Accident Analysis and Prevention 59 (2013) 309–318

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Random parameter models of interstate crash frequencies by severity, number of vehicles involved, collision and location type

Narayan Venkataraman[a,1], Gudmundur F. Ulfarsson[a,*], Venky N. Shankar[b,2]

[a] University of Iceland, Faculty of Civil and Environmental Engineering, Hjardarhagi 2-6, IS-107 Reykjavik, Iceland
[b] Pennsylvania State University, Department of Civil and Environmental Engineering, 226C Sackett Building, University Park, PA 16802, USA

ARTICLE INFO

Article history:
Received 4 September 2012
Received in revised form 4 June 2013
Accepted 14 June 2013

Keywords:
Crash frequency aggregations
Heterogeneity
Random parameters
Roadway geometrics
Crash severities
Interchange

ABSTRACT

A nine-year (1999–2007) continuous panel of crash histories on interstates in Washington State, USA, was used to estimate random parameter negative binomial (RPNB) models for various aggregations of crashes.

A total of 21 different models were assessed in terms of four ways to aggregate crashes, by: (a) severity, (b) number of vehicles involved, (c) crash type, and by (d) location characteristics. The models within these aggregations include specifications for all severities (property damage only, possible injury, evident injury, disabling injury, and fatality), number of vehicles involved (one-vehicle to five-or-more-vehicle), crash type (sideswipe, same direction, overturn, head-on, fixed object, rear-end, and other), and location types (urban interchange, rural interchange, urban non-interchange, rural non-interchange). A total of 1153 directional road segments comprising of the seven Washington State interstates were analyzed, yielding statistical models of crash frequency based on 10,377 observations. These results suggest that in general there was a significant improvement in log-likelihood when using RPNB compared to a fixed parameter negative binomial baseline model. Heterogeneity effects are most noticeable for lighting type, road curvature, and traffic volume (ADT). Median lighting or right-side lighting are linked to increased crash frequencies in many models for more than half of the road segments compared to both-sides lighting. Both-sides lighting thereby appears to generally lead to a safety improvement. Traffic volume has a random parameter but the effect is always toward increasing crash frequencies as expected. However that the effect is random shows that the effect of traffic volume on crash frequency is complex and varies by road segment. The number of lanes has a random parameter effect only in the interchange type models. The results show that road segment-specific insights into crash frequency occurrence can lead to improved design policy and project prioritization.

1. Introduction

Several research efforts have occurred in the last decade in the field of traffic crash analysis typology. Typological analysis (Wells-Parker et al., 1986; Gundy, 1990; Massie et al., 1993; Corbett, 2000; Summala, 2000; Kim et al., 2002; Corbett and Caramlau, 2006; Najim et al., 2007; Stradling et al., 2009; Council et al., 2010; Minikel, 2010) in the traffic safety area has typically been conducted to provide insight into human behavior such as speed-related crashes, alcohol-related crashes, or crashes with driving violations. Najim et al. (2007) demonstrate multiple pre-crash scenarios for crashes involving at least one light vehicle such as a passenger car or light truck. Corbett (2000) performed typological analysis of driver responses to speed cameras and Corbett and Caramlau (2006) added an exploration of gender differences. Wells-Parker et al. (1986) develop a driving-tendency based classification for DUI offenders. Kim et al. (2002) explored crash models of various classifications of motorcycle crashes. Massie et al. (1993) classify collision scenarios to explore effect of crash avoidance technologies. Summala (2000) examined relationships between behavior, age and gender, with crash information such as speed controls and fatality rates. Minikel (2010) evaluates street classifications and their relationship to bicycling safety. Gundy (1990) evaluated the usefulness of crash classification whereas Council et al. (2010) and Stradling et al. (2009) develop classifications for speeding behavior.

Count models, including the negative binomial and alternative models accommodating overdispersion, have been used extensively for crash modeling purposes. This is due to the well-known problem of overdispersion of traffic crash counts (Engel, 1984; Lawless, 1987; Miaou, 1994; Maher, 1991; Shankar et al., 1995, 1997; Lord et al., 2007; Malyshkina et al., 2009). The evolution of

* Corresponding author. Tel.: +354 525 4907; fax: +354 525 4632.
E-mail addresses: nsv1@hi.is (N. Venkataraman), gfu@hi.is (G.F. Ulfarsson), shankarv@engr.psu.edu (V.N. Shankar).
[1] Tel.: +354 525 4907.
[2] Tel.: +1 814 865 9434; fax: +1 814 863 7304.

0001-4575/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2013.06.021

count modeling since these studies has resulted in a recent emphasis on random parameter approaches, multivariate approaches, and Bayesian methods (see for example, Ma et al., 2008; El-Basyouny and Sayed, 2009; Anastasopoulos and Mannering, 2009; Malyshkina et al., 2009; Lord and Mannering, 2010). In spite of these methodological advances, the issue of heterogeneity remains as a substantial methodological issue in the field of traffic safety. Heterogeneity affects traffic crash analysis in two fundamental ways. The first is unobserved heterogeneity which occurs because of the problem that the variables cannot fully describe the differences between two or more roadway segments so they appear identical in the observed variables. The second is the omitted heterogeneity that occurs when a single variable is forced to have the same effect on all roadway segments through a fixed parameter.

In a recent study (Venkataraman et al., 2011), the issue of heterogeneity was addressed with respect to the impact of roadway geometrics on total interstate crash frequency. This was done by testing if variables describing the geometrics of road segments varied in their impact on crash frequency for separate road segments. The results indicate a road segment geometric feature, such as lighting or shoulder width, can in one location increase the likelihood of crashes, while at another location the same feature can contribute to safety benefits. This can be useful in aiding public agencies to target safety improvements at location contexts where the expected outcome is a decrease in crash likelihood. The findings emphasize the need for identifying the source of geometric-specific heterogeneity, rather than handling all the heterogeneity in terms of one single overdispersion parameter as is conventionally done in negative binomial models.

Frequency models of crash outcome type can provide substantial insights into the effects of roadway geometry. For example, one can explore the influence of a particular geometry and if it is statistically significant in multiple outcome cases, such as crashes aggregated by severity type, or by number of vehicles involved, or crash type. Furthermore it is possible to investigate which parameters are fixed across crash outcome types, and which are potentially random due to a persistent heterogeneity effect.

The objective of this paper is to shed further light on heterogeneity effects in roadway geometric features with respect to the type of traffic crash. This is because the effect of a roadway geometric feature can be different, for example, for fatal injury crashes than property damage only crashes. There are a number of other ways to classify traffic crashes than by severity. In this paper several different such aggregations are tested to explore how heterogeneity in geometric features varies based on how crashes are aggregated.

