ANALYSIS OF ZONAL CHARACTERISTICS OF DRIVERS IN VOLVED IN TRAFFIC CRASHES

Mohamed Abdel-Aty[1], Jaeyoung Lee[1], Keechoo Choi[2], Chowdhury Siddiqui[1]

[1]University of Central Florida, Orlando, Florida, USA

[2]Ajou University, Suwon, South Korea

Email for correspondence: m.aty@ucf.edu

Mohamed Abdel-Aty, Jaeyoung Lee, Keechoo Choi, Chowdhury Siddiqui

University of Central Florida

Email for correspondence: m.aty@ucf.edu

ABSTRACT

It is known that the demographic, socioeconomic and traffic characteristics of crash locations have been used as important factors in macroscopic crash analysis. In this study, we focus on the residential ZIP code areas of drivers who were involved in crashes. The objective of this study is to identify the origin's characteristics of drivers involved in traffic crashes so better targeted education and awareness could be designed and delivered.

Various characteristic factors of the postal ZIP code area of a driver's residence were used in the study. The postal codes were collected from police crash reports for the year 2006 and demographic, socioeconomic and travel pattern data were retrieved from US Census Bureau. Several negative binomial (NB) models were estimated for specific types of crashes such as, total number of crashes (for at-fault drivers), severe crashes (for at-fault drivers), pedestrian crashes, and bicycle crashes.

It was found that demographic characteristics such as gender, ethnic group, socioeconomic characteristics including family income and unemployment, and travel patterns as commute mode and travel time to work are significant factors for specific types of crashes.

The findings from the study implied that several demographic, socioeconomic and traffic factors of zones can influence the crash frequency of the resident. In the planning phases we

can forecast the crash frequency with these models and predicted independent factors in the future. From operational perspective, the results from the study can be used to identify zones that have residents with higher chances to be involved in crashes, thus we can concentrate on these specific zones for education and stricter enforcement.

Keywords: Traffic Crash, Negative Binomial, Residence Analysis, ZIP Code

INTRODUCTION

It is known that the demographic, socioeconomic and traffic characteristics of the crash location is important factors for the crash occurrence in the macroscopic crash analysis. Thus, many traffic safety researchers have studied the characteristics of crash location to find out significant factors for the crash occurrence. Nevertheless, there have been few studies that have focused on residence characteristics of people involved in crashes in traffic safety studies. We assumed that residence characteristics also have effect on the crash occurrence. Therefore, the objective of this study is to find out the significant characteristics associated with the origin of the drivers involved in traffic crashes.

We used the data of nine counties in Central Florida. Crash data and the corresponding census data based on ZIP codes were collected from Florida Department of Transportation and the United States Census Bureau, respectively. We investigate four types of crashes: total crashes, severe crashes, bicycle crashes and pedestrian crashes in Florida. For total and severe crashes, the crash frequency was counted by the residence of only at-fault drivers whereas the residences of bicyclists and pedestrians were considered for bicycle and pedestrian crashes, respectively. Negative binomial (NB) models were developed for these four types of crashes.

Several previous studies have been conducted at the macro-level safety analysis. In macroscopic safety analysis, the crash frequency and corresponding independent variables are aggregated based on the specific spatial units. These spatial units such as block groups or BGs (Kim and Nitz, 1995; Siddiqui and Abdel-Aty, 2012), census tracts or CTs (Siddiqui and Abdel-Aty, 2012; LaScala et al, 2000), counties (Aguero-Valverde and Jovanis, 2006; Siddiqui, Abdel-Aty and Choi, 2012), states (Noland, 2003), an extensive region including multiple states (Stamatiadis and Puccini, 2000), and traffic analysis zones or TAZs (Siddiqui and Abdel-Aty, 2012; Abdel-Aty et al., 2011; Siddiqui and Abdel-Aty, 2012) were used.

