PREDICTING PEDESTRIAN CRASH OCCURRENCE AND INJURY SEVERITY IN TEXAS

Mashrur Rahman
Community and Regional Planning, School of Architecture
The University of Texas at Austin
mashrur@utexas.edu

Kara M. Kockelman
(Corresponding Author)
Dewitt Greer Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu Tel: 512-471-0210

Kenneth A. Perrine
Center for Transportation Research
The University of Texas at Austin
kperrine@utexas.edu

Published in *Traffic Injury Prevention* in May 2022.

ABSTRACT

This study investigates pedestrian-involved crashes across Texas from 2010 through 2019. Crashes were mapped to over 708,738 road segments, along with road design, land use, transit, hospital, rainfall and other location features. Negative binomial model results show how total and fatal pedestrian-crash rates and counts rise with a segment's number of lanes, transit stops, population and job densities, as well as proximity to schools and hospitals, while greater median and shoulder widths provide some protection. Higher speed limits are associated with lower crash frequencies but more fatalities. A heteroskedastic ordered probit (HOP) model for injury severity demonstrates how pedestrian crashes are more likely to be severe and fatal at night (8 PM – 5 AM), without overhead lighting, and when the pedestrians or drivers are intoxicated. Use of light-duty trucks (including SUVs, pickup trucks, CUVs, and vans) also significantly increases the risk of pedestrians being severely injured or killed. While newer vehicle safety features may be argued to lower crash severity, newer crash-involved vehicles in Texas are not found to deliver less pedestrian injury. However, being a younger or female pedestrian, on a straight segment, off the state (and interstate) highway system, in the presence of a traffic control device (stop sign or signal) lowers the likelihood of pedestrian injury, when one does become involved in such a crash.

Keywords: Pedestrian safety; crash counts; injury severity; Negative Binomial (NB) model; Ordered Probit (OP) model; Heteroskedastic Ordered Probit (HOP) model.

INTRODUCTION

Increasing numbers of U.S. pedestrian injuries and deaths have become a major issue in traffic safety. The number of U.S. pedestrian fatalities rose 53% between 2009 and 2018, while total U.S. traffic deaths rose 8%. The share of pedestrian deaths, as a percentage of all U.S. crash fatalities, rose from 12% to 17% (GHSA, 2020), even though pedestrians make up less than 1% of all person-miles traveled in the nation (NHTS 2017). In the State of Texas, pedestrian fatalities rose by a stunning 86%, and their share of deaths went from 12% to 19%. While Americans are walking more, their walking distances cannot explain these numbers: National Household and Travel Survey (NHTS) data suggest that from 2009-2017, walking-miles traveled (WMT) per capita rose 13% and walking-trips per capita rose 6%. In contrast, pedestrian fatalities per capita rose 46%. In 2017, 10.4% of U.S. person-trips were walking-related, but pedestrian deaths were 16% of all traffic fatalities (FHWA, 2018). The soft, 25-lb to 250-lb frame of a pedestrian cannot compete with the higher speed, 2500-lb (and up) mass, and hard metal of motorized vehicle bodies. So, pedestrians experience dramatically higher risk than those seated inside such vehicles.

Development of effective crash countermeasures requires a comprehensive understanding of factors that influence both crash frequency and severity. Previous studies have found that certain roadway attributes, demographic and land use characteristics influence pedestrian crash frequency (Wang and Kockelman, 2013; Weir et al., 2009; Ukkusuri et al., 2012; Ukkusuri et al., 2008; Schneider et al., 2010). The spatial unit of analysis of those studies ranges from zone-level counts (at the census tract, zip code, county, or state level, for example) to segment and intersection counts. Weir et al. (2009) estimated how commercial land use shares, employment, population, and persons living below the poverty line have a positive impact on pedestrian crash frequency, at the U.S. Census tract level, while higher shares of persons over 65 years in age comes with lower counts of pedestrian crashes (presumably, in large part, because older persons tend to walk less distance outside). Ukkusuri et al. (2012) used both Census tract and zip code-level data to estimate how the shares of commercial and industrial land uses, and the numbers of schools and transit stops increase pedestrian crash frequency. The authors found different results depending of the level of data aggregation (census tract vs zip code) and concluded that more disaggregate data (for census tracts, in their case) provides more consistent results.

While zone-level data sets readily capture certain land use and built environment characteristics at the same scale of aggregation, micro-level studies can more effectively control for local design details and presumably better assess the benefits of many different countermeasure or safety improvement options. Schneider et al. (2010) analyzed pedestrian crash risk at 81 intersections in Alameda County, California and found that those with more right-turn-only lanes and those without raised medians on intersecting streets had more pedestrian crashes. While several studies have analyzed segment-level data for motor vehicle crashes (Xu et al., 2014; Aguero-Valverde and Jovanis, 2008; Ma et al., 2008; Kockelman et al., 2006), no such studies for pedestrian crashes were identified in this work.

Another important issue considering pedestrian safety is injury severity. Previous studies show that the variables associated with injury severity include: pedestrian and driver characteristics such as age, gender, intoxication, vehicle characteristics, roadway, and environmental factors (Lee and Abdel-Aty, 2005; Siddiqui et al., 2006; Kim et al., 2008; Kim et al., 2010; Aziz et al., 2013; Mohamed et al., 2013; Halem et al., 2015; Pour-Rouholamin and Zhou, 2016; Islam et al., 2016; Liu et al., 2019). Lee and Abdel-Aty (2005) used an ordered probit model for analyzing pedestrian crash data from Florida over 4 years (1999-2002). The study found that older (age 65 and over) and intoxicated pedestrians, high vehicle speed, heavy vehicles (van, pick up, bus) and reduced visibility increases the likelihood of injury severity. Kim et al. (2008) used a heteroskedastic model to address the individual-specific variance in crash severity analysis. Compared with a Multinomial Logit Model (MNL), the study showed a better fit for the heteroskedastic model. The unobserved effect (error term) varies more widely as the age of pedestrians increases over 65. Notable factors that increase the risk of pedestrian fatalities include pedestrian age, a driver that is male and intoxicated, speeding vehicles, dark conditions without streetlights, and vehicle types – particularly, SUVs and trucks. The study shows that intoxicated drivers increase the likelihood of pedestrian fatalities by 2.7 times.

Although previous studies have dealt with different pedestrian safety issues, those studies are few in number compared to the large volume of research devoted to crashes that only involve motor vehicles. No studies have been conducted on pedestrian crashes specifically in Texas. This study investigates 78,497 pedestrian-involved crashes in Texas over a 10-year period of time from 2010 to 2019. The study analyzes the relationship between segment-wise pedestrian crash counts and a variety of factors such as roadway characteristics, traffic attributes, demographic and environmental factors using a negative binomial (NB) model. Furthermore, the ordered probit models also investigate various driver, pedestrian, traffic, temporal and environmental characteristics that influence pedestrian injury severity. Findings from this research predict risk factors, help in understanding mitigations in infrastructure and vehicle design, motivate better data collection, and can be used to prioritize micro-level studies.

