INTERNATIONAL JOURNAL OF ARCHITECTURAL ENGINEERING AND URBAN RESEARCH

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77– 92. www.egyptfuture.org/ojs/

Cars, Community, and the Urban Landscape: Exploring the Spatial Patterns of Car Ownership and Associated Factors Using Geographically Weighted Regression

Mohammed Mansour Gomaa

Department of Architectural Engineering, Faculty of Engineering, Aswan University, Aswan 81542, Egypt.

Email: mmgomaa@aswu.edu.eg

Abstract:

A growing body of literature has delved into the relationships between the number of cars in households and a range of relative factors. Most previous research has focused on the relationships between car ownership and individuals' travel patterns, socio-demographics, and economic factors. However, sufficient studies on simultaneously identifying the spatial patterns and associated factors are lacking. This paper examines the relationship between the number of cars households own and various factors, focusing specifically on spatial patterns and variables that may influence car ownership. Using data from the 2017 National Household Travel Survey, the impact of housing and population density, as well as transit accessibility, on car ownership has been analyzed using OLS and GWR regression models. Moreover, to test the spatial autocorrelation of car ownership and examine its associative factors, statistical methods have been used. Global Moran's I and Getis-Ord general G tests were used to analyze car ownership data's spatial autocorrelation. Five key variables - housing density, population density, proximity to bus stops and bus stations, and distance to key locations – significantly impact car ownership, suggesting that areas with a high rate of car ownership tend to be clustered together.

Keywords: Urban Landscape, Car ownership, Geographically Weighted Regression, spatial autocorrelation, GIS.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77 – 92. www.egyptfuture.org/ojs/

1. Introduction

The research on the correlation between the number of cars in households and various other factors is familiar (Golob, 1990; Raphael & Rice, 2002). A growing body of literature has investigated these relationships. These studies can be broken down into two major categories. Several studies have examined the relationships between various variables and vehicle ownership, viewed as an intermediary factor or a key driver behind other variables (Clark, 2007; Giuliano & Dargay, 2006; Golob, 1990; Raphael & Rice, 2002). Golob (1990) analyzed the relationship between vehicle ownership, weekly travel time by car, weekly travel time by public transit, and weekly travel time by nonmotorized modes. Using panel data, Golob found that car ownership and the other three variables have mutually causal relationships (Figure 1). This indicates that increased car ownership is driven by households' desire to reduce time expenditure and cost and benefit factors. The study also notes that a shift to a more expensive but less time-consuming mode of transportation would likely increase the number of cars households own in the short term.

Previous research has demonstrated that, in the long run, changes in car ownership can impact where households choose to live (Giuliano & Dargay, 2006; Raphael & Rice, 2002). For instance, Raphael et al. (2002) examined whether the positive correlation between vehicle ownership and employment outcomes is a result of vehicle ownership. They found that having access to a vehicle positively impacted employment status, work hours, and wages by using these three factors as dependent variables in their study. This was shown in all three OLS regression models, where the coefficients for auto ownership were significant.

Without a comprehensive understanding of the built environment's effects, most previous research has focused on the relationships between car ownership and individuals' travel patterns, socio-demographics, and economic factors. Additionally, numerous articles investigate factors associated with car ownership as the dependent

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77_92. www.egyptfuture.org/ojs/

variable (Bhat & Guo, 2007; Cao et al., 2007; Chin & Smith, 1997; Clark, 2007; Giuliano & Dargay, 2006; Raphael & Rice, 2002; Sheller & Urry, 2000). Some literature examined the relationship between car ownership and the built environment using ordered probit and static-score models. They found that the number of vehicles in households is primarily determined by socio-demographic factors, with minimal influence from the built environment (Cao et al., 2007).

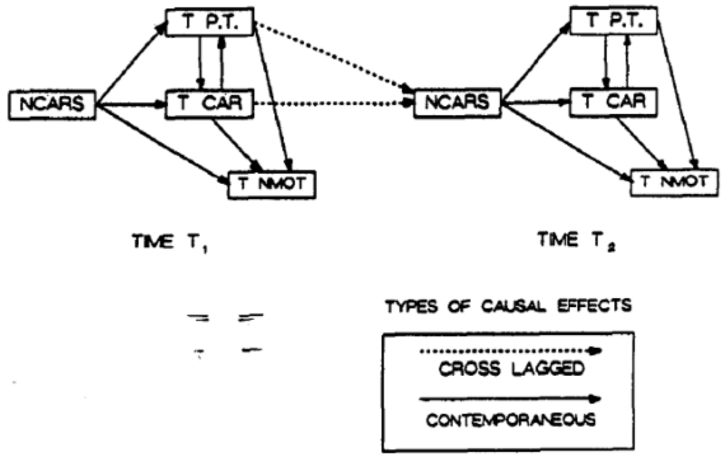


Figure 1. Flow diagram of causal linkages between endogenous variables (Golob, 1990).

