Civil Engineering and Architecture 11(3): 1487-1497, 2023 http://www.hrpub.org
DOI: 10.13189/cea.2023.110330

Model Willingness to Use Public Transport in the USA Based on Socio-Economic and Demographic Characteristics

Ahmed Elkafoury[1], Maged Zagow[2,3], Khaled Saeed[4], Ahmed Mahmoud Darwish[5,*]

[1]Department of Public Works Engineering, Faculty of Engineering, Tanta University, Egypt
[2]Department of Architecture, Faculty of Architecture, Galala University, Galala 43713, Egypt
[3]Department of Architecture, Faculty of Engineering, Tanta University, Tanta 31111, Egypt
[4]Department of Civil Engineering, Faculty of Engineering, Aswan University, Egypt
[5]Department of Transportation Engineering, Faculty of Engineering, Alexandria University, Alexandria, Egypt
[*]Corresponding Author: Adarwish@alexu.edu.eg

Received December 26, 2022; Revised February 6, 2023; Accepted March 12, 2023

Cite This Paper in the Following Citation Styles

(a): *[1] Ahmed Elkafoury, Maged Zagow, Khaled Saeed, Ahmed Mahmoud Darwish, "Model Willingness to Use Public Transport in the USA Based on Socio-Economic and Demographic Characteristics," Civil Engineering and Architecture, Vol. 11, No. 3, pp. 1487 - 1497, 2023. DOI: 10.13189/cea.2023.110330.*

(b): *Ahmed Elkafoury, Maged Zagow, Khaled Saeed, Ahmed Mahmoud Darwish (2023). Model Willingness to Use Public Transport in the USA Based on Socio-Economic and Demographic Characteristics. Civil Engineering and Architecture, 11(3), 1487 - 1497. DOI: 10.13189/cea.2023.110330.*

Abstract Promoting public transport can increase the role of transport in sustainable development. Thus, studying the determinants of choosing public transport by travelers is crucial for transportation planning purposes, where developing an accurate model can help in examining any proposed scenarios. This paper aims to develop a multivariable regression model to describe the willingness to use public transport (W) represented as the percentage of people who use public transport in United States cities. First, census data of socio-economic and demographic characteristics are analyzed to identify significant factors for W to develop the model. Then, the regression technique is utilized to develop the model. The model is statistically assessed, in which the significance of all independent variables is examined and represented by a p-value. Moreover, the correlation between variables is examined. Then, the most statistically appropriate model for W is identified based on a set of performance measures such as coefficient of determination, average error, geometrical mean, Theil's inequality coefficient, and frictional bias. Finally, sensitivity analysis is conducted to assess the elasticity of the model to changes in the significant variables by considering a 10% change (increase or decrease) in the average of each variable.

Keywords Travel Demand Management (TDM), Willingness to Use Public Transport, Regression, Socio-Economic Characteristics

1. Introduction

The concentration of traffic can generate serious social, environmental, and economic problems. Traffic jams are constantly increasing due to increasing the demand for mobility resulting from population growth and economic progress [1]–[3]. Recently, the concept of sustainable development has been primarily applied in different cities through the introduction of Travel Demand Management (TDM) measures [4], which include restricting the use of private vehicles and promoting public transport [5], [6]. Indeed, having organized and efficient public transport would achieve many positive impacts, especially in the efficient use of public space, individual life, and the environment [7], [8]. Thus, modeling public transport demand is crucial for planning purposes, where an accurate model can be used as a tool to examine any proposed scenarios.

The term "transit-oriented development" (TOD) was created in the early 1990s by New Urbanist Peter Calthorpe [9], as well as TOD guidelines that were developed in Sacramento, San Diego, and Portland. TOD principles include prioritizing the walkability requirements, such as allocating mixed-use developments within a ten-minute walk of a transit station and having limited and managed parking inside the ten-minute walking area [10], [11]. Public transport demand relates to many attributes including the number of transfers, the volume of passengers, and the independent conditions of transport activity [12]. Generally, the factors that affect the traveler's choice of transport mode can be arranged into three groups [13]–[15]:

- Trip maker characteristics. These include the socio-economic characteristics of individuals, such as gender, age, income, car ownership, driving license availability, and household structure.
- Trip characteristics. These include the features and target of the demanded trip, such as trip purpose, trip origin, trip destination, and trip time throughout the day.
- Transportation mode characteristics. These include the features and availability of the transport mode, such as travel time, trip cost, availability, parking cost, frequency, comfort, and safety.

Several studies focused on the possible correlation between transport demand and related factors [16]. These include two approaches: the first approach depends on calibrating a function between the actual travel demand and the independent variables, such as demographic and socio-economic data of the study area, while the second approach is the third step in the transportation planning process (i.e., Transport Mode Choice Modeling), which is calibrating choice models based on travelers' interview surveys [17]. In the first approach, the identification of transport demand and its related factors can be expressed by creating regression models. It is an easy, applicable, and appropriate statistical tool for modeling the behavior of travelers in choosing their transport mode [18]. The second approach (i.e., choice models) has been traditionally developed based on travel behavior data obtained by direct observation or surveys (e.g., household surveys). A comparison of the chosen transport mode and the non-chosen modes reveals the travelers' preferences. These models reflect the behavior in choosing a transport mode among several available modes that connect to the exact origin and destination. By using appropriate statistical techniques (e.g., likelihood maximization), the implicit utility functions of the travelers can be figured out [19].