Previous studies focused on crash locations aggregated by specific geographic units. On the contrary, some researcher focused on the residence, instead of the crash location. Most of these studies used ZIP codes as geographical units for the analysis because the residence information is typically provided as a form of ZIP code. For example, FARS (Fatality Analysis Reporting System) offers ZIP codes of drivers involved in crashes. Blat and Furman (1998) examined the residence types of drivers involved in fatal crashes using ZIP codes of drivers from FARS based on county-level aggregation. They concluded that not only the majority of fatal crashes occurred in rural area but also rural residents are more likely to be involved in fatal crashes. Lener et al. (2001) conducted a retrospective chart review from patients of a trauma center for injuries from traffic crashes. Age, gender, race and ZIP code were used to identify significant factors of seatbelt use. ZIP code was a proxy for socioeconomic status by using census data. At last, a logistic model revealed that younger people, male, African American, people with lower income and passengers are less likely to use seatbelts.

Moreover, Clark (2003) found out the population density of drivers' residence (using ZIP code), populations at crash location, age, seat belt use, vehicle speed and rural locations significantly affect the mortality after crashes. Romano et al. (2006) investigated the effect of race/ethnicity, language skills, income levels and education levels on alcohol-related fatal crashes. They collected fatal crash data including drivers' ZIP code and socioeconomic data from FARS and the US Census Bureau, respectively. The authors confirmed that people with lower income and less education are more vulnerable to alcohol-related fatal crashes. Males (2009) focused on the relationship between poverty and young drivers' fatal crashes. The author revealed that driver age itself is not a significant predictor of fatal crash risk once other factors associated with high poverty condition such as more occupants per vehicle; smaller vehicle size, older vehicle, lower state per-capita income and so forth were controlled. These factors were significantly associated with each other and with higher crash involvement among drivers from other age groups as well.

Furthermore, Stamatiadis and Puccini (2000) concentrated on the Southeast United States which has higher fatality rates compared to other regions using ZIP codes and corresponding census data from FARS and the US Census Bureau, respectively. Authors showed that higher percentage of the population below poverty levels, rural area and lower educated people affected the fatal crash rates in the Southeast. These socioeconomic factors were found significant for single vehicle fatal crash rates; however, they were not significant for multi vehicle fatal crash rates. Girasek and Taylor (2010) looked into the relationship between socioeconomic status based on ZIP code and vehicle characteristics such as crash test rating, electronic stability control, side impact air bags, vehicle age and weight. Specific vehicle data were collected from the Insurance Institute for Highway Safety using vehicle identification numbers (VINs). Authors revealed that lower income groups experience more risk since it is more likely that their vehicles are not safe enough.

Aside from the traffic safety field, there have been many efforts to find out the effect of demographic and socioeconomic characteristics of the residence in medical studies (Smith et al., 1996; Sundquist et al., 2004), psychology (Cutrona et al., 2006; Ross, 2000) and criminology (Gruenewald et al., 2006; Gyimah-Brempong, 2006). These studies commonly suggested that the lower income and/or lower education level of the residence are significant factors for higher rates of mortality, specific diseases including the disease, depression and crime as well. It is interesting to note that several common socioeconomic characteristics of the residence are significant in various fields.

DATA PREPARATION

Data of nine counties in Central Florida were used in this study. In order to examine the residence characteristics of drivers involved in traffic crashes, two types of data are required. First, we need the aggregated number of crashes by ZIP code of people involved in the crash. Second, we also need corresponding demographic and socioeconomic data, In the crash report, the address of each person involved in the crash is recorded; however, detailed address is not coded because of privacy concerns. Fortunately, ZIP codes were coded by the Florida Department of Transportation (FDOT). Therefore, the ZIP code was possible to be used as a spatial unit for the analysis.

The crash data of 2006 was used in this study due to the data availability of ZIP information. However, only 2000 demographic and socioeconomic are available from US Census Bureau currently, therefore, 2006 crash data and corresponding 2000 census data were used. The number of crashes by the ZIP code of people involved in crashes was combined with the census data. After that, ZIP code areas with unreasonable values were eliminated, for instance, the ZIP code with zero population or critical missing values were removed.