2. Methods

This paper uses the random parameter negative binomial model to account for road segment heterogeneity and crash count overdispersion. Although, negative binomial (NB) models have greater explanatory power in terms of relationship between site geometrics and crashes, their major limitation rests in that they cannot incorporate time variation or road segment-specific effects, thereby resulting in an underestimation of standard errors in the regression coefficients and subsequently inflated t-ratios. Some studies (see for example, Shankar et al., 1998) have attempted to fix this problem, by introducing a trend variable in the crash model whereas Hausman et al. (1984) demonstrate a general approach to count data applications via fixed and random effects models, an application of which has been seen in the safety literature as well (Shankar et al., 1998; Chin and Quddus, 2003).

In the area of classical models, incorporating randomness in parameters of traffic safety counts, Milton et al. (2008), show the first documented application through an evaluation of geometric and traffic variables and their association with the occurrence of severity proportions at the road segment level. Anastasopoulos and Mannering (2009) and El-Basyouny and Sayed (2009) are other notable examples of modeling at the road segment-corridor level. Washington et al. (2010), Ulfarsson and Shankar (2003), and Sittikariya and Shankar (2009) underscore the importance of heterogeneity due to time-varying effects.

To begin with, a generalized representation of the conditional density function for crash counts y_{it} in the ith road segment in year t is as follows:

$$P(y_{it}|x_{it}, \beta_i) = g(\cdot), \quad \forall i = 1, \ldots, I; \quad \forall t = 1, \ldots, T; \tag{1}$$

where $g(\cdot)$ is the density function of the appropriate count distribution. The data vary with both time and space, thereby working to capture changes across road segments and over time. In a negative binomial model this density is (Greene, 1997):

$$g(y_{it}|x_{it}, \beta_i, \theta) = \frac{\theta^{\theta}\lambda_{it}^{y_{it}}\Gamma(y_{it}+\theta)}{\Gamma(\theta)y_{it}!(\lambda_{it}+\theta)^{y_{it}+\theta}} \tag{2}$$

where the mean crash rate is $\lambda_{it} = \exp(\beta_i x_{it})$, β_i is a vector of estimable parameters, x_{it} is a vector of observed variables describing each road segment in each year, such as lighting, geometric, and traffic characteristics, θ is an overdispersion parameter. The random parameter negative binomial model is introduced by adding a heterogeneity term and a random term to the estimable parameters:

$$\beta_i = \beta + \Delta z_i + \Gamma v_i, \tag{3}$$

where the first term, β, is the mean of the random parameter, the second term introduces heterogeneity (z_i is a vector of observed variables inducing road segment-specific heterogeneity and Δ are estimable parameters on the heterogeneity variables), and the third term is a random deviation from the mean (Γ is an estimable diagonal covariance matrix capturing spatial and temporal parameter correlations, v_i are unobservable normally distributed random error terms with zero mean and variance one). The likelihood contribution of the ith road segment to the sample likelihood is conditioned on the unobserved random heterogeneity v_i and denoted by:

$$L_i(\beta, \Delta, \Gamma, \theta | y_{i1}, \ldots y_{iT}, x_{it}, z_i, v_i) = \prod\nolimits_{t=1}^{T} g(\cdot). \tag{4}$$

The likelihood for the ith road segment takes a non-closed form when integrating over v_i. It is therefore necessary to approximate the resulting integral through simulation by drawing R Halton draws for the random heterogeneity. Each draw is denoted with an index r, v_{ir}, and is inserted into the likelihood function and its value calculated. From the series of simulated likelihood values the expected value of the likelihood unconditioned on v_i is found using the relationship (Greene, 2007),

$$\begin{aligned} &E(L_i(\beta, \Delta, \Gamma, \theta | y_{i1}, \ldots, y_{iT}, x_{it}, z_{it}, z_{it})) \\ &\quad \approx \frac{1}{R}\sum\nolimits_{r=1}^{R} L_i(\beta, \Delta, \Gamma, \theta | y_{i1}, \ldots, y_{iT}, x_{it}, z_{it}, v_{itr}). \end{aligned} \tag{5}$$

The above-mentioned procedure is generally called simulated maximum likelihood and its accuracy relies in the number of Halton draws R, (see Venkataraman et al., 2011, for a recent prior traffic safety application).

The random parameters have a normal distribution, Φ_{β}, and to interpret the effect of each random parameter it is most straightforward to calculate the percentage of the density of the distribution of each random parameter that is positive, the so called positive parameter density:

$$P_{+\beta} = \int_0^{+\infty} \Phi_\beta(x)dx, \quad (6)$$

which shows the percentage of road segments where a variable with a random parameter takes a positive sign and thereby increases the crash frequency in that model.

3. Crash frequency aggregations

As discussed in the introduction, the way crash frequencies are aggregated can matter in terms of explaining the effect of observed variables on crash frequency. This is both due to unobserved heterogeneity in variables but also due to heterogeneity in terms of driver responses to roadway features. This paper proposes crash frequency aggregations that account for the major outcome categories of crash severity, number of vehicles involved, collision type, and geographical location and operational type. Within each of these categories, crashes are classified further as shown in Fig. 1.

Models are therefore developed for four different outcome types: (1) all injury severity classes based on the most severe injury in the crash (property damage only, possible injury, evident injury, disabling injury, and fatality); (2) for number of vehicles involved in the crash (one-vehicle to five-or-more-vehicle); (3) for various common collision types (sideswipe, same direction, overturn, head-on, fixed object, rear-end, and other); and (4) for basic location types (urban interchange, rural interchange, urban non-interchange, rural non-interchange). This results in 21 different models for crash frequency which are estimated and examined in this paper.

Fig. 1 shows there can be interaction between the four major outcome types. In addition, the four major outcome types can themselves be interpreted differently in terms of nested models. For example, geographical location and operational type can include interchange types. Non-interchange models can include finer resolutions such as non-high-occupant-lane segments versus high-occupant-inclusive segments.

Similarly, crash severity type can include finer resolutions in terms of driver injuries as opposed to most severe outcome of the crash. In terms of number of vehicles involved, a different interpretation might be to involve vehicle types within the vehicle count nesting. For example, one-vehicle categorization can include one-vehicle truck only, one-vehicle passenger car only, one-vehicle sports-utility vehicle only, etc. As the number of vehicles included increases, the combination of vehicle types in the involvement also increases.