Other researchers studied the same theme using a different definition of the built environment, considering factors such as land-use types, urban forms, street networks, and land-use diversity index (Bhat & Guo, 2007). Their findings are consistent with those of Cao et al. (2007), they discovered that these characteristics influence not only car ownership rates but also where households choose to live. They also found that socio-demographic and built environment characteristics are significant determinants of car ownership decisions (Bhat & Guo, 2007; Cao et al., 2007; Chin & Smith, 1997). However, these studies have some limitations of not considering the potential spatial pattern of car ownership, which could be a valuable addition to the modeling process.

How land-use development and commuting patterns affect car ownership has been the subject of extensive research for decades (Bhat & Guo, 2007; Cao et al., 2007; Clark, 2007; Giuliano & Dargay, 2006; Paleti et al., 2013; Pinjari et al., 2011). However, there has been a lack of research examining spatial patterns and car ownership factors. This paper seeks to address this limitation by investigating the spatial nature of car ownership and its relationship to housing, population density, and transit accessibility.

As car ownership remains a concern for the general public and the academic community, it is essential to comprehend the patterns and factors influencing the number of vehicles in households. The economic benefits of car ownership for the automobile industry and local governments are well known. However, excessive car ownership can also adversely affect traffic congestion and air pollution (Smith, 1992). Previous research has demonstrated a correlation between high car ownership rates and these problems. Consequently, it is essential to investigate whether car ownership is concentrated in certain areas, what factors contribute to high car ownership rates, and where the "hot spots" and "cool zones" of car ownership are located.

Therefore, this study employs a Graphically Weighted Regression (GWR) model to incorporate the spatial patterns of car ownership into the modeling process. Previous research has utilized GWR models to investigate factors associated with car ownership (Brunsdon et al., 1996; Chiou et al., 2015; Wang & Chen, 2017; Zhao & Park, 2004). This method incorporates spatial patterns of automobile ownership into the modeling procedure. However, the GWR methodology is limited because it needs to provide meaningful global results (Brunsdon et al., 1996; Chiou et al., 2015; Ma & Gopal, 2018; Wang & Chen, 2017).

Despite extensive research on the factors associated with automobile ownership, little effort has been devoted to identifying its spatial patterns. In addition, these studies do not simultaneously examine car ownership statistically and spatially. This study addresses

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77_92. www.egyptfuture.org/ojs/

these knowledge gaps by comparing the results of spatial autocorrelation tests of car ownership across the study area and exploring the effect of population density, housing density, and transit accessibility on the number of cars in households using multiple regression models (OLS, explanatory and GWR models).

2. Data and Methods

This study analyzes data from three South Florida counties: Broward, Palm Beach, and Miami-Dade. The data used in this analysis is from the 2017 National Household Travel Survey (NHTS), which provides insights into the participating households' travel patterns and socio-demographic characteristics. The focus group of the NHTS used in this study is the civilian and non-institutionalized population, comprised of households. Furthermore, the characteristics of the built environment are obtained from land-use parcel data from the Florida Department of Revenue and processed by the University of Florida's Institute of Transportation Engineers. This data focuses primarily on land use categories and transit accessibility. Following the data cleaning process and overlaying with Census Tract map files, a sample of 3,980 households has been determined. This study's analysis and data organization are based on this sample.

Statistical methods are utilized to test spatial autocorrelation of car ownership and explore associative factors affecting the number of vehicles per household. Specifically, analyzing patterns using global Moran’s I and Getis-Ord general G tests were used to examine the spatial autocorrelation of car ownership data. In addition, we investigated the hot spots or cool zones using Local Moran’s I statistics. The specifications of these models or methods are introduced in each part of the analysis of the next section, empirical results. Table 1 shows the model structure, including the dependent and independent variables used.