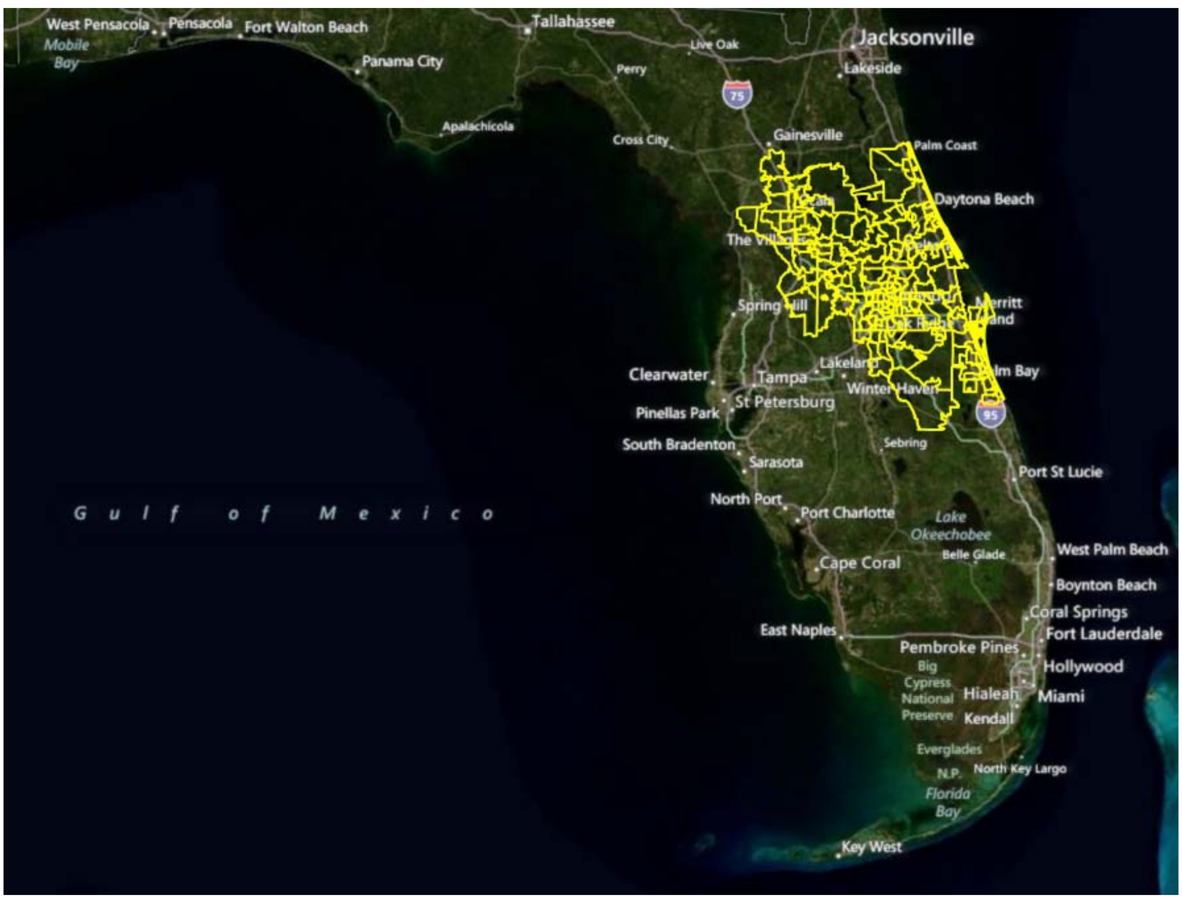

Figure 1 – The Study Area in Central Florida

Table I – Variable Description

Variables	Mean	Std Dev	Min	Max
Number of observations (N=168)				
Crashes				
Total crashes by the at-fault driver's ZIP code	147	117	3	584
Severe crashes by the at-fault driver's ZIP code	25	19	0	95
Bicycle crashes by the bicyclist's ZIP code	8	8	0	42
Pedestrian crashes by the pedestrian's ZIP code	4	5	0	28
Demographic Variables				
Male population	9283	6566	107	25601
Female population	9702	6939	14	27013
Population under 18 years old	4240	3449	9	14519
Population 18-24 years old	1705	1676	21	8419
Population 25-44 years old	5030	4113	68	17122
Population 45-64 years old	4854	3269	23	13563
Population over 65 years old	3155	2553	0	15591
Population of white people	12478	8656	68	40139
Population of African American	2513	4093	2	28879
Population of Hispanic people	3165	4548	12	25866
Socioeconomic Variables				
Median Family Income	45764	12396	24856	106908
Employed people	4645	4039	17	17625
Unemployed people	237	261	0	1679
People without high school diploma	1824	1422	7	7228
People only with high school diploma	7563	5431	27	22389
People with bachelor's degree or higher	8124	6512	0	29116
Commute Mode Variables				
Workers commuting by passenger car	7001	5768	33	24563
Workers commuting by public transportation	92	197	0	1399
Workers commuting by bicycle	35	41	0	177
Workers commuting by walking	103	123	0	900
Commute Time Variables				
Workers with commute time less than 10 min	781	662	0	2701
Workers with commute time 10-19 min	2146	1839	0	8526
Workers with commute time 20-29 min	1652	1518	0	6891
Workers with commute time 30-39 min	1416	1360	0	6394
Workers with commute time 49-59 min	836	884	0	4520
Workers with commute time 60 min or more	484	405	0	2507
Location / Building Age				
Housing units in urban area	6515	5608	0	23473
Housing units in rural area	945	1166	0	5846
Median year of structures built	1982	8	1956	1996

DATA EXPLORATION

According to the crash report data from FDOT, 24,750 drivers were retained as at-fault crash involvements after excluding observations with missing ZIP codes or from out-of-state ZIP codes in the study area in 2006. Among them, 4,129 drivers were responsible for the fatal or severe injury crashes. We also looked into bicycle and pedestrian involved crashes. A total of 1,266 bicycle crashes and 718 pedestrian crashes had occurred in the same year.