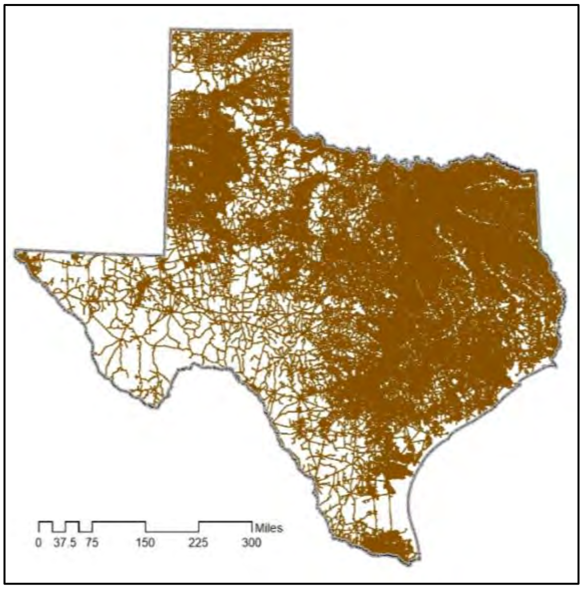
DATA DESCRIPTION

A key source of data for this study is the Texas Department of Transportation (TxDOT) Crash Records Information System (CRIS) (Texas Department of Transportation, 2020). These records come from police reports among all 254 Texas counties and hundreds of municipalities therein. Variables within the database characterize crashes according to time, location, severity, and road conditions. Crash records are not guaranteed to have all variables defined, and many of these data are not provided. A relevant aspect not captured by CRIS records involving pedestrians is whether each pedestrian is experiencing homelessness.

Although these characteristics of CRIS provide challenges when performing an analysis on crashes, CRIS remains a valuable resource, and offers suitable sample sizes for creating useful prediction models. From the year 2010 through 2019:

- 5,631,223 crash records exist
- 9,875,257 roadway vehicles are explicitly recorded among all crashes
- 4,756,671 crash records have geographic coordinates, either from GPS latitude/longitude written in the crash record, or geocoded from street names or addresses
- 78,497 are determined to involve collisions or avoidances of pedestrians
- 72,243 total pedestrians are explicitly recorded among all crash records
- 5,674 pedestrian fatalities are reported

Road-specific attributes were obtained from the TxDOT Roadway Inventory database (Texas Department of Transportation, 2018). The horizontal curves (GEO-HINI) database was spatially matched with the road inventory database to map road geometry. Census tract level population and job data were obtained from the 2010 population census and Longitudinal Employer-Household Dynamics (LEHD), respectively. Road segments were matched with the closest census tract centroid using the ArcGIS spatial join routine. All data were normalized by the area of census tracts. Other data sources include annual rainfall data (1981-2010) from the Texas Water Board, school locations from the Texas Education Agency, hospital locations from the Homeland Infrastructure Foundation-Level Data and transit stop locations from OpenStreetMap (OSM). Numbers of transit stops and Euclidean distances from each road segment to the nearest schools and hospitals were calculated using ArcGIS Spatial Analysis tools.



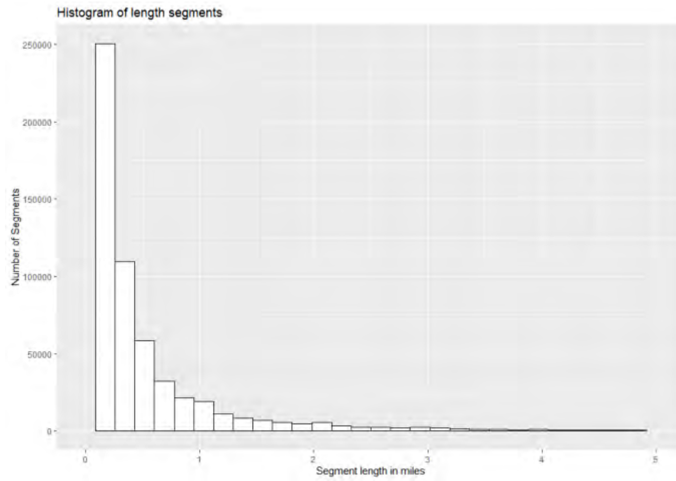


FIGURE 1: MAP SHOWING TEXAS ROADWAY SEGMENTS (LEFT); HISTOGRAM SHOWING THE DISTRIBUTION OF SEGMENT LENGTH (RIGHT)

TABLE 1: SUMMARY STATISTICS OF VARIABLES FOR ROAD SEGMENTS ACROSS TEXAS

	Mean	Std. dev	Min	Median	Max
Number of pedestrian crashes	0.0796	0.6530	0	0	115
Number of fatal pedestrian crashes	0.0068	0.1024	0	0	10
Segment length (in miles)	0.4338	0.8142	0.001	0.186	44.24
Number of lanes	2.2341	0.7835	1	2	14
Median width (in feet)	1.7407	11.789	0	0	519
Average shoulder width (in feet)	1.4066	3.6213	0	0	42
On system road	0.2246	0.4173	0	0	1
Indicator of curvature	0.1098	0.3126	0	0	1
Curve length (in meter)	21.676	125.77	0	0	9630.572
Curve angle (degrees)	3.5376	12.954	0	0	331.8
ADT per lane	888.35	2366.1	0	165	92090
Percentage of truck ADT	5.9598	7.2173	0	3.200	95.8
DVMT	1035.4	7319.4	0	54.418	793941.6
Speed limit (mph)	20.998	28.687	0	0	85
Rural (pop <5000)	0.4072	0.49130	0	0	1
Small urban (pop:5000-49999)	0.0977	0.2970	0	0	1
Urbanized (Pop:50000-199999)	0.0915	0.2883	0	0	1
Large urbanized (Pop: 200000+)	0.4036	0.4906	0	0	1
Population density (per sq mile)	1671.5	2274.9	0	635.830	55239.7
Job density (per sq mile)	805	3285.3	0	139.642	130011.1
Average yearly precipitation (1981-2010) (inches)	36.481	11.516	8	37	61
Distance to nearest hospital (miles)	6.8216	7.2760	0.0018	3.968	98.208
Distance to nearest school (miles)	2.0839	3.0864	0.01	0.741	53.952
Presence of transit stop within 100-meter buffer	0.0057	0.0753	0	0	1
Number of transit stops within 100-meter buffer	0.0114	0.2003	0	0	27

METHODOLOGY

ANALYSIS OF PEDESTRIAN CRASH COUNTS

The CRIS data were spatially matched with the road segments along with land use, population, job, rainfall and other location features (schools, hospitals, transit stops) to examine the association between pedestrian crash counts and various contributing factors along Texas roads. A total of 708,738 road segments were included in the analysis (Figure 1). Table 1 shows the summary statistics of the roadway segments.