This paper will consider and apply the first approach (i.e., regression models). These models have many advantages over traditional mode choice modeling due to their simplicity in calibration and application, as mentioned before, results interpretation, direct response, low cost, and required information [20].

Numerous studies have applied this approach to model the public transport demand, where different factors that affect the public transport demand were addressed and various degrees of sophistication for mixtures of surveys and field data were considered. For instance, Kuby et al. [21] developed a multiple regression model in the USA to define factors that affect the demand for public transport (i.e., light-rail transit "LRT"), where 268 stations in nine different cities were chosen to collect the data regarding the weekday boarding travelers. The results indicated that the LRT ridership was mainly linked with land use characteristics, accessibility, employment condition, population, renters' percentage, bus lines condition, park-and-ride availability, and centrality.

Taylor et al. [22] calibrated a regression model to determine the possible correlation between public transit ridership and its related factors in urbanized areas, where the data were collected over 265 US areas. The analysis results indicated that the transit ridership among different areas was significantly linked with geography, economic condition, population, and the auto/highway structure. Cervero et al. [23] developed a direct ridership model to study the relationship between the demand for Bus Rapid Transit (BRT) in Southern California and its main related factors. The results indicated that increasing the BRT ridership was mainly related to frequency, intermodal connectivity, feeder system condition, population, and employment densities.

Cardozo et al. [20] developed a model using geographically weighted regression (GWR) to figure out the relation between the demand for the metro line in Madrid and the main related attributes. The results showed that the metro line was significantly affected by four main factors: the number of lines, workers, and employers, and the condition of suburban bus lines. Moreover, Ingvardson & Nielsen [24] modeled the function between the demand for public transport across 48 European cities and the related factors. To this end, a multiple regression model was calibrated, where the results showed that the main elements are the catchment area of the railway network, number of transfers, employment, Gross Domestic Product (GDP), population density, and job intensity. Guzman & Cardona [25] focused on Bus Rapid Transit (BRT) demand in Bogotá and its correlation with the built environment factors. They revealed that the BRT System is significantly affected by the characteristics of population and job densities.

This research aims to calibrate a regression model to determine the possible relation between the willingness of people to use public transport and the demographic and socio-economic characteristics of metropolitan areas in the USA. This model is crucial in the transport-oriented urban planning of new cities. This research paper is organized as follows: Section 2, after this introduction, presents details of the research methodology. Section 3 discusses the case study and data collection. The data is statistically described in Section 4. Section 5 comprises the details of the model development and performance assessment. The sensitivity of the model for significant variables is tested in Section 6. Finally, in Section 7, the work is concluded.

2. Research Methodology

The research tends to study the relationship between willingness to use public transport represented in the percentage of the population that uses public transportation (W) and the socio-economic and demographic data. The research methodology is explained in Figure 1.

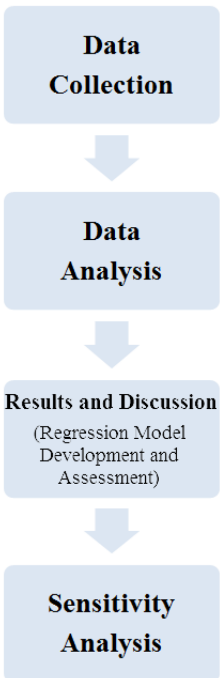

Figure 1. Research Methodology

The first stage is data collection for model development. The information includes the percentage of the population that uses public transportation in each TAZ, the percentage of occupied housing units to total housing units, the median age, the total number of business activities, and so on.

The collected data are statistically analyzed in terms of maximum and minimum values, range, standard deviation, median value, and average value. Then, the correlation matrix of variables considered in the analysis is figured out.

After the data are analyzed, the regression technique is utilized to develop the model that describes the willingness to use public transport as a function of significant socio-economic and demographic factors. The model is statistically assessed, in which the significance of all independent variables is examined and represented by a p-value, which has to be less than 0.05 for any variable considered in the model. Moreover, the correlated variables should be eliminated from the regression model. Then, the best model is statistically examined in terms of the coefficient of determination, average error, frictional bias (FB) of the regression model, etc. Finally, sensitivity analysis (W) is conducted to assess the elasticity of the model to changes in the significant variables by considering a 10% change (increase or decrease) in the average of each variable.

3. Data Collection

Data handled in this research is driven by the United States Census Bureau (US Census 1990, 2000, and 2010). This data is demographics and block census by zip code. Each zip code will be noted as a Transportation Analysis Zone (TAZ). USA Office of Management and Budget (OMB) classified the zones into metropolitan and micropolitan statistical zones using Census Bureau data. Zoning to metropolitan or micropolitan statistical areas depends mainly on the population nucleus and adjacent communities that are economically and socially connected with that nucleus. The Core-Based Statistical Area (CBSA) is determined to include at least one city with a population higher than 10,000, according to the 2010 criteria. At least one urbanized region with a population of 50,000 people must exist in each metropolitan. Each micropolitan statistical region should have an urban cluster of greater than 10,000 people but no more than 50,000. A Central County is an area with at least half of the residents living in urban areas with populations higher than 10,000. Also, any area with greater than 5,000 people living in a single 10,000-person urban area is defined as a Central County. For 2020 Census, it includes "outlying counties", which travel to or from the core counties. The data in this study include 15,290 TAZs, as shown in Table 1.