Table 1. Model Structure and Specifications

Dependent Variable	*Independent Variables*
Car Ownership (Number of Vehicles per household)	1- Transit Accessibility: Number of the bus station in 1-mile buffer Distance to the nearest bus stop (1000 meters) Total length of bus route in 0.5-mile buffer (1000 meters) 2- Density: Job density block level (1000/sq mile) House density block level (1000/sq mile) Population density block level (1000/sq mile) 3- Location Attributes: Distance to Regional Activity center (miles) Distance to closest residential center (miles)

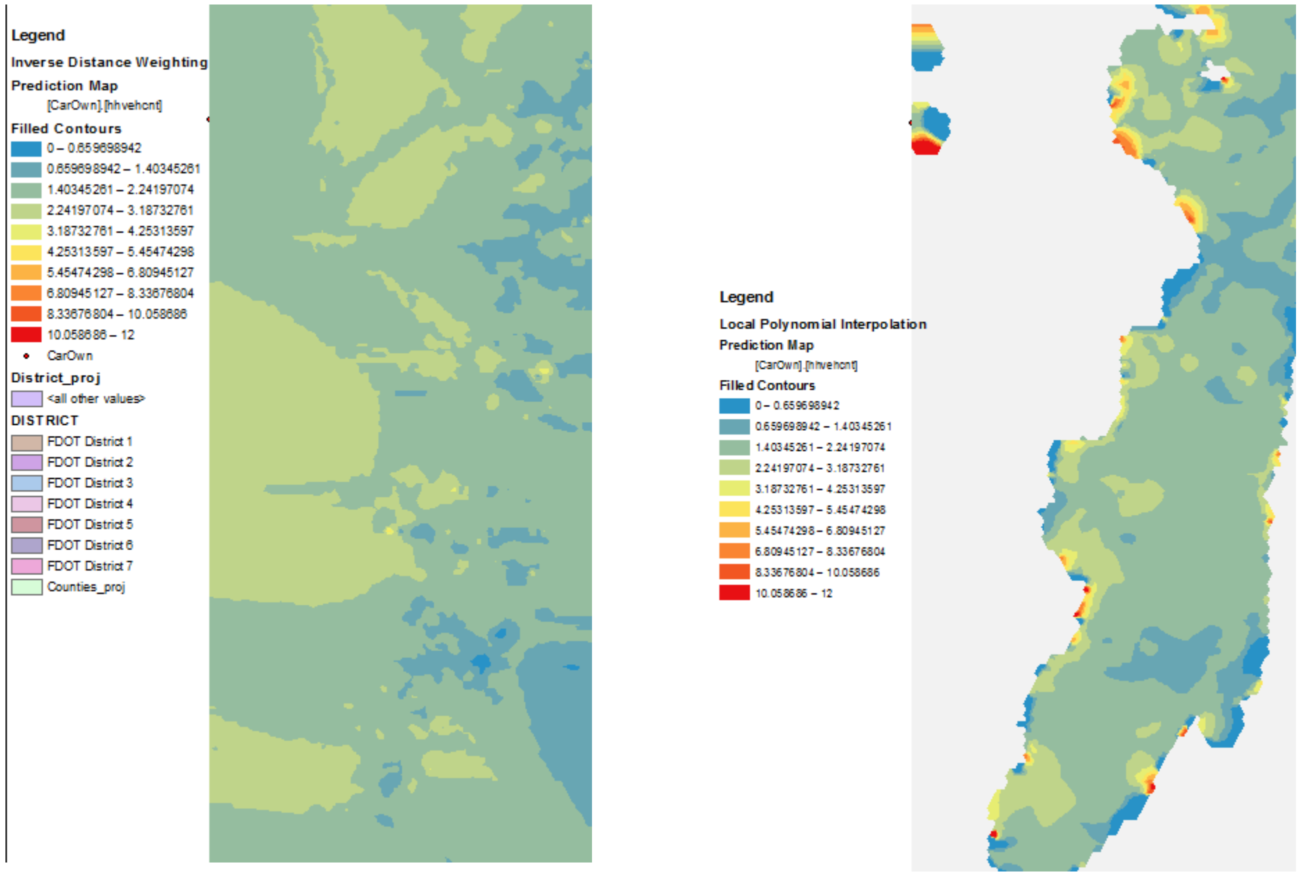
3. Results and Discussion

Data Interpolation Processes

Different iterations have been conducted using IDW and local polynomials to find the best prediction model for car ownership. Figure 2 shows two models developed for this purpose during the iteration process.

INTERNATIONAL JOURNAL OF ARCHITECTURAL ENGINEERING AND URBAN RESEARCH

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77– 92. www.egyptfuture.org/ojs/



a) b)

Figure 2. Interpolating the Data; a) using IDW and b) using local polynomial.

