Transportation Research Record
Two-Way Stop-Controlled Intersection Analysis with Zero-inflated Models
--Manuscript Draft--

Full Title:	Two-Way Stop-Controlled Intersection Analysis with Zero-inflated Models
Abstract:	Intersection safety continues to be a crucial issue throughout the United States. In 2016, 27 percent of the 37,461 traffic fatalities on U.S. roadways occurred at or near intersections. Nearly 70 percent of intersection-related fatalities occured at unsignalized intersections. At such intersections, vehicles stopping or slowing to turn create speed differentials between vehicles traveling in the same direction. This is particularly problematic on two-lane highways. Research was performed to analyze safety performance for intersections on rural, two-lane roadways, with stop control on the minor roadway. Roadway, traffic, and crash data were collected from 4,148 stop-controlled intersections of all 64 Parishes (counties) statewide in Louisiana, for the period of 2013 to 2017. Four count approaches, Poisson, Negative Binomial (NB), Zero-inflated Poisson (ZIP) and Zero-inflated Negative Binomial (ZINB) were used to model the number of intersection crashes for different severity levels. The results indicate thatZIP models provide a better fit than all other models. In addition to traffic volume, greater curve radius of major and minor roads, greater curve lengths of major roads, and greater lane widths of minor roads led to significantly smaller crash occurrences. However, intersections close to the beginning of curves, or located in the middle of curves, higher speed limits of minor roads, and urban areas led to significantly greater crash occurrences. Four-leg stop-controlled (4ST) intersections have 35 percent greater total crashes, 49 percent greater fatal and injury crashes, and 25 percent greater property damage only (PDO) crashes, relative to three-leg (3ST) intersections.
Manuscript Classifications:	Data and Information Technology; Safety Data, Analysis, and Evaluation ANB20; Safety Data; Design; Roadways, Right of Way, Users; Geometric Design AFB10; Intersections Geometric Design
Manuscript Number:	20-03303
Article Type:	Presentation and Publication
Order of Authors:	Ming Sun
	Xiaoduan Sun
	M. Ashifur Rahman
	Mousumy Akter
	Subasish Das

Two-Way Stop-Controlled Intersection Analysis with Zero-inflated Models

Ming Sun, Ph.D. candidate
(Corresponding author)
Department of Civil Engineering
University of Louisiana at Lafayette, Lafayette, Louisiana, LA 70504
Email: ming.sun1@louisiana.edu

Xiaoduan Sun, Ph.D., P.E.
Department of Civil Engineering
University of Louisiana at Lafayette, Lafayette, Louisiana, LA 70504
Email: xsun@louisiana.edu

M. Ashifur Rahman, Ph.D. candidate
Department of Civil Engineering
University of Louisiana at Lafayette, Lafayette, Louisiana, LA 70504
Email: ashifur@louisiana.edu

Mousumy Akter
Department of Civil Engineering
University of Louisiana at Lafayette, Lafayette, Louisiana, LA 70504
Email: C00419623@louisiana.edu

Subasish Das, Ph.D.
Texas A&M Transportation Institute, Texas A&M University System
3135 TAMU, College Station, TX 77843
E-mail: s-das@tti.tamu.edu

Word Count: 5,107 words + 4 tables = 6,107 words

Submitting to the 2020 TRB Annual Meeting for Presentation and Publication

August 1, 2018

ABSTRACT

Intersection safety continues to be a crucial issue throughout the United States. In 2016, 27 percent of the 37,461 traffic fatalities on U.S. roadways occurred at or near intersections. Nearly 70 percent of intersection-related fatalities occured at unsignalized intersections. At such intersections, vehicles stopping or slowing to turn create speed differentials between vehicles traveling in the same direction. This is particularly problematic on two-lane highways. Research was performed to analyze safety performance for intersections on rural, two-lane roadways, with stop control on the minor roadway. Roadway, traffic, and crash data were collected from 4,148 stop-controlled intersections of all 64 Parishes (counties) statewide in Louisiana, for the period of 2013 to 2017. Four count approaches, Poisson, Negative Binomial (NB), Zero-inflated Poisson (ZIP) and Zero-inflated Negative Binomial (ZINB) were used to model the number of intersection crashes for different severity levels. The results indicate that ZIP models provide a better fit than all other models. In addition to traffic volume, greater curve radius of major and minor roads, greater curve lengths of major roads, and greater lane widths of minor roads led to significantly smaller crash occurrences. However, intersections close to the beginning of curves, or located in the middle of curves, higher speed limits of minor roads, and urban areas led to significantly greater crash occurrences. Four-leg stop-controlled (4ST) intersections have 35 percent greater total crashes, 49 percent greater fatal and injury crashes, and 25 percent greater property damage only (PDO) crashes, relative to three-leg (3ST) intersections.