As potential independent variables for the models, overall 26 variables were prepared. They include demographic data such as gender, age, race/ethnicity, socioeconomic data such as employment status, household income, travel pattern related data including major transportation mode to work, average travel time to work, and so forth. Table 1 summarizes the variable description of the data. All data used in model were log-transformed since they have large values.

METHODOLOGY

Negative Binomial Model

The number of crashes is non-negative integers which are not normally distributed. Poisson or Negative Binomial (NB) models have the ability to estimate the crash frequencies with explanatory variables; however the Poisson mode has been criticized because of its implicit assumption that the variance equals mean. This assumption is often violated especially in the crash data. Most of crash data have a greater variance than their mean and therefore the data is over-dispersed. NB models can relax the over-dispersion problem. The mean-variance relationship in the negative binomial distribution is as follows:

$$Var\ (Y) = \mu + \alpha\mu^2$$

Thus, if the dispersion parameter α is near to zero, the variance is also near to the mean, which is the basic assumption of the Poisson distribution. The existence of over-dispersion is adjusted by the log-linear relationship between the expected number of crashes and covariates.

$$\ln(\mu_i) = X_i\beta + \varepsilon_i$$

,where, *i* is an observation unit, μ_i is the expected number of the crash, X_i is covariates, β is the estimated coefficient vector and ε_i is the random error term. The following function is the probability of mass function of the negative binomial distribution, where $\Gamma(x)$ is gamma function and over-dispersion parameter α should be greater than 0.

$$Pr(Y = y_i) = \frac{\Gamma(y_i + \frac{1}{\alpha})}{\Gamma\left(\frac{1}{\alpha}\right)\Gamma(y_i + 1)}\left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i}\left(\frac{1}{1 + \alpha\mu_i}\right)^{\frac{1}{\alpha}}$$

Since negative binomial model has been broadly used in traffic safety studies (Siddiqui and Abdel-Aty, 2012; Siddiqui, Abdel-Aty and Choi, 2012), we determined that the application of the NB model is suitable in this study.

DISCUSSION OF RESULTS

Several demographic factors such as age and ethnicity, socioeconomic factors including homeownership and household income, commuting travel factors such as commuting modes to work and average travel time to work, and also household characteristics including location and types of industry fields were found significant for NB models. Significant factors in each model are shown in Table 2.

