A home-based approach to understanding the effect of spatial autocorrelation on seat belt non-use

Amin Mohamadi Hezaveh[1], Christopher R. Cherry[1*] Trond Nordfjærn[2],

[1] Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, United States

[2] Norwegian University of Science and Technology, Department of Psychology

[*] Corresponding author at Civil and Environmental Engineering, University of Tennessee, Knoxville, TN, United States, E-mail address: Cherry@utk.edu

Abstract

Roadside observations indicate that seat belt use rates are often spatially correlated with nearby areas. However, very few studies have examined the effects of spatial autocorrelation on seat belt use. This study used exploratory spatial data analysis (ESDA) to explore spatial autocorrelation in Tennessee, which has a lower seat belt use than the United States national average. We geocoded home-addresses of vehicle occupants involved in traffic crashes between 2014-16 (n = 1,251,901) and projected them to the census tract corresponding to their home address. This projection reveals information about the spatial distribution of seat belt non-use and socioeconomics of the areas surrounding the crash victim's home. The presence of highly spatially correlated observations (i.e., a significant positive Moran's I) suggests that seat belt non-use is not produced solely by the internal structural factors represented in the non-spatial models. ESDA reveals a distinctive regional imprint for spatial autocorrelation, in which Southern-metropolitan areas' (Southern-MPOs) in Tennessee census tracts have higher than average seat belt non-use compared to Non-Southern-MPOs (16% vs. 9%). The spatial error model was suitable for Non-Southern-MPOs, whereas the spatial lag model was more suitable for Southern MPOs. Comparison of the estimated models indicates that in the Non-Southern MPOs, percentage of the White population, percentage of the population with Bachelor's degree, median household income, vehicle ownership, and population density are significant predictors of seat belt non-use. On the other hand, median household income, vehicle ownership, and percentage of population aged between 16-42 years old predict seat belt non-use in Southern MPOs. The study results could be used to identify seat belt non-use clusters at the state level and identify seat belt non-use hot zones. Furthermore, this analysis indicates that the relationship between demographic variables and seat belt non-use varies across regimes. Failing to consider the spatial regimes in the analysis would lead to falsely prioritizing groups prone to seat belt non-use.

Keywords: Seat Belt Non-Use; ESDA; Spatial Lag Model; Zonal Model

Introduction

It is well-established that seat belt use could reduce severe injuries and fatalities from car occupants' traffic crashes (Blincoe *et al.* 2015). There are mandatory seat belt laws in the United States, and not wearing a seat belt violates the state law, which could lead to a fine in many jurisdictions. In 34 States, the District of Columbia, and Puerto Rico, seat belt laws are primary, which enable law enforcement officers to stop vehicles and write citations when they observe seat belt non-use (IIHS 2018). In 15 States, the laws have specified secondary Enforcement, meaning that law enforcement officers are permitted to issue a seat belt citation only after they stop a vehicle for another primary violation. Nevertheless, some occupants do not use their seat belt. In Tennessee, seat belt use is also compulsory and is a primary law (i.e., secured shoulder and lap belts) when riding in the front seat of a vehicle (IIHS 2018). Meanwhile, roadside observations of 190 sites in 2017 revealed that, on average, 88.5% of the vehicle occupants in Tennessee used their seat belt (CTR 2018), which is 1.2% lower than the national average in the United States (NHTSA 2017).

There is compelling evidence suggesting spatial dependency of seat belt use (Majumdar *et al.* 2004). Spatial dependence may reflect variations in a wide range of factors, including demographics, economic, historical, geographical background, enforcement level, or traffic culture. As presented in Figure 1, roadside observations imply the presence of spatial dependence on seat belt non-use at the state level. Visual screening of this map indicates the presence of spatial clusters of seat belt use (e.g., a state with high seat belt use rates shares borders with other states with high seat belt use and vice versa). Spatial clusters could indicate the presence of spatial autocorrelation. Spatial autocorrelation exists when a variable displays interdependence over space. The presence of spatial autocorrelation in seat belt use was also reported by Majumdar *et al.* (2004) at the state level.

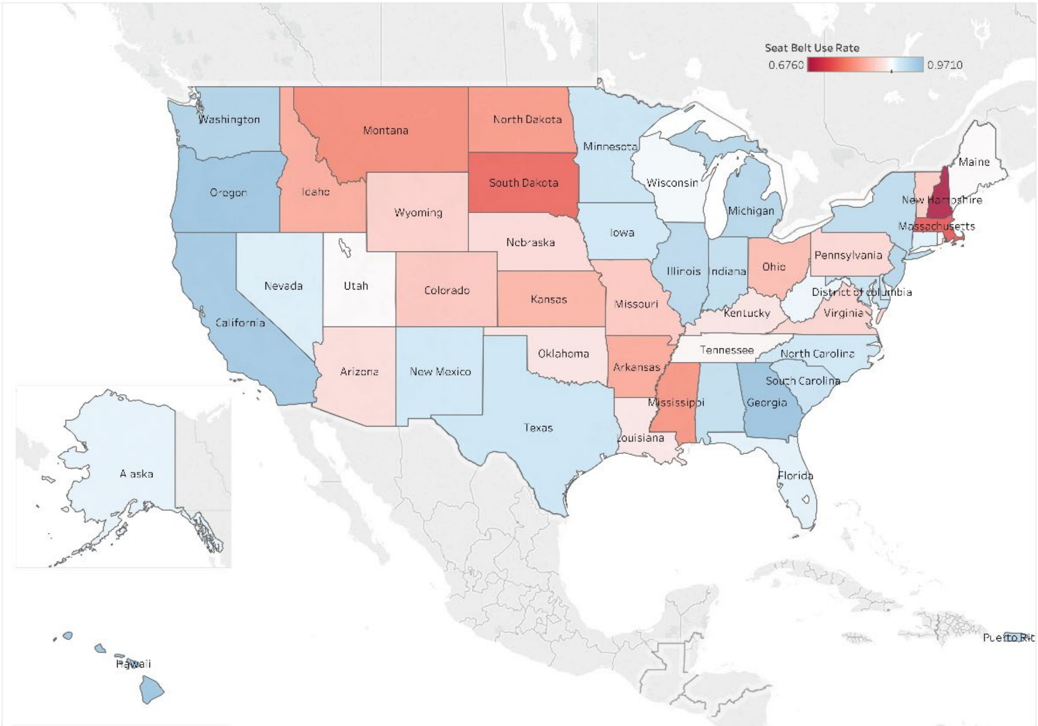


Figure 1 Seat belt use distribution at the state level –2017; adopted from NHTSA (2017)

Wearing a seat belt is also a decision-making problem, and as could be expected, social-psychological factors among vehicle occupants affect seat belt use (Calisir and Lehto 2002, Şimşekoğlu and Lajunen 2008). Several studies used self-reported questionnaires to explore factors influencing seat belt use. These studies highlight the role of local effects, such as regulation enforcement, on seat belt use. Subjective norms (i.e., perceived social pressure to perform or not to perform the behavior) (Ajzen 1991), and normative beliefs (i.e., an individual's perception of social normative pressures, or relevant others' beliefs that he or she should or should not perform such behavior) may represent the effect of social interaction and pressure for seat belt use. Furthermore, subjective norms (Şimşekoğlu and Lajunen 2008, Ali *et al.* 2011, Torquato *et al.* 2012), attitude (positive or negative evaluations of seat belt use) (Şimşekoğlu and Lajunen 2008, Ali *et al.* 2011, Torquato *et al.* 2012), or cues to action (e.g., seeing other drivers wearing a seat belt) (Şimşekoğlu and Lajunen 2008, Ali *et al.* 2011) have a significant impact on seat belt use. Negative attitudes and beliefs about the effectiveness of seat belt use may adversely affect seat belt use (Fockler and Cooper 1990, Begg and Langley 2000). Many of these psychological factors may reflect the safety culture in certain areas (Şimşekoğlu *et al.* 2013, Nordfjærn *et al.* 2014b) and may drive spatial dependency in seat belt use. Several studies showed that personality traits related to road safety vary across countries (Nordfjærn *et al.* 2011, Şimşekoğlu *et al.* 2013, Nordfjærn *et al.* 2014a, Nordfjærn *et al.* 2014b). Besides, deterrence theory argues that people avoid violating regulations through fear of punishment by Enforcement (Homel, 1986).