Keywords: Zero-inflated models, intersections, two-way stop-controlled, two-lane highways, severity

INTRODUCTION

Intersection safety continues to be a crucial issue throughout the United States. In 2016, 27 percent of the 37,461 traffic fatalities on U.S. roadways occurred at or near intersections. Nearly 70 percent of intersection-related fatalities occured at unsignalized intersections (*1*). Unsignalized intersections are of particular concern, because majority of intersections along low- to moderate-volume roads in rural and suburban areas are unsignalized. Unsignalized intersections represent potential hazards not present at signalized intersections. At such intersections, vehicles stopping or slowing to turn create speed differentials between vehicles traveling in the same direction. This is particularly problematic on two-lane highways (*2*). Having an intersection on a horizontal curve could double the crash risk, because of combined challenges. Although the AASHTO states that "an intersection on a sharp curve should be avoided or designed to compensate for reduced sight distance", in design practice, it is often allowed to have intersections on horizontal curves, if other solutions are prohibitively expensive (*3*). Many such intersections were constructed after, or long after, the major roadway was built, in order to provide accessibility to a minor street. There are many intersections on horizontal curves, located on state-owned and locally-owned roads in Louisiana, based on our preliminary investigation. **Figure 1** shows a collision that occurred at a T-intersection (a popular intersection type on rural, two-lane roadways controlled by a stop sign) between a right-turning vehicle and a ROR vehicle, trying to negotiate the curve.

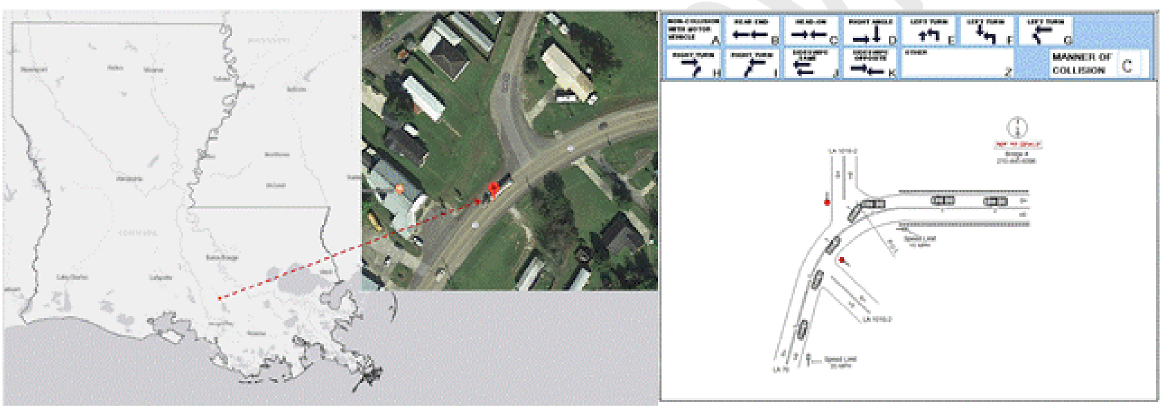

Figure 1 A typical example of an intersection on a curve

Intersection safety has been a long-standing problem in Louisiana. Unfortunately, in 2016, in Louisiana, intersection-related fatalities and severe injuries accounted for 19.1 percent of total fatalities and 39.9 percent of total severe injuries. More than 55 percent of intersection-related fatalities occured at unsignalized intersections (*4*). To get a better understanding of the influential factors on intersection crash frequency and injuries, count-data modelings are widely used to relate the number of crashes of different types or severities to site characteristics. These models always include traffic volume (Average Annual Daily Traffic, AADT), but also include site characteristics, such as lane width, radius/degree of horizontal curves, and presence of turn lanes at intersections. To fulfill the hefty goal established by Louisiana's Strategic Highway Safety Plan (SHSP) to reduce roadway departure, intersection, and non-motorized user fatalities and severe injuries by 50 percent by 2030, there is a clear need to further investigate intersection safety performance on two-lane roads with geometric characteristics.

LITERATURE REVIEW

Unsignalized intersection safety has been conducted by many studies in the past. Investigating characteristics of intersection crashes, identifying the risk factors related to the injury severity levels, and predicting crash frequency were the key focuses of these studies.

Various models were developed to study the relationship between crashes at unsignalized intersections and contributing factors. Bauer and Harwood (*5*) applied multiple linear regression analysis in developing crash prediction models for at-grade intersections in California, using three years of crash data (1990 to 1992), as well as geometric design, traffic control, and traffic volume data. The multiple linear regression was used for urban four-leg stop-controlled and signalized intersections, while Poisson and Negative Binomial (NB) regression were used for the remaining intersection types. Since crash occurrences are more likely random events, Poisson and NB models have been used extensively in prior studies (*6-10*). The results indicated that roadway geometric, vehicular, and operational features had an effect on crash frequency. Therefore, those factors that significantly affect crashes should be given more attention in crash analyses at intersections (*11*).

The presence of horizontal curves adds complexity to intersections. Kuciemba and Cirillo's studies have shown that safety is affected by the presence of horizontal curves in close vicinity of intersections (*12*). Vogt and Bared Vogt (1999) describes the development of a NB regression model for three types of intersections on rural roads in California and Michigan, for the period of 1993 to 1995. The study involved 84 three-leg intersections, 72 four-leg intersections, and 49 signalized intersections. Degree of curve was found to increase the total number of crashes on three-leg intersections between four-lane major roads, and two-lane stop-controlled minor roads (*7*). Savolainen and Tarko conducted a study for the Indiana Department of Transportation (INDOT) and found that curvature was a significant factor in the relative safety of intersections, where the intersection is two-lane two-way stop controlled, and the major road is a rural four-lane divided highway. NB models were developed to determine the statistical relationship between the crash occurrence and intersection geometric characteristics. The same study stated that full curvature and superelevation increased crashes by 30 percent, in comparison to tangent intersections (*3*).