Explanatory variables were chosen based on their p-values less than 0.05. In order to confirm the goodness-of-fit of the model, the log-likelihood (LR) ratio[1], Akaike information criterion (AIC)[2], Bayesian information criterion (BIC)[3], McFadden's pseudo R^2 [4] and R^2_a [5] were calculated and suggested.

---

[1] $LR\ ratio = -2\ ln\left(\frac{likelihood\ for\ intercept\ model}{likelihood\ for\ full\ model}\right)$

[2] $AIC = -2\ ln(likelihood\ for\ full\ model) + 2p$, where p is the number of parameters

[3] $BIC = -2\ ln(likelihood\ for\ full\ model) + p\ ln(n)$

[4] *McFadden's pseudo* $R^2 = \left(1 - \frac{likelihood\ for\ full\ model}{likelihood\ for\ intercept\ model}\right)$

[5] $R^2_\alpha = 1 - \frac{\alpha(full\ model)}{\alpha(intercept\ model)}$

Total At-Fault Crash Model

Table II presents the result of the NB model estimation for total at-fault crashes.
As a result, the female population, the Hispanic population, the median family income and the number of workers with long commute time variables were found statistically significant at the 95% level.

It was found that the female population has a positive effect on total at-fault crash occurrence. It is interpreted that female drivers are more likely to contribute to total traffic crashes compared to male drivers. This result is consistent with the study of Massie et al. (1995), the authors concluded that female drivers have higher probabilities of involvements in all police reported crashes.

The estimated coefficient of number of Hispanic people has positive sign. Several earlier studies suggested that Hispanic drivers are involved in fatal crashes more frequently (Harper et al., 2000); nevertheless, the authors did not investigate the effect of the ethnicity on total crashes. The result from our total at-fault crash model implies the Hispanic ethnic group is more vulnerable to not only fatal crashes but also to total crashes including property damage only and minor injury crashes.

The household income is an important indicator of the economic status. It was shown that areas with lower income are more vulnerable to total at-fault crash involvements. Moreover, it was shown that the ZIP area with increased workers with commute time more than 60 minutes has more number of total at-fault crashes. It seems reasonable because people with longer trip distance are more exposed to the traffic.

Table II – NB Model for the total at-fault crash

Variables	**Estimate**	**Std Error**	**Pr > χ^2**
Intercept	3.7612	0.7740	<0.0001
Female population	0.6319	0.0514	<0.0001
Population of Hispanic people	0.1581	0.0220	<0.0001
Median family income	-0.6063	0.0744	<0.0001
Workers with commute time 60+ min	0.1171	0.0391	<0.0028
Dispersion	0.0399	0.0060	
Log-likelihood ratio	435.0506		
AIC	1579.5328		
BIC	1598.2766		
McFadden's pseudo R^2	0.2172		
R^2_a	0.9424		

Severe At-Fault Crash Model

Table III summarizes the result of the NB model estimation for the severe at-fault crashes. Statistically significant variables are as follow: the male population, the number of young people, the median family income and the education level.

The male population was positive associated with severe at-fault crash model, which means males are more to cause severe crashes. As stated previously, Massie et al. (1995) claimed that females are overrepresented in total number of crashes; however, they also found that males have higher risks of experiencing fatal crashes compared to females. The results found from this study have same outcomes.

The population of young people aged 18-24 years old variable was significant and positively associated with the severe crash frequency. It can be interpreted that young drivers can cause more severe crashes compared to other age group. This result is also consistent with the study of Broughton (1988) and Williams and Carsten (1989). They asserted that fatal crash rates of young drivers are much higher than mid and older drivers.

The coefficient of median family income has negative sign, which indicates poorer communities cause severe crashes more frequently. There are several studies that suggested that residents within the lower income communities are more likely to be involved in fatal crashes (Romano et al., 2006; Males, 2009; Stamatiadis and Puccini, 2000; Girasek and Taylor, 2010; Kristensen et al., 2011). The community with higher education level has lower severe at-fault crash since people with bachelor's degree or higher has negative value and statistically significant. In recent, Kristensen et al. (2011) focused on only young drivers but they found similar result that the increasing crash fatality of the young driver was found in association with decreasing parental education level.