Then, we compared those two models to develop the best prediction model. Based on the comparison, the second model using a local polynomial performs better than the first IDW one as it has a smaller root mean square (Figure 3).

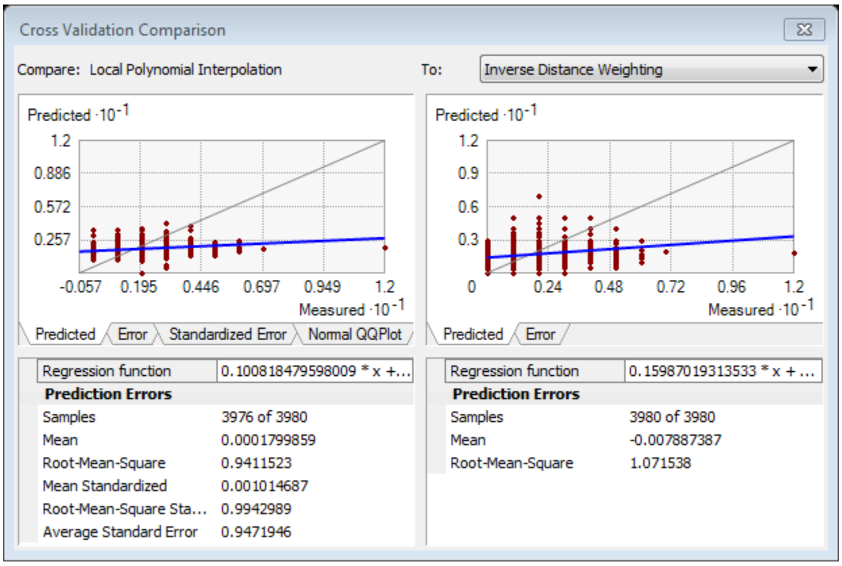


Figure 3. Comparing the two models between the IDW and local polynomial.

Analyzing Patterns and Clusters of the Data

As shown in Figure 4, a global Moran's I and Getis-Ord test have been conducted. The spatial autocorrelation (Moran's I) tool has been used to analyze whether the data are clustered. The null hypothesis is that there is no spatial autocorrelation between the variables. From the results below, Moran's I is 0.227, close to +1, and Z-score is 19.66, indicating we can reject the null hypothesis at 95% confidence level. That means the residuals are clustered and not randomly distributed. We got the same results when we ran the high/low clustering (Getis-Ord general G test).