When there is a zero crash record over a period of time, it may indicate either that the intersection is nearly safe, or that the zero record is a chance occurrence or crashes are not reported. Since the standard Poisson and NB models do not help to identify crash contributory factors in this case, it becomes necessary to model the two states (*11*). To handle count data with excess zeros, the zero-inflated models (ZIP and ZINB) have been used in many traffic safety studies. Miaou et al. first used ZIP structure for traffic crash analysis (*13*). Shankar et al. presented an empirical review into the applicability of zero-inflated count data modeling to roadway segment crash frequencies. The findings show that the ZIP structure models have great flexibility in uncovering processes affecting crash frenquencies on roadway sections observed with zero crash and those with observed crash occurrence (*14*). A study by Lee et al. used zero-inflated count models and nested logit models for developing crash frequency models and severity models. The findings also showed significant potential in applying these two techniques to single vehicle crash analysis (*15*). Empirical models based on ZIP were presented and discussed in terms of their applicability to pedestrian crash in two studies (*16, 17*). The results showed that ZIP is effective enough to provide explanatory insights into the causality behind pedestrian-traffic crashes. Lord et al. used ZIP and ZINB to account for the dominance of excessive zeroes observed in crash count data of vehicle crashes (*18*). Zero-inflated models have also been used to analyze crash severtity on rural two-lane roadway segments (*19*).

There has been limited research using zero-inflated models to model the expected crashes with large localized intersection database. In Louisiana, proximately one-third of total two-way stop-controlled (stop sign on minor roadway) intersections have zero crashes during the most recent five years. Consider that factor and address the gap, research was performed to analyze safety performance for intersections on two-lane roadways, with stop control on the minor roadway. Roadway, traffic, and crash data were collected for 4,148 stop-controlled intersections of all 64 Parishes (counties) statewide in Louisiana, covering the period of 2013 to 2017. A series of models were generated for different crash severities using conventional models (Poisson and NB) and zero-inflated models (ZIP and ZINB).

DATA

Prior to the crash modeling, a huge effort was made to develop a comprehensive database based on the intersection crashes, traffic volume, and other relevant roadway characteristics for two-way stop-controlled intersections on Louisiana rural two-lane highways. The most recent five-year crash data (2013 to 2017) were used to reflect the latest highway safety conditions. As shown in **Figure 2,** there are several important steps in retrieving and merging different data files from Louisiana Department of Transportation and Development (LADOTD).