Measuring these psychological attributes is very difficult, and the data are usually not readily available. A possible solution to overcome these challenges could be to incorporate any available information about vehicle occupants' residential locations into their seat belt use choices. This information may serve as a proxy for those attributes where data are not available. In addition, one may intuitively assume that vehicle occupants' seat belt use is influenced by the geographic location where they reside because geography highly influences behaviors, attitudes, and social norms (Rentfrow 2010). The geography of vehicle occupants' residential location might also serve as a proxy for their behavioral patterns when such data are not available (Foster 1999, Van Acker *et al.* 2010, Kamruzzaman and Hine 2013).

Considering the demographics of vehicle occupants, males have lower seat belt use rates compared to females (Preusser *et al.* 1991, Reinfurt *et al.* 1997, Nelson *et al.* 1998, Calisir and Lehto 2002, Wells *et al.* 2002, Glassbrenner *et al.* 2004, Gkritza and Mannering 2008, Pickrell and Ye 2009, Afghari *et al.* 2020). This also holds for younger drivers than older adults (Reinfurt *et al.* 1997, Calisir and Lehto 2002, Glassbrenner *et al.* 2004). Individuals with higher education and/or income tend to have higher seat belt compliance (Preusser *et al.* 1991, Reinfurt *et al.* 1997, Wells *et al.* 2002, Houston and Richardson 2005). Studies in the United States have also shown that African-Americans are less likely to use a seat belt than Whites or Hispanics (Vivoda *et al.* 2004, Gkritza and Mannering 2008, Pickrell and Ye 2009).

In another study, Afghari et al. (2020) studied the effect of residential location characteristics of vehicle occupants and crash location characteristics on seat belt use. (Afghari et al. 2020) used a latent class binary logit model to capture the unobserved heterogeneity in data. Their finding indicated that seatbelt use determinants have varied effects across drivers in Tennessee (Afghari et al. 2020). Afghari et al. (2020) reported that males, young age, and drug consumption are negatively associated with seatbelt use, whereas population density, travel time, and income per capita contribute to seatbelt use.

Enforcement and education also have a significant impact on seat belt use. Several studies showed the effectiveness of Enforcement and education on seat belt use (Dee 1998, Eby *et al.*

2000, Majumdar *et al.* 2004, Reinfurt 2004, Solomon *et al.* 2004, Thomas *et al.* 2008, National Highway Traffic Safety Administration 2010, Tison and Williams 2010, Thomas *et al.* 2011). Notably, there are several campaigns ongoing in Tennessee and the United States that target seat belt use. However, each campaign's extent, such as study area, message, and the targeted population, is unknown (Hezaveh *et al.* 2019b). This is also the case for enforcement activities.

Like Macroscopic Crash Prediction Models, seat belt use can also be measured at the aggregate level (Hezaveh and Cherry 2019, Hezaveh and Cherry 2020). Macroscopic Crash Prediction Models are a set of methods that provide information regarding the association between road safety at the zonal level and data elements at an aggregate level, such as sociodemographic factors, network characteristics, and travel behavior. By using a wide range of safety outcomes, researchers explored the association between travel behavior, socioeconomic factors, transportation network characteristics at the zonal level and crash frequency (by road user type or injury severity), the burden of traffic crash (i.e., monetize the value of traffic injury), and seat belt use. In a recent study, Hezaveh and Cherry (2019) used the vehicle occupant's home address involved in traffic crashes from police crash reports. The study showed that seat belt use rates vary at a fine geographic level (i.e., census tract) within a state. Moreover, the authors showed several demographic factors besides ethnicity, gender, and age cohorts that influence seat belt use rates at the zonal level; for instance, population density, age, household vehicles' ownership, and householdsize (Hezaveh and Cherry 2019). Nonetheless, Hezaveh and Cherry (2019) did not consider spatial autocorrelation's effect in their analysis.

In the current study, we explore the issue of spatial autocorrelation in seat belt use analysis. Failing to incorporate spatial autocorrelation is likely to cause biased estimates and unreliable statistical inferences (Azimi *et al.* 2019, Xie *et al.* 2019). To reach this goal, we measure the seat belt non-use rates at the zonal level and evaluate the relationship between seat belt use rates and sociodemographic variables based on the home address of the individuals involved in traffic crashes at the zonal level. Evidence of spatial autocorrelation may provide promising opportunities for understanding the local factors contributing to seat belt use.

The next section presents the methods used in this study. In the methodology section, we discuss the geocoding process and measuring seat belt non-use at the zonal level by incorporating spatial effects. In the last section, we present and discuss the findings of the analysis.

Methodology

Database and geocoding process

This study's data were provided by the Tennessee Integrated Traffic Analysis Network (TITAN), a statewide repository for traffic crashes and surveillance reports completed by Tennessee law enforcement agencies. For the years 2014-16, the TITAN records include 694,276 crashes and information about 1,607,995 vehicle occupants involved in traffic crashes. The Bing API was used in this study for geocoding the residential address of the individuals. Only those addresses with an accuracy level of the premise (e.g., property name, building name), address-level accuracy, or intersection level accuracy were used in the analysis (Hezaveh and Cherry 2019, Mohamadi Hezaveh 2019, Merlin *et al.* 2020, Mohamadi Hezaveh and Cherry 2020). A sample of addresses was verified by manual inspection. After geocoding the home-addresses, we retrieved home-addresses' coordinates of 1,510,506 individuals (94% success rate), which met the address quality filter criterion. Among geocoded addresses, 162,447 individuals lived out of Tennessee. After controlling for seat belt use type (i.e., excluding child seat boosters), 1,252,139 observations with a Tennessee address were selected for assignment to the census tract data.

Following the MMUCC (2012), TITAN provides information regarding occupants' restraint use at the time of the crash. For this study, we defined seat belt non-use as vehicle occupants who did not restrain both lap and shoulder seat belt at the time of a traffic crash. Accordingly, we estimated seat belt non-use rates at the zonal level as the percentage of seat belt non-use cases over a total number of observations in a specific geographic area.

Census data from the US survey in 2010 were also used for obtaining sociodemographic data elements. We also used Highway Performance Monitoring System data for Tennessee in 2015 to obtain Average Annual Daily Traffic for each road segment and calculate the total Vehicle Miles Travelled (VMT) for high-speed and low-speed roads for each census tract (Hezaveh *et al.* 2019a). Table 1 summarizes the sample characteristics of the selected variables considered as input for model estimation. To prevent outliers, we only considered the census tracts that had more than 20 observations.