Table III – NB Model for the severe at-fault crash

Variables	**Estimate**	**Std Error**	**Pr > χ^2**
Intercept	3.5503	1.0188	0.0005
Male population	0.9911	0.1117	<.0001
Age group: 18-24 years old	0.1197	0.0609	0.0494
Median family income	-0.8330	0.0932	<.0001
People with bachelor's degree or higher	-0.1522	0.0599	0.0111
Dispersion	0.0195	0.0073	
Log-likelihood ratio	368.5762		
AIC	1047.676		
BIC	1066.419		
McFadden's pseudo R^2	0.2625		
R^2_a	0.9698		

Bicycle Crash Model

Different from previous two models, we concentrate on the residence of bicyclists involved in the crashes in the bicycle crash model. Table IV shows the result of the bicycle model estimation. It was found that coefficients of the male population, housing units in the urban area and the median year of the structure built were positively related to the number of bicycle crashes whereas the median family income has negative association with the bicycle crash frequency.

The male population variable was positive, thus it suggests males are more vulnerable to the bicycle crashes. In previous study, Rivara et al. (1997) also found male bicyclists have increased risks of hospital treatment of bicycle related crashes compared to female bicyclists (OR=1.3 , 95% CI 1.04 to 1.8). It was shown that the median family income has a negative relationship with bicycle crash frequency. It could be explained that the people from low income families is more likely to use bicycles instead of passenger cars for economic reasons. Thus, the community with lower income families has more bicycle crashes. Noland and Quddus (2004) also investigate the factors affecting to bicycle crashes, and they found the income (GDP per capita) was found negatively associated with bicycle crashes using negative binomial models. In addition, bicyclists in the urban area are more likely to be involved in crashes compared to those in rural areas. It could be interpreted that bicyclists in urban areas have more exposed to traffic compared to those in rural areas.

The median year of structure built factor is only significant in the bicycle and the pedestrian models (Table IV and Table V). The variable was defined if median year of the structure built is before 1980, the value was unity whereas if it is 1980 or later, the value was zero. Since it has positive coefficient, it can be interpreted ZIP code areas with many older buildings also have many bicyclists involved in the crash. It can be explained that the area with older buildings could be less likely to have enough safety related facilities such as wider sidewalk width, guardrails, or bicycle lanes. It is also possible that older towns have limited sight distance at intersections.

Table IV – NB Model for the bicycle crash

Variables	**Estimate**	**Std Error**	**Pr > χ^2**
Intercept	0.2252	2.5531	0.9297
Male population	1.0261	0.1472	<0.0001
Median family income	-0.9429	0.2282	<0.0001
Housing units in urban area	0.2025	0.0698	0.0037
Median year of structure built (1: <1980, 0: other)	0.3433	0.1248	0.0060
Dispersion	0.1820	0.0479	
Log-likelihood ratio	194.7100		
AIC	674.8211		
BIC	693.5649		
McFadden's pseudo R^2	0.2271		
R^2_a	0.8576		

Pedestrian Crash Model

The result of the pedestrian model estimation is presented in Table V. The number of unemployed people, lower educated people, workers using public transportation and median year of structure built were found significant.

Regarding unemployed people, since its coefficient has a negative sign, unemployed people are contributing to the increased number of pedestrian crashes. McMahon et al. (1999) found similar result that the lower percentage of the unemployment has a significant negative effect on pedestrian crashes. Concerning the education level, it was found that the area with the increased number of less educated people is more likely to have frequent pedestrian crashes. Chakravarthy et al. (2010) showed the proportion of people who completed less than a high school diploma also increased the pedestrian crashes.

It is interesting that number of workers commuting by walking was not significant for the pedestrian crash. Instead, the number of workers commuting by the public transportation was found significant and positively associated with pedestrian crashes. It is reasonable because many public transportation users need to access the public transportation by walking. Clifton and Kreamer-Fults (2007) found that the transit access, commercial access and population density have positive association with the number of pedestrian crashes, which is consistent with the result from the pedestrian model in this study.