A negative binomial (NB) model was used to predict pedestrian crash count along roadway segments. The expected number of counts $E(Y_i)$ along i_{th} segment is expressed as follows:

$$E(Y_i) = VMT_i^{\alpha} \exp\,(\beta_0 + \sum_{K} x_{ik}\beta_k + \varepsilon_i) \qquad (1)$$

VMT denotes vehicle miles traveled along i_{th} segment; parameter α shows potential non-linear relation between crash count and VMT. β_k is k_{th} covariates, ε_i is random error which follows gamma distribution $\varepsilon_i \sim \text{gamma}(\gamma, \gamma)$. Y_i represents crash counts with mean $E(Y_i) = \mu_i = VMT_i^{\alpha} \exp(\beta_0 + \sum_K x_{ik}\beta_k + \varepsilon_i)$ and variance $\text{Var}(Y_i) = \mu_i + \rho {\mu_i}^2$. Here, ρ is the dispersion parameter which collapses to a Poisson model when $\rho = 0$.

ANALYSIS OF PEDESTRIAN INJURY SEVERITY

Injury severity was analyzed at the individual crash level. Both standard ordinal probit (OP) and heteroskedastic ordered probit (HOP) models were used to account for the ordinal nature of injury severity. The model specification follows a latent variable framework:

$$y_i^* = \beta X_i + \varepsilon_i \qquad (2)$$

y_i^* is the underlying continuous latent variable representing injury severity of the i_{th} pedestrian. X_i is the vector (k×1) of explanatory variables; β is the vector (k×1) of unknown parameters to be estimated associated with explanatory variables; ε_i is the random error term which is unobserved. In probit, ε_i is assumed to be normally distributed with mean zero and unit variance.

In any given pedestrian crash, we only observe the injury severity y_i as reported by police in crash records. The relationship between the observed discrete variable y_i and the latent variable y_i^* is expressed as follows:

$$y_i = \begin{cases} 0, if\ y_i^* \leq 0 \ \ \text{(Not injured)} \\ 1, if\ 0 < y_i^* \leq \mu_1 \ \text{(Possible injury)} \\ 2, if\ \mu_1 < y_i^* \leq \mu_2 \quad \text{(Non-Incapacitating Injury)} \\ 3, if\ \mu_2 < y_i^* \leq \mu_3 \ \ \text{(Suspected serious injury)} \\ 4, if\ \mu_3 < y_i^* \leq \infty \ \ \text{(Killed)} \end{cases}$$

$\mu_0 = 0$ and μ_j $(j = 1,2,3)$ are threshold parameters (to be estimated) which determines among five observed values of injury severity, y_i. In general, the probability of y_i taking on injury severity j on i_{th} pedestrian can be expressed as follows:

$$\Pr(y_i = j \mid X_i) = \Phi\left(\frac{\mu_j - \beta X_i}{\sigma_i}\right) - \Phi\left(\frac{\mu_{j-1} - \beta X_i}{\sigma_i}\right) \qquad (3)$$

Φ is the standard normal cumulative distribution function, and σ_i is variance of the error term. In standard ordered probit models, it is assumed that variance of error term is constant across all observations. However, error term can vary across observations: for instance, there can be unobserved heterogeneity in terms of vehicle attributes such as vehicle type, weight and footprint (Wang and Kockelman, 2005; Chen and Kockelman, 2012; Lemp, Kockelman and Unnikrishnan, 2011) and in terms of pedestrian characteristics (health, weight and initial response to crashes) (Kim et al., 2010). Failure to account for heteroskedasticity can lead to biased parameter estimates in probit analysis. To overcome this limitation, a heteroskedastic ordered probit (HOP) was used where variance of the error term is allowed to vary. We follow a flexible specification for HOP model where $\sigma_i$is determined as a function of observed attributes associated with variance as the following equation (Wang and Kockelman, 2005):

$$\sigma_{i=} \exp(Z_i\gamma) \qquad (4)$$

γ is the coefficient for variable Z_i. If γ is not significantly different from zero for all Z_i, then it implies no heteroskedasticity and HOP takes the form of OP. On the other hand, if γ is significantly different from zero, it shows the presence of heteroskedasticity for that particular variable.

The parameters in Equation 3 were estimated by maximizing the log-likelihood function, that for a sample consisting of n observations:

$$\mathrm{L}\,(\beta, \mu, \gamma) = \sum_{i=1}^{n}\sum_{j=0}^{j=J} I(y_i = j)\ln\left(\Phi\left(\frac{\mu_j - \beta X_i}{\exp(Z_i, \gamma)}\right) - \Phi\left(\frac{\mu_{j-1} - \beta X_i}{\exp(Z_i, \gamma)}\right)\right) \qquad (5)$$

RESULTS AND DISCUSSION

PEDESTRIAN CRASH OCCURRENCE

Table 2 shows the parameter estimates of the NB models. Two models were estimated, one for all pedestrian crashes, and another for fatal pedestrian crashes. The dispersion parameters, ρ for both models are greater than zero, implying that the data are over-dispersed (the variance exceeds the mean of crash counts), and the NB model is preferred over the Poisson regression model.

The association between VMT and pedestrian crash frequencies is positive and non-linear (exponents α =0.7390 for all pedestrian crashes and α= 0.8730 for fatal pedestrian crashes), consistent with the expectation that crash frequencies increase with VMT but crash rate effectively falls as VMT of the segment rises. Among highway design variables, on-system roads (state-maintained arterials), median width, shoulder width and speed limit were found to be practically significant. On-system roads show strong association with fatal crashes: 42.81% increase of all pedestrian crashes vs 136.53% increase of fatal crashes only. As per CRIS data, two-thirds of all fatal pedestrian crashes in Texas (2010-2019) occurred on on-system roads. Other variables, such as shoulder width, median width and speed limit are negatively associated with pedestrian crashes. Higher speed limit roadways usually have fewer pedestrian activities which might contribute to lower numbers of pedestrian crashes; however pedestrian crashes on high speed segments are associated with more severe injuries, discussed later in the injury severity analysis.

Surprisingly, ADT per lane is estimated to have negative effects on pedestrian crashes when other variables are controlled (population and job density). Percentage of Truck ADT, however, shows positive association. This might be due to the fact that the impact of high ADT per lane is captured by population density and job density. Previous studies also found weak effect of ADT on pedestrian crashes when other variables are controlled (Huang et al., 2017; Pandey and Abdel-Aty; 2009; Zajac and Ivan, 2003).

Population density, job density and types of urban areas were used as proxies of land use. All of these variables were found to be strong predictors of pedestrian crashes. Pedestrian crashes including fatal crashes increase with population and job density, with very high crash rate percentage change (35.78% for population density and 11.06% for job density). This might be partly due to high variance-to-mean ratios for both of these variables; thus one-SD change implies a substantial shift. The effect of urbanization should be interpreted with urbanized areas having a population of 50,000-200,000 as a baseline. Compared to the baseline, large urban areas with populations greater than 200,000 are expected to have 23.05% and 14.63% more pedestrian crashes and fatal pedestrian crashes, respectively. By contrast, small urban areas and rural areas have fewer numbers of crashes. This is consistent with expectations because more dense locations in large urbanized areas usually have higher traffic volumes and pedestrian activities, thus increasing the exposure of pedestrian crashes.