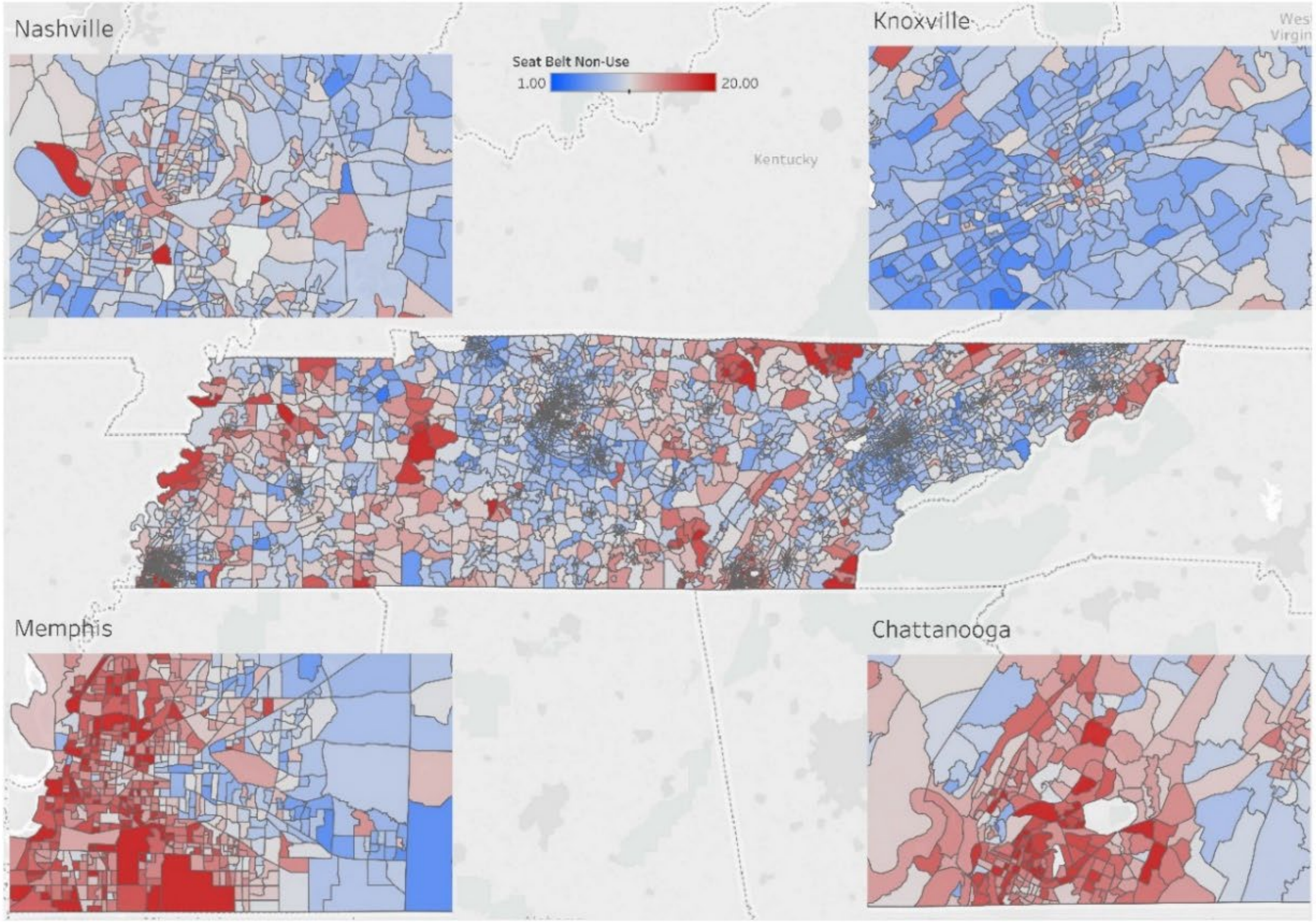
Figure 2 exhibits the geographical distribution of seat belt non-use at the census tract level. Red colors indicate a higher level of seat belt non-use, while blue colors show a higher compliance level. Visual inspection indicates that census tracts are clustered together, meaning that blue colors are usually surrounded by blue neighbors and vice versa. Moreover,

Figure 2 shows that the four major metropolitan areas in the Tennessee, Chattanooga, and Memphis metropolitan areas (here defined as Southern-MPOs) have higher seat belt non-use rates than the Knoxville and Nashville metropolitan areas (here defined as Non-Southern-MPOs).

Table 1 Sample statistics for the state of Tennessee at the census tract

Variable	**Mean**	**Std. Deviation.**	**[95% Conf. Interval]**	
Total Population	1,530	789	1,506	1,554
Age Cohort %				
<16 Years	0.23	0.08	0.22	0.23
16-42 Years	0.32	0.11	0.32	0.33
43-59 Years	0.25	0.08	0.24	0.25
> 59 Years	0.20	0.10	0.20	0.20
Race %				
White	0.77	0.30	0.76	0.78
Non-White	0.23	0.23	0.22	0.24
Means Of Transportation To Work Proportion				
Motorized Modes	0.92	0.11	0.92	0.93
Non-Motorized modes	0.10	0.08	0.10	0.11
Education Degree %				
High School And Lower	0.52	0.20	0.51	0.53
Some College Degree	0.20	0.08	0.20	0.21
Bachelors' Degree	0.20	0.12	0.19	0.20
Other Degrees	0.08	0.08	0.07	0.08
Median Household Income ($10,000)	45.9	25.1	45.2	46.7
Household Vehicles' Ownership %	0.93	0.01	0.92	0.93
VMT				
High-Speed Roads				
Low-Speed Roads				

174



175 ***Figure 2 Seat belt non-use map***

Spatial Analysis

Our methodology consists of several parts. First, we will examine the spatial clustering of seat belt non-use at the census tract level, searching for distinctive spatial regimes in the data. With the assumption of spatial regimes' presence, we will estimate separate models for each regime to learn whether the models are significantly varying across each regime. Next, we estimate the effects of structural variables on seat belt non-use with adjustments for spatial dependency with the assumption of substantially different models. Finally, we will assess the extent to which any observed spatial dependence is best described with reference to the effects of unmeasured predictor variables (the spatial error model) or with reference to the spatial lag effect in neighboring census tracts (the spatial lag model).

Exploratory Spatial Data Analysis (ESDA)

ESDA discovers spatial association patterns or clusters and suggests spatial regimes or other forms of spatial heterogeneity (Anselin 1990, 1999, Baller *et al.* 2001). First, we use Global Moran's I statistics (Moran 1950) to investigate the presence of spatial autocorrelation. Moran's I values range from -1 (Perfect dispersion) to +1 (Perfect clustering). The extreme values are indicators of significant spatial autocorrelation where values close to 0 indicate a random pattern between residuals. Moran's I can be written as:

$$I = \frac{\sum_i \sum_j w_{ij}(y_i - \mu)(y_j - \mu)}{\sum_i (y_i - \mu)^2} \qquad \text{Equation 1}$$

where w_{ij} is an element of a row-standardized spatial weights matrix, y_i is the seat belt non-use, and μ is the average seat belt non-use in the sample. Moran's I's statistical significance is based on the Z-score (Andrew and Ord 1981).

Next, we will test the Local Indicator of Spatial Association (LISA) statistics. Local Moran's I is helpful to identify regimes that could be targeted by separate models. The LISA statistics check for local spatial autocorrelation by applying local Moran's statistics, which indicates to what extent the pattern of the seat belt non-use rates in one geographic unit is compatible with spatial randomness. Rejection of the null hypothesis indicates that local clustering of high-high (units with high values surrounded by units with high values), low-low (units with low values surrounded by units with low values), high-low (units with high values surrounded by units with low values), and low-high (units with low values surrounded by units with high values) exist. The Local Moran's I can be calculated by the following equation:

$$I_i = \left(\frac{z_i}{\sum_i {z_i}^2}\right) \sum_j w_{ij} z_j \qquad \textbf{\textit{Equation 2}}$$

where z refers to seat belt use in the mean-deviation form. For more details about the Moran's I please see Anselin and Florax (1995).