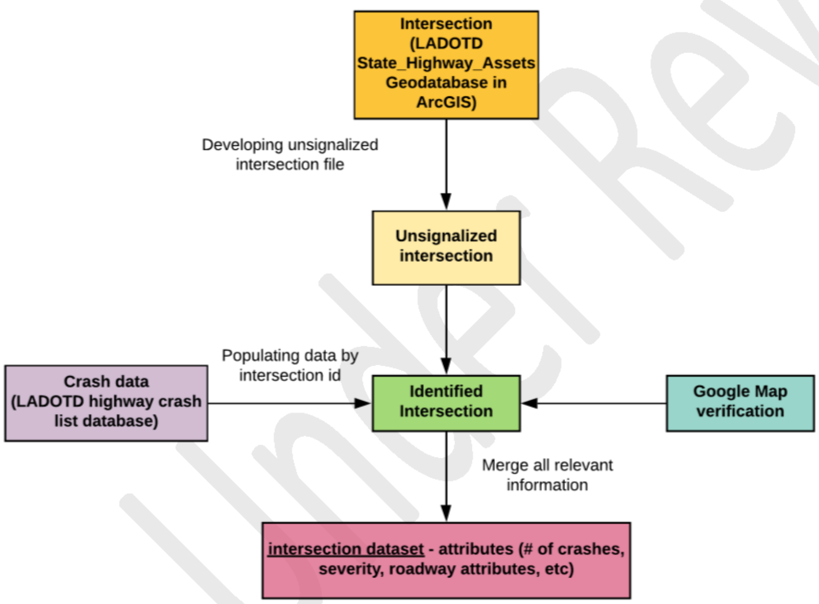

Figure 2 Database development flow chart

To assure the data accuracy, a significant effort was made to verify and correct the information presented in the data files shown in **Figure 2**. The shapfiles from the LADOTD *State_Highway_Assets Geodatabase* in ArcGIS format provide the spatial basis for gathering the necessary roadway attributes for the intersections, such as horizontal curve radius for major and minor roads, lane widths, number of intersection legs, speed limits, etc. The AADT data was obtained system-wide, for each rural intersection, from the LADOTD roadway inventory file. Satellite imagery and street-level imagery in Google Maps were utilized to exam the data accuracy intersection by intersection such as turning radiuses and type of traffic control. Further, additional information that was not otherwise included in the existing datasets, was manually collected, which includes the presence of left- and right-turn lane, and intersection locations (beginning of curve, or middle of curve). The LADOTD's crash database uses the ABCDE scale to describe severity level of crashes. 'A' indicates fatal injury, 'B', 'C', and 'D' indicate incapacitating or severe injury, non-incapacitating or moderate injury, and possible or compliant injury, respectively. 'E' represents that no injuries occurred in the crash. In this study, A, B, C, and D levels of severity were combined as fatal and injury crashes. The end product from the process illustrated in **Figure 2** is a comprehensive database containing information on intersection, roadway, and crash characteristics. The data process was facilitated by the intersection identification number and intersection name, used in the LADOTD system, which allowed data from different sources to be identified. Incomplete and, obviously, incorrect data (such as AADT entered as zero or blank) were removed. **Table 1** gives an overview of descriptive statistics of the intersections used in the model development.

TABLE 1 Overview of Two-Way Stop-Controlled Intersections (4,148 intersections)

Variable	Mean	SD	Min	Max
ADT (Average daily traffic)				
Major road	4,297	4,192	55	29,400
Minor road	769	1,366	10	17,000
Radius (Horizontal curve radius in ft.)				
Major road	474.7	716.8	0	3,222.0
Minor road	20.8	155.8	0	2,438.8
Curve length (Horizontal curve length in ft.)				
Major road	164.7	370.9	0	2,701.8
Minor road	12.1	97.1	0	2,388.3
Lane width				
Major road	11.6	0.8	8	16
Minor road	9.5	1.9	6	14
Speed limit				
Major road	47.0	9.2	25	65
Minor road	30.9	13.5	5	55
Intersection location (0 if on tangent, 1 if on beginning of curve, 2 if on middle of curve)	0.577	0.770	0	2
Number of leg (0 if 3-leg, 1 if 4-leg)	0.176	0.381	0	1
Left-turn lane presence (0 if no, 1 if yes)	0.022	0.148	0	1
Right-turn lane presence (0 if no, 1 if yes)	0.015	0.122	0	1
Setting (0 if rural, 1 if urban)	0.359	0.480	0	1
Total crashes	2.739	4.988	0	79
Fatal and injury crashes	0.983	1.840	0	20
Property damage only crashes	1.757	3.508	0	65

METHODOLOGY

Given the nature of random, discrete, and non-negative crash data, the Poisson distribution has been shown to provide a better fit and has been used widely to model crash frequency data (*7*, *8*, *17-19*). The probability of y_i crashes occurring at a given intersection i, $P(y_i|\lambda_i)$, is shown in **Equation 1**:

$$P(y_i|\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!};\ y_i = 0,1,2,3 \ldots \quad (1)$$

The relationship between the number of crashes at intersection i and the q parameter (X_{i1}, X_{i2},

X_{iq}) is shown in **Equation 2**:

$$\lambda_i = \exp\left(\beta_0 + \sum_{j=1}^{q} X_{ij}\beta_j\right) \tag{2}$$

Where, $\lambda_i =$ expected number of crashes per year at intersection i, X_i is the independent variables at intersection i, and β_j is a vector of estimable regression coefficients.

The Poisson regression model assumes that the mean of crash counts is equal to its variance (equal-dispersion). However, in much of the crash data, the variance is greater than the mean, well known as over-dispersion. For these cases, applying a Poisson regression model for intersection crash data would result in underestimation of the standard error of the regression parameters, which can, ultimately, lead to a biased selection of covariates. In some cases, excess zeros in crash data exist, considered to be a result of over-dispersion. In this study, 1,382 intersections had zero crashes during 2013 to 2017.The Poisson model cannot be used for these cases, as it cannot handle the over-dispersion, due to these high amount of zeros. To address this challenge, the ZIP model can be alternatively used. The ZIP model serves as a dual-state method for modeling data, characterized by a significant amount of zeros, or more zeros than one would expect in a traditional Poisson distribution. The ZIP model assumes that all zero counts come from two different processes: (i) the process generating excess zero count (zero-crash state) derived from a binary model and (ii) the process generating non-negative counts for intersection crashes including zero values, which estimated from the Poisson distribution (*17*). Suppose π_i is the probability that intersection i will exist in the zero-crash state, and $1 - \pi_i$ is the probability that crash counts are generated according to a Poisson model. Therefore, the probability distribution of the ZIP random variable is shown in **Equation 3**:

$$P(Y = y_i) = \begin{cases} \pi_i + (1 - \pi_i)e^{-\lambda_i}; & y_i = 0 \\ (1 - \pi_i)\frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}; & y_i > 0 \end{cases} \tag{3}$$

The probability of being in the zero-crash state, P_i, is often fitted using the losgistic regression model, as follows in **Equation 4**:

$$logit(P_i) = ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \sum_{j=1}^{q} Z_{ij}\beta_j \tag{4}$$