Table V – NB Model for the pedestrian crash

Variables	**Estimate**	**Std Error**	**Pr > χ^2**
Intercept	-3.0160	0.4461	<.0001
Number of unemployed people	0.3481	0.0719	<.0001
People without high school diploma	0.3351	0.0885	0.0002
Workers commuting by public transportation	0.1584	0.0277	<.0001
Median year of structure built	0.1733	0.0710	0.0146
Dispersion	0.0341	0.0213	
Log-likelihood ratio	247.1894		
AIC	800.6973		
BIC	819.4410		
McFadden's pseudo R^2	0.2386		
R^2_a	0.9671		

CONCLUSION

The main objective of this study is to investigate the effect of residence characteristics of people involved in traffic crashes. The idea is to identify the factors associated with the origin of the drivers involved in crashes rather than the crash location. We concentrated on four types of crashes: total at-fault crashes, severe at-fault crashes, bicycle crashes and pedestrian crashes. In this study the ZIP code was used as a base areal unit for the analysis. Residence ZIP codes of people involved in the crash are recorded in the police crash report, we used these ZIP codes and classified them into several types of crashes. Demographic and socioeconomic data based on the ZIP code were collected from US Census Bureau. Four NB models were estimated for the each type of the crash.

It was shown that the family income factor were significant for total at-fault, severe at-fault and bicycle crashes and they had negative coefficients. It implies drivers from low-income areas are more likely to cause traffic crashes and bicyclists from low-income area have higher probabilities to be involved in crashes. On the other hand, for the pedestrian crashes, the income factor was not significant but education level and unemployment variables were significant, which should show the deprived socioeconomic status. Concerning the median year of structure built, it was shown that the areas with older buildings have more victims involved in bicycle and pedestrian crashes. It is possible that the older towns do not have enough safety related facilities. Moreover, gender factors also play important roles both in total at-fault and severe at-fault crashes. In total at-fault crashes, it was shown that female drivers are more likely to cause total crashes whereas male drivers have more chance to cause severe crashes compared to female drivers. Lastly, it was found that commute pattern variables were significant for two models. The more number of long distance commuters contribute to the increased total at-fault crashes, and the number of public transportation commuters also increase probabilities of pedestrian crashes. Lastly, it was discovered that bicycle crashes are more likely to occur in the urban area because bicyclists in the urban area are more exposed to the traffic.

Key findings from the study implied that several demographic and socioeconomic as well as travel characteristics of residence zones contribute to crash occurrence. The findings could be used to identify residence areas with people who are more likely to be involved in certain types of crashes. Therefore, we can concentrate on these specific zones for safety treatments. The results are important for designing and tailoring specific education and awareness campaigns and stricter enforcement. The limitation of this study is that we only focused on the residence characteristics. Admittedly, the residence characteristics of at-fault drivers, bicyclists and pedestrians involved in the crash play key roles in the crash occurrence as shown in this study. However, the crash occurrence is also equally affected by the physical characteristics of the crash location. A combined analysis should be addressed in follow-up studies.

REFERENCES

Abdel-Aty, M., C. Siddiqui, and H. Huang (2011) Integrating Trip and Roadway Characteristics in Managing Safety at Traffic Analysis Zones. In Transportation Research Board 90th Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C.

Aguero-Valverde, J., and P. P. Jovanis (2006) Spatial Analysis of Fatal and Injury Crashes in Pennsylvania. Accident Analysis and Prevention 38, 618-625.

Chakravarthy B., C. L. Anderson, J. Ludlow, S. Lotfipour, and F. E. Vaca (2010) The Relationship of Pedestrian Injuries to Socioeconomic Characteristics in a Large Southern California County. Traffic Injury Prevention, 11 (5), 508-513.

Clark, D. E. (2003) Effect of Population Density on Mortality after Motor Vehicle Collision. Accident Analysis and Prevention, 35, 965-971.

Clifton, K. J., and K. Kreamer-Fults (2007) An Examination of the Environmental Attributes Associated with Pedestrian–Vehicular Crashes near Public Schools. Accident Analysis and Prevention, 39 (4),708-715.

Girasek, D. C., and B. Taylor (2010) An Exploratory Study of the Relationship between Socioeconomic Status and Motor Vehicle Safety Features. Traffic Injury Prevention, 11 (2), 151-155.