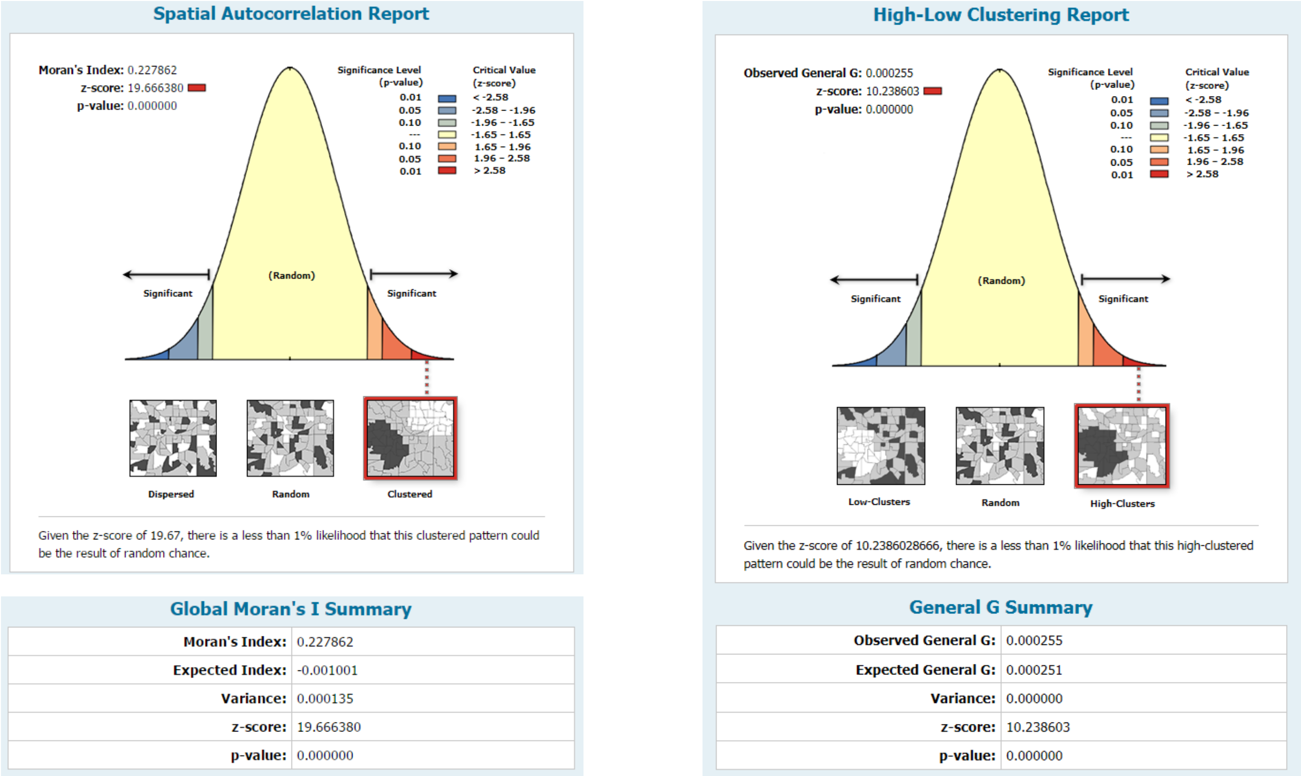


Figure 4. Analyzing the patterns of the data (Moran's I and G* statistics).

Modeling Specification and Regression analysis

In this step, we conduct regression analysis to understand the impact of population, housing density, and transit accessibility on car ownership. The overall model structure and procedures are presented in Figure 5. In the first step, we conduct explanatory regression to have different iterations and understand which variables are statistically significant. Then, we ran OLS regression using only the statistically significant variables from the previous step. After analyzing the OLS regression results, we found that we still have an issue with spatial dependence between the dependent and independent variables. Therefore, in the final step, we conduct GWR to capture the spatial component in the model. In the following sections, we will present the outputs for each step. Then, we will deliver the final model we got and the discussion of the variable effects of this model.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77– 92. www.egyptfuture.org/ojs/

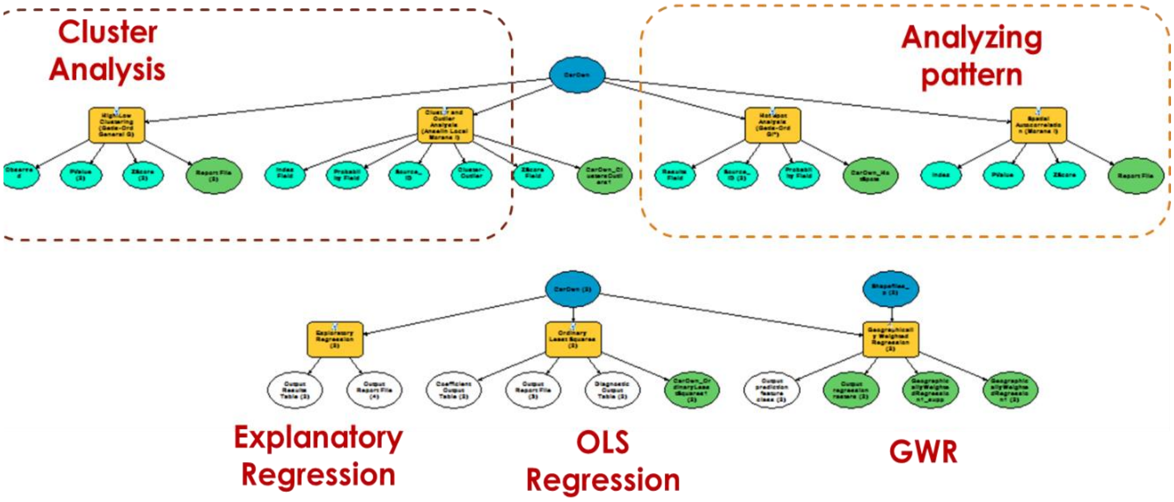


Figure 5. Overall model structure.

