-GWR model.

Table 7: Comparison of model results.

Indicator	MLR	GWR	S-GWR
Residual sum of squares	112.160	89.645	75.429
−2 log-likelihood	645.349	547.657	472.374
Classic AIC	661.349	612.292	591.082
AICc	661.687	617.640	610.155
R-squared	0.831	0.865	0.886
Adjusted R-squared	0.828	0.851	0.862

before COVID-19. To capture the spatially varying relationship between built environment and ridership, GWR and S-GWR models are established.The MLR model is also developed to be comparable. We found that S-GWR has the highest goodness-of-fit from many perspectives.

We also found that the total bike-sharing ridership declined by half after the outbreak of COVID-19. The decline of the ridership of each station is different; the spatial distribution of usage of bike-sharing during COVID-19 is different from that before COVID-19. This observation indicates that transportation planners and bike-sharing operators should pay attention to this change and could adjust the capacity and location of the stations as well as the rebalancing scheme according to the current ridership pattern.

In terms of the relationship between the built environment and change in ridership, some variables are local variables (i.e., household density and office employment density), whileother variables are global variables such as education employment density and distance to the nearest subway station. The complex relationship should be fully considered when estimating the change in ridership of bike-sharing stations.

There are also some limitations in this study. First, although the results obtained from this study may not be applied to all cities, the analysis framework could be applied to other cities. Each city should develop policies based on its own condition. Secondly, because we used cross-sectional data, the revealed relationship between the independent variables and the response variable should be regarded as a correlation instead of a causal relationship. Although some causal relationships could be inferred from the results, this inference should be made with caution. In the future, panel data could be used to deal with this issue. Of course, traditional Poisson and negative binomial models are designed for count variables. However, it would be more appropriate to use linear regression models when the values of the response variable do not contain zero or small values [44, 45]. Moreover, when the area of the intersection of the buffer and the CBG is not the whole block, we assume that the independent variables are uniformly distributed in the CBG. However, the ground truth may not be the case. Another limitation of this study is that there are different ways to construct the regression models, and no theoretical justification for which model is more suitable. Each of the two ways has its own advantage. The model using the ridership during COVID-19 as the dependent variable is adopted because the results are more intuitive, and the goodness of fit is better. But it should be noted that the high goodness of fit is probably because the ridership before COVID-19 is highly correlated with the dependent variable, and this high correlation could overshadow the effect of other independent variables. As such, modeling other related variables, such as the ratio of peri-COVID-19 and pre-COVID-19 usage, is also meaningful because it could avoid this issue and is worth investigating. Finally, we use four-week bike-sharing ridership data during COVID-19 as the response variable. Though the travel volume is relatively stable during COVID-19 (as indicated in Figure 2)makes it possible to capture bike-sharing usage patterns under COVID-19, using a longer time horizon may generate more reliable modeling results.

Data Availability

The data used in this paper are available at https://www.divvybikes.com/system-data, https://www.epa.gov/smartgrowth/smart-location-mapping, https://data.cityofchicago.org/Health-Human-Services/COVID-19-Cases-Tests-and-Deaths-by-ZIP-Code/yhhz-zm2v, and https://data.cityofchicago.org/browse?tags=gis.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

The authors confirm their contribution to the paper as follows. Hongtai Yang contributed to conceptualization, methodology, formal analysis, and writing, particularly review and editing. Zishuo Guo contributed to data processing, formal analysis, and writing, particularly original draft preparation. Guocong Zhai contributed to writing, review, and editing. Linchuan Yang contributed to writing, review, and editing. Jinghai Huo contributed to review and editing. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments

This study was funded by the National Natural Science Foundation of China (Grant nos. 51774241 and 71704145), China Postdoctoral Science Foundation, and Sichuan Youth Science and Technology Innovation Research Team Project (Grant nos. 2019JDTD0002 and 2020JDTD0027).

References

[1] H. Wang and R. B. Noland, "Bikeshare and subway ridership changes during the COVID-19 pandemic in New York City," *Transport Policy*, vol. 106, pp. 262–270, 2021.

[2] L. Yang, Y. Liang, B. He, Y. Lu, and Z. Gou, "COVID-19 effects on property markets: the pandemic decreases the implicit price of metro accessibility," *Tunnelling and Underground Space Technology*, vol. 125, Article ID 104528, 2022.

[3] A. Nikiforiadis, G. Ayfantopoulou, and A. Stamelou, "Assessing the impact of COVID-19 on bike-sharing usage: the case of thessaloniki, Greece," *Sustainability*, vol. 12, no. 19, 2020.

[4] P. Bucsky, "Modal share changes due to COVID-19: the case of Budapest," *Transportation Research Interdisciplinary Perspectives*, vol. 8, Article ID 100141, 2020.