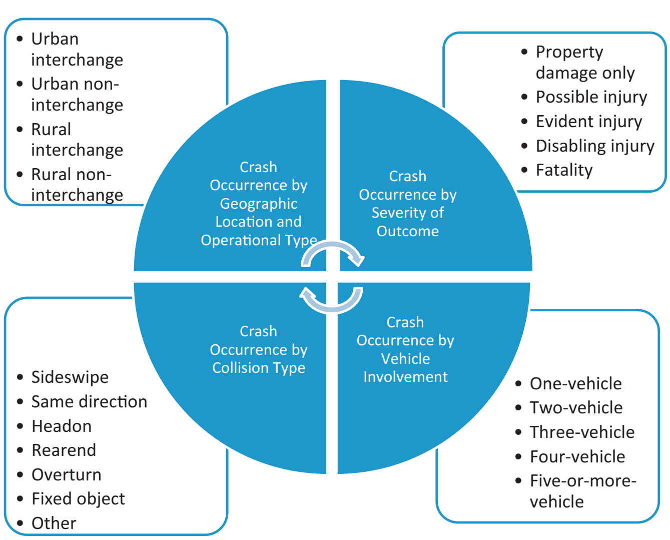

Fig. 1. Scheme of major aggregations of crash frequency.

The same logic can be applied to crash type as well, with vehicle types nested within crash types. It is clear from the above discussion that the combinatorial possibilities from just the base of major categories of crashes presented in this paper are significant. While these are theoretical outcomes, practical limitations may determine the level of resolution. For example, if enough observations are not obtained in terms of head-on crashes involving trucks, then, that category cannot be an independent model.

Multivariate model alternatives may be thought of, e.g. to simultaneously model the various injury severity type frequencies. However, the main advantage of multivariate alternatives, especially in the literature (see for example, Ye et al., 2013), is that there are statistical efficiency gains. Especially, when using severity types jointly, this advantage comes to the forefront in noticeable ways since improved efficiencies can help identify geometric effects accurately. In the present case, the vast number of models estimated involves varying specifications in a random parameter form. To incorporate these, while allowing for specifications unique to each element of the joint density, would make computation infeasible, and secondly, strong significance of the variables in the resulting models will indicate that efficiency is not lacking and that potential related effects have not been ignored.

4. Data description

The data in this research consisted of panel data with nine years (1999–2007 inclusive) of annual crash counts on road segments which comprise Washington State's interstate system in the United States, namely I-5, I-205, I-405, I-705, I-82, I-182, and I-90. The data also include roadway geometric variables describing each road segment during each of the years, and traffic volumes in each year for the road segments from the Washington State Department of Transportation (WSDOT).

A total of 1153 directional road segments were assembled based on interchange and non-interchange definitions. The use of directional road segments means that, crash, traffic, and geometric data varies by direction. Interchange segments were defined by the farthest merge/diverge ramp limits for each direction. The limits are determined on the basis of detailed WSDOT design drawings that show merge/diverge mileposts for each ramp at every interchange and these mileposts specify the beginning and end of an interchange segment. Since the location of merge/diverge ramp limits can vary by directions, a single interchange will have at least two different segments, one for each direction. Non-interchange segments were defined as continuous travel segments between two interchanges. Data was not missing for any year of the 1153 road segments during the nine years, which leads to the panel data being balanced, i.e. 9 observations for each road segment. The total number is therefore 10,377 observations of crash frequency by year and road segment.

Table 1 shows the descriptive statistics for key variables used in the estimation of the random parameter negative binomial models. WSDOT maintains a detailed database of roadway geometric information which was made available for this study. The data included: roadway cross section information such as widths of traveled way, shoulders and auxiliary lanes, horizontal and vertical alignment data, average daily traffic information, geographic descriptors such as urban/rural, speed limits, as well as route-specific information such as increasing and decreasing milepost indicators for geometric features. WSDOT maintains a video library showing footage (side and front views) of interstates and state routes. Footage is recorded every two years for main routes, such as the interstates in this study.

Table 1
Summary statistics of crash type, geometric, and traffic variables for Washington State interstate segments.

Variable	Mean	SD	Min	Max
Crash characteristics				
Number of crashes in a year	10.69	18.91	0	388
Number of property damage only crashes in a year	6.68	11.94	0	221
Number of possible injury crashes in a year	2.85	6.148	0	136
Number of disabling injury crashes in a year	0.13	0.397	0	5
Number of evident injury crashes in a year	0.98	1.611	0	31
Number of fatal injury crashes in a year	0.04	0.217	0	2
Number of one-vehicle crashes in a year	2.94	3.857	0	40
Number two-vehicle crashes in a year	5.89	12.352	0	267
Number three-vehicle crashes in a year	1.41	3.767	0	79
Number four-vehicle crashes in a year	0.34	1.117	0	23
Number five-vehicle crashes in a year	0.11	0.450	0	8
Number of sideswipe crashes in a year	1.56	3.35	0	74
Number of same direction crashes in a year	0.53	1.099	0	17
Number of head-on crashes in a year	0.009	0.097	0	2
Number of rear-end crashes in a year	5.266	13.22	0	277
Number of overturn crashes in a year	0.59	1.296	0	22
Number of fixed object crashes in a year	1.99	2.94	0	34
Number of other type crashes in a year	0.74	1.238	0	14
Geometric characteristics				
Portion of segment with median roadway lighting	0.12	0.287	0	1
Portion of segment with right-side roadway lighting	0.097	0.258	0	1
Portion of segment with tunnel lighting	0.001	0.027	0	0.64
Portion of segment with both-side roadway lighting	0.005	0.061	0	1
Portion of segment with point roadway lighting	0.011	0.039	0	0.75
Portion of segment with no roadway lighting	0.78	0.359	0	1
Portion of segment with one lane	0.001	0.019	0	0.54
Portion of segment with two lanes	0.52	0.494	0	1
Portion of segment with three lanes	0.34	0.462	0	1
Portion of segment with four lanes	0.13	0.331	0	1
Portion of segment with five or more lanes	0.002	0.043	0	1
Two-foot left shoulder width segment proportion	0.17	0.336	0	1
Three and four-foot left shoulder width proportion	0.199	0.484	0	1
Five-to-nine foot left shoulder width proportion	0.1321	0.5966	0	1
Ten-foot left shoulder width proportion	0.472	0.478	0	1
Greater-than-ten foot left shoulder width proportion	0.0187	0.262	0	1
Two-foot right shoulder width segment proportion	0.179	0.341	0	1
Three and four-foot right shoulder width proportion	0.2065	0.489	0	1
Five-to-nine foot right shoulder width proportion	0.119	0.56	0	1
Ten-foot right shoulder width proportion	0.463	0.477	0	1
Greater-than-ten foot right shoulder width proportion	0.02855	0.4346	0	1
Number of horizontal curves in segment	1.805	2.458	0	37
Minimum horizontal curve length in miles	0.182	0.193	0	1.2193
Maximum horizontal curve length in miles	0.275	0.276	0	2.4021
Minimum horizontal curve radius in feet	4076.73	5261.746	0	70,000
Maximum horizontal curve radius in feet	6510.35	7823.842	0	70,000
Minimum horizontal curve central angle	1217.59	1469.028	0	9842.6
Maximum horizontal curve central angle	2106.17	2100.06	0	11,129.4
Number of vertical curves	3.12	3.210	0	30
Minimum vertical curve length in miles	0.11	0.115	0	1.117424
Maximum vertical curve length in miles	0.21	0.185	0	1.268939
Minimum vertical curve gradient in percent	1.23	1.385	0	7.43
Maximum vertical curve gradient in percent	2.78	2.075	0	10
Minimum vertical curve length in feet	575.16	605.744	0	5900
Maximum vertical curve length in feet	1123.35	979.538	0	6700
Minimum vertical curve rate of curvature, k	432.43	731.095	0	11,133.33
Maximum vertical curve rate of curvature, k	1446.04	3421.252	0	60,000
Minimum vertical curve parabolic parameter, a	−9.52e−06	0.00001	−0.000103	0.0000496
Maximum vertical curve parabolic parameter, a	0.00001	0.000017	−0.000056	0.000115
Minimum vertical curve parabolic parameter, b	−0.01066	0.0179	−0.0517	0.0554
Maximum vertical curve parabolic parameter, b	0.01144	0.016784	−0.0452	0.0554
Minimum vertical curve parabolic parameter, c	77.599	314.848	−562.13	801.9278
Maximum vertical curve parabolic parameter, c	92.0368	314.356	−539.73	811.5778
Other segment characteristics				
Segment length in miles	1.307	1.753	0.1	20.380
Average daily traffic	13,058.64	8391.396	916.45	44,223.78