Climate, proximity and transit-related variables such as rainfall, distance to the closest schools and hospitals, and the number of transit stops were also included in the model. Among these variables, distance to the closest schools, distance to the closest hospitals and the presence of transit offer practical significance although these variables are rarely considered in pedestrian safety literature. Results from the model estimation show that 1 SD decrease in nearest school distance (1 SD= 2.72 miles) is associated with a 52.45% increase in pedestrian crashes and a 22.92% increase in fatal pedestrian crashes. Similarly, hospital distance also shows strong association (except fatal crashes) but less significant than school distance. Finally, the presence of transit stops along the segments was found to be strongly significant (95.54% increase in pedestrian crashes and 53.46% increase in fatal pedestrian crashes), presumably due to high pedestrian activity near transit stops.

TABLE 2: ESTIMATION RESULTS OF NB FOR ALL PEDESTRIAN CRASHES AND FATAL PEDESTRIAN CRASHES

	All Ped Crashes			**Fatal Ped Crashes**			**% Of Change**	
	Coeff	Std. Error	Pr>|z|	Coeff	Std. Error	Pr>|z|	All ped crashes	Fatal ped crashes
Ln (VMT)	0.7390	0.0039	0.000	0.8730	0.0115	0.000		
Highway Design Variables								
Number of lanes	0.0316	0.0060	0.000	0.0459	0.0121	0.000	2.50%	3.60%
Median width	-0.0052	0.0005	0.000	-0.0033	0.0007	0.000	**-5.93%**	-3.86%
Shoulder width	-0.0187	0.0020	0.000	-0.0164	0.0036	0.000	**-6.55%**	**-5.76%**
On system roads	0.3564	0.0273	0.000	0.8678	0.0617	0.000	**42.81%**	**136.53%**
Indicator of curvature	0.0064	0.0281	0.820	-0.0576	0.0524	0.272	0.64%	-3.65%
Curve angle	-0.0047	0.0008	0.000	-0.0028	0.0014	0.044	-5.95%	-2.88%
Speed limit	-0.0093	0.0004	0.000	-0.0024	0.0012	0.037	**-23.46%**	**-6.43%**
Traffic Attributes								
ADT per lane	-5.5E-05	2.25E-06	0.000	-3E-05	3.84E-06	0.000	-12.26%	-6.95%
% of truck AADT	0.0054	0.0012	0.000	0.0056	0.0024	0.020	3.95%	4.14%
Land Use Variables								
Population density	0.0001	0.0000	0.000	0.0001	4.89E-06	0.000	**35.78%**	**17.46%**
Job density	3.19E-05	7.35E-07	0.000	0.0000	2.07E-06	0.001	**11.06%**	2.35%
Rural (pop<5000)	-0.6061	0.0321	0.000	-0.6200	0.0746	0.000	**-45.45%**	**-46.20%**
Small urban (pop:5000-49999)	-0.1213	0.0278	0.000	-0.1917	0.0774	0.000	**-11.42%**	**-17.44%**
Large urbanized (Pop: 200000+)	0.2074	0.0199	0.000	0.1366	0.0545	0.000	**23.05%**	**14.63%**
Ref: Urbanized (pop: 50000- 199999)								
Climate And Proximity Factors								
Rainfall	-0.0041	0.0005	0.000	0.0024	0.0014	0.000	-4.63%	0.098 2.80%
Distance to the nearest school	-0.2730	0.0083	0.000	-0.0958	0.0137	0.604	**-52.45%**	**-22.92%**
Distance to nearest hospital	-0.0227	0.0021	0.000	0.0022	0.0043	0.000	**-15.24%**	1.70%
Transit stop indicator	0.6706	0.0484	0.014	0.4290	0.1116	0.339	**95.54%**	**53.46%**
Number of transit Stops	0.0372	0.0151	0.000	0.0269	0.0281	0.000	0.75%	0.53%
(Intercept)	-7.3860	0.0448	0.000	-11.7900	0.1237	0.000		
No. of observations	708738							
Dispersion Parameter: ρ	2.01			1.39				
McFadden's R2:	0.278			0.335				
LR chi2	89206			17945				
Prob > chi2	0.0000			0.0000				
2 x log-likelihood	-231909.99			-35603.96				

Continuous variables show the % change for 1 SD increase. Binary variables show the % change from 0 to 1.
Bolded percentages indicate more practically significant variables

PEDESTRIAN INJURY SEVERITY

Both the ordered probit (OP) and heteroskedastic ordered probit (HOP) were estimated using the "oglmx" package in R (Carroll, 2017). Results from the likelihood ratio test suggest that heteroskedasticity exists ($\chi 2 = 2561.7$; $P<0.0001$), and therefore the HOP model was preferred over the OP model (Table 3). The coefficients of both models show consistent estimates; however, the main difference is observed in terms of variance components. The HOP model shows significant variance for pedestrian age, gender, speed limit, vehicle type, traffic control type, population of the area, time of day and lighting condition, suggesting that these variables can affect the spread of latent severity y_i^*. Other variables which do not show significant impacts are discarded from the variance equation. The following section discusses details about the impacts of explanatory variables on pedestrian injury severity.

Among different vehicle types, pick-up trucks, sports utility vehicles (SUVs), vans, heavy-duty trucks and buses significantly increase pedestrian injury severity in pedestrian-motor vehicle crashes (Table 3). Previous studies also reported similar findings, particularly high injury severity associated with light-duty trucks (SUVs, pickup trucks and vans) (Lefler & Gabler, 2004; Pour-Rouholamin and Zhou, 2016; Anarkooli et al., 2017; Liu et al., 2019). These vehicles pose higher risks due to heavy mass, higher bumpers and a more geometrically blunt frontal profile (Lefler & Gabler, 2004). The model also predicts significant variance for vehicle types, suggesting that impacts of unobserved attributes are associated with vehicle types (e.g. shape, stiffness, frontal profile) which increase the range of injury severity prediction. Marginal effects (Table 4) show that compared to passenger cars, light-duty trucks (pickup trucks, SUVs and vans) increase the probability of being killed or seriously injured by 13.9%. According to CRIS data, the number of light-duty vehicles involved in pedestrian deaths is increasing at a fast rate in Texas: during 2010-2018, the number of cars involved in fatal pedestrian crashes increased by 64.7%, while the number of SUVs and pickup trucks involved in fatal pedestrian crashes increased by 98.6% and 92.9%, respectively. Growing popularity of SUVs, pickup trucks and vans partly explains high injury severity associated with these vehicles. From 2009 to 2016, the share of cars to the total number of light duty vehicles purchased in the USA dropped from 60.5% to 43.8%, while during the same time period, share of SUVs, pickup trucks and vans increased from 39.4% to 56.2% (EPA, 2017).