Gruenewald, P. J., B. Freisthler, L. Remer, E. A. LaScala, and A. Treno (2006) Ecological Models of Alcohol Outlets and Violent Assaults: Crime Potentials and Geospatial Analysis. Addiction 101 (5), 666-677.

Gyimah-Brempong, K. (2006) Neighborhood Income, Alcohol Availability, and Crime Rates. The Review of Black Political Economy, 33 (3), 21-44.

Harper, J. S., W. M. Marine, C. J. Garrett, D. Lezotte, and S. R. Lowenstein (2000) Motor Vehicle Crash Fatalities: A Comparison of Hispanic and Non-Hispanic Motorists in Colorado. Annals of Emergency Medicine, 36 (6), 589-596.

Kristensen, P., T. Kristiansen, M. Rehn, H. M. Gravseth, and T. Bjerkedal (2011) Social Inequalities in Road Traffic Deaths at Age 16-20 Years among All 611,654 Norwegians Born between 1967-1976: A Multilevel Analysis. Injury Prevention 18 (1), 3-9.

LaScala, E. A., D. Gerber, and P.J. Gruenewald (2000) Demographic and Environmental Correlates of Pedestrian Injury Collision: A Spatial Analysis. Accident Analysis and Prevention 32, 651-658.

Males, M. A. (2009) Poverty as a Determinant of Young Drivers' Fatal Crash Risks. Journal of Safety Research, 40, 443-448.

Massie, D. L., K. L. Campbell, and A. F. Williams (1995), Traffic Accident Involvement Rates by Driver Age and Gender. Accident Analysis and Prevention, 27 (1), 73-87.

McMahon, P. J., C. Ducan, J. R. Stewart, C. V. Zegger, and A. Khattak (1999) Analysis of Factors Contributing to "Walking Along Roadway" Crashes. Journal of the Transportation Research Board, No. 1674, 41-48.

Noland, R. B. (2003) Traffic Fatalities and Injuries: The Effect of Changes in Infrastructure and Other Trends. Accident Analysis and Prevention 35, 599-611.

Noland, R. B., and M. A. Quddus (2004) Analysis of Pedestrian and Bicycle Casualties with Regional Penal Data. Journal of the Transportation Research Board, No. 1897, 28-33.

Rivara, F. P., D. C. Thompson, and R. S. Thompson (1997) Epidemiology of Bicycle Injuries and Risk Factors for Serious Injury. Injury Prevention 3, 110-114.

Romano, E., S. Tippetts, K. Blackman, and R. Voas. (2006) Language, Income, Education and Alcohol-Related Fatal Motor Vehicle Crashes. Journal of Ethnicity in Substance Abuse, 5 (2), 119-137.

Ross, C. E. (2000) Neighborhood Disadvantage and Adult Depression. Journal of Health and Social Behavior, 41 (2), 177-187.

Siddiqui, C., and M. Abdel-Aty (2012) A Comparison of Geographical Unit-Based Macro-level Safety Modeling. In Transportation Research Board 91th Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C.

Siddiqui, C., and M. Abdel-Aty (2012) On the Nature of Modeling Boundary Pedestrian Crashes at Zones. In Transportation Research Board 91th Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C.

Siddiqui, C., M. Abdel-Aty, and K. Choi (2012) Macroscopic Spatial Analysis of Pedestrian and Bicycle Crashes. Accident Analysis and Prevention 45, 382–391.

Smith, G. D., J. D. Neaton, D. Wentworth, R. Stamler, and J. Stamler (1996) Socioeconomic Differentials in Mortality Risk among Men Screened for the Multiple Risk Factor Intervention Trial: I. White Men. American Journal of Public Health, 86 (4), 486-496.

Stamatiadis, N., and G. Puccini (2000) Socioeconomic Descriptors of Fatal Crash Rates in the Southeast USA. Injury Control and Safety Promotion, 7 (3), 165-173.

Sundquist, K., M. Winkleby, H. Ahlen, and S. E. Johansson (2004) Neighborhood Socioeconomic Environment and Incidence of Coronary Heart Disease: A Follow-up Study of 25,319 Women and Men in Sweden. America Journal of Epidemiology, 159 (7), 655-662.