This video library was used to confirm information from the electronic database such as roadway cross-sections, interchange types, and ramp limits. The video library was also used to develop new variables such as lighting data, presence of roadside features such as medians, and the presence of overpasses.

Table 1 lists the different geometric variable and crash statistics at the road segment level. The geometric variable list is an exhaustive one including lighting type, lane cross sections, shoulder widths, horizontal curve parameters, and vertical curve parameters. We also use vertical curvature parabolic parameters to determine if they have an influence on crash frequency. The upshot of using variables as mentioned above in a random parameter model is that given the multitude of geometric variable types, one would expect these variables to capture road segment-specific heterogeneity with greater resolution than traditional dummy variable definitions. Due to the exhaustive list of geometric variables, their statistical characteristics are not discussed in detail due to space constraints.

There are 21 different crash categorizations varying by crash outcome in this paper and the distributional aspects can be substantially different. For example, it is shown that out of an average of 10.69 crashes per segment per year, 6.68 crashes account for property damage only outcomes, while the distribution tapers off from 2.85 possible injury crashes to 0.04 fatal crashes per year. The mean number of one-vehicle crashes is 2.94 per year, while that for two-vehicle crashes is 5.89, and the number of five-or-more vehicle crashes is 0.11 crashes per year on the average across the 9-year panel. In terms of collision types, rear-ends are the most common with an average of 5.266 crashes per year per segment, while head-on crashes are the rarest with an average of 0.009 crashes per year. Head-on crashes are observed to occur when wrong-way driving result in a crash, for example, due to a driver driving under the influence of alcohol, or being visually impaired, or entering a reversible lane when entry for lane use is violated.

5. Model results and findings

Since a total of 21 models were estimated, a summary of their overall log-likelihood at convergence and log-likelihood at starting values, which is a baseline negative binomial fixed parameter model. Table 2 shows the relative improvement in likelihood due to random parameters varies from a high value of 0.124 for the urban non-interchange model to a low of 0.001 for disabling injury in terms of rho-squared, which is a measure of the relative improvement in likelihood from starting values to convergence. When considering that the starting model is the fixed parameter negative binomial, a rho-squared of 0.124 represents a non-trivial improvement, considering the degrees of freedom given up in estimating the off-diagonal elements for the correlations of the random parameters. Furthermore, when considering the full model results the significance of the standard deviations, as well as the contributions of the off-diagonal elements from the Cholesky decomposition indicate randomness in the parameters which should not be ignored and further favors the RPNB over the NB.

5.1. Model performance by crash severity, number of vehicles involved, crash type, and geographical location and operational type

First, an examination of the improvements in the likelihoods of the random parameter models with respect to the fixed parameter counterparts indicates that the fixed parameter assumption can be rejected using a likelihood ratio test, accounting for parameter penalties. As a conservative threshold, a chi-squared threshold of 140.169 is exceeded at $p = 0.005$ for a substantially large number of degrees of freedom (100). Considering the degrees of freedom in the above table, and the smallest difference in log-likelihood multiplied by a factor of two, it is evident that the calculated chi-squared will exceed the chi-squared threshold of 140.169 for a p-value of 0.005 for all models. In fact, the fixed object model shows a chi-squared (based on LL) of 8.73, which exceeds the table value of 7.879 for 1 degree of freedom at $p = 0.005$.

Among the interchange-specific models, the urban non-interchange random parameters negative binomial (RPNB) model shows the best improvement over the fixed parameter negative binomial baseline with a rho-squared of 0.124 while the rural interchange model shows the least improvement, in terms of a rho-squared of 0.052. Among the severity models, the property damage only (PDO) model shows the best improvement over its fixed parameter baseline with a rho-squared of 0.089 while the disabling injury model shows the least improvement with a rho-squared of 0.001. Among the number of vehicles involved models, the two-vehicle model performs the best with a rho-squared of 0.093, while the five-or-more-vehicle model performs the worst with a rho-squared of 0.039. Finally, among the crash type models, the rear-end model performs the best with a rho-squared of 0.105 while the overturn model performs the worst with a rho-squared of 0.011. The specifications at the interchange level are models of total crash frequencies and therefore include heterogeneity across severities, crash types, and number of vehicles involved subsumed in the parameters. Breaking these specifications down into sub categories by interchange level, and possibly by interchange type is beyond the scope of this paper. Heterogeneous effects in interchange level parameters are worthy of further investigation based on the initial results discussed here.

5.2. Model interpretations and findings

The model results are discussed in terms of the nature of the various geometric variables in their association with the crash aggregation, the interpretation of random parameter effects, and fixed parameters. The models include coefficients which were found statistically significantly different from zero at the 0.05 level or better with a few exceptions which are noted in the tables. A comparison between these model results and the fixed parameter negative binomial models showed that the signs on coefficients are the same as for the RPNB models. However, the RPNB models allow a richer interpretation because they allow road segment-specific heterogeneity through the random parameters.

Tables 3–5 show the positive sign density, $P_{+\beta}$, of the distributions of random parameters in the crash severity, number of vehicles involved, collision type, and the interchange level models. The results for the fixed parameter estimates are presented in Tables 6–8. The discussion is focused on the positive sign density, $P_{+\beta}$, of the distributions of random parameters by crash aggregation and on the sign of the fixed parameter values across models. Further research is necessary to explore parameter elasticities and marginal effects.