Improved vehicle safety features contribute to pedestrian safety, and thus can reduce injury severity sustained by pedestrians in motor-vehicle crashes. These features include vehicle shape and stiffness, particularly, car front-end design – bumper height, bonnet leading edge, bonnet length and windscreen (Liu et al., 2002; Nie and Zhou, 2016; Li et al., 2018). Studies show that after the New Car Assessment Programs (Euro NCAP) in Europe, the newer car models exhibit safer front design (less bottom depth, flatter and wider bumpers) which significantly reduced pedestrian injury severity (Nie and Zhou, 2016; Li et al., 2018). In this study, we included vehicle model year in the injury severity model to understand if newer car models lead to less severe injury for pedestrians when struck by vehicles. However, the result does not show any significant impact of newer models (model year 2011 or later compared to those earlier than 2005) on pedestrian injury severity. This indicates that although newer vehicles models in the USA have succeeded in reducing injury severity for drivers and occupants (Chen and

Kockelman, 2010; Islam et al., 2016; Anarkooli et al., 2017), safety technology features have not improved much for pedestrians.

Pedestrian characteristics – both age and gender are found to be significant. Injury severity increases with pedestrians' age, suggesting that older people are vulnerable for more consequential outcomes. An increase of pedestrian age by one SD increases the risk of fatality by 1.69% and serious injury by 3.16%. Male pedestrians are also more likely to sustain severe injury than female counterparts. CRIS data shows that 72.38% of the pedestrians killed in motor-vehicle crashes in Texas from 2010-2019 were male. The effect of pedestrian age and gender on injury severity is consistent with the previous findings of Kim et al. (2008), Zhu et al. (2013), Pour-Rouholamin and Zhu (2016). The model also predicts significant heteroskedasticity for pedestrian gender and age. The unobserved effects of pedestrians on injury severity vary more widely as the age of the pedestrian increases.

Drivers' characteristics also affect pedestrian injury severity. Younger drivers (aged less than 24) significantly increase the risk of pedestrian injury compared to drivers of the middle-age group (25-64). Male drivers are also more likely to be involved in pedestrian crashes than female drivers. Previous studies also had similar findings regarding male and younger drivers (Kim et al., 2008, Kim et al., 2010; Pour-Rouholamin and Zhu, 2016); however, the effect of older drivers (aged 65 or above) is mixed (Kim et al., 2008; Siddiqui et al., 2006; Mohamed et al., 2013). The results show that drivers aged 65 or above increase injury severity for pedestrians; however, it should be noted that the effect size is small. Wood et al (2014) found that older drivers (age range 63–80) recognize pedestrians at approximately half the distance required for younger drivers (age range 18-38) which gives less response time to pedestrians.

Among different explanatory variables in the model, intoxication (in drivers and pedestrians) is found to have the strongest effect on pedestrian injury severity. Alcohol- or drug- related crashes are more likely to result in serious injury or deaths for pedestrians. According to CRIS data, alcohol and/or drugs were involved in 37.6% of pedestrian deaths. In most of these cases (33.38% of pedestrian deaths), pedestrians were tested positive in alcohol and/or drug screens. 88.84% of alcohol/drug-related pedestrian deaths were at dark. Walking under the influence, particularly at night, is one of the major causes of pedestrian fatalities.

With regard to time of day, crashes occurring from 8:00 PM – 5AM showed an increase in the probability of severe pedestrian injuries. 79.22% of pedestrian deaths occur at nighttime. This finding is consistent with previous studies (Pour-Rouholamin, 2016; Aziz et al., 2013; Kim et al., 2008). The results also show higher risk of severe injuries in early morning hours (5AM-7AM). There might be several possible explanations: during these time periods (late night and early morning hours), traffic is lighter than usual which might cause both pedestrians and drivers to ignore safety rules (drivers might travel at reckless speeds while pedestrian might choose to cross roads abruptly). Moreover, pedestrian activities early in the morning (walking, jogging, physical exercise) and alcohol/drug involvements at night (discussed earlier) combined with darkness might also contribute to high injury severity during overnight hours. Although the effect of darkness is controlled by the time of the day, lighting conditions also have a separate and significant influence. It is found that compared to daylight conditions, dark conditions increase the probability of severe injuries, however, a difference in probabilities of severe injuries between lighted roads and unlighted roads is also observed. Roads without streetlights at dark significantly increase the risk of pedestrian fatalities.

Roads with higher speed limits lead to more severe pedestrian injuries. Table 4 shows the change in predicted probabilities by injury severity levels due to one SD increase of speed limit. The positive association between speed limit and injury severity is consistent with previous studies (e.g. Halem et al., 2015; Chen and Fan, 2019). Although the posted speed limit usually influences vehicle speed on roads, a more appropriate indicator would be the actual speed of the vehicle at impact, which is difficult to obtain for a large number of cases. Speed limit increases the variance and outcome uncertainty: the unobserved effect varies more widely as the speed limit increases.

Hit-and-run crashes increase injury severity levels. 19.4% of pedestrian deaths are hit-and-run cases. Fleeing drivers increase the risk of pedestrian fatality because this often causes a delay in emergency service arrival and there is also the possibility that a pedestrian might get hit again by another vehicle after the first impact.

With regard to roadway characteristics, it is found that compared to city streets, there is a higher risk of severe pedestrian injury if a crash takes place on Interstate, US and State highways, county roads and other types of roads not classified. Generally, city streets accommodate speed limits and traffic controls, which reduces pedestrian crash severity. Analyzing CRIS data, we find that Interstate highways account for 5.5% of pedestrian crashes but 20.6% of pedestrian fatalities in Texas. This percentage becomes higher when restricted to major urban areas. For instance, IH-35 alone accounts for 28.2% of pedestrian deaths in Austin over the last ten years. Higher speeds, poor lighting conditions, pedestrians entering onto the highways, and lack of countermeasures might contribute to the severity of crashes on highways. Road geometry also affects crash severity. It is found that curved roads are more likely to result in severe injuries than straight roads at level. The marginal effect shows that curved roads increase the probability of fatal crashes by 4.7% and serious injury by 8.1%.

The location of the crash affects the type of injury. Crashes that occur at an intersection are associated with less severe injuries. Most pedestrian fatalities (89.16%) occur at non-intersection locations. The probability of less severe injury increases when the crash takes place off-roadways (e.g. parking lots, driveways), shoulders and medians, compared to on-roadways. Vehicle impact speed is usually lower in these locations, therefore there is less likelihood of severe injury.

The presence of traffic controls, such as traffic signals, reduces the probability of fatal and severe injuries. Pedestrians and drivers are better informed of each other's right of way and expected movements when there are traffic signals or traffic signs. As seen in studies on traffic calming in urban areas, drivers are usually more cautious and drive at lower speeds compared to places where there are no such controls (Ewing, 1999).