Explanatory Regression

As shown in Figure 6, the model's fitness indicates that the R2 is about .05, and VIF is 5.35. On the other hand, the Jarque-bera p-value is significant, which means the model has an issue with spatial dependence among the variables. The summary of variable significance also shows that four variables turned out to be statistically significant: housing density, bus stop within 1 mile, distance to nearest bus stations, and the distance to action centers. Population density is also statistically significant at 77% confidence level. Therefore, we decide to proceed with those five significant variables to include them in the final model. As shown in the multicollinearity summary, all the variables have a VIF score lower than 7.5, which means we do not have an issue with multicollinearity in our model.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77_92. www.egyptfuture.org/ojs/

```
                                                     Highest Adjusted R-Squared Results
AdjR2      AICc    JB K(BP)  VIF   SA   Model
 0.05 10918.17 0.00  0.27 5.35 0.00  -HOUDENBLK***  +POPDENBLK***  -JOBDENSBLK  -BUS1MILE***  +DISTONEARM***
       Passing Models
AdjR2 AICc JB K(BP) VIF SA   Model

******************************************************************************
************** Exploratory Regression Global Summary (HHVEHCNT) **************

              Percentage of Search Criteria Passed
                   Search Criterion Cutoff Trials # Passed % Passed
             Min Adjusted R-Squared > 0.50     255        0     0.00
            Max Coefficient p-value < 0.05     255       93    36.47
                        Max VIF Value < 7.50    255      255   100.00
             Min Jarque-Bera p-value > 0.10     255        0     0.00
Min Spatial Autocorrelation p-value > 0.10      25        0     0.00

------------------------------------------------------------------------------
                                                    Summary of Multicollinearity
      Summary of Variable Significance            Variable     VIF Violations Covariates
Variable    % Significant % Negative % Positive    HOUDENBLK   4.66      0       --------
HOUDENBLK          100.00     100.00       0.00    POPDENBLK   5.35      0       --------
BUS1MILE           100.00     100.00       0.00    JOBDENSBLK  1.63      0       --------
DISTONEARM         100.00       0.00     100.00    BUS1MILE    2.37      0       --------
DISTTOACTI         100.00       0.00     100.00    DISTONEARM  1.28      0       --------
POPDENBLK           77.34      50.00      50.00    FTBUSROUT   2.16      0       --------
FTBUSROUT           50.00      91.41       8.59    DISTTOACTI  1.90      0       --------
DISTANCETO          50.00      20.31      79.69    DISTANCETO  1.70      0       --------
JOBDENSBLK          30.47      90.62       9.38
```

Figure 6. Explanatory regression outputs.

OLS Regression

As shown in Figure 7, the fitness of the OLS regression model shows that the R2 is about .048. The summary of variable significance also indicates that all five variables turned out to be statistically significant: housing density, population density, bus stop within 1 mile, distance to nearest bus stations, and the distance to action centers. Moreover, all the variables have a VIF score lower than 7.5, so we don't have an issue with multicollinearity in our model.

On the other hand, we tried to check the distribution of the standardized residuals. Figure 8 shows the distribution curve for the standardized residuals; we still have an issue with the skewed distribution of our residuals. That means we still violate one of the central assumptions for the OLS regression analysis. That is why we might need to go further for the GWR model to overcome this issue.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77– 92. www.egyptfuture.org/ojs/

Summary of OLS Results - Model Variables

Variable	Coefficient [a]	StdError	t-Statistic	Probability [b]	Robust_SE	Robust_t	Robust_Pr [b]	VIF [c]
Intercept	1.619680	0.044949	36.033832	0.000000*	0.047417	34.157889	0.000000*	--------
HOUDENBLK	-0.000037	0.000007	-5.238249	0.000000*	0.000008	-4.855080	0.000002*	4.478195
POPDENBLK	0.000014	0.000004	3.444892	0.000593*	0.000004	3.370298	0.000774*	4.516236
BUS1MILE	-0.002603	0.000391	-6.656387	0.000000*	0.000396	-6.575860	0.000000*	1.283607
DISTONEARM	0.000024	0.000007	3.179939	0.001500*	0.000007	3.421831	0.000645*	1.265102
DISTTOACTI	0.015591	0.003162	4.930367	0.000001*	0.003350	4.654074	0.000005*	1.206181

OLS Diagnostics

Input Features:	CarOwn	Dependent Variable:	HHVEHCNT
Number of Observations:	3980	Akaike's Information Criterion (AICc) [d]:	10914.670241
Multiple R-Squared [d]:	0.048952	Adjusted R-Squared [d]:	0.047756
Joint F-Statistic [e]:	40.909931	Prob(>F), (5,3974) degrees of freedom:	0.000000*
Joint Wald Statistic [e]:	187.869276	Prob(>chi-squared), (5) degrees of freedom:	0.000000*
Koenker (BP) Statistic [f]:	1.637813	Prob(>chi-squared), (5) degrees of freedom:	0.896639
Jarque-Bera Statistic [g]:	4687.562109	Prob(>chi-squared), (2) degrees of freedom:	0.000000*

Figure 7. OLS regression summary outputs.

