

EVALUATING THE EFFECTS OF ROAD GEOMETRY, ENVIRONMENT, AND TRAFFIC VOLUME ON ROLLOVER CRASHES

Mehdi Hosseinpour¹, Ahmad Shukri Yahaya², Ahmad Farhan Sadullah³,
Noriszura Ismail⁴, Seyed Mohammad Reza Ghadiri⁵

^{1,2,3}*School of Civil Engineering, University of Science, Malaysia*

⁴*School of Mathematical Science, National University of Malaysia, Malaysia*

⁵*Dept of Transportation Engineering, Malaysia University of Science and Technology, Malaysia*

Submitted 30 April 2014; resubmitted 11 January 2015; accepted 10 July 2015

Abstract. There are a number of factors that cause motor vehicles to rollover. However, the impacts of roadway characteristics on rollover crashes have rarely been addressed in the literature. This study aims to apply a set of crash prediction models in order to estimate the number of rollovers as a function of road geometry, the environment, and traffic conditions. To this end, seven count-data models, including Poisson (PM), negative binomial (NB), heterogeneous negative binomial (HTNB), zero-inflated Poisson (ZIP), zero-inflated negative binomial (ZINB), hurdle Poisson (HP), and hurdle negative binomial (HNB) models, were developed and compared using crash data collected on 448 segments of Malaysian federal roads. The results showed that the HTNB was the best-fit model among the others to model the frequency of rollovers. The variables Light-Vehicle Traffic (LVT), horizontal curvature, access points, speed limit, and centreline median were positively associated with the crash frequency, while UnPaved Shoulder Width (UPSW) and Heavy-Vehicle Traffic (HVT) were found to have the opposite effect. The findings of this study suggest that rollovers could potentially be reduced by developing road safety countermeasures, such as access management of driveways, straightening sharp horizontal curves, widening shoulder width, better design of centreline medians, and posting lower speed limits and warning signs in areas with higher rollover tendency.

Keywords: rollover; crash prediction models; over-dispersion; zero-altered models.

Introduction

Globally, over 1.2 million people are killed in traffic crashes every year, and as many as 50 millions are injured. The global economic losses from road crashes are estimated to be more than US\$ 500 billion annually (WHO 2009). In Malaysia, 414421 road crashes were reported in 2010, resulting in 6872 deaths and more than 9 billion ringgit of loss to the country's economy (RMP 2011); of which, rollovers accounted for nearly 1.4% of the total fatal crashes (ITF 2012).

Rollovers occur when a vehicle rotates at least one-quarter turn about its lateral or longitudinal axis (Conroy *et al.* 2006). According to the National Automotive Sampling System – Crashworthiness Data System (NASS–CDS), there are eight types of rollover crashes based on the cause and configuration of the collision

(Fig.). For more details on each rollover type, the reader is referred to Thomson *et al.* (2006). As noted in prior studies, despite the relative rarity of rollovers in comparison to other collision types, they account for a considerable number of serious injuries and fatalities (Khattak *et al.* 2003; Pape *et al.* 2008; Keall, Newstead 2009; Funk *et al.* 2012). For example, Conroy *et al.* (2006) reported

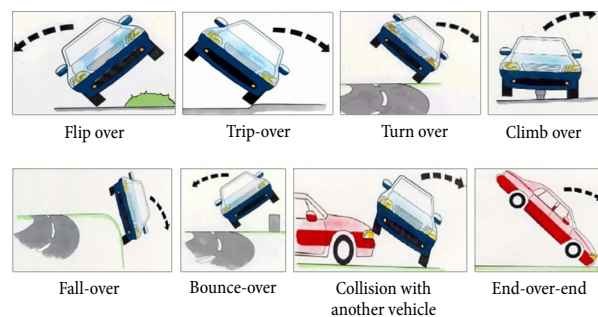


Fig. NASS-CDS classification of rollover crashes (Thomson *et al.* 2006)

This article has been corrected since first published. Please see the statement of correct (DOI:10.3846/16484142.2016.1235833 of the erratum).

that rollovers account for only 3% of all crashes, yet they are responsible for 33% of fatal crashes. Pape *et al.* (2008) indicated that rollovers comprise two-thirds of serious single-vehicle crashes of cargo tank motor vehicles. Keall and Newstead (2009) reported that rollover crashes accounted for nearly one third of all passenger vehicle occupant deaths for the year 2000 in the US.

These statistics justify the need for more attention to rollover incidents as one of the major causes of injuries and fatalities. To do this, it is necessary to first identify factors that contribute to the risk of rollovers. This can assist in allocating limited funds for safety improvement programs in a more effective and efficient way. Statistical modelling techniques are routinely used tools for evaluating the effects of various roadway factors on the accident occurrence (Abdel-Aty, Haleem 2011).