TABLE 3: INJURY SEVERITY RESULTS: ORDERED PROBIT VS HETEROSKEDASTIC ORDERED PROBIT (OP vs HOP) MODELS

	OP		HOP	
	Estimate	P-value	Estimate	P-value
Vehicle Type				
Pickup trucks	0.0945	0.000	0.1559	0.000
SUV	0.1042	0.000	0.1566	0.000
Heavy-Duty Truck	0.0479	0.029	0.1054	0.001
Van	0.0927	0.000	0.1435	0.000
Bus	0.1883	0.000	0.2665	0.001
Motorcycle	-0.1497	0.011	-0.1452	0.124
Others (ambulance, fire truck, police vehicle etc.)	0.0159	0.404	0.0262	0.270
(Reference vehicle = Passenger Car)				
Model Year				
After 2016	0.0268	0.200	0.0268	0.315
2011-2015	0.0245	0.045	0.0296	0.056
2005-2010	0.0818	0.000	0.1099	0.000
Unknown	0.0492	0.000	0.0579	0.001
(Reference Data = Before 2005)				
Pedestrian Age	0.0071	0.000	0.0083	0.000
Pedestrian Gender (1=Male)	0.1218	0.000	0.1537	0.000
Driver Age				
Driver Age (<24 years)	0.1550	0.000	0.2139	0.000
Driver Age (>65 years)	0.0357	0.013	0.0493	0.006
Driver Gender (1=Male)	0.1477	0.000	0.1861	0.000
Pedestrian/Driver Intoxicated	1.4382	0.000	2.8614	0.000
Speed Limit (mi/hr)	0.0171	0.000	0.0215	0.000
Hit-and-Run (1=Yes)	0.1353	0.000	0.1381	0.000
Crash Took Place At Intersection (1=Yes)	-0.1146	0.000	-0.1369	0.000
Road Type				
County Road	0.1097	0.000	0.1560	0.000
Farm To Market	0.1247	0.000	0.1597	0.000
Interstate	0.1087	0.000	0.1556	0.000
Non Trafficway	0.1005	0.000	0.1846	0.000
Other Roads	0.4114	0.000	0.5482	0.000
Tollway/Toll bridge	-0.4073	0.000	-0.3737	0.011
US State	0.1460	0.000	0.1867	0.000
(Reference type = City Streets)				
Crash Location				
Off Roadway	-0.1564	0.000	-0.0758	0.005
Shoulder	-0.1876	0.000	-0.1338	0.024
Median	-0.4384	0.000	-0.4544	0.000
(Reference location = On Roadway)				
Road Geometry				
Straight Grade	0.1426	0.000	0.2149	0.000
Curved	0.1939	0.000	0.2763	0.000
(Reference = Straight & Level)				
Control Type				
Traffic Sign	0.0224	0.044	0.0423	0.003
Traffic Signal	-0.0786	0.000	-0.0887	0.000
Other (human control, rail gate etc.)	-0.0131	0.556	-0.0034	0.896
(Reference = No Control)				

	OP		HOP	
	Estimate	P-value	Estimate	P-value
Area Population				
<5000	0.2085	0.000	0.2833	0.000
5000-9999	0.1466	0.000	0.1942	0.000
10000-24999	0.1394	0.000	0.2009	0.000
25000-49999	0.1132	0.000	0.1474	0.000
50000-99999	0.1012	0.000	0.1389	0.000
(Reference = 100000+)				
Crash Time				
5AM-7AM	0.3164	0.000	0.3959	0.000
7AM-11AM	0.1837	0.000	0.2190	0.000
4PM-8PM	0.1963	0.000	0.2349	0.000
8PM-11PM	0.2559	0.000	0.3166	0.000
11PM-5AM	0.2863	0.000	0.3799	0.000
(Reference = 11 AM-4PM)				
Lighting Condition				
Dark Lighted	0.1152	0.000	0.1329	0.000
Dark Not Lighted	0.2721	0.000	0.3599	0.000
(Reference = Daylight)				
HOP's Variance Equation				
Pedestrian Age (years)			0.0008	0.000
Pedestrian Gender (Male)			0.0515	0.000
Crash Speed Limit (mi/hr)			0.0052	0.000
Pickup Truck Indicator			0.0601	0.000
SUV			0.0337	0.000
Heavy-Duty Truck			0.1458	0.000
Van			0.0277	0.079
Bus			0.1966	0.000
Motorcycle			0.1717	0.000
Other Vehicle Type			-0.0161	0.223
Intersection			-0.0506	0.000
Traffic Sign			-0.0186	0.037
Traffic Signal			-0.0450	0.000
Other Control Type			-0.0258	0.140
Population: <5000 persons			0.0814	0.008
Population: 5000-9999			0.0678	0.004
Population: 10000-24999			0.0535	0.001
Population: 25000-49999			0.0007	0.966
Population: 50000-99999			-0.0666	0.000
Time: 5 AM- 7 AM			0.0687	0.000
Time: 7 AM-11 AM			-0.0246	0.024
Time: 4 PM-8 PM			0.0061	0.532
Time: 8 PM-11 PM			0.0217	0.132
Time: 11 PM-5 AM			0.0415	0.006
Dark & Lighted			0.0456	0.000
Dark & Not Lighted			0.0972	0.000
Threshold Parameters				
μ_0	0	-	0	-
μ_1	1.1813	0.000	1.4569	0.000
μ_2	2.2264	0.000	2.7943	0.000
μ_3	3.1568	0.000	4.1406	0.000

	OP		HOP	
	Estimate	P-value	Estimate	P-value
Number Of Observations	66,419		66,419	
Model Fit Statistics	**OP**		**HOP**	
Log-Likelihood	-88505.78		-87224.93	
Mcfadden's R2:	0.0601		0.0737	
AIC	177111.6		174603.9	
LR Test	**X_2 = 2561.7 (P<0.0001)**			

TABLE 4: MARGINAL EFFECTS (HOP)

	No Injury	Possible Injury	Non-Incapacitating Injury	Suspected Serious Injury	Killed
Car vs Vehicle Type					
Pickup Truck	-0.0034	-0.0305	-0.0172	0.0277	0.0234
SUV	-0.0084	-0.0299	-0.0080	0.0277	0.0186
Heavy-Duty Truck	0.0190	-0.0256	-0.0470	0.0168	0.0368
Van	-0.0081	-0.0271	-0.0068	0.0251	0.0169
Bus	0.0139	-0.0500	-0.0632	0.0328	0.0665
Motorcycle	0.0541	0.0049	-0.0612	-0.0170	0.0192
Others	-0.0056	-0.0042	0.0065	0.0039	-0.0007
Model Year: 2005/Older Model Vs Newer Model					
After 2016	-0.0044	-0.0085	0.0014	0.0079	0.0035
2011-2015	-0.0049	-0.0094	0.0015	0.0088	0.0039
2005-2010	-0.0170	-0.0348	0.0038	0.0327	0.0154
Unknown	-0.0093	-0.0183	0.0027	0.0172	0.0078
Pedestrian Age (One SD Increase)	-0.0148	-0.0338	0.0000	0.0316	0.0169
Pedestrian Gender (1=Male)	-0.0053	-0.0315	-0.0121	0.0294	0.0196
Driver Age: 25-65 Years Vs Other Age Groups					
Driver Age (<24)	-0.0303	-0.0677	0.0015	0.0635	0.0330
Driver Age (>65)	-0.0079	-0.0156	0.0023	0.0146	0.0066
Driver Gender (1=Male)	-0.0288	-0.0588	0.0063	0.0552	0.0261
Pedestrian/Driver Intoxicated	-0.0497	-0.2467	-0.2673	0.0059	0.5578
Crash Speed Limit (One SD Increase)	-0.0174	-0.0682	-0.0174	0.0634	0.0397
Hit And Run (1=Yes)	-0.0208	-0.0438	0.0037	0.0411	0.0199
Crash Took Place At Intersection	0.0040.	0.0289	0.0117	-0.0271	-0.0175
City Street Vs Road Types					
County Road	-0.0226	-0.0495	0.0022	0.0464	0.0234
Farm To Market	-0.0231	-0.0507	0.0022	0.0475	0.0240