Where Z_{ij} is a function of at intersection i, and β_j is a vector of estimable regression coefficients. The mean and variance of ZIP are given as follows in **Equation 5 and 6**:

$$E(Y) = \lambda_i(1 - \pi_i) \tag{5}$$

$$Var(Y) = \lambda_i(1 - \pi_i)(1 + \lambda_i\pi_i) \tag{6}$$

Similar to ZIP model, the probability density function for the ZINB model is given by **Equation 7**:

$$P(Y = y_i) = \begin{cases} \pi_i + (1 - \pi_i)\frac{1}{(1+\alpha\lambda_i)^{1/\alpha}}; & y_i = 0 \\ (1 - \pi_i)\frac{\Gamma(\frac{1}{\alpha}+y_i)}{\Gamma(1+y_i)\Gamma(\frac{1}{\alpha})}\frac{(\alpha\lambda_i)^{y_i}}{(1+\alpha\lambda_i)^{\frac{1}{\alpha}+y_i}}; & y_i > 0 \end{cases} \tag{7}$$

Where α is the dispersion parameter. As α increases, this indicates that the data are more dispersed, which leads to higher standard error values.

RESULTS AND DISCUSSIONS

Model Selection

Four count approaches (Poisson, Negative Binomial, ZIP and ZINB) were used to model the number of intersection crashes occurring on the identified intersections. In a general sense, if partial observability and overdispersion are suspected, negative binomial variants of the ZIP model are plausible. In statistically validating any zero-altered model, one has to distinguish between the count models. A statistical test to make this distinction has been proposed by Vuong (1989). The Vuong-test is based on the t-statistic and has reasonable power in count-data applications (see Greene, 1994). The Vuong-statistic (V-statistic) is computed as **Equation 8**:

$$V = \frac{\bar{m}\sqrt{N}}{S_m}$$

Where $\bar{m}$ is the mean of $m = ln\left[\frac{f_1(.)}{f_2(.)}\right]$, $f_1(.)$ is the density function of the ZIP/ZINB distribution, and $f_2(.)$ is the density function of the parent-Poisson/Negative Binomial distribution, and S_m and N are the standard deviation and sample size, respectly. A value greater than 1.96 (the 95 percent confidence level for the t-test) for the V-statistic favors the ZIP/ZINB, while a value less than -1.96 favors the parent-Poisson/ Negative Binomial, with values in between 1.96 and -1.96, meaning that the test is inconclusive.

Table 2 summarized the results of considered models. In the NB and ZINB model, α is not statistically significant (p-value is 0.48 and 0.29 for the NB and ZINB models, respectively) which implies overdisperion is not due to unobserved heterogeneity in the crash data, but it may caused by excess zeros. The Vuong statistic for the Poisson versus ZIP (p-value = 0.0000) and NB versus ZIP (p-value < 0.05) favors the ZIP model for total, injury and property damage only (PDO) crashes, which as shown in **Table 3**.

TABLE 2 Model Goodness-of-fit Results

Model	Poisson	NB	ZIP	ZINB
Log-likelihood	-3743.99	-3351.96	-3080.72	-3090.35
α p-value	-	0.48	-	0.29
AIC	6277.15	6259.73	6221.67	6236.69

TABLE 3 Vuong Test Statistic

Severity Types	ZIP vs Poisson	Vuong Test Statistic	p-value	ZIP vs NB	Vuong Test Statistic	p-value
Total Crashes	ZIP > Poisson	12.5763	0.0000	ZIP > NB	2.5763	0.0052
Fatal and Injury Crashes	ZIP > Poisson	9.0161	0.0000	ZIP > NB	2.0161	0.0137
PDO Crashes	ZIP > Poisson	11.0527	0.0000	ZIP > NB	3.0527	0.0005

Model Estimation Results

The objective of this research is to quantify safety performance at unsignalized intersections in Louisiana, identifying relative effects of the roadway characteristics (e.g. curve radius, presence of turning lanes, traffic volume, speed, etc.) on crashes. The variable selection is based on extensive literature review and preliminary analysis of the dataset. Different coefficients specified for three different traffic severity levels (total, fatal and injury, and PDO crashes) are established for intersections on two-lane roadways, with stop control on the minor roadway, respectively. NCSS data analysis statistical software was used for ZIP model estimation. The Vuong test results favor the ZIP model over all other models. The ZIP model estimation results for total, fatal and injury, and PDO crashes are shown in **Table 2**.

TABLE 4 Estimation of Zero-Inflated Poisson Regression Model of Intersection Crashes