Table 1. Variables considered in the analysis and Data Source

Variables	Definition	Data Source
Dependent variables		
% Population uses public transportation (Y)	Percentage of the population using public transportation in each in each TAZ	'American Community Survey 2009-2013 "American Community Survey 5-year (ACS 5-yr) dataset" US Census Bureau, 2017. [26]
Independent variable (*Xi*)		
% Occupied Housing Units	Percentage of occupied housing units to total housing units in each TAZ	U.S. Census Bureau, 2016 [27]
Median age	Median age	
No. of Business Places	Total number of Businesses activities in each TAZ	
Transport (tCO2e/yr)	Amount of Carbon dioxide emission by public transportation in each TAZ	CoolClimate Network, 2022 [28]
Population density	Number of people per square mile in each TAZ	U.S. Census Bureau, 2016 [27]
No. of Retail	Number of Retail facilities in each TAZ	
No. of Education Places	Number of education facilities in each TAZ	
No. of Accommodation and Food Services Places	Number of accommodation and food services facilities in each TAZ	
No. of Arts, Entertainment, and Recreation Places	Number of arts, entertainment, and recreation facilities in each TAZ	
No. of Finance Places	Number of Finance facilities in each TAZ	
Mix factor	Index of the proportion of jobs to population and diversity of employment types	2006 national data set of street centerlines generated by TomTom that ships for ArcGIS [29]
No. of Health Places	Number of health building facilities in each TAZ	U.S. Census Bureau, 2016 [27]
Average Monthly Household Income	Average monthly income of households in each TAZ	
No. of Employees	Number of employees inhabiting each TAZ	

4. Data Analysis

Table 2 describes the statistics of the variables involved in the willingness to use public transport regression model development. It is noticed that the Average Monthly Household Income data achieves the highest standard deviation with 23,739.4 USD/month, while the lowest standard deviation value records 0.2 USD/month that is noticed for the percentage Occupied Housing Units data. The standard deviation of the data regarding the percentage of choosing public transportation (W) by travelers in each TAZ is 7.6%, with maximum, minimum, and average values of 77.9%, 0.1%, and 3.5%, respectively. The analysis of the correlation between all variables, as explained in Table 3, indicates that there is a high correlation (Correlation Factor (CC) > 0.5) between some pairs of independent variables. For example, the CC between No. of Educational Places and No. of Health Places is 0.63. A correlation factor of 0.73 holds between No. of Retail and No. of Employees in the TAZ. The highest correlation factor between two independent variables holds between No. of Finance Places and No. of Employees in the TAZ with a CC of 0.8. Correlation factors higher than 0.5 are grayed in Table 3.

Table 2. Descriptive statistics of variables considered in the analysis

Variables	Maximum	Minimum	Range	Standard Deviation	Median	Average
Mix Factor	316.0	0.0	316.0	82.5	114.0	100.5
No. of Retail	934.0	0.0	934.0	70.3	31.0	57.3
No. of Education Places	286.0	0.0	286.0	9.5	2.0	5.6
No. of Accommodation and Food Services Places	704.0	0.0	704.0	45.8	19.0	36.1
No. of Health Places	757.0	0.0	757.0	62.7	20.0	46.4
No. of Arts, Entertainment, and Recreation Places	848.0	0.0	848.0	17.7	3.0	6.7
No. of Finance Places	1274.0	0.0	1274.0	40.9	11.0	25.2
Population Density	143761.1	0.1	143761.1	6290.6	303.9	2193.3
No. of Business Places	7273.0	0.0	7273.0	484.0	229.0	405.3
% Occupied Housing Units	1.0	0.0	1.0	0.2	0.9	0.9
Transport (tCO_2e/yr)	45.6	0.0	45.6	4.3	16.3	16.5
Median Age	74.3	18.8	55.5	6.1	39.9	39.5
W	77.9	0.1	77.8	7.6	1.0	3.5
Average Monthly Household Income	228487.0	0.0	228487.0	23739.4	49390.5	54715.6
No. of Employees	150697.0	0.0	150697.0	9531.2	2636.0	6253.9