	No Injury	Possible Injury	Non-Incapacitating Injury	Suspected Serious Injury	Killed
Interstate	-0.0226	-0.0494	0.0023	0.0463	0.0233
Non Trafficway	-0.0269	-0.0585	0.0030	0.0549	0.0276
Other Roads	-0.0531	-0.1612	-0.0541	0.1470	0.1214
Tollway/Tollbridge	0.0881	0.1023	-0.0592	-0.0995	-0.0318
Us State	-0.0269	-0.0592	0.0024	0.0555	0.0282
On Roadway Vs Other Location					
Off Roadway	0.0136	0.0236	-0.0061	-0.0221	-0.0090
Shoulder	0.0254	0.0409	-0.0130	-0.0386	-0.0148
Median	0.1143	0.1175	-0.0795	-0.1168	-0.0355
Curvature + Grade + Traffic Control					
Straight Grade	-0.0295	-0.0680	-0.0004	0.0637	0.0341
Curved	-0.0355	-0.0869	-0.0058	0.0813	0.0469
Traffic Sign	-0.0076	-0.0071	0.0079	0.0067	0.0002
Traffic Signal	0.0002	0.0194	0.0120	-0.0183	-0.0132
Other (human control, rail gate etc.)	-0.0046	0.0019	0.0090	-0.0019	-0.0044
Population					
<5000	-0.0101	-0.0509	-0.0265	0.0449	0.0425
5000-9999	-0.0051	-0.0363	-0.0208	0.0323	0.0299
10000-24999	-0.0084	-0.0374	-0.0157	0.0341	0.0274
25000-49999	-0.0133	-0.0275	0.0028	0.0258	0.0122
50000-99999	-0.0268	-0.0221	0.0282	0.0212	-0.0005
Crash Time: 11 AM-4 PM vs Other Times of Day					
5 am-7 am	-0.0210	-0.0701	-0.0245	0.0633	0.0523
7 am-11 am	-0.0240	-0.0411	0.0129	0.0384	0.0138
4 pm-8 pm	-0.0208	-0.0438	0.0034	0.0411	0.0201
8 pm-11 pm	-0.0240	-0.0587	-0.0038	0.0548	0.0317
11 pm-5 am	-0.0253	-0.0694	-0.0123	0.0642	0.0427
Lighting Conditions					
Dark + Lighted	-0.0041	-0.0262	-0.0124	0.0241	0.0185
Dark + Not Lighted	-0.0142	-0.0652	-0.0305	0.0581	0.0518

CONCLUSION

This study identified major risk factors associated with pedestrian crashes in Texas. Crash frequencies were analyzed at road segment level and injury severities were analyzed at the individual crash level. A rich database was constructed, including 10 years of crash records from CRIS along with road attributes, road geometry, and an extensive list of demographics, job, land use, climate variables, and important location features such as schools, hospitals and transit stops.

Findings from the NB model indicate the practical significance of micro-level variables in predicting pedestrian crashes. Proximity to schools, hospitals and presence of transit are associated with higher crash frequencies, although these variables are rarely included in pedestrian crash frequency models. Total crash rates and fatal crash counts rise with number of lanes, population and job densities, while greater median and shoulder widths provide some protection. Higher speed limits are associated with lower crash frequencies, but increase the likelihood of more severe injuries, as shown by the HOP model.

Results from the HOP model identified several risk factors at pedestrian, driver, roadway and vehicle levels that significantly affect pedestrian injury severity. Crashes occurring at night (8 PM – 5 AM), without overhead lighting, involving intoxicated pedestrians or drivers, and light-duty trucks (SUVs, pickup trucks, CUVs, and vans) are associated with more severe injuries. In contrast, being a younger and female pedestrian, on a straight segment off the state (and Interstate) highway systems, in the presence of a traffic control device (stop sign or signal) lowers the likelihood of pedestrian injury. Vehicles from more recent model years were not found to lower pedestrian injury, rather growing numbers of SUVs and CUVs being purchased in recent years further raises concerns about pedestrian safety. Findings from this study underscore the importance of enhanced vehicle safety features for pedestrians, campaigns against driving and walking while intoxicated, improved roadway design, enforcement of safety countermeasures near schools and bus stops and installment of additional traffic controls and streetlights where there are more pedestrian activities.

The study is not without some limitations. These data rely on reported and recorded crashes only; crashes with no injury or light injury often go unreported or unrecorded. Moreover, injury severities rely on police officers' initial assessments. Publicly available crash records do not include certain crash details due to privacy issues. Detailed police reports and hospital records may offer useful information about victims and motorists, including blood-alcohol levels, for example. More in-depth case studies, by specific crash site, vehicle dimensions and weight, hospital records, prior health issues, vehicle movements, pedestrians' position and action, homelessness and other unobserved factors are relevant, but require more digging.

AUTHOR CONTRIBUTION

The authors confirm contribution to the paper as follows: writing-original draft preparation: M. Rahman; conceptualization and design: K. M. Kockelman, M. Rahman; methodology: K. M. Kockelman, M. Rahman, K.A. Perrine; supervision: K. M. Kockelman; data assemble and analysis: M. Rahman, K.A. Perrine; writing-reviewing and editing: K. M. Kockelman, K.A. Perrine. All authors have reviewed the results and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

Funding for this research comes from TxDOT Research and Technology Innovation Project 0-7048. The authors are thankful for Jade (Maizy) Jeong's editing and submission support.

REFERENCES

Aguero-Valverde, J., & Jovanis, P. P. (2008). Analysis of road crash frequency with spatial models. *Transportation Research Record*, *2061*(1), 55-63.

Anarkooli, A. J., Hosseinpour, M., & Kardar, A. (2017). Investigation of factors affecting the injury severity of single-vehicle rollover crashes: a random-effects generalized ordered probit model. *Accident Analysis & Prevention*, *106*, 399-410.

Aziz, H. A., Ukkusuri, S. V., & Hasan, S. (2013). Exploring the determinants of pedestrian–vehicle crash severity in New York City. *Accident Analysis & Prevention*, *50*, 1298-1309.

Carroll, N. (2017). oglmx: Estimation of Ordered Generalized Linear Models. *R package version*, *2*(0.3).

Chen, Z., & Fan, W. D. (2019). A multinomial logit model of pedestrian-vehicle crash severity in North Carolina. *International Journal of Transportation Science and Technology*, *8*(1), 43-52.

Chen, T. D., & Kockelman, K. M. (2012). Roles of vehicle footprint, height, and weight in crash outcomes: application of a heteroscedastic ordered probit model. *Transportation Research Record*, 2280(1), 89-99.

Ewing, R. (1999). *Traffic Calming: State of the Practice*, ITE/FHWA, August 1999 (No. FHWA-RD-99-135). United States. Federal Highway Administration.