The interpretation of the parametric effects begins with an analysis of model results for crash severity outcomes. In terms of the size of the random effects, the greatest number of random effects occur in the property damage only model, in some part due to the larger number of property damage crashes. Property damage crashes have been known to have unobserved effects shared with possible injury outcomes (see for example, Shankar et al., 1996), and for this reason, some of the geometric random effects are also observed in the possible injury model.

Median continuous lighting proportion by length of road segment and largest degree of curvature are found to influence possible injury crash frequency via random effects due to road segment-level heterogeneity. The degree to which the effect varies in terms of sign is indicated by a lower percentage of positive values in the possible injury model. Median continuous lighting proportion almost uniformly tends to increase property damage crash frequency; the increasing effect is estimated to occur in 62.1 percent of road segments. This implies there are some productive effects of median side lighting for more severe injuries, such as possible injuries. In a similar vein, it can be observed that right side lighting tends to influence evident injury frequencies through an increase in 67.5% of the road segments, whereas 32.5% of the road segments are expected to have a decrease in evident injury occurrences.

Table 2
Model estimation summaries.

Model	Number of fixed parameters	Number of random parameters	Log-*L* at convergence of RPNB model	Log-*L* for fixed parameter NB model	Rho-squared	Number of observations
Property damage only crashes	16	6	−22,891.34	−25,127.33	0.089	10,377
Possible injury crashes	16	3	−16,077.37	−17,478.63	0.080	10,377
Evident injury crashes	8	2	−12,159.17	−12,535.29	0.030	10,377
Disabling injury crashes	8	1	−3888.176	−3892.541	0.001	10,377
Fatality injury crashes	4	0	−1761.983	−1763.838	0.001	10,377
One-vehicle involved crashes	10	2	−19,023.34	−19,921.80	0.045	10,377
Two-vehicles involved crashes	13	4	−20,410.21	−22,491.88	0.093	10,377
Three-vehicles involved crashes	12	3	−10,380.03	−11,316.83	0.083	10,377
Four-vehicles involved crashes	11	3	−4949.760	−5314.407	0.069	10,377
Five-or-more vehicles involved crashes	5	2	−2785.441	−2898.346	0.039	10,377
Sideswipe crashes	11	1	−12,710.27	−13,364.79	0.049	10,377
Same direction crashes	8	1	−8471.919	−8717.768	0.028	10,377
Head-on crashes	3	0	−549.1418	−550.6772	0.003	10,377
Rear-end crashes	11	1	−17,780.28	−19,864.55	0.105	10,377
Overturn crashes	7	2	−8751.492	−8849.701	0.011	10,377
Fixed object crashes	17	2	−16,111.49	−17,017.41	0.053	10,377
Other type crashes	10	2	−10,485.17	−10,626.83	0.013	10,377
Urban interchange crashes	7	11	−6956.392	−7976.973	0.128	2754
Urban non-interchange crashes	6	10	−6878.783	−7820.289	0.120	2718
Rural interchange crashes	6	11	−4128.975	−4412.115	0.064	2421
Rural non-interchange crashes	9	10	−5434.075	−5872.757	0.075	2484

Table 3
Positive sign density of the random parameter distributions for crash severity and number of vehicles involved models.

Model	Property damage only	Possible injury	Evident injury	Disabling injury	One-vehicle	Two-vehicle	Three-vehicle	Four-vehicle	Five-plus vehicle
Logarithm of ADT	100%	100%	100%	100%	100%	100%	100%	100%	100%
Logarithm of length of segment	–	–	–	–	–	–	–	–	–
Median continuous segment proportion	99.9%	62.1%	–	–	–	97.8%	72.1%	70.9%	–
Right-side lighting segment proportion	–	–	67.5%	–	100%	71.1%	65.3%	–	–
Point lighting segment proportion	–	–	–	–	–	–	–	60.1%*	–
Ten-foot left shoulder width proportion	0%	–	–	–	–	–	–	–	–
Largest degree of curvature in segment	84.1%	61.5%	–	–	–	–	–	–	–
Number of vertical curves in segment	–	–	–	–	–	–	–	–	68.8%*
Smallest vertical curve gradient in segment	49.9%	–	–	–	–	51.4%	–	–	–
Largest vertical curve gradient in segment	65.5%	–	–	–	–	–	–	–	–

* The parameter is significant at the 0.1 level or better. All other parameters are significant at the 0.05 level or better.

Table 4
Positive sign density of the random parameter distributions for crash collision type models.

Model	Sideswipe	Same direction	Rear-end	Overturn	Fixed object	Other type
Logarithm of ADT	100%	100%	–	100%	100%	100%
Logarithm of length of segment	–	–	–	99.4%	–	–
Median continuous segment proportion	–	–	–	14.6%	–	30.6%
Right-side lighting segment proportion	–	–	–	–	–	–
Point lighting segment proportion	–	–	–	–	–	–
Ten-foot left shoulder width proportion	–	–	–	–	–	–
Largest degree of curvature in segment	–	–	–	–	63.5%	–
Number of vertical curves in segment	–	–	–	–	–	–
Smallest vertical curve gradient in segment	–	–	–	–	–	–
Largest vertical curve gradient in segment	–	–	–	–	–	–

Table 5
Positive sign density of the random parameter distributions for interchange and non-interchange level models.

Variable	Urban interchange	Urban non-interchange	Rural interchange	Rural non-interchange
Two-lane cross section segment proportion	–	–	100%	99.99%
Three-lane cross section segment proportion	80.9%	71.7%	100%	99.99%
Four-lane cross section segment proportion	99.5%	81.2%	–	–
Smallest degree of curvature in segment	59.8%	–	2.1%	64.62%
Minimum horizontal distance	–	21.7%	–	–
Smallest vertical curve gradient in segment	68.5%	58.2%	–	–
Largest vertical curve parabolic parameter, c	69.2%	69.2%	–	–
Largest vertical curve rate of curvature, k	–	–	50%	–
Largest vertical curve gradient in segment	–	–	–	93.3%

Table 6
Fixed parameters for crash severity and number of vehicles involved models.