[5] K. Wang, G. Akar, and Y. J. Chen, "Bike sharing differences among millennials, gen xers, and baby boomers: lessons learnt from New York city's bike share," *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 1–14, 2018.

[6] R. A. Rixey, "Station-level forecasting of bikesharing ridership: station network effects in three US systems," *Transportation Research Record*, vol. 2387, no. 1, pp. 46–55, 2013.

[7] H. Yang, Y. Zhang, L. Zhong, X. Zhang, and Z. Ling, "Exploring spatial variation of bike sharing trip production and attraction: a study based on Chicago's Divvy system," *Applied Geography*, vol. 115, Article ID 102130, 2020.

[8] L. Yang, X. Chu, Z. Gou, H. Yang, Y. Lu, and W. Huang, "Accessibility and proximity effects of bus rapid transit on housing prices: h," *Journal of Transport Geography*, vol. 88, Article ID 102850, 2020.

[9] L. Yang, K. W. Chau, W. Y. Szeto, X. Cui, and X. Wang, "Accessibility to transit, by transit, and property prices: spatially varying relationships," *Transportation Research Part D: Transport and Environment*, vol. 85, Article ID 102387, 2020.

[10] T. Cox and R. Hurtubia, "Compact development and preferences for social mixing in location choices: results from revealed preferences in Santiago, Chile," *Journal of Regional Science*, vol. 62, 2021.

[11] Y. Guo, L. Yang, Y. Lu, and R. Zhao, "Dockless bike-sharing as a feeder mode of metro commute? The role of the feeder-related built environment: analytical framework and empirical evidence," *Sustainable Cities and Society*, vol. 65, Article ID 102594, 2021.

[12] A. Faghih-Imani, N. Eluru, A. M. El-Geneidy, M. Rabbat, and U. Haq, "How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal," *Journal of Transport Geography*, vol. 41, pp. 306–314, 2014.

[13] J. Huo, H. Yang, C. Li, R. Zheng, L. Yang, and Y. Wen, "Influence of the built environment on E-scooter sharing ridership: a tale of five cities," *Journal of Transport Geography*, vol. 93, Article ID 103084, 2021.

[14] H. Yang, Y. Bao, J. Huo, S. Hu, L. Yang, and L. Sun, "Impact of road features on shared e-scooter trip volume: a study based on multiple membership multilevel model," *Travel Behaviour and Society*, vol. 28, pp. 204–213, 2022.

[15] A. Faghih-Imani and N. Eluru, "Analysing bicycle-sharing system user destination choice preferences: Chicago's Divvy system," *Journal of Transport Geography*, vol. 44, pp. 53–64, 2015.

[16] S. Kaviti, M. M. Venigalla, and K. Lucas, "Travel behavior and price preferences of bikesharing members and casual users: a Capital Bikeshare perspective," *Travel Behaviour and Society*, vol. 15, pp. 133–145, 2019.

[17] R. Cervero, O. L. Sarmiento, E. Jacoby, L. F. Gomez, and A. Neiman, "Influences of built environments on walking and cycling: lessons from bogotá," *International journal of sustainable transportation*, vol. 3, no. 4, pp. 203–226, 2009.

[18] H. Li, H. Ding, G. Ren, and C. Xu, "Effects of the lcslch," *Transportation Research Part A: Policy and Practice*, vol. 111, pp. 304–315, 2018.

[19] A. Faghih-Imani and N. Eluru, "Determining the role of bicycle sharing system infrastructure installation decision on usage: case study of montreal BIXI system," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 685–698, 2016.

[20] W. El-Assi, M. Salah Mahmoud, and K. Nurul Habib, "Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto," *Transportation*, vol. 44, no. 3, pp. 589–613, 2017.

[21] H. Yang, J. Huo, Y. Bao, X. Li, L. Yang, and C. R. Cherry, "Impact of e-scooter sharing on bike sharing in Chicago," *Transportation Research Part A: Policy and Practice*, vol. 154, pp. 23–36, 2021.

[22] H. Yang, G. Zhai, L. Yang, and K. Xie, "How does the suspension of ride-sourcing affect the transportation system and environment?" *Transportation Research Part D: Transport and Environment*, vol. 102, Article ID 103131, 2022.

[23] H. Yang, Y. Liang, and L. Yang, “Equitable? Exploring ridesourcing waiting time and its determinants,” *Transportation Research Part D: Transport and Environment*, vol. 93, Article ID 102774, 2021.

[24] H. Yang, J. Huo, R. Pan, K. Xie, W. Zhang, and X. Luo, “Exploring built environment factors that influence the market share of ridesourcing service,” *Applied Geography*, vol. 142, Article ID 102699, 2022.

[25] S. Hu and P. Chen, “Who left riding transit? Examining socioeconomic disparities in the impact of COVID-19 on ridership,” *Transportation Research Part D: Transport and Environment*, vol. 90, Article ID 102654, 2021.