Variable	Total Crashes				Fatal and Injury Crashes				PDO Crashes			
	Coefficient	Std. Error	Z Value	p-value	Coefficient	Std. Error	Z Value	p-value	Coefficient	Std. Error	Z Value	p-value
Non-zero crash probability state as Poisson function												
(Intercept)	-6.02253	0.20197	-29.82	0.0000	-5.25584	0.38727	-13.57	0.0000	-6.32672	0.26526	-23.85	0.0000
Ln(ADT)												
Major road	0.50028	0.01551	32.25	**0.0000**	0.37361	0.02955	12.64	**0.0000**	0.50946	0.02067	24.65	**0.0000**
Minor road	0.26573	0.00994	26.73	**0.0000**	0.21310	0.01873	11.38	**0.0000**	0.26182	0.01316	19.90	**0.0000**
Radius												
Major road	-0.00017	0.00003	-6.38	**0.0000**	-0.00018	0.00006	-3.29	**0.0010**	-0.00015	0.00004	-4.21	**0.0000**
Minor road	-0.00027	0.00009	-2.92	**0.0035**	-0.00021	0.00018	-1.17	0.2421	-0.00035	0.00013	-2.76	**0.0058**
Curve length												
Major road	-0.00014	0.00004	-3.88	**0.0001**	-0.00003	0.00007	-0.47	0.6365	-0.00020	0.00005	-4.39	**0.0000**
Minor road	-0.00007	0.00012	-0.57	0.5690	-0.00041	0.00027	-1.47	0.1404	-0.00026	0.00014	-1.87	0.0611
Lane width												
Major road	-0.02198	0.01299	-1.69	0.0905	-0.00373	0.02492	-0.15	0.8809	-0.02934	0.01666	-1.76	0.0782
Minor road	-0.04296	0.00579	-7.42	**0.0000**	-0.02641	0.01106	-2.39	**0.0169**	-0.04858	0.00733	-6.63	**0.0000**
Speed limit												
Major road	-0.00184	0.00133	-1.38	0.1676	-0.00213	0.00264	-0.81	0.4203	-0.00671	0.00172	-3.91	**0.0001**
Minor road	0.01569	0.00109	14.39	**0.0000**	0.01596	0.00206	7.74	**0.0000**	0.01498	0.00143	10.46	**0.0000**
Intersection location												
Tangent	-	-	-	-	-	-	-	-	-	-	-	-
Begin of curve	0.23269	0.03896	5.97	**0.0000**	0.31771	0.07865	4.04	**0.0001**	0.14621	0.05076	2.88	**0.0040**
Middle of curve	0.16745	0.04285	3.91	**0.0001**	0.19248	0.08736	2.20	**0.0276**	0.12358	0.05614	2.20	**0.0277**
Number of leg												
3-leg	-	-	-	-	-	-	-	-	-	-	-	-
4-leg	0.29975	0.02303	13.01	**0.0000**	0.39967	0.04323	9.24	**0.0000**	0.22638	0.03014	7.51	**0.0000**
Left-turn lane presence												
No	-	-	-	-	-	-	-	-	-	-	-	-
Yes	-0.06334	0.05034	-1.26	0.2083	-0.09263	0.10012	-0.93	0.3549	-0.01479	0.06295	-0.23	0.8142
Right-turn lane presence												
No	-	-	-	-	-	-	-	-	-	-	-	-
Yes	0.05830	0.06212	0.94	0.3480	0.16279	0.11703	1.39	0.1642	0.00308	0.07854	0.04	0.9687
Setting												
Rural	-	-	-	-	-	-	-	-	-	-	-	-
Urban	0.45988	0.02721	16.90	**0.0000**	0.41938	0.05135	8.17	**0.0000**	0.45502	0.03679	12.37	**0.0000**

Zero crash probability state as logistic function												
(Intercept)	7.40854	0.88637	8.36	0.0000	9.14421	1.04036	8.79	0.0000	6.11792	0.93004	6.58	0.0000
Ln(ADT)												
Major road	-0.49595	0.06477	-7.66	**0.0000**	-0.61565	0.07720	-7.97	**0.0000**	-0.43049	0.06939	-6.20	**0.0000**
Minor road	-0.14102	0.05038	-2.80	**0.0051**	-0.17049	0.05649	-3.02	**0.0025**	-0.19458	0.05185	-3.75	**0.0002**
Radius												
Major road	0.00015	0.00013	1.18	0.2363	0.00015	0.00016	0.93	0.3533	0.00021	0.00014	1.53	0.1253
Minor road	0.00050	0.00044	1.15	0.2483	0.00059	0.00076	0.78	0.4347	0.00007	0.00047	0.16	0.8740
Curve length												
Major road	-0.00007	0.00018	-0.38	0.7018	-0.00010	0.00023	-0.44	0.6573	-0.00005	0.00018	-0.27	0.7898
Minor road	-0.00041	0.00076	-0.54	0.5921	-0.00161	0.00179	-0.90	0.3699	0.00025	0.00066	0.38	0.7022
Lane width												
Major road	-0.13968	0.06752	-2.07	**0.0386**	-0.14863	0.07809	-1.90	0.0570	-0.02190	0.06928	-0.32	0.7519
Minor road	-0.02715	0.03376	-0.80	0.4213	-0.02985	0.03645	-0.82	0.4129	-0.01046	0.03378	-0.31	0.7567
Speed limit												
Major road	-0.02101	0.00630	-3.33	**0.0009**	-0.02646	0.00753	-3.52	**0.0004**	-0.02151	0.00659	-3.26	**0.0011**
Minor road	-0.01563	0.00547	-2.86	**0.0043**	-0.00869	0.00632	-1.38	0.1690	-0.01416	0.00572	-2.47	**0.0134**
Intersection location												
Tangent	-	-	-	-	-	-	-	-	-	-	-	-
Begin of curve	-0.35587	0.19703	-1.81	0.0709	-0.27670	0.23615	-1.17	0.2413	-0.54936	0.21185	-2.59	**0.0095**
Middle of curve	-0.25024	0.21138	-1.18	0.2365	-0.28708	0.26199	-1.10	0.2732	-0.40572	0.22739	-1.78	0.0744
Number of leg												
3-leg	-	-	-	-	-	-	-	-	-	-	-	-
4-leg	-0.55714	0.14039	-3.97	**0.0001**	-0.35475	0.14331	-2.48	**0.0133**	-0.57199	0.14294	-4.00	**0.0001**
Left-turn lane presence												
No	-	-	-	-	-	-	-	-	-	-	-	-
Yes	-0.28496	0.50647	-0.56	0.5737	0.15767	0.46152	0.34	0.7326	-0.07998	0.42068	-0.19	0.8492
Right-turn lane presence												
No	-	-	-	-	-	-	-	-	-	-	-	-
Yes	-0.34280	0.59267	-0.58	0.5630	-0.48043	0.56386	-0.85	0.3942	-0.00027	0.47666	0.00	0.9995
Setting												
Rural	-	-	-	-	-	-	-	-	-	-	-	-
Urban	-0.73866	0.12854	-5.75	**0.0000**	-0.20928	0.15142	-1.38	0.1669	-0.84780	0.13355	-6.35	**0.0000**