Stability of the coefficients

In this step, we apply a spatial regime regression using two separate ordinary least squares (OLS) models, which allow the coefficients to be different in each regime (High-high regime vs. others). A spatial Chow test on the stability of these coefficients across regimes (Chow 1960, Anselin 1990, Myers *et al.* 2017) produces a statistic similar to the F-statistic, which detects differences in selected covariates between census tracts across two regimes. If regional stability is rejected, the modeling allows for varying spatial processes to be considered in each region (Baller *et al.* 2001).

Regression models

The Spatial lag model (SLM) and the spatial error model (SEM) are two common methods to address spatial dependency. The methodological distinction between the two models is how they consider spatial dependency (Figure 3) (Doreian 1980, 1982). The SLM considers spatial dependency as a spatial lag, which is a weighted average of values for the dependent variable in neighboring locations. The SEM incorporates the spatial dependency in the error term.

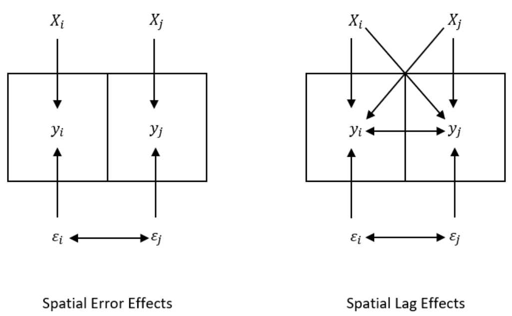


Figure 3 Spatial Error vs. Spatial Lag effect

Spatial error

A satisfactory spatial error model implies that it is unnecessary to posit the distinctive effects of the lagged dependent variable (Anselin 1990). In the SEM, the error term is treated as a spatially structured random effect vector. The SEM is similar to linear regression models with an additional term for the spatial dependency of errors in neighboring units:

$$y = X\beta + \varepsilon \qquad \text{Equation 3}$$

$$\varepsilon = \lambda W_{\varepsilon} + u = (I - \lambda W)^{-1} u \qquad \text{Equation 4}$$

$$y = \lambda W_{y} + X\beta + \lambda WX\beta + u \qquad \text{Equation 5}$$

where y is a vector of seat belt non-use, X is a vector of independent variables presented in Table 1, β is the corresponding vector of estimated coefficients (X). In this model, ε is the error term, which contains two parts: W_{ε} and u. W_{ε} presents the spatially lagged error term corresponding to a weight matrix W and u refers to the spatial uncorrelated error term that satisfies the normal regression assumption ($u \sim \mathrm{N}(0, \sigma^2 I)$). Finally, λ presents the spatial error term parameters; if the value of the spatial error parameters equals zero, the SEM is similar to the standard linear regression model.

Spatial lag

The spatial lag model incorporates the spatial influence of unmeasured independent variables but also stipulates an additional effect of neighbors' seat belt non-use via the lagged dependent variable:

$$y = \rho W_{y} + X\beta + \varepsilon \qquad \text{Equation 6}$$

where y is a vector of seat belt non-use, where ρ presents the spatial lag parameter, W_y is a spatially lagged variable corresponding to W matrix, X is a vector of independent variables, β is the vector of estimated coefficients. Last, ε is assumed to be a vector of independent and identically distributed error terms. The model is appealing since it integrates the effect of both independent variables on the outcome with the network (interdependence) effect of W_y (Marsden and Friedkin 1993) i.e., a strategic interaction. The corresponding "reduced form" of equation 6 is

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \qquad \text{Equation 7}$$

Equation 7 illustrates how the dependent variable at each location is not only determined by X, but also by the X at all other locations through the “Leontief inverse” $(I - \rho W)^{-1}$. This is the model most compatible with common notions of influence processes because it implies an influence of neighbors' seat belt non-use that is not simply a result of measured or unmeasured independent variables (Marsden and Friedkin 1993, Leenders 2002, Vitale *et al.* 2016).

Weight matrix

Different types of weighting matrices were considered in this analysis to obtain the most suitable model; namely, rook, queen order 1 and 2, and distance-based weight matrix were used for the analysis. The optimal weighting matrix selection could be based on the AICc (Hurvich and Tsai 1989); the weight matrix with the lowest AICc is preferred (Fotheringham and Brunsdon, Nakaya *et al.* 2005, Hadayeghi *et al.* 2010, Nakaya 2014).

Model comparison and assessment

A Lagrange Multiplier principle was also used to test the specifications against SEM and SLM. These tests are based on the regression residuals obtained from model estimates under the null hypothesis regression (i.e., OLS). SLM and SEM models have their own specific LM statistics, which offer the opportunity to exploit the values of these statistics to suggest the likely alternative. The LM statistics against SEM (LM_{SEM}) and SLM (LM_{SLM}) models take the following forms:

$$LM_{SEM} = \frac{\left(\frac{e'W_e}{s^2}\right)^2}{T} \qquad \text{Equation 8}$$

$$LM_{SLM} = \frac{\left(\frac{e'W_e}{s^2}\right)^2}{\frac{(WXb)'M(WXb)}{s^2} + T} \qquad \text{Equation 9}$$

where e is a vector of OLS residuals, s^2 its estimated standard error, $T = tr[(W + W')W]$, tr as the matrix trace operator, and $M = I - X(X'X)^{-1}X'$. Both LM_{SEM} and LM_{SLM} are asymptotically distributed as $\chi^2(1)$ under the null. Several researchers illustrate the relative power of these tests by using extensive simulation studies (Anselin and Rey 1991, Anselin and Florax 1995, Anselin *et al.* 1996).

It is possible that in some cases both LM_{SEM} and LM_{SLM} statistics turn out to be highly significant. Anselin *et al.* (1996) developed a robust form of the LM statistics to deal with this issue. The robust tests perform well in a wide range of simulations and form the basis of a practical specification search, as illustrated by (Anselin and Florax 1995, Anselin *et al.* 1996).

In this study, we used the GeoDa software to estimate models and perform the LM tests (Anselin 2003). Furthermore, we used the White statistics to check the presence of heteroscedasticity (White 1980). Variance Inflation Factors (VIF) were also used to control potential multicollinearity in each step (O'brien 2007).

Results

Average seat belt non-use rates for selected census tracts (n = 4,097) is 10.1% (seat belt use rates = 89.9%) (SD = 4.1); which is close to the average roadside observations (88.1%.) for the same period in Tennessee (NHTSA 2017). Figure 4presents the seat belt non-use histograms at the zonal level.