Table 3. Correlation matrix of variables considered in the analysis

Variables	Mix Factor	No. of Retail	No. of Educational Places	No. of Accommodation and Food Services Places	No. of Health Places	No. of Arts, Entertainment, and Recreation Places	No. of Finance Places	Population Density	No. of Business Places	% Occupied Housing Units	Transport (tCO2e/yr)	Median Age	% Population uses public transportation	Average Monthly Household Income	No. of Employees
Mix Factor	1.00														
No. of Retail	0.41	1.00													
No. of Educational Places	0.44	0.61	1.00												
No. of Accommodation and Food Services Places	0.47	0.86	0.66	1.00											
No. of Health Places	0.45	0.78	0.63	0.77	1.00										
No. of Arts, Entertainment, and Recreation Places	0.26	0.40	0.40	0.49	0.44	1.00									
No. of Finance Places	0.36	0.68	0.58	0.72	0.72	0.43	1.00								
Population Density	0.46	0.35	0.34	0.41	0.32	0.27	0.23	1.00							
No. of Business Places	0.51	0.86	0.73	0.89	0.84	0.55	0.83	0.37	1.00						
% Occupied Housing Units	0.27	0.14	0.16	0.13	0.18	0.07	0.12	0.10	0.18	1.00					
Transport (tCO2e/yr)	-0.21	-0.29	-0.15	-0.33	-0.24	-0.14	-0.17	-0.37	-0.25	0.15	1.00				
Median Age	-0.31	-0.25	-0.18	-0.26	-0.20	-0.05	-0.14	-0.21	-0.24	-0.22	0.23	1.00			
% Population uses public transportation	0.47	0.25	0.30	0.32	0.22	0.18	0.14	0.78	0.28	0.08	-0.42	-0.19	1.00		
Average Monthly Household Income	0.30	0.04	0.24	0.09	0.13	0.17	0.18	0.04	0.19	0.23	0.58	0.18	0.06	1.00	
No. of Employees	0.44	0.73	0.62	0.80	0.70	0.44	0.80	0.30	0.89	0.13	-0.27	-0.27	0.25	0.12	1.00

5. Results and Discussion

To understand the effect of related variables on willingness to use public transport (W), a statistical regression model is utilized as a modeling tool, where W is considered the dependent variable. The independent variables, utilized for each TAZ, include the number of businesses, mix factor, population density…etc. The regression model takes the following general form:

$$W = \sum_{i=1}^{n}(a_i X_i) + \varepsilon \qquad (1)$$

Where:

W : willingness to use public transport (% of the population uses public transport) in the TAZ.

X_i: independent explanatory variables *i*.

ε: error term.

a_i: corresponding coefficient.

The significance of W in each TAZ has been tested for all independent variables. As shown in Table 4, the significance value is represented by the p-value, which needs to be less than 0.05 for any Xi considered in the model, indicating that there is less than a 5% likelihood that the null hypothesis could have been true [30]. It is found that five factors; No. of Accommodation and Food Service, No. of Business Places, Median age, Transport (tCO2e/yr), and % Occupied Housing Units have p-values greater than 0.05. Thus, these five factors are eliminated from the modeling process as they have no significance in the model.

Table 4. Significance of independent variables with respect to willingness to use public transport

Independent variable	P-Value
Mix Factor	< 0.0001
No. of Retail	< 0.0001
No. of educational Places	< 0.0001
No. of Accommodation and Food Service	0.43
No. of Health Places	< 0.0001
No. of Arts, Entertainment, and Recreation Places	< 0.0001
No. of Finance Places	< 0.0001
Population Density	< 0.0001
No. of Business Places	0.2
% Occupied Housing Units	0.831
Transport (tCO2e/yr)	0.0865
Median Age	0.968
Average Monthly Household Income	< 0.0001
No. of Employees	< 0.0001

Further assessment for the included variables is conducted by examining the multi-collinearity, where the correlated variables should be eliminated from the regression model to avoid the multi-collinearity effect that may distort the results obtained from regression analysis. Accordingly, the adopted methodology for each highly correlated independent variable is based on keeping the variable with a higher correlation with the independent variable for further modeling steps, while the other variable is excluded from the model. This is measured by the Variance Inflation Factor (VIF), which measures the correlation and strength between independent variables.

The minimum value of VIF for any variable is 1, where this value indicates that this independent variable does not correlate with any others. If the value of VIF lies between 1 and 5, then there is a moderate correlation, while significant multi-collinearity levels of a variable exist if it has a VIF of more than 5. The VIF results are shown in Table 5. The results indicate that all variable has a VIF of less than 5. The maximum VIF value is 3.61 noticed for No. of Employees in the TAZ. Thus, all significant variables will be included in the model.

Table 5. Variance Inflation Factor of significant variable

Term	VIF
Mix Factor	1.65
No. of Retail	3.23
No. of Education Places	2.06
No. of Health Places	3.31
No. of Arts, Entertainment, and Recreation Places	1.36
No. of Finance Places	3.32
Population Density	1.39
Average Monthly Household Income	1.21
No. of Employees	3.61

The stepwise regression technique is used to model the willingness to use public transport as a function of significant socio-economic and demographic data. The results of the models are shown in Table 5. It is found that the mix factor and no. of education places have the highest two positive effects on willingness to use public transport in sequence. This is reasoned that as the mix factor increases, a high proportion of the working-age population is employed. This raises the daily job-type trips, which are related to the usage of public transport [31]. On the other hand, although the absolute number of employees affects the willingness of people to use public transport, it records the lowest positive effect. The number of finance places in the TAZ has the maximum negative effect on the willingness to use public transport in the zone.