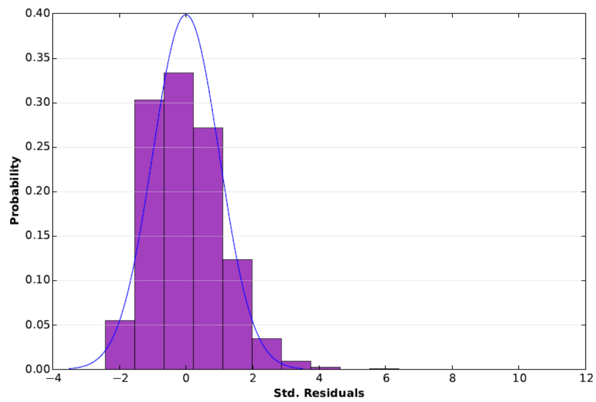


Figure 8. OLS regression Standardized residuals distribution.

Geographically Weighted Regression

The fitness of the GWR regression model shows that the R2 is much higher than in our previous trials, which is about 0.48. Figures 9 and 10 also indicate that Moran's I test shows a small value for Moran's I index and a z-value of 1.015. That means we, at a 95% confidence level, fail to reject the null hypothesis. That means the data are not different from the random distribution.

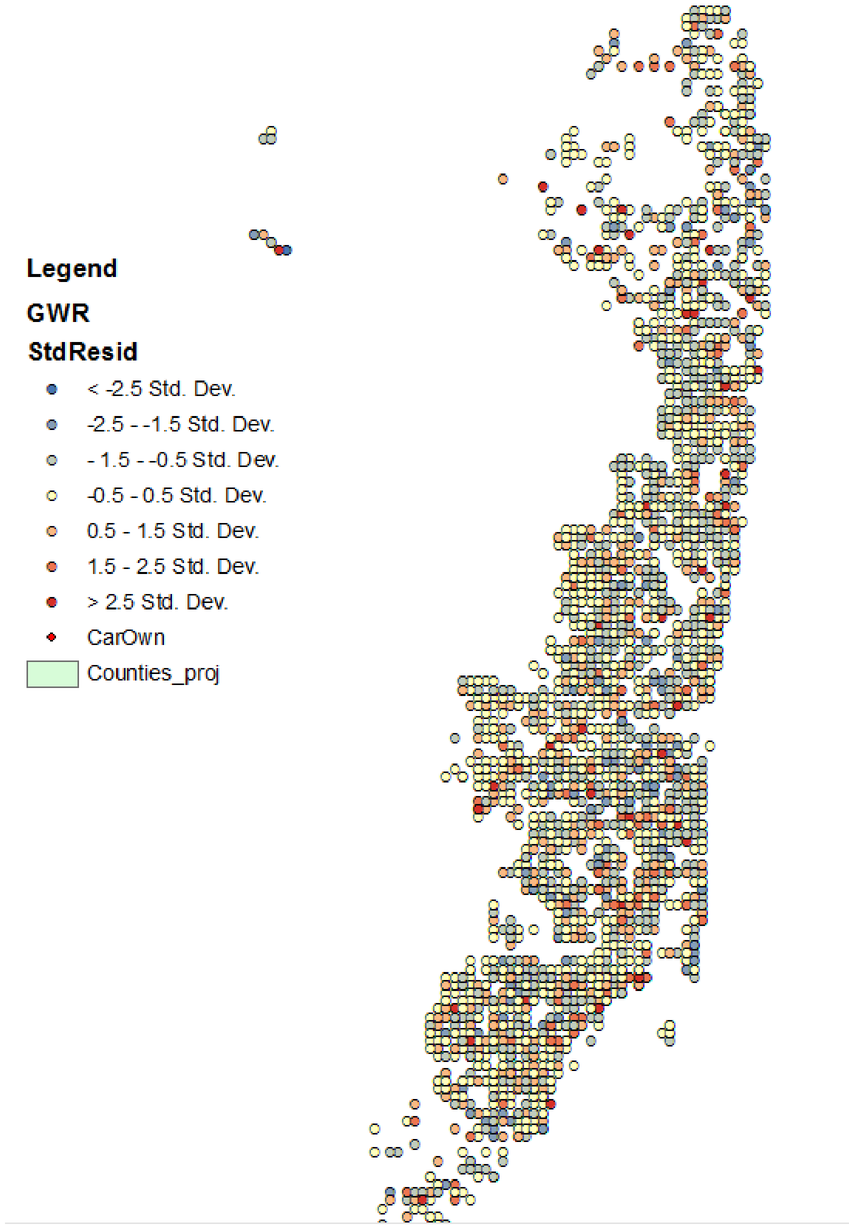


Figure 9. GWR regression map.

```
Neighbors             : 30
ResidualSquares       : 106.80755980240208
EffectiveNumber       : 70.989123317782528
Sigma                 : 1.0334212304136092
AICc                  : 597.93038794547465
R2                    : 0.48481011041472144
R2Adjusted            : 0.12427243780905672
```

```
 Global Moran's I Summary
Moran's Index:     0.002233
Expected Index:   -0.000251
Variance:          0.000006
z-score:           1.015818
p-value:           0.309716

Distance measured in Meters
```

Figure 10. GWR regression outputs and Moran’s I summary.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77_92. www.egyptfuture.org/ojs/

4. Conclusion

This study investigates both global and local household car ownership patterns, as well as the associated factors. The following are the principal findings and contributions of the paper: Using techniques such as clustering analysis and Global/Local Moran's I statistics, this paper first determines whether there is global or local spatial autocorrelation of car ownership in the study area. This research shows that the global distribution of households with a high rate of car ownership is not random when the sample population in each census tract is considered. Furthermore, the study needs to provide more evidence that households with high car ownership rates are concentrated in a specific area within the study region. Using OLS and GWR regression models, five key variables influence car ownership significantly. The most statistically significant factors are housing density, population density, proximity to bus stops, distance to the closest bus stations, and distance to essential locations.

References

1. Bhat, C. R., & Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transportation Research Part B: Methodological, 41(5), 506-526.
2. Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. Geographical analysis, 28(4), 281-298.