Variable	Property damage only	Possible injury	Evident injury	Disabling injury	Fatality injury	One-vehicle	Two-vehicle	Three-vehicle	Four-vehicle	Five-plus vehicle
Constant	−9.29	−10.15	−4.89	−6.47	−7.03	−3.08	12.70	−19.83	−24.69	−24.68
Urban segment indicator (1 if segment is in an urban location; 0 otherwise)	0.02	−0.14	–	–	–	0.16	–	–	–	–
Logarithm of length of segment	–	0.66	0.80	0.93	1.07	0.83	0.66	0.56	0.40	0.67
Logarithm of ADT	–	–	–	–	0.38	–	–	–	–	–
No lighting segment proportion	0.68	–	–	−0.37	–	–	–	–	–	–
Median continuous lighting segment proportion	–	–	0.20	–	–	–	–	–	–	–
Right-side lighting segment proportion	0.77	0.14	–	–	–	–	–	–	–	–
Two–lane cross section segment proportion	–	–	−0.80	–	–	–	–	–	–	–
Three-lane cross section segment proportion	0.43	0.58	−0.50	–	0.27	0.41	0.39	0.52	0.56	–
Four-lane cross section segment proportion	–	1.26	–	0.57	0.92	0.76	1.01	1.25	1.25	–
Five-lane cross section segment proportion	0.65	1.09	–	1.37	–	0.82	0.78	–	–	1.04
Two-foot left shoulder width proportion	–	–	0.41	–	–	0.38	0.50	0.52	0.50	0.60
Three and four-foot left shoulder width proportion	−0.20	−0.25	0.11	–	–	–	0.10	–	–	–
Five-to-nine foot left shoulder width proportion	−0.38	−0.57	–	−0.16	–	–	–	–	–	–
Ten-foot left shoulder width proportion	–	−0.64	–	–	–	–	–	–	–	–
Greater-than-ten foot left shoulder width proportion	–	–	–	–	–	–	0.21	0.54	0.89	–
Two-foot right shoulder width proportion	–	–	0.32	–	–	0.17	0.38	0.47	0.57	0.77
Three and four-foot right shoulder width proportion	–	–	–	–	–	–	–	–	−0.26	–
Five-to-nine foot right shoulder width proportion	−0.34	−0.42	–	–	–	−0.15	0.11	0.23	–	–
Ten-foot right shoulder width proportion	−0.37	−0.51	–	–	–	–	–	–	–	–
Greater-than-ten foot right shoulder width proportion	–	–	–	–	–	–	0.33	0.30	–	–
Minimum horizontal distance	−0.25	−0.27	–	–	–	–	−0.40	−0.34	−0.56	–
Number of horizontal curves	0.03	0.05	0.04	–	–	0.04	0.41	0.03	0.09	–
Maximum horizontal curve radius	–	–	–	0.00001	–	–	–	–	–	–
Smallest degree of curvature	–	–	–	0.10	–	0.12	–	–	–	–
Smallest horizontal curve length (ft)	–	–	–	–	–	–	–	–	–	–
Largest degree of curvature	–	–	–	–	–	–	0.05	0.06	–	0.03[*]
Largest horizontal curve central angle	–	–	–	–	–	–	–	–	0.00003	–
Number of vertical curves	0.02	–	–	–	–	–	–	–	–	–
Smallest vertical curve gradient	–	−0.003	–	–	–	–	–	–	–	–
Largest vertical curve gradient	–	0.02	–	–	–	–	–	–	–	–
Largest vertical curve length in miles	–	–	–	–	–	–	–	–	–	–
Smallest vertical curve rate of curvature, *k*	–	–	–	–	–	–	–	–	–	–
Smallest vertical curve parabolic parameter, *a*	–	–	–	–	–	–	–	–	–	–
Largest vertical curve parabolic parameter, *a*	–	–	–	–	–	–	–	5315.25	8537.08	–
Largest vertical curve parabolic parameter, *b*	–	–	–	–	–	–	–	–	–	–
Smallest vertical curve parabolic parameter, *c*	–	–	–	–	–	–	–	–	–	–

Table 7
Fixed parameters for crash collision type models.

Variable	Sideswipe	Same direction	Head-on	Rear-end	Overturn	Fixed object	Other type
Constant	−13.71	−9.23	−9.73	−16.01	−0.86	−6.38	−4.70
Urban segment indicator (1 if segment is in an urban location; 0 otherwise)	–	–	–	–	0.22	0.06	0.10
Logarithm of length of segment	0.45	0.65	–	0.51	–	0.69	0.88
Logarithm of ADT	–	–	0.49	–	–	–	–
No lighting segment proportion	–	–	–	–	–	–	–
Median continuous segment proportion	0.31	–	–	0.41	–	–	–
Right-side lighting segment proportion	–	–	–	–	–	0.13	–
Two-lane cross section segment proportion	–	–	–	–	–	–	–
Three-lane cross section segment proportion	0.40	0.60	–	0.37	–	0.32	0.38
Four-lane cross section segment proportion	1.20	1.22	–	1.01	0.51	0.69	0.90
Five-lane cross section segment proportion	–	–	–	–	–	0.51	1.23
Two-foot left shoulder width proportion	0.26	0.33	–	0.64	−0.45	0.49	–
Three and four-foot left shoulder width proportion	–	–	–	–	–	–	–
Five-to-nine foot left shoulder width proportion	–	–	–	–	–	–	–
Ten-foot left shoulder width proportion	–	–	–	–	–	–	–
Greater-than-ten foot left shoulder width proportion	–	–	–	0.67	–	–	–
Two-foot right shoulder width proportion	0.28	0.35	–	0.55	–	0.35	–
Three and four-foot right shoulder width proportion	–	–	–	–	–	–	–
Five-to-nine foot right shoulder width proportion	–	–	–	–	–	–	–
Ten-foot right shoulder width proportion	–	–	–	–	–	–	–
Greater-than-ten foot right shoulder width proportion	–	–	–	–	–	–	–
Minimum horizontal distance	−0.41	−0.46	–	−0.60	–	–	–
Number of horizontal curves	0.05	0.04	–	0.04	0.01	0.04	−0.01
Maximum horizontal curve radius	–	–	–	–	–	–	–
Smallest degree of curvature	–	–	–	–	–	–	–
Smallest horizontal curve length (ft)	–	–	–	–	–	−0.00009	–
Largest degree of curvature	–	–	–	–	–	–	–
Largest horizontal curve central angle	0.00006	0.00006	–	–	0.00002	0.00006	0.00003
Number of vertical curves	–	–	0.06*	–	−0.04	–	0.03
Smallest vertical curve gradient	–	–	–	–	–	–	–
Largest vertical curve gradient	0.02	–	0.09	0.02	−0.03	–	–
Largest vertical curve length in miles	–	–	–	–	–	–	−0.4
Smallest vertical curve rate of curvature, k	–	–	–	–	–	−0.00007	–
Smallest vertical curve parabolic parameter, a	–	–	–	–	–	−1201.39	–
Largest vertical curve parabolic parameter, a	3843.39	–	–	8223.87	–	2953.78	–
Largest vertical curve parabolic parameter, b	–	–	–	–	–	−1.25	–
Smallest vertical curve parabolic parameter, c	–	–	–	–	–	0.0004	–

* The parameter is significant at the 0.1 level or greater. All other parameters are significant at the 0.05 level or greater.

Other effects of lighting type have fixed parameters. For example, median lighting proportions uniformly increase evident injury occurrences while no lighting and right side lighting proportions tend to increase property damage occurrences uniformly. No lighting proportion is estimated to be associated with a decrease in disabling injury occurrences, perhaps a reflection of vehicle speed as well as that the distribution is shifted toward an increase in property damage occurrences.