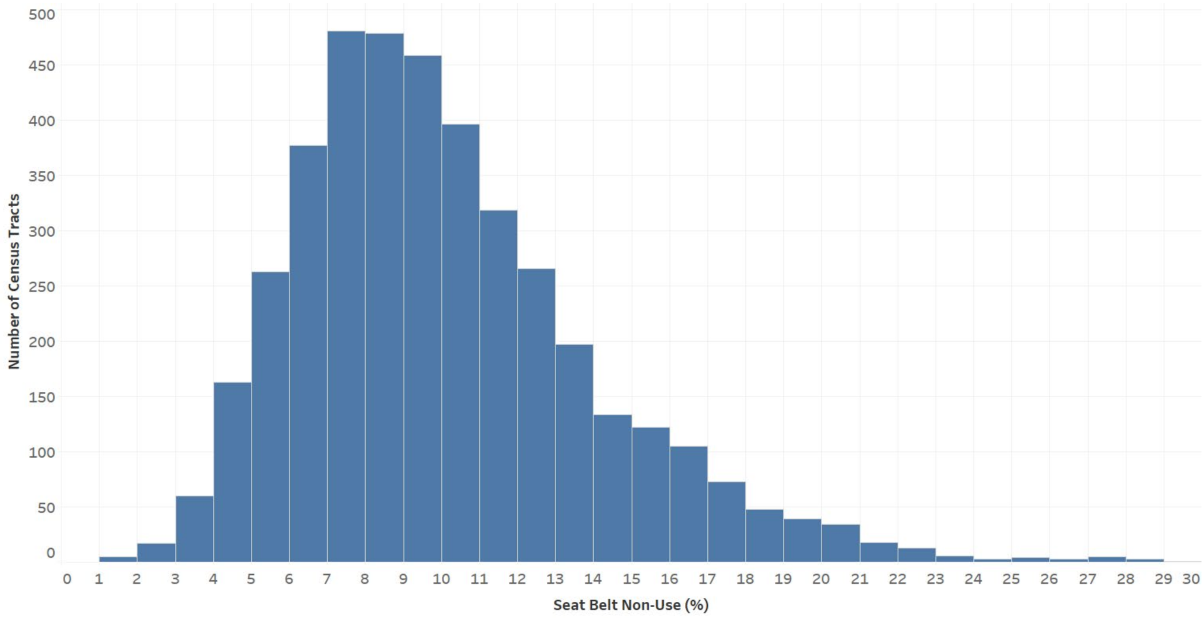


Figure 4 Distribution of seat belt non-use at the zonal level

Spatial diagnosis

A significant Global Moran's I value (I = 0.56) based on the queen contiguity matrix indicates the presence of substantial spatial dependency. The Moran's I statistic indicates that there is spatial autocorrelation in the OLS model, and the positive sign of the Moran's I shows that the neighborhoods with higher seat belt non-use are clustered together vice versa.

Figure 5 presents the visual map of local Moran's I. The clusters with high rates (i.e., high-high) are located in Chattanooga and Memphis metropolitans' areas as well as some scattered clusters in the rural areas. Alternatively, the clusters with low rates (i.e., low-low) are located in other metropolitan areas in Tennessee, namely the suburban areas surrounding the Nashville metropolitan area (except the urban core of Nashville), Knoxville, Clarksville, and Kingsport.

Based on Figure 5 and LISA, we conclude two regimes in Tennessee: Southern metropolitans and rural areas (i.e., Memphis and Chattanooga) –Southern MPOs – and other metropolitan areas, i.e., Non-Southern MPOs. The average seat belt non-use in the Southern metropolitan areas is 16% (90^{th} percentile range between 12-21%). On the other hand, seat belt non-use in the Non-Southern metropolitan areas is substantially lower, with average seat belt non-use of 9% (90^{th} percentile range between 5-13%).

Regression estimation

Table 2 presents the separate OLS models for seat belt non-use in Tennessee by considering a dummy variable for the regional effect to capture Southern MPOs. Positive significant values of the Moran's I (0.169, $p < 0.001$) and White test (409.03, $p < 0.001$), reveal a strong presence of both spatial dependency and heteroscedasticity in the model.

Table 2 also presents the results of the Chow test. The Chow test rejects the null hypothesis of coefficient stability. A closer examination of the individual tests on coefficient stability across regimes supports the conclusion that the proportions of the white population, population with a bachelor's degree, and age cohorts (i.e., percentages of the population aged 16-42 years and 43-59 years) exert significantly different effects across regions. Therefore, we estimate separate models for each regime and will scrutinize the presence of spatial dependence.

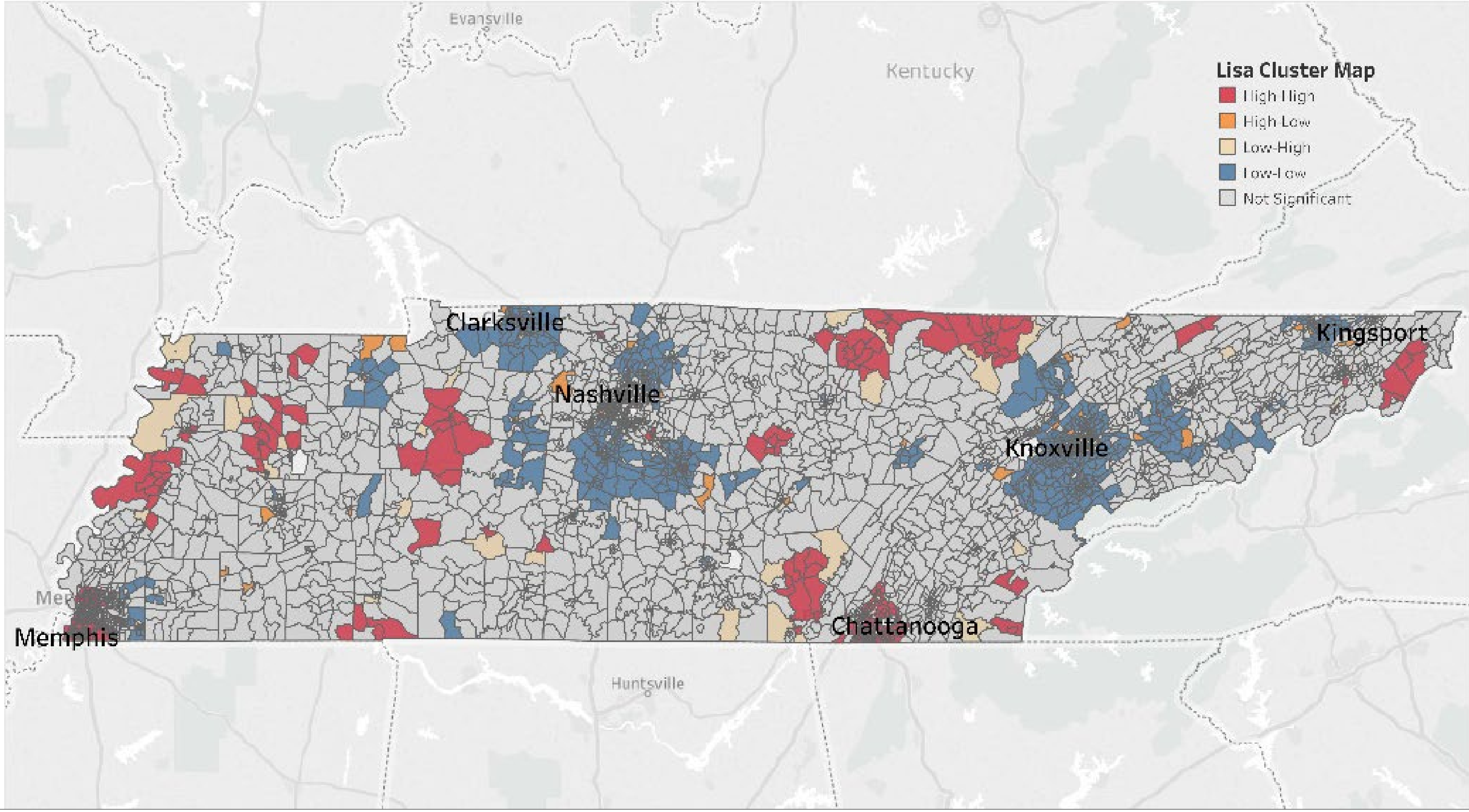


Figure 5 Moran's scatterplot map

Table 2 Ordinary Least Square regression of seat belt non-use and Chow test statistics