FHWA. (2018). Summary of travel trends: 2017 National household travel survey (Report FHWA-PL-18-019). *Federal Highway Administration, US Department of Transportation, Washington, DC*.

GHSA (2020). Pedestrian Traffic Fatalities by State: 2019 Preliminary Data. Governors Highway Safety Association. Retrieved from https://www.ghsa.org/sites/default/files/2020-02/GHSA-Pedestrian-Spotlight-FINAL-rev2.pdf

Haleem, K., Alluri, P., & Gan, A. (2015). Analyzing pedestrian crash injury severity at signalized and non-signalized locations. *Accident Analysis & Prevention*, *81*, 14-23.

Huang, H., Zhou, H., Wang, J., Chang, F., & Ma, M. (2017). A multivariate spatial model of crash frequency by transportation modes for urban intersections. *Analytic Methods In Accident Research*, *14*, 10-21.

Islam, S., Hossain, A. B., & Barnett, T. E. (2016). Comprehensive injury severity analysis of SUV and pickup truck rollover crashes: Alabama case study. *Transportation Research Record*, *2601*(1), 1-9.

Kim, J. K., Ulfarsson, G. F., Shankar, V. N., & Kim, S. (2008). Age and pedestrian injury severity in motor-vehicle crashes: A heteroskedastic logit analysis. *Accident Analysis & Prevention*, *40*(5), 1695-1702.

Kim, J. K., Ulfarsson, G. F., Shankar, V. N., & Mannering, F. L. (2010). A note on modeling pedestrian-injury severity in motor-vehicle crashes with the mixed logit model. *Accident Analysis & Prevention*, *42*(6), 1751-1758.

Kockelman, K., Bottom, J., Kweon, Y. J., Ma, J., & Wang, X. (2006). *Safety impacts and other implications of raised speed limits on high-speed roads* (Vol. 90). Washington, DC, USA: Transportation Research Board.

Lee, C., & Abdel-Aty, M. (2005). Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis & Prevention*, *37*(4), 775-786.

Lefler, D. E., & Gabler, H. C. (2004). The fatality and injury risk of light truck impacts with pedestrians in the United States. *Accident Analysis & Prevention*, *36*(2), 295-304.

Lemp, J. D., Kockelman, K. M., & Unnikrishnan, A. (2011). Analysis of large truck crash severity using heteroskedastic ordered probit models. *Accident Analysis & Prevention*, 43(1), 370-380.

Li, G., Wang, F., Otte, D., Cai, Z., & Simms, C. (2018). Have pedestrian subsystem tests improved passenger car front shape?. *Accident Analysis & Prevention*, *115*, 143-150.

Liu, J., Hainen, A., Li, X., Nie, Q., & Nambisan, S. (2019). Pedestrian injury severity in motor vehicle crashes: an integrated spatio-temporal modeling approach. *Accident Analysis & Prevention*, *132*, 105272.

Liu, X. J., Yang, J. K., & Lövsund, P. (2002). A study of influences of vehicle speed and front structure on pedestrian impact responses using mathematical models. *Traffic Injury Prevention*, *3*(1), 31-42.

Ma, J., Kockelman, K. M., & Damien, P. (2008). A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis & Prevention*, 40(3), 964-975.

Mohamed, M. G., Saunier, N., Miranda-Moreno, L. F., & Ukkusuri, S. V. (2013). A clustering regression approach: A comprehensive injury severity analysis of pedestrian–vehicle crashes in New York, US and Montreal, Canada. *Safety Science*, *54*, 27-37.

Nie, B., & Zhou, Q. (2016). Can new passenger cars reduce pedestrian lower extremity injury? A review of geometrical changes of front-end design before and after regulatory efforts. *Traffic Injury Prevention*, *17*(7), 712-719.

Pande, A., & Abdel-Aty, M. (2009). A novel approach for analyzing severe crash patterns on multilane highways. *Accident Analysis & Prevention*, *41*(5), 985-994.

Pour-Rouholamin, M., & Zhou, H. (2016). Investigating the risk factors associated with pedestrian injury severity in Illinois. *Journal of Safety Research*, *57*, 9-17.

Schneider, R. J., Diogenes, M. C., Arnold, L. S., Attaset, V., Griswold, J., & Ragland, D. R. (2010). Association between roadway intersection characteristics and pedestrian crash risk in Alameda County, California. *Transportation Research Record*, *2198*(1), 41-51.

Siddiqui, N. A., Chu, X., & Guttenplan, M. (2006). Crossing locations, light conditions, and pedestrian injury severity. *Transportation Research Record*, *1982*(1), 141-149.

Texas Department of Transportation (2018). Roadway Inventory. Accessed July, 2020: https://www.txdot.gov/inside-txdot/division/transportation-planning/roadway-inventory.html

Texas Department of Transportation (2020). Crash Data Analysis and Statistics. Accessed July, 2020: https://www.txdot.gov/government/enforcement/crash-statistics.html

U.S. Environmental Protection Agency (2017). Light-Duty Automotive Technology, Carbon Dioxide Emissions, and Fuel Economy Trends: 1975 Through 2017. Appendix D. U.S. EPA: Washington DC. https://nepis.epa.gov/Exe/ZyPDF.cgi?Dockey=P100TGDW.pdf

Ukkusuri, S., Miranda-Moreno, L. F., Ramadurai, G., & Isa-Tavarez, J. (2012). The role of built environment on pedestrian crash frequency. *Safety Science*, *50*(4), 1141-1151.

Wier, M., Weintraub, J., Humphreys, E. H., Seto, E., & Bhatia, R. (2009). An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. *Accident Analysis & Prevention*, *41*(1), 137-145.

Wang, X., & Kockelman, K. M. (2005). Use of heteroscedastic ordered logit model to study severity of occupant injury: distinguishing effects of vehicle weight and type. *Transportation Research Record*, *1908*(1), 195-204.

Wang, Y., & Kockelman, K. M. (2013). A Poisson-lognormal conditional-autoregressive model for multivariate spatial analysis of pedestrian crash counts across neighborhoods. *Accident Analysis & Prevention*, 60, 71-84.

Wood, J. M., Lacherez, P., & Tyrrell, R. A. (2014). Seeing pedestrians at night: effect of driver age and visual abilities. *Ophthalmic and Physiological Optics*, *34*(4), 452-458.

Xu, J., Kockelman, K. M., & Wang, Y. (2014). Modeling crash and fatality counts along mainlanes and 1 frontage roads across texas: 2 the roles of design, the built environment, and weather 3. In *93rd Annual Meeting of the Transportation Research*.

Zajac, Sylvia S., and John N. Ivan. "Factors influencing injury severity of motor vehicle–crossing pedestrian crashes in rural Connecticut." *Accident Analysis & Prevention* 35.3 (2003): 369-379.

Zhu, M., Zhao, S., Coben, J. H., & Smith, G. S. (2013). Why more male pedestrians die in vehicle-pedestrian collisions than female pedestrians: a decompositional analysis. *Injury Prevention*, *19*(4), 227-231.