3. Cao, X., Mokhtarian, P. L., & Handy, S. L. (2007). Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership. Environment and planning A, 39(4), 830-847.
4. Chin, A., & Smith, P. (1997). Automobile ownership and government policy: The economics of Singapore's vehicle quota scheme. Transportation Research Part A: Policy and Practice, 31(2), 129-140.

ISSN: 2785-9673 VOLUME 3, ISSUE 1, 2020, 77_92. www.egyptfuture.org/ojs/

5. Chiou, Y.-C., Jou, R.-C., & Yang, C.-H. (2015). Factors affecting public transportation usage rate: Geographically weighted regression. Transportation Research Part A: Policy and Practice, 78, 161-177.
6. Clark, S. D. (2007). Estimating local car ownership models. Journal of transport geography, 15(3), 184-197.
7. Giuliano, G., & Dargay, J. (2006). Car ownership, travel and land use: a comparison of the US and Great Britain. Transportation Research Part A: Policy and Practice, 40(2), 106-124.
8. Golob, T. F. (1990). The dynamics of household travel time expenditures and car ownership decisions. Transportation Research Part A: General, 24(6), 443-463.
9. Ma, Y., & Gopal, S. (2018). Geographically weighted regression models in estimating median home prices in towns of Massachusetts based on an urban sustainability framework. Sustainability, 10(4), 1026.
10. Paleti, R., Bhat, C. R., & Pendyala, R. M. (2013). Integrated model of residential location, work location, vehicle ownership, and commute tour characteristics. Transportation Research Record, 2382(1), 162-172.
11. Pinjari, A. R., Pendyala, R. M., Bhat, C. R., & Waddell, P. A. (2011). Modeling the choice continuum: an integrated model of residential location, auto ownership, bicycle ownership, and commute tour mode choice decisions. Transportation, 38(6), 933-958.
12. Raphael, S., & Rice, L. (2002). Car ownership, employment, and earnings. Journal of Urban Economics, 52(1), 109-130.
13. Sheller, M., & Urry, J. (2000). The city and the car. International journal of urban and regional research, 24(4), 737-757.
14. Smith, P. (1992). Controlling traffic congestion by regulating car ownership: Singapore's recent experience. Journal of Transport Economics and Policy, 26(1), 89-95.
15. Wang, C.-H., & Chen, N. (2017). A geographically weighted regression approach to investigating the spatially varied built-environment effects on community opportunity. Journal of transport geography, 62, 136-147.

16. Zhao, F., & Park, N. (2004). Using geographically weighted regression models to estimate annual average daily traffic. Transportation Research Record, 1879(1), 99-107.

Received: March 2020
Accepted: May 2020