	Tennessee			Non-Southern MPOs			Southern MPOs			Chow statistics	
Variable	**Coef.**	**S. E.**	**P-value**	**Coef.**	**S. E.**	**P-value**	**Coef.**	**S. E.**	**P-value**	**Value**	**P-value**
Household with Vehicle	-3.064	0.743	0.000	-2.606	0.719	0.000	-3.000	1.538	0.051	0.054	0.817
VMT High-Speed Roads	-0.005	0.003	0.061	-4.86E-03	2.25E-03	0.031	1.20E-02	7.51E-03	0.109	4.646	0.031
VMT Low-Speed Roads	-0.018	0.019	0.343	0.015	0.016	0.361	-0.048	0.075	0.519	0.677	0.411
Population Density (per Square miles)	1.06E-04	7.00E-05	0.129	1.43E-04	6.39E-05	0.025	8.18E-05	1.78E-04	0.646	0.106	0.745
% with College Education	-5.453	0.658	0.000	-5.447	0.599	0.000	-0.874	1.694	0.606	6.479	0.011
% with Bachelor Education	-2.097	0.742	0.005	-2.085	0.689	0.002	-3.339	1.713	0.052	0.461	0.497
Median Household Income ($10,000)	-0.142	0.034	0.000	-0.108	0.031	0.000	-0.254	0.103	0.013	1.868	0.172
% Non-Motorized Road Users	2.104	0.961	0.029	-0.381	1.003	0.704	2.126	1.725	0.218	1.577	0.209
% Population 16-42 Years	-0.657	0.862	0.446	1.050	0.807	0.193	-4.812	2.034	0.018	7.174	0.007
% Population 43-59 Years	1.773	0.979	0.070	1.634	0.921	0.076	-2.977	2.156	0.168	3.869	0.049
% Population > 59 Years	1.159	0.804	0.150	0.971	0.743	0.191	-1.912	1.973	0.333	1.870	0.171
% White Population	-4.575	0.252	0.000	-1.676	0.279	0.000	-0.124	0.520	0.811	6.928	0.009
Constant	18.119	0.833	0.000	13.963	0.826	0.000	22.601	1.681	0.000	21.266	0.000
Global Chow Test										1140.04	0.000
AIC	21954.4			17288.4			3418.4				
Log-likelihood (Full)	-10964.2			-8631.2			-1696.2				
Adjusted R-squared	0.249			0.122			0.098				
Number of observations	4125			3463			634				

The estimated models based on different weight matrices were broadly in agreement. By comparing the AICc, we learned that the queen contiguity matrix has the lowest value of the AICc and therefore is more suitable than other models for spatial analysis.

As presented in Table 3, Moran's significant values indicate that spatial dependency exists in both regimes. Interestingly, White test statistics indicate that heteroscedasticity is present in the Southern MPOs, whereas there is heterogeneity in the Non-Southern MPOs.

The Lagrange Multiplier test (Table 4) suggests that for the Southern MPOs area, a spatial lag model is more suitable, whereas, in the rest of the study area, a spatial error model is more suitable. Table 5 presents the estimated SLM and SEM model for each region.

In both regimes, median household income and percentage of households with vehicles have a significant negative association with seat belt non-use. In the Non-Southern MPOs percentage, the white population at the census tract and percentage of the population with Bachelor's degree have a significant positive association with seat belt non-use. These variables did not have a significant association with seat belt non-use in the Southern MPOs. The significant association between income, race, and education-related variables are consistent with previous research (Preusser *et al.* 1991, Reinfurt *et al.* 1997, Wells *et al.* 2002, Vivoda *et al.* 2004, Houston and Richardson 2005, Gkritza and Mannering 2008, Pickrell and Ye 2009, Hezaveh and Cherry 2019).

Population density is correlated with lower seat belt use in Non-Southern MPOs. This negative impact could be attributed to the shorter distances in the urban areas and a relatively lower travel speed in general. As a result, vehicle occupants may decide not to use their seat belt in urban areas. Findings regarding the effect of vehicle ownership and population density are in agreement with Hezaveh and Cherry (2019).

Table 3 Moran's I and White test statistics for each regime

Test	Non-Southern MPOs		Southern MPOs	
	Value	P-value	Value	P-value
Moran's' I	0.2756	0.000	0.139	0.000
White Test	60.253	0.000	7.122	0.624

Table 4 LM test statistics for each regime

Test	Non-Southern MPOs		Southern MPOs	
	Value	P-value	Value	P-value
Lagrange Multiplier (lag)	464.93	0.00	33.74	0.00
Lagrange Multiplier (error)	586.99	0.00	23.35	0.00
Robust LM (lag)	0.43	0.51	11.62	0.00
Robust LM (error)	122.49	0.00	1.23	0.27

Table 5 Results of the spatial models

	SLM (Southern MPOs)			**SEM (Non-Southern MPOs)**		
Variable	**Coef.**	**S. E.**	**P-value**	**Coef.**	**S. E.**	**P-value**
% White Population	0.130	0.501	0.795	-1.090	0.335	0.001
% with College Education	-2.738	1.647	0.096	-1.102	0.657	0.094
% with Bachelor Education	-0.414	1.621	0.798	-4.100	0.589	0.000
Median Household Income ($10,000)	-0.249	0.099	0.012	-0.125	0.031	0.000
% Non-Motorized Road Users	3.967	3.891	0.308	-0.382	0.929	0.681
Population Density (per Square miles)	1.24E-04	1.71E-04	0.470	2.02E-04	6.19E-05	0.001
Household with Vehicle	-3.045	1.295	0.019	-2.642	0.663	0.000
VMT High-Speed Roads	0.009	0.008	0.227	-0.001	0.002	0.692
VMT Low-Speed Roads	-0.038	0.072	0.596	0.006	0.015	0.677
% Population 16-42 Years	-4.165	1.926	0.031	0.380	0.755	0.615
% Population 43-59 Years	-2.725	2.116	0.198	0.502	0.830	0.545
% Population > 59 Years	-1.866	1.897	0.325	0.008	0.691	0.991
Constant	17.405	1.803	0.000	13.736	0.786	0.000
Lag Coeff (Rho)	0.300	0.055	0.000			
Lag Coeff (LAMBDA)				0.493	0.021	0.000
AIC	**3392.740**			**16770.100**		
Log-likelihood (Full)	**-1682.370**			**8372.040**		
Adjusted R-squared	**0.167**			**0.282**		
Number of observations	**634**			**3463**		

Conclusion

In this study, we used seat belt use reported by police officers at crash sites to explore the spatial dependency of seat belt non-use at the zonal level. We found that seat belt non-use rates are not randomly distributed in space. Southern-MPOs census tracts have higher-than-average seat belt non-use rates that form statistically significant clusters.

ESDA and Chow statistics reveal distinct regional imprints in Tennessee. The LM test results indicate that SLM and SEM are more suitable in Southern MPOs and Non-Southern MPOs, respectively. A comparison of the coefficients of the estimated models indicates that the models behave differently. Consequently, non-consideration of the spatial regimes in large scale models (i.e., at the state level) yields unreliable statistical inferences.