Table 5. Regression modeling for willingness to use public transport

Term	Coefficient	R^2
Constant	0.777	0.72
Mix Factor	0.014751	
No. of Retail	-0.005459	
No. of Education Places	0.05248	
No. of Health Places	-0.00621	
No. of Arts, Entertainment, and Recreation Places	-0.01635	
No. of Finance Places	-0.02023	
Population Density	0.000867	
Average Monthly Household Income	-0.000003	
No. of Employees	0.000076	

The statistical indicators are utilized as follows:

- The model has a relatively high correlation coefficient (R2= 0.72), which indicates a high representativeness of the model.
- The average error in estimating W values is 0.023 with a standard deviation of 4.61.
- The statistical relation between modeled and measured values of willingness to use public transport (W) (Figure 2) shows a good coefficient of determination of 0.63, which indicates good fitness of the model [32].
- The relative error between the average of surveyed and modeled values of W can be expressed as:

$$100 * \left(\frac{\overline{W_{modeled}} - \overline{W_{surveyed}}}{\overline{W_{surveyed}}}\right) \quad (2)$$

It shows a small difference (0.64%), which indicates that the regression model generally describes well the surveyed data.

- The frictional bias (FB) of the regression model is 0.0064. This indicates a relatively very small overall overestimation of the model as this value ranges between 2 and -2 [33]. FB can be estimated as follows:

$$2\left(\frac{\overline{W_{modeled}} - \overline{W_{surveyed}}}{\overline{W_{modeled}} + \overline{W_{surveyed}}}\right) \quad (3)$$

- The existence of systematic biased can be assessed statistically by the value of the Geometrical Mean, as follows:

$$\mathrm{GM} = \exp(\overline{\ln W_{surveyed}} - \overline{\ln W_{modeled}}) \quad (4)$$

GM of the model has a value of 0.04 while the ideal value for it is 1.0 [33]. This means that the regression model accurately simulates the surveyed data with acceptable relative biases.

- Theil's inequality (T) coefficient is also used in the study to determine how well the model's estimated results explain the actual data. Theil's inequality coefficient is a statistic performance measure related to the root mean square forecast error, which ranges between zero and one. Equation 5 describes T factor estimation, where N is the number of observations. The value of T for the developed model is 0.15 indicating a reasonable representation of actual data. The perfect representation is for T=0, while the worst is for T=1 [34].

$$T = \frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(w_{modeled} - w_{surveyed}\right)^2}}{\sqrt{\frac{1}{N}\sum_{n=1}^{N}(w_{modeled})^2} + \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(w_{surveyed}\right)^2}} \quad (5)$$

6. Sensitivity Analysis

Socio-economic and demographic data affect significantly the travelers' willingness to use public transport on their daily trips. To study the effect of the change of one indicator on W, the average value of W is calculated using Equation (1) based on the average value of independent variables. Then, the effect of a 10% change (increase and decrease) of the average of each variable on the average value of W is assessed. The sensitivity analysis results, presented in Figure 3, show that the factor that has a significant effect on changes in the average value of W is population density, since a 10% change (increase or decrease) in average population density changes the average W absolute value by about 5.4%. The second most remarkable change in average W can occur by increasing or decreasing the mix factor by 10%. Changes in average values of the number of arts, entertainment, and recreational places; and changes in average monthly household income have minor effects on the average value of W.

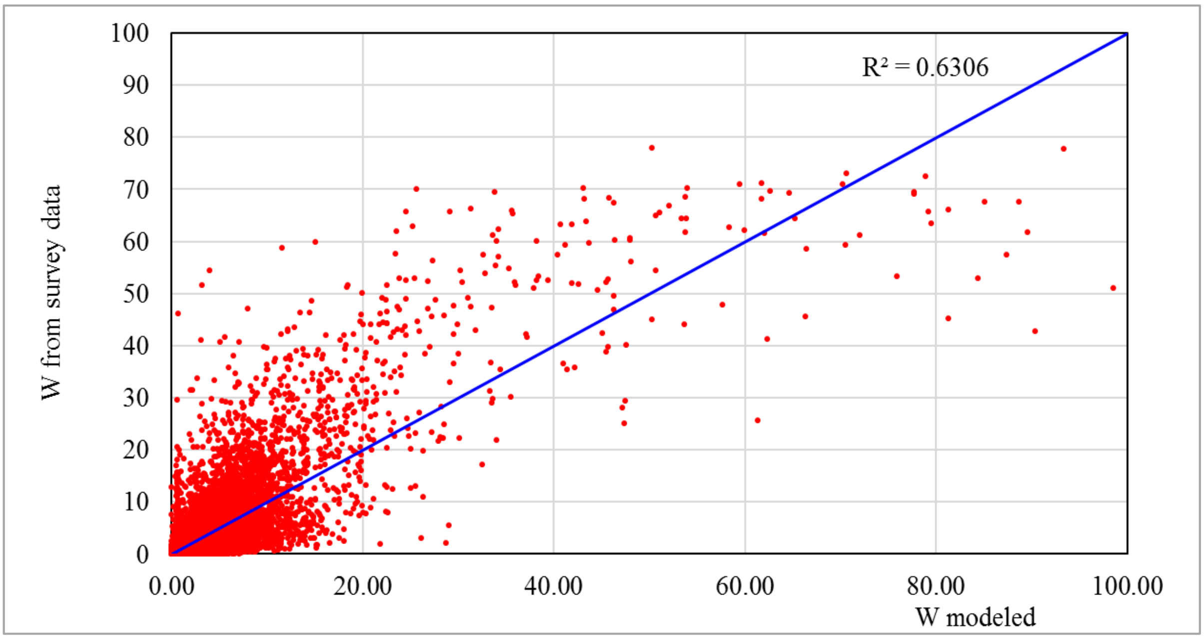

Figure 2. Plot of modeled values of willingness to use public transport against surveyed values