[26] S. Hu, C. Xiong, M. Yang, H. Younes, W. Luo, and L. Zhang, “A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic,” *Transportation Research Part C: Emerging Technologies*, vol. 124, Article ID 102955, 2021.

[27] S. Hu, C. Xiong, Z. Liu, and L. Zhang, “Examining spatiotemporal changing patterns of bike-sharing usage during COVID-19 pandemic,” *Journal of Transport Geography*, vol. 91, no. 11, Article ID 102997, 2021.

[28] R. B. Noland, M. J. Smart, and Z. Guo, “Bikeshare trip generation in New York City,” *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 164–181, 2016.

[29] J. Wang and G. Lindsey, “Neighborhood socio-demographic characteristics and bike share member patterns of use,” *Journal of Transport Geography*, vol. 79, Article ID 102475, 2019.

[30] M. Verma and A. Awasthi, “Evaluating bikesharing service quality: a case study for BIXI, Montreal,” *International Journal of Productivity and Quality Management*, vol. 29, no. 1, pp. 45–61, 2020.

[31] T. D. Tran, N. Ovtracht, and B. F. d’Arcier, “Modeling bike sharing system using built environment factors,” *Procedia Cirp*, vol. 30, pp. 293–298, 2015.

[32] Y. Zhang, T. Thomas, and M. Brussel, “Exploring the impact of built environment factors on the use of public bikes at bike stations: case study in Zhongshan, China,” *Journal of Transport Geography*, vol. 58, pp. 59–70, 2017.

[33] K. Wang and Y. J. Chen, “Joint analysis of the impacts of built environment on bikeshare station capacity and trip attractions,” *Journal of Transport Geography*, vol. 82, Article ID 102603, 2020.

[34] Y. Shen, X. Zhang, and J. Zhao, “Understanding the usage of dockless bike sharing in Singapore,” *International Journal of Sustainable Transportation*, vol. 12, no. 9, pp. 686–700, 2018.

[35] D. Singhvi, S. Singhvi, and P. I. Frazier, “Predicting bike usage for New York city’s bike sharing system,” in *Proceedings of the Workshops at the twenty-ninth AAAI conference on artificial intelligence*, Austin, TX, USA, January 2015.

[36] X. Wang, G. Lindsey, J. E. Schoner, and A. Harrison, “Modeling bike share station activity: effects of nearby businesses and jobs on trips to and from stations,” *Journal of Urban Planning and Development*, vol. 142, no. 1, Article ID 04015001, 2016.

[37] R. An, R. Zahnow, D. Pojani, and J. Corcoran, “Weather and cycling in New York: the case of Citibike,” *Journal of Transport Geography*, vol. 77, pp. 97–112, 2019.

[38] F. Sun, P. Chen, and J. Jiao, “Promoting public bike-sharing: a lesson from the unsuccessful Pronto system,” *Transportation Research Part D: Transport and Environment*, vol. 63, pp. 533–547, 2018.

[39] M. Hyland, Z. Hong, H. K. R. D. Pinto, and Y. Chen, “Hybrid cluster-regression approach to model bikeshare station usage,” *Transportation Research Part A: Policy and Practice*, vol. 115, pp. 71–89, 2018.

[40] P. Lin, J. Weng, Q. Liang, D. Alivanistos, and S. Ma, “Impact of weather conditions and built environment on public bikesharing trips in beijing,” *Networks and Spatial Economics*, vol. 20, no. 1, pp. 1–17, 2020.

[41] R. Buehler and J. Pucher, “COVID-19 impacts on cycling, 2019-2020,” *Transport Reviews*, vol. 41, no. 4, pp. 393–400, 2021.

[42] J. M. Wooldridge, *Introductory Econometrics: A Modern Approach*, Cengage Learning, Massachusetts, MA, USA, 2015.

[43] M. G. Turner, R. H. Gardner, and R. V. O’neill, *Landscape Ecology in Theory and Practice*, Springer, New York, NY, USA, 2001.

[44] G. Zhai, K. Xie, and D. Yang D, “Assessing the safety effectiveness of citywide speed limit reduction: A causal inference approach integrating propensity score matching and spatial difference-in-differences,” *Transportation Research Part A: Policy and Practice*, pp. 94–106, 2022.

[45] J. M. Wooldridge, “Distribution-free estimation of some nonlinear panel data models,” *Journal of Econometrics*, vol. 90, no. 1, pp. 77–97, 1999.

[46] J. M. C. S. Silva and S. Tenreyro, “The log of gravity,” *The Review of Economics and Statistics*, vol. 88, no. 4, pp. 641–658, 2006.

[47] H. Yang, X. Lu, C. Cherry, X. Liu, and Y. Li, “Spatial variations in active mode trip volume at intersections: a local analysis utilizing geographically weighted regression,” *Journal of Transport Geography*, vol. 64, pp. 184–194, 2017.