The spatial lag effect implies that seat belt non-use in Southern MPOs is not produced solely by the internal structural factors, and it is influenced by their neighboring units. The spatial lag model depicts a spatial imprint at a given instant that would be expected to emerge if the phenomenon under investigation was characterized by a diffusion process (or social influence) (Baller *et al.* 2001). However, a diffusion process ultimately requires vectors of transmission. The observation of spatial effects thus indicates that further inquiry into diffusion is warranted. In contrast, the failure to observe such effects implies that such inquiry is likely to be unfruitful (Baller *et al.* 2001). Understanding the social influence process and its underlying mechanisms would help design an effective road safety campaign, such as communication methods with the campaign recipients.

Implications

The current practice for selecting seat belt campaigns rely on blanket coverage for areas with lower seat belt use rate. The method presented in this study could help practitioners to more efficiently reach groups that are more prone to seat belt non-use. First, the developed methodology helps practitioners decide on seat belt campaign's geographic scopes instead of statewide blanket coverage. This method can effortlessly identify seat belt non-use hot zones, and agencies could prioritize these areas for focusing on Enforcement or educational resources.

Second, by identifying different spatial regimes, we estimated two separate models to predict seat belt non-use. Furthermore, this analysis indicates that the relationship between demographic variables and seat belt non-use varies across regimes. For example, the assumption that education and percentage of the White population positively impact seat belt use does not hold in the Southern-MPOs. Failing to consider the spatial regimes in the analysis would lead to falsely prioritizing groups that are more prone to seat belt non-use; this issue raises from the biased estimation of the aspatial models.

Moreover, one needs to consider the features of the area where the crash occurred; this could be achieved in disaggregate modeling that explores factors affecting seat belt non-use for each vehicle occupant. However, in this study, we used an aggregated approach, making it impossible to explore this matter. This issue could be explored in future studies.

Limitations

Having information about Enforcement and driving exposure at the census tract level could better understand seat belt use's spatial distribution. Unfortunately, this information was not available at the time of the study. Instead, we used VMT and vehicle ownership as proxies for driving exposure in our analysis.

The present study population consists of vehicle occupants with a home address in Tennessee involved in a traffic crash in Tennessee during 2014-16. This study's population is likely skewed towards those who are more prone to unsafe behavior (i.e., they were involved in crashes). Nevertheless, the sample used in this study consists of 1.25 million observations or about 19% of the state population. These findings present a sample of Tennessean vehicle occupants. Careful consideration is needed when transferring these findings to other settings. Nevertheless, the method and results that this study presents could be generalizable to other contexts as well.

Acknowledgments

The authors would like to express gratitude to the Tennessee Department of Safety & Homeland Security. This project was supported by the Collaborative Sciences Center for Road Safety, www.roadsafety.unc.edu, a U.S. Department of Transportation National University Transportation Center promoting safety.

References

Afghari, A.P., Hezaveh, A.M., Haque, M.M., Cherry, C., 2020. A home-based approach to understanding seatbelt use in single-occupant vehicles in tennessee: Application of a latent class binary logit model. Accident Analysis & Prevention 146, 105743.

Ajzen, I., 1991. The theory of planned behavior. Organizational behavior and human decision processes 50 (2), 179-211.

Ali, M., Haidar, N., Ali, M.M., Maryam, A., 2011. Determinants of seat belt use among drivers in sabzevar, iran: A comparison of theory of planned behavior and health belief model. Traffic injury prevention 12 (1), 104-109.

Andrew, C., Ord, J.K., 1981. Spatial processes: Models and applications. London: Pion.

Anselin, L., 1990. Spatial dependence and spatial structural instability in applied regression analysis. Journal of Regional Science 30 (2), 185-207.

Anselin, L., 1999. Interactive techniques and exploratory spatial data analysis. Geographical Information Systems: principles, techniques, management and applications 1, 251-264.

Anselin, L., 2003. Geoda 0.9 user's guide. Urbana 51, 61801.

Anselin, L., Bera, A.K., Florax, R., Yoon, M.J., 1996. Simple diagnostic tests for spatial dependence. Regional science and urban economics 26 (1), 77-104.

Anselin, L., Florax, R.J., 1995. Small sample properties of tests for spatial dependence in regression models: Some further results. New directions in spatial econometrics. Springer, pp. 21-74.

Anselin, L., Rey, S., 1991. Properties of tests for spatial dependence in linear regression models. Geographical analysis 23 (2), 112-131.

Azimi, G., Asgari, H., Rahimi, A., Jin, X., 2019. Investigation of heterogeneity in severity analysis for large truck crashes.

Baller, R.D., Anselin, L., Messner, S.F., Deane, G., Hawkins, D.F., 2001. Structural covariates of us county homicide rates: Incorporating spatial effects. Criminology 39 (3), 561-588.

Begg, D.J., Langley, J.D., 2000. Seat-belt use and related behaviors among young adults. Journal of Safety Research 31 (4), 211-220.

Blincoe, L., Miller, T.R., Zaloshnja, E., Lawrence, B.A., 2015. The economic and societal impact of motor vehicle crashes, 2010. (revised) (report no. Dot hs 812 013). Washington, dc: National highway traffic safety administration.

Calisir, F., Lehto, M.R., 2002. Young drivers' decision making and safety belt use. Accident Analysis & Prevention 34 (6), 793-805.

Chow, G.C., 1960. Tests of equality between sets of coefficients in two linear regressions. Econometrica: Journal of the Econometric Society, 591-605.

Ctr, 2018. 2017 survey of safety belt usage in tennessee final report. The University of Tennessee Center for Transportation Research.

Dee, T.S., 1998. Reconsidering the effects of seat belt laws and their enforcement status. Accident Analysis & Prevention 30 (1), 1-10.

Doreian, P., 1980. Linear models with spatially distributed data: Spatial disturbances or spatial effects? Sociological Methods & Research 9 (1), 29-60.

Doreian, P., 1982. Maximum likelihood methods for linear models: Spatial effect and spatial disturbance terms. Sociological Methods & Research 10 (3), 243-269.

Eby, D.W., Molnar, L.J., Olk, M.L., 2000. Trends in driver and front-right passenger safety belt use in michigan: 1984–1998. Accident Analysis & Prevention 32 (6), 837-843.

Fockler, S.K., Cooper, P.J., 1990. Situational characteristics of safety belt use. Accident Analysis & Prevention 22 (2), 109-118.

Foster, S.A., 1999. The geography of behaviour: An evolutionary perspective. Trends in Ecology & Evolution 14 (5), 190-195.

Fotheringham, A., Brunsdon, C., M. Charlton. 2002. Geographically weighted regression-the analysis of spatially varying relationships. Chichester, UK: John Wiley & Sons.