Shoulder width effects are generally fixed parameters, and the results show that ten foot left shoulder widths are expected to almost uniformly decrease the occurrence of property damage outcomes. A similar effect is also estimated for possible injury occurrences. In comparison to ten foot shoulder widths, road segments with greater than ten foot left shoulder widths were not found to be statistically significant in any of the severity models. However, three-to-four foot and five-to-nine foot left shoulder

Table 8
Fixed parameters for urban, rural, interchange and non-interchange level models.

Variable	Urban interchange	Urban non-interchange	Rural-interchange	Rural non-interchange
Constant	−7.20	−7.22	−9.19	−8.49
Logarithm of length of segment	0.75	0.49	0.59	0.74
Logarithm of ADT	0.98	0.97	0.82	0.91
Four-lane cross section segment proportion	–	–	5.35	2.02
Five-lane cross section segment proportion	–	0.44	–	–
Three and four-foot left shoulder width proportion	−0.20	–	−0.18	−0.56
Five-to-nine foot left shoulder width proportion	–	–	–	−0.28
Ten-foot left shoulder width proportion	−0.35	−0.36	−1.43	–
Greater-than-ten foot left shoulder width proportion	1.46	–	−0.90	0.47
Three and four-foot right shoulder width proportion	−0.42	−0.38	0.58	−0.32
Five-to-nine foot right shoulder width proportion	−0.37	−0.30	0.46	–
Ten-foot right shoulder width proportion	−0.60	−0.48	0.71	0.26
Greater-than-ten foot right shoulder width proportion	−0.78	–	0.51	–
Number of horizontal curves	0.11	−0.19	−0.026	0.03
Minimum horizontal distance	−0.37	–	–	−0.32
Largest degree of curvature in segment	–	–	0.23	–
Largest vertical curve gradient	–	–	0.06	–
Largest vertical curve parabolic parameter, c	–	–	–	0.0002

widths were estimated to uniformly decrease the occurrence of property damage and possible injury outcomes, with the five-to-nine foot width effect being larger. The effect of five-to-nine foot shoulder widths in a road segment is less negatively associated in comparison to the ten foot width effect, which underscores the importance of considering smaller left shoulder widths with respect to right-of-way-safety tradeoff. The evident injury occurrence tapers off from two foot left shoulder widths to three-to-four foot left shoulder widths with larger widths not being significant. Right shoulder widths show a similar pattern of effect on higher severities in terms of two foot widths. Evident injury occurrences are more likely to occur in association with two foot right shoulder widths. Five-to-nine foot right shoulder widths tend to be negatively associated in property damage and possible injury occurrences in comparison to two foot right shoulder widths. The effect's size improves marginally with ten foot right shoulder widths. This once again emphasizes the tradeoff gains in the consideration of shoulder width policies from a severity perspective.

In terms of capacity effects on crash severity frequencies, which mostly are found with fixed parameters, two lane road segments tend to be associated with a decrease in evident injury outcomes, in some part due to locational effects such as transition areas where two-lane cross-sections precede expansions to three- or four-lane mainlines. Comparatively, three-lane cross-sections show an ambiguous effect on severity types, with estimated increases in property damage and possible injury crashes but a comparably strong decrease in evident injury outcomes and an increase in fatal outcomes. The disbenefits are more noticeable as lane cross-sections increase to four lanes and five lanes or more, four- and five-lane cross-sections are positively associated with possible injury occurrences, and showing a respective increase in disabling injury outcomes while fatal outcomes tend to increase in four-lane cross-sections. Five-lane cross-sections are also positively associated with property damage occurrences. This is expected since larger cross-sections are associated with greater mixing effects such as lane changing, speed differentials and urban driving effects, even after accounting for average daily traffic (ADT).

In terms of horizontal curve length, minimum horizontal curve length is associated with estimated decreases in property damage and possible injury occurrences, whereas the number of horizontal curves in a segment is associated with an increase in the expected number of property damage, possible injury, or evident injury occurrences. Maximum horizontal curve radius is associated with an increase in disabling injury outcomes, which appears to indicate a behavioral tendency to drive faster. Perhaps fatal occurrences would have been associated with the same effect as well but such an effect is not observed due to the observation and scale of this dataset (at the interchange level). The finding that maximum horizontal curve radius is positively associated with disabling injury occurrences should be viewed with some caution as it may also apply potentially to the counterproductive effect of fatal occurrences although not statistically significant in this study. The effect is in part capturing the asymptotic influence of near tangent or tangential road segment configurations. The smallest degree of curvature is positively associated with disabling injury occurrences, with increase in degree of curve. This implies that as the smallest radius of curve in the road segment decreases by the order of one-half, one-third, one-fourth and so on compared to a 5730 foot radius, the corresponding increase in the estimated number of disabling injury outcomes is marginal for each step reduction.

The number of vertical curves in a road segment is positively associated with an increase in property damage occurrences, while the smallest vertical gradient is associated with a decrease in possible injury occurrences. Comparatively, the largest vertical gradient is associated with an increase in possible injury occurrences.

Other fixed-effects relating to vertical curvature parabolic variables appear influential across the crash type spectrum more so than the injury spectrum. The heterogeneity capturing effects of these variables seem to be most significant in the interchange level models, where they appear as random parameters as well.

6. Conclusions and recommendations

This paper presents an analysis of random parameter negative binomial models for various traffic crash aggregations. The aggregations are based on crash severity, number of vehicles involved, collision type, and geography and operational type. A total of 21 models were estimated. The improvement in likelihood in 19 of the 21 models was due to some parameters being random, and as a result contributing to a better likelihood compared to the baseline fixed effect negative binomial models. It was observed that the heterogeneity in parameters is strongest in the interchange type models for property damage only, with the number of random parameters ranging from four to six. Two models, namely the head-on collision type and the fatality severity model, had no random parameters, and furthermore followed the Poisson specification, meaning that overdispersion did not exist either. Among the number of vehicles involved models, the two-vehicle and three-vehicle crash models have four and three random parameters each and correspondingly greater improvement in likelihood.

Some other notable results come to light. First, it appears that heterogeneity effects are most noticeable via the significance of the standard deviations of lighting type variables, curvature variables, and ADT. None of the cross-sectional variables appear to have any heterogeneity effects in the crash severity, collision type, and number of vehicles involved models. However, cross-sectional variables such as number of lanes do appear as random parameters in the interchange type models. This may be attributed to the impact of auxiliary lanes at interchanges and near-interchange portions of road segments. That curvature effects are also random in the interchange level models bears importance.