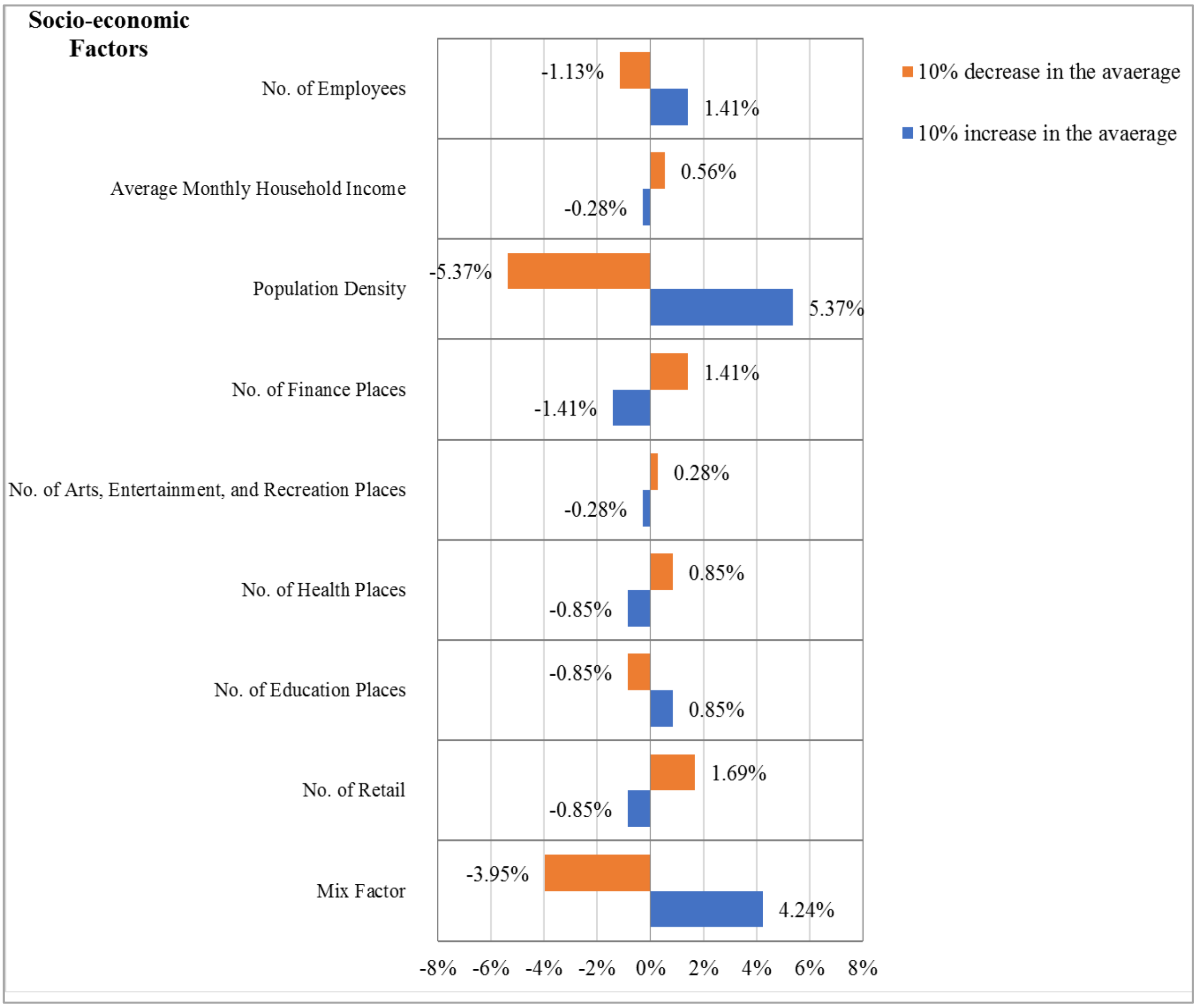

Figure 3. Effect of changes of average values of socio-economic and demographic variables on the average value of willingness to use public transport estimated using the regression model