Gkritza, K., Mannering, F.L., 2008. Mixed logit analysis of safety-belt use in single-and multi-occupant vehicles. Accident Analysis & Prevention 40 (2), 443-451.
Glassbrenner, D., Carra, J.S., Nichols, J., 2004. Recent estimates of safety belt use. Journal of safety research 35 (2), 237-244.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using geographically weighted poisson regression. Accident Analysis & Prevention 42 (2), 676-688.
Hezaveh, A.M., Arvin, R., Cherry, C.R., 2019a. A geographically weighted regression to estimate the comprehensive cost of traffic crashes at a zonal level. Accident Analysis & Prevention 131, 15-24.
Hezaveh, A.M., Cherry, C.R., 2019. Neighborhood-level factors affecting seat belt use. Accident Analysis and Prevention 122, 153-161.
Hezaveh, A.M., Cherry, C.R., Year. Identifying seat belt non-use hot zones. In: Proceedings of the Transportation Research Board Annual Meeting 99th Annual Meeting.
Hezaveh, A.M., Nordfjærn, T., Everett, J., Cherry, C.R., 2019b. The correlation between education, engineering, Enforcement, and self-reported seat belt use in tennessee: Incorporating heterogeneity and time of day effects. Transportation Research Part F: Traffic Psychology and Behaviour 66, 379-392.
Houston, D.J., Richardson, L.E., 2005. Getting americans to buckle up: The efficacy of state seat belt laws. Accident Analysis & Prevention 37 (6), 1114-1120.
Hurvich, C.M., Tsai, C.-L., 1989. Regression and time series model selection in small samples. Biometrika 76 (2), 297-307.
Iihs, 2018. Safety belts. Insurance Institute for Highway Safety, Insurance Institute for Highway Safety.
Kamruzzaman, M., Hine, J., 2013. Self-proxy agreement and weekly school travel behaviour in a sectarian divided society. Journal of Transport Geography 29, 74-85.
Leenders, R.T.A., 2002. Modeling social influence through network autocorrelation: Constructing the weight matrix. Social networks 24 (1), 21-47.
Majumdar, A., Noland, R.B., Ochieng, W.Y., 2004. A spatial and temporal analysis of safety-belt usage and safety-belt laws. Accident Analysis & Prevention 36 (4), 551-560.
Marsden, P.V., Friedkin, N.E., 1993. Network studies of social influence. Sociological Methods & Research 22 (1), 127-151.
Merlin, L.A., Cherry, C.R., Mohamadi-Hezaveh, A., Dumbaugh, E., 2020. Residential accessibility's relationships with crash rates per capita. Journal of Transport and Land Use 13 (1), 113-128.
Mmucc, 2012. Model minimum uniform crash criteria. DOT HS 811, 631.
Mohamadi Hezaveh, A., 2019. Incorporating the home address of road users involved in traffic crashes in road safety analysis. University of Tennessee, Knoxville.
Mohamadi Hezaveh, A., Cherry, C.R., 2020. Applying a home-based approach to the understanding distribution of economic impacts of traffic crashes. Transportation Research Record, 0361198120953431.
Moran, P.A., 1950. Notes on continuous stochastic phenomena. Biometrika 37 (1/2), 17-23.
Myers, C.A., Slack, T., Broyles, S.T., Heymsfield, S.B., Church, T.S., Martin, C.K., 2017. Diabetes prevalence is associated with different community factors in the diabetes belt versus the rest of the united states. Obesity 25 (2), 452-459.
Nakaya, T., 2014. Gwr4 user manual. WWW Document. Available online: http://www. st-andrews. ac. uk/geoinformatics/wp-content/uploads/GWR4manual_201311. pdf (accessed on 4 November 2013).

Nakaya, T., Fotheringham, A.S., Brunsdon, C., Charlton, M., 2005. Geographically weighted poisson regression for disease association mapping. Statistics in medicine 24 (17), 2695-2717.

National Highway Traffic Safety Administration, 2010. Nigttime Enforcement of seat belt laws: An evaluation of three community programs. Traffic Safety Facts. Traffic Tech -Technology Transfer Series.

Nelson, D.E., Bolen, J., Kresnow, M.-J., 1998. Trends in safety belt use by demographics and by type of state safety belt law, 1987 through 1993. American Journal of Public Health 88 (2), 245-249.

Nhtsa, 2017. Seat belt use in 2017—use rates in the states and territories. In: Analysis, N.S.N.C.F.S.A. ed.

Nordfjærn, T., Jørgensen, S., Rundmo, T., 2011. A cross-cultural comparison of road traffic risk perceptions, attitudes towards traffic safety and driver behaviour. Journal of Risk Research 14 (6), 657-684.

Nordfjærn, T., Şimşekoğlu, Ö., Rundmo, T., 2014a. Culture related to road traffic safety: A comparison of eight countries using two conceptualizations of culture. Accident Analysis & Prevention 62, 319-328.

Nordfjærn, T., Şimşekoğlu, Ö., Zavareh, M.F., Hezaveh, A.M., Mamdoohi, A.R., Rundmo, T., 2014b. Road traffic culture and personality traits related to traffic safety in turkish and iranian samples. Safety science 66, 36-46.

O'brien, R.M., 2007. A caution regarding rules of thumb for variance inflation factors. Quality & quantity 41 (5), 673-690.

Pickrell, T., Ye, T., 2009. Traffic safety facts (research note): Seat belt use in 2008– demographic results. National Highway Traffic Safety Administration/Department of Transportation, Washington DC.

Preusser, D.F., Williams, A.F., Lund, A.K., 1991. Characteristics of belted and unbelted drivers. Accident Analysis & Prevention 23 (6), 475-482.

Reinfurt, D., Williams, A., Wells, J., Rodgman, E., 1997. Characteristics of drivers not using seat belts in a high belt use state. Journal of Safety Research 27 (4), 209-215.

Reinfurt, D.W., 2004. Documenting the sustainability of a mature click it or ticket program: The north carolina experience. J Safety Res 35 (2), 181-8.

Rentfrow, P.J., 2010. Statewide differences in personality: Toward a psychological geography of the united states. American Psychologist 65 (6), 548.

Şimşekoğlu, Ö., Lajunen, T., 2008. Social psychology of seat belt use: A comparison of theory of planned behavior and health belief model. Transportation Research Part F: Traffic Psychology and Behaviour 11 (3), 181-191.

Şimşekoğlu, Ö., Nordfjærn, T., Zavareh, M.F., Hezaveh, A.M., Mamdoohi, A.R., Rundmo, T., 2013. Risk perceptions, fatalism and driver behaviors in turkey and iran. Safety science 59, 187-192.

Solomon, M.G., Compton, R.P., Preusser, D.F., 2004. Taking the click it or ticket model nationwide. J Safety Res 35 (2), 197-201.

Thomas, A.M., Cook, L.J., Olson, L.M., 2011. Evaluation of the click it or ticket intervention in utah. Accident Analysis & Prevention 43 (1), 272-275.

Thomas, F.D., Blomberg, R.D., Peck, R.C., Cosgrove, L.A., Salzberg, P.M., 2008. Evaluation of a high visibility enforcement project focused on passenger vehicles interacting with commercial vehicles. J Safety Res 39 (5), 459-68.

Tison, J., Williams, A.F., 2010. Analyzing the first years of the click it or ticket mobilizations. In: Administration, N.H.T.S. ed.

Torquato, R., Franco, C., Bianchi, A., 2012. Seat belt use intention among brazilian undergraduate students. Revista Colombiana de Psicología 21 (2), 253-263.

Van Acker, V., Van Wee, B., Witlox, F., 2010. When transport geography meets social psychology: Toward a conceptual model of travel behaviour. Transport Reviews 30 (2), 219-240.
Vitale, M.P., Porzio, G.C., Doreian, P., 2016. Examining the effect of social influence on student performance through network autocorrelation models. Journal of Applied Statistics 43 (1), 115-127.
Vivoda, J.M., Eby, D.W., Kostyniuk, L.P., 2004. Differences in safety belt use by race. Accident Analysis & Prevention 36 (6), 1105-1109.
Wells, J.K., Williams, A.F., Farmer, C.M., 2002. Seat belt use among african americans, hispanics, and whites. Accident Analysis & Prevention 34 (4), 523-529.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: Journal of the Econometric Society, 817-838.
Xie, K., Ozbay, K., Yang, H., 2019. A multivariate spatial approach to model crash counts by injury severity. Accident Analysis & Prevention 122, 189-198.