This is one plausible set of crash frequency aggregations and it is not comprehensive. For example, it does not include roadside crashes as an element. Inclusion of a roadside element, along with the broad findings shown in this paper, indicates that the pursuit of a geometric basis for crash frequency is more complex than previously thought. The richness of the geometry appearing to be significant for multiple crash type aggregations indicates that one type of aggregation at which comprehensive safety analysis can begin is at the level of interchange densities. This is the scale at which this analysis is done, to be explicit about the separation of interchange and non-interchange road segments. A broad and rich set of parameters at this scale of analysis indicates that the pursuit of a "common denominator" of defining geometric criteria for the evaluation of substantive safety as opposed to nominal safety is feasible, given this study as one example.

While substantial insight has resulted in terms of exploring the estimated parameters for road geometrics and traffic volumes on the various crash outcome aggregations, complete insight needs to be developed via the determination of elasticities. The elasticities of independent variables in a conventional negative binomial model are straightforward, being computed as the product of the estimated parameter and the variable mean. However, in a random parameter model, the mean itself has a distribution, which implies a plausible distribution of road segment-specific marginal effects needs to be determined. Another fruitful area for future research is the examination of finer resolution within the aggregations as mentioned earlier in this paper. This will of course be decided by available sample sizes. However, the finer aggregations can also allow for more complex model development in terms of treatment of the random parameters.

Acknowledgments

The authors gratefully thank The University of Iceland Research Fund for their support of this research. The funding agent had no involvement in the study.

References

Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis and Prevention 41 (1), 153–159.
Chin, H., Quddus, M., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accident Analysis and Prevention 35 (2), 253–259.
Corbett, C., 2000. A typology of drivers' responses to speed cameras: implications for speed limit enforcement and road safety. Psychology, Crime and Law 6 (4), 305–326.
Corbett, C., Caramlau, I., 2006. Gender differences in responses to speed cameras: typology findings and implications for road safety. Criminology and Criminal Justice 6 (4), 411–433.
Council, F.M., Reurings, M., Srinivasan, R., Masten, S., Carter, D., 2010. Development of a Speeding-Related Crash Typology 1. Federal Highway Administration, FHWA-HRT-10-024, pp. 112.
El-Basyouny, K., Sayed, T., 2009. Accident prediction models with random corridor parameters. Accident Analysis and Prevention 41 (5), 1118–1123.
Engel, J., 1984. Models for response data showing extra-Poisson variation. Statistica Neerlandica 38 (3), 159–167.
Greene, W.H., 1997. Econometric Analysis, 3rd ed. Prentice Hall, New Jersey, pp. 1075.
Greene, W.H., 2007. Limdep Version 9.0. Econometric Modeling Guide, vol. 2. Econometric Software, Inc.
Gundy, C.M., 1990. Accident Typology. SWOV Institute for Road Safety Research, The Netherlands, Leidschendam, pp. 161.
Hausman, J.A., Hall, B.H., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52 (4), 909–938.
Kim, K., Boski, J., Yamashita, E., 2002. Typology of motorcycle crashes: rider characteristics, environmental factors, and spatial patterns. Transportation Research Record 1818, 47–53.
Lawless, J., 1987. Negative binomial and mixed Poisson regression. Canadian Journal of Statistics 15 (3), 209–225.
Lord, D., Washington, S.P., Ivan, J.N., 2007. Further notes on the application of zero inflated models in highway safety. Accident Analysis and Prevention 39 (1), 53–57.
Lord, D., Mannering, F.L., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternative. Transportation Research Part A – Policy and Practice 44 (5), 291–305.
Maher, M., 1991. A new bivariate negative binomial model for accident frequencies. Traffic Engineering and Control 32 (9), 422–443.
Ma, J., Kockelman, K.M., Damien, P., 2008. Multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accident Analysis and Prevention 40 (3), 964–975.
Malyshkina, N., Mannering, F., Tarko, A., 2009. Markov switching negative binomial models: an application to vehicle accident frequencies. Accident Analysis and Prevention 41 (2), 217–226.
Massie, D.L., Campbell, K.L., Blower, D.F., 1993. Development of a collision typology for evaluation of collision avoidance strategies. Accident Analysis and Prevention 25 (3), 241–257.
Miaou, S., 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention 26 (4), 471–482.
Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.
Minikel, E., 2010. Street typology and bicyclist safety: a systems approach. Massachusetts Institute of Technology (M.Sc. thesis).
Najim, W.G., Smith, J.D., Yanagisawa, M., 2007. Pre-Crash Scenario Typology for Crash Avoidance Research (Report No. DOT HS 810767). National Highway Traffic Safety Administration, Washington, D.C.
Shankar, V., Mannering, F., Barfield, W., 1995. Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. Accident Analysis and Prevention 27 (3), 542–555.
Shankar, V.N., Mannering, F.L., Barfield, W., 1996. Statistical analysis of rural freeway accident severity. Accident Analysis and Prevention 28 (3), 391–401.
Shankar, V., Milton, J., Mannering, F.L., 1997. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. Accident Analysis and Prevention 29 (6), 829–837.
Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1998. Evaluating median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. Transportation Research Record, Journal of the Transportation Research Board (1635), 44–48.
Sittikariya, S., Shankar, V., 2009. Modeling Heterogeneity: Traffic Accidents. VDM-Verlag, Saarbrücken, Germany.
Stradling, S., Gormley, M., Fuller, R., Broughton, P., Kinnear, N., O'Dolan, C., Hannigan, B.,2009. A typology of speeding drivers: extent of and motives for, exceeding the speed limit. In: Behavioural Research in Road Safety, Seventeenth Seminar Proceedings. Department for Transport, London, pp. 236–252.
Summala, H., 2000. Behavior and Risk Typology: Disaggregation of Accident Statistics and Behavior, Transportation (Netherlands). Martinus Nijhoff Publishers, pp. 117–132.
Ulfarsson, G.F., Shankar, V.N., 2003. Accident count model based on multiyear cross-sectional roadway data with serial correlation. Transportation Research Record: Journal of the Transportation Research Board (1840), 193–197, Transportation Research Board of the National Academies, Washington, DC.
Venkataraman, N.S., Ulfarsson, G.F., Shankar, V., Oh, J., Park, M., 2011. Model of relationship between interstate crash occurrence and geometrics: exploratory insights from a random parameters negative binomial approach. Transportation Research Record: Journal of the Transportation Research Board (2236), 41–48.
Washington, S., Karlaftis, M.G., Mannering, F.L., 2010. Statistical and Econometric Methods for Transportation Data Analysis, 2nd ed. Chapman and Hall/CRC, Boca Raton, FL.
Wells-Parker, E., Cosby, P.J., Landrum, J.W., 1986. A typology for drinking driving offenders: methods for classification and policy implications. Accident Analysis and Prevention 18 (6), 443–453.
Ye, X., Pendyala, R.M., Shankar, V.N., Konduri, K., 2013. A simultaneous equations model of accident frequency by severity level for freeway sections. Accident Analysis and Prevention 57, 140–149.