7. Conclusions

Trip generation rates are essential knowledge for planners and engineers working in urban planning and transportation to evaluate the implications of land-use policies. Based on the assumption that trip generation and willingness to use public transport are correlated with the socio-economic and demographic characteristics of the area, a multivariable regression model is developed to describe the percentage of people who uses public transport service (W) in United States cities. Different characteristics are analyzed to identify significant and controlling factors for willingness to use public transport services. The multivariable regression technique is utilized to produce the most statistically appropriate mathematical model for willingness to use public transport. Mix factor, no. of retail, no. of education places, no. of health places, no. of arts, entertainment, and recreation places, no. of finance places, population density, average monthly household income, and no. of employees are involved in the model with a coefficient of determination (R^2) of 0.72. The assessment of the model utilized statistical parameters to assess the predicted values against census values. It is found that the average error in estimating W values is 0.023 with a standard deviation of 4.61 and a relation between modeled and measured values of willingness to use public transport of 0.63 coefficient of determination. The relative error between the average of surveyed and modeled values of W is 0.64% showing that the regression model generally describes well the surveyed data. The model indicates a relatively very small overall overestimation of the model with a geometrical mean of 0.04 and Theil's inequality (T) coefficient of 0.15. The sensitivity of W to changes in the average value of controlling parameters is measured. It is found that the factor that significantly affects the average value of W is population density followed by mix factor, since a 10% change (increase or decrease) in the average population density changes the average W absolute value by about 5.4%. In comparison, a 10% increase and decrease in the average mix factor changes the average W value by about 4.24% and -3.95%, respectively.

Funding

This research received no external funding.

Data Availability Statement

“Not applicable”.

Conflicts of Interest

The authors declare no conflict of interest.

REFERENCES

[1] K. Goldmann and G. Sieg, “Economic implications of phantom traffic jams: evidence from traffic experiments,” *Transportation Letters*, vol. 12, no. 6, pp. 386–390, Jul. 2019, doi: 10.1080/19427867.2019.1611077.

[2] V. Palevičius, G. M. Paliulis, J. Venckauskaite, and B. Vengrys, “Evaluation of the requirement for passenger car parking spaces using multi-criteria methods,” *Journal of Civil Engineering and Management*, vol. 19, no. 1, pp. 49–58, Feb. 2013, doi: 10.3846/13923730.2012.727463.

[3] J. Čarský, A. Mačerinskienė, J. Carsky, and A. Macerinskiene, “Modern Ways of Designing Roads Through Urban Areas,” *Journal of Civil Engineering and Management*, vol. 9, no. 3, pp. 208–213, 2012, doi: 10.1080/13923730.2003.10531328.

[4] N. Huan, S. Hess, and E. Yao, “Understanding the effects of travel demand management on metro commuters' behavioural loyalty: a hybrid choice modelling approach,” *Transportation (Amst)*, vol. 49, no. 2, pp. 343–372, Apr. 2022, doi: 10.1007/S11116-021-10179-3/FIGURES/7.

[5] N. T. Ratrout, U. Gazder, and K. J. Assi, “Effect of public transportation in reducing passenger car trips to schools in Al-Khobar–Dhahran metropolitan area, Saudi Arabia,” *Transportation Letters*, vol. 10, no. 1, pp. 43–51, Jan. 2016, doi: 10.1080/19427867.2016.1223927.

[6] M. Diao, “Towards sustainable urban transport in Singapore: Policy instruments and mobility trends,” *Transp Policy (Oxf)*, vol. 81, pp. 320–330, Sep. 2019, doi: 10.1016/J.TRANPOL.2018.05.005.

[7] D. Toro-González, V. Cantillo, and V. Cantillo-García, “Factors influencing demand for public transport in Colombia,” *Research in Transportation Business & Management*, vol. 36, p. 100514, Sep. 2020, doi: 10.1016/J.RTBM.2020.100514.

[8] A. Melkonyan, T. Gruchmann, F. Lohmar, and R. Bleischwitz, “Decision support for sustainable urban mobility: A case study of the Rhine-Ruhr area,” *Sustain Cities Soc*, vol. 80, p. 103806, May 2022, doi: 10.1016/J.SCS.2022.103806.

[9] P. Calthorpe, *The Next American Metropolis: Ecology, Community, and the American Dream - Peter Calthorpe - Google Books*. Princeton architectural press, 1993.

[10] R. Cervero, *Transit-oriented Development in the United States: Experiences, Challenges ... - Robert Cervero, Transit Cooperative Research Program - Google Books*, Vol. 102. Transportation Research Board, 2004.

[11] M. Ušpalyte - Vitkūniene, R. and Burinskiene, “Analysis of the dynamics of walking distances to public transport routes and its influence on housing prices,” *Journal of Civil Engineering and Management*, vol. XII, no. 3, pp. 261–267, 2006, doi: 10.1080/13923730.2006.9636401.

[12] T. van Vuren, “Modeling of transport demand – analyzing, calculating, and forecasting transport demand,” *https://doi.org/10.1080/01441647.2019.1635226*, vol. 40, no. 1, pp. 115–117, Jan. 2019, doi:

10.1080/01441647.2019.1635226.

[13] J. de D. Ortúzar and L. G. Willumsen, *Modelling transport*, 4th ed. Chichester: Wiley-Blackwell, 2011. doi: 10.1002/9781119993308.

[14] M. C. J. Bliemer, C. Mulley, and C. J. Moutou, *Handbook on transport and urban planning in the developed world*. Edward Elgar Publishing Ltd., 2016. doi: 10.4337/9781783471393.