The predictors with positive coefficients indicate an increase in the likelihood of crash occurrences. Logarithm of ADT (Ln(ADT)) of major and minor roads, curve radius of major and minor road, curve length of major roads, lane width of minor roads, speed limit of major and minor roads, intersection locations, number of legs, and settings were found to be statistically significant with *p* value less than 5 percent. Out of these variables, Ln(ADT) of major and minor roads, speed limit of minor roads, intersection locations, and settings are positively related to intersection crashes. On the contrary, curve radius of major and minor roads, curve length of major roads, lane width of minor roads, and speed limit of major roads have negative association.

Intersection crashes increase as traffic flow increases in all three different traffic severity levels, as expected. The results show that intersections with higher speed limits on minor roads tend to have higher probability of crashes occurring. However, higher speed limits on major roads are negatively related to PDO crashes. Major road with higher speed limits have safer designs and environmental conditions. Notably, higher speed limits should not be interpreted as a measure to reduce PDO intersection crashes. Intersections close to beginning of a curve had 26 percent greater total crashes compared to intersections on a tangent, while it is 37 percent and 16 percent for fatal and injury crashes and PDO crashes, resectively. In case of intersections on the middle of a curve, the probability of having total crashes is 18 percent higher than intersections on a tangent, while it is 21 percent and 13 percent for fatal and injury crashes and PDO crashes, resectively. The probability of having total crashes, fatal and injury crashes, and PDO crashes decreases as curve radius and curve length on major roads increase, as well as lane widths of minor roads. Four-leg stop-controlled intersections were found to have 35 percent greater total crashes, 49 percent greater fatal and injury crashes, and 25 percent greater PDO crashes, relative to three-leg intersections. As for settings, stop-controlled intersections in urban areas were found to have 58 percent greater total crashes, 52 percent greater fatal and injury crashes, and 58 percent greater PDO crashes, relative to rural intersections.

CONCLUSIONS

This study presents the estimation of safety performance for intersections on two-lane roadways, with stop control on the minor roadway in Louisiana. Data was collected for 4,148 intersections from 64 Parishes statewide. A comprehensive dataset of roadway characteristic data, including traffic crashes, traffic volumes, geometry, speed limits, and other site characteristics were developed for the study period of 2013 to 2017.

After the data was assembled for the stop controlled intersectios, four count approaches (Poisson, NB, ZIP and ZINB) models were developed to estimate annual crash occurrences on these intersections. To account for the excess zeros in crash data existence, separate models were developed for total crashes, fatal and injury crashes, and PDO crashes. Based on the test statistic, ZIP models provided a better fit than all other models. The models were specified considering factors such as curve radius of major and minor roads, curve length of major and minor roads, lane width of major and minor roads, speed limit of major and minor roads, intersection locations, left/right-turn lane presence, and settings, in addition to traffic volume.

The ZIP estimation results show that, except for traffic volume, greater curve radius of major and minor roads, greater curve length of major roads, greater lane width of minor roads, and higher speed limit of major roads, there were significantly smaller crash occurrences for stop controlled intersections. Intersections close to the beginning of a curve, or located in the middle of a curve, and higher speed limit of minor roads, led to significantly greater crash occurrences. Comparison of the three-leg to the four-leg stop controlled intersections showed that four-leg stop-controlled intersections have 35 percent greater total crashes, 49 percent greater fatal and injury crashes, and 25 percent greater PDO crashes, relative to three-leg intersections. As for settings, stop-controlled intersections in urban areas were more likely to have crashes in all three different traffic severity levels, relative to rural intersections.