[15] C. A. O'Flaherty, *Transport Planning and Traffic Engineering*. 2018.

[16] T. K. Ojo, "Quality of public transport service: an integrative review and research agenda," *Transportation Letters*, vol. 11, no. 2, pp. 104–116, Mar. 2017, doi: 10.1080/19427867.2017.1283835.

[17] O. Petrik, J. De Abreu e Silva, and F. Moura, "Stated preference surveys in transport demand modeling: disengagement of respondents," *Transportation Letters*, vol. 8, no. 1, pp. 13–25, 2016, doi: 10.1179/1942787515Y.0000000003.

[18] V. Konecný, M. Brídziková, and P. Marienka, "Research of bus transport demand and its factors using multicriteria regression analysis," *Transportation Research Procedia*, vol. 55, pp. 180–187, Jan. 2021, doi: 10.1016/J.TRPRO.2021.06.020.

[19] F. Heiss, "Discrete Choice Methods with Simulation," *http://dx.doi.org/10.1080/07474938.2014.975634*, vol. 35, no. 4, pp. 688–692, Apr. 2016, doi: 10.1080/07474938.2014.975634.

[20] O. D. Cardozo, J. C. García-Palomares, and J. Gutiérrez, "Application of geographically weighted regression to the direct forecasting of transit ridership at station-level," *Applied Geography*, vol. 34, pp. 548–558, May 2012, doi: 10.1016/J.APGEOG.2012.01.005.

[21] M. Kuby, A. Barranda, and C. Upchurch, "Factors influencing light-rail station boardings in the United States," *Transp Res Part A Policy Pract*, vol. 38, no. 3, pp. 223–247, Mar. 2004, doi: 10.1016/J.TRA.2003.10.006.

[22] B. D. Taylor, D. Miller, H. Iseki, and C. Fink, "Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas," *Transp Res Part A Policy Pract*, vol. 43, no. 1, pp. 60–77, Jan. 2009, doi: 10.1016/J.TRA.2008.06.007.

[23] R. Cervero, J. Murakami, and M. Miller, "Direct Ridership Model of Bus Rapid Transit in Los Angeles County, California:," *Transp Res Rec*, no. 2145, pp. 1–7, Jan. 2010, doi: 10.3141/2145-01.

[24] J. B. Ingvardson and O. A. Nielsen, "How urban density, network topology and socio-economy influence public transport ridership: Empirical evidence from 48 European metropolitan areas," *J Transp Geogr*, vol. 72, pp. 50–63, Oct. 2018, doi: 10.1016/J.JTRANGEO.2018.07.002.

[25] L. A. Guzman and S. Gomez Cardona, "Density-oriented public transport corridors: Decoding their influence on BRT ridership at station-level and time-slot in Bogotá," *Cities*, vol. 110, p. 103071, Mar. 2021, doi: 10.1016/J.CITIES.2020.103071.

[26] US Census Bureau, "Now available: 2009-2013 ACS 5 Year Summary Files and Data Profiles," 2017. https://www.census.gov/data/developers/updates/acs-5-yr-summary-available-2009-2013.html (accessed Mar. 25, 2022).

[27] U.S. Census Bureau, "North American Industry Classification System (NAICS)." 2016.

[28] CoolClimate Network, "Smart Tools for a Cooler Planet," 2022. https://coolclimate.berkeley.edu/index (accessed Jan. 25, 2022).

[29] S. Ewing, R. and Hamidi, "Measuring Urban Sprawl and Validating Sprawl Measures," 2010.

[30] M. A. Khan, S. F. Bokhari, A. Khan, M. S. Amjad, A. M. Butt, and M. Z. Rafique, "Clean and sustainable transportation through electric vehicles — a user survey of three-wheeler vehicles in Pakistan," *Environmental Science and Pollution Research*, vol. 1, pp. 1–18, Feb. 2022, doi: 10.1007/S11356-022-19060-X/TABLES/8.

[31] Y. C. Chiou, R. C. Jou, and C. H. Yang, "Factors affecting public transportation usage rate: Geographically weighted regression," *Transp Res Part A Policy Pract*, vol. 78, pp. 161–177, Aug. 2015, doi: 10.1016/J.TRA.2015.05.016.

[32] A. Elkafoury, A. M. Negm, M. H. Aly, M. F. Bady, and T. Ichimura, "Develop dynamic model for predicting traffic CO emissions in urban areas," *Environmental Science and Pollution Research*, vol. 23, no. 16, pp. 15899–15910, Aug. 2016, doi: 10.1007/S11356-015-4319-8/FIGURES/6.

[33] Misra, M. J. Roorda, and H. L. MacLean, "An integrated modelling approach to estimate urban traffic emissions," *Atmos Environ*, vol. 73, pp. 81–91, Jul. 2013, doi: 10.1016/J.ATMOSENV.2013.03.013.

[34] T. Y. Jang, "Count data models for trip generation," *J Transp Eng*, vol. 131, no. 6, pp. 444–450, Jun. 2005, doi: 10.1061/(ASCE)0733-947X(2005)131:6(444).