The study shows that ZIP models provided a better fit than conventional Poisson or NB models for total, fatal and injury, and PDO crashes for two-lane highway stop controlled intersections. With the safety strategy plan emphasizing the Destination Zero Deaths objective, Louisiana must pay close attention to intersection safety. To reduce intersection crashes, particularly the fatal and injury crashes, it is critical to select countermeasures that target identified risk factors. Avoiding intersection on sharp curves, or design to compensate for reduced sight distance, enhancing sign and pavement markings, providing travel infrastructures, such as bypass lanes on shoulder at T-intersections, wide shoulder on two-lane highways, and roadway lightings, could work better for intersections safety.

ACKNOWLEDGMENTS

This research was sponsored by the Louisiana Transportation Research Center. The assistance and guidance provided by the project review committee is appreciated. However, the authors are fully responsible for the results and discussions presented here.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: study conception: Ming Sun and Xiaoduan Sun; methodology: Ming Sun; data collection: Ming Sun and Mousumy Akter; analysis and interpretation of results: Ming Sun and Ashifur Rahman; draft manuscript preparation: Ming Sun, Xiaoduan Sun and Subasish Das. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

1. *Fatality Analysis Report System (FARS)*. NHTSA, U.S. Department of Transportation. https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars. Accessed June 15, 2019.

2. Neuman, T. R., R. Pfefer, K. L. Slack, D. W. Hardwood, I. B. Potts, D. J. Torbic, and E. R. K. Rabbani. *NCHRP Report 500: Volume 5: A Guide for Addressing Unsignalized Intersection Collisions*. Transportation Research Board, Washington, D. C., 2003. http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_500v5.pdf

3. Savolainen, P. T. and A. P. Tarko. Safety of Intersections on High-Speed Road Segments with Superelevation. Report No. FHWA/IN/JTRP-2004/25, Purdue University, IN, 2004.

4. *Crash Data*. Louisiana Department of Transportation and Development. wwwsp.dotd.la.gov/Inside_LaDOTD/Divisions/Multimodal/Highway_Safety/Pages/Crash_Data.aspx. Accessed September 18, 2018.

5. Bauer, K. M. and D.W. Harwood. *Statistical Models of at-grade Intersections Accidents-addendum.* Publication No. FHWA-RD-99-094, FHWA, VA, 2000.

6. Poch, M. and F. Mannering. Negative Binomial Analysis of Intersection Accident Frequencies. *Journal of Transportation Engineering*, 1996. 122(2): 105-113.

7. Vogt, A. and J. Bared. *Accident Models for Two-Lane Rural Roads: Segments and Intersections.* Publication No. FHWA-RD-98-133, FHWA, VA, 1998.

8. Sayed, T. and F. Rodriguez. Accident Prediction Models for Urban Unsignalized Intersections in British Columbia. *Transportation Research Record: Journal of the Transportation Research Board*, 1999. 1665: 93-99.

9. Harnen, S., U. R. S. Radin, S. V. Wong, and W. W. I. Hashim. Motorcycle Crash Prediction Model for Non-Signalized Intersections. *IATSS Research*, 2003. 27(2): 58-65.

10. Salifu, M. Accident prediction models for unsignalized urban junctions in Ghana. *IATSS Research*, 2004. 28(1): 68-81.

11. Nambuusi, B. B., T. Brijs, and E. Hermans. *A review of accident prediction models for road intersections.* Report No. RA-MOW-2008-004, Steunpunt Mobiliteit & Openbare Werken – Spoor Verkeersveiligheid, 2008.

12. Kuciemba, S. R. and J. A. Cirillo. *Safety Effectiveness of Highway Design Features, Volume V: Intersections*. Report No. FHWA/RD-91/048, FHWA, VA, 1992.

13. Miaou, S. The Relationship between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions. *Accident Analysis & Prevention*, 1994. 26: 471–482.

14. Shankar, V., J. Milton, and F. Mannering. Modeling Accident Frequency as Zero-altered Probability Processes: An Empirical Inquiry. *Accident Analysis & Prevention*, 1997. 29: 829–837.

15. Lee, J. and F. Mannering. Impact of Roadside Features on the Frequency and Severity of Run-Off-Roadway Accidents: An Empirical Analysis. *Accident Analysis & Prevention*, 2002. 34: 149–161.

16. Shankar, V., G. Ulfarsson, R. Pendyala, and M. Nebergall. Modeling Crashes Involving Pedestrians and Motorized Traffic. *Safety Science*, 2003. 41: 627–640.

17. Pour, M. H., J. Prasetijo, A. S. Yahaya, and S. M. R. Ghadiri. Modeling Vehicle-Pedestrian Crashes with Excess Zero along Malaysia Federal Roads. *Procedia-Social and Behavioral Sciences*, 2012. 53: 1218-1227.

18. Lord, D., S. P. Washington, and J. N. Ivan. Poisson, Poisson-Gamma and Zero-Inflated Regression Models of Motor Vehicle Crashes: Balancing Statistical Fit and Theory. *Accident Analysis & Prevention*, 2005. 37: 35-46.

19. Das, S. and X. Sun. Zero-Inflated Models for Different Severity Types in Rural Two-Lane Crashes. Presented at 94th Annual Meeting of the Transportation Research Board, Washington, D.C., 2015.