There are a large number of studies in the literature that have focused on road safety modelling (Poch, Manering 1996; Chin, Quddus 2003; Dinu, Veeraragavan 2011). However, research on rollover crashes is still limited in the literature compared to those devoted to total or other crash types. Furthermore, most of these studies have focused on the impacts of vehicle and human factors associated with rollover crashes, while the effects of roadway geometry and the surrounding environment on the frequency of rollover crashes have rarely been addressed in the literature. One reason for this might be due to the rarity of rollover crashes and also the lack or unavailability of detailed information on roadway attributes; these can be a barrier to the development of models for establishing the relationship between road-specific features and rollover crashes. In addition, from a methodological viewpoint, rollover crashes were treated as a dichotomous variable and fitted by binary models in most prior studies, whilst the frequency of rollovers has not yet been considered. Finally, to the authors' best knowledge, no study has been published to address rollover safety on Malaysia's road network, especially the federal road system, where it is responsible for over 40% of all fatal crashes throughout the country (ITF 2012).

In response, this study aims to fill that gap in the existing literature by estimating the number of rollovers as a function of roadway features. The objective of this study is to provide insights into the effects of road geometry, the environment, and traffic flows on rollover crashes to identify which set of explanatory variables related to roadway traits are significantly associated with the occurrence of rollovers. To accomplish the objective of this study, a set of count-data models were developed and compared using data collected on the Malaysian federal roads. This paper is organised as follows. Section 1 gives a review of previous studies that have focused on rollover crashes. Section 2 describes the methodology applied for developing and comparing the count models as well as the characteristics of the data used in this study. Section 3 provides the results of comparison analysis, the interpretations of parameter estimates, and the implications of preventive strategies. Last section summarises the findings and conclusions of the study and provides recommendations for further research.

1. Literature Review

There are few studies in the literature that have investigated the effects of factors on rollover crashes. Viner (1995) examined rollovers in run-off-the-road crashes using Illinois roadway and crash data. He found that slopes and hitting fixed objects were the critical tripping mechanisms involved in rollovers. In addition, higher speed limits were related to rollover frequencies. Farmer and Lund (2002) evaluated the effects of driver, vehicular, and environmental factors on the likelihood of vehicle rollover. The authors also compared rollover risk of cars and light trucks. The findings showed that young drivers were more likely to be involved in rollover crashes. In addition, rural curves were more exposure to the event. Within vehicle types, the likelihood of being involved in rollovers was higher for smaller vehicles than larger vehicles. Using 3 years (1996–1998) of crash data occurred in North Carolina, Khattak *et al.* (2003) investigated the impacts of driver behaviours, vehicle, and roadway factors on truck rollovers and occupant injury severity. Two binary probit and ordered models were used to analyse rollover propensity and injury severity, respectively. The results showed that truck-driver behaviours, sharp curves, and turning manoeuvres were associated with higher rollover risk.

Khattak and Rocha (2003) evaluated the rollover intensity and severity of driver injury of Sport Utility Vehicles (SUVs). A binary logit model was applied to investigate the likelihood of rollover occurrence. The number of quarter turns in rollover crashes, rollover intensity, and injury severities were modelled by using weighted negative binomial models and weighted ordered logit models, respectively. The results indicated that SUVs were more likely than passenger cars to rollover and also experienced more intense rollovers, though they protect their drivers due to their greater crashworthiness. Using 128 injury cases in three rollover coach crashes in Sweden, Albertsson *et al.* (2006) analysed injury outcome, injury mechanisms and the possible effects of 2-point or 3-point seat belts on injury reduction for occupants when wearing a seatbelt. The analysis emphasised the use of seat belts to reduce the risk of injury in rollover coach crashes as well as the effects of other measures, such as higher side window panels and retentive glazing to prevent the occupants from being ejected. McKnight and Bahouth (2009) conducted a study to identify causes underlying 239 truck rollover crashes. The results showed that truck rollovers were associated with loads, brake condition, road surface, and failing to adjust speed at curves. Moreover, lack of attention, misdirected attention, falling asleep, and distraction were found to be other contributing factors to rollover crashes.

Keall and Newstead (2009) analysed the risk of rollover crashes in three Australian states – Victoria, Queensland, and Western Australia – and New Zealand for the years from 1993 to 2004. Separate logistic regression models were fitted for these four data sets to identify the factors associated with the likelihood of rollo-

vers. The findings showed that vehicles with a relatively higher centre of gravity had higher rollover risk. Furthermore, female drivers, higher speed limit areas, and older model vehicles appeared to have a higher rollover risk. Using five-year data during the years from 2002 to 2006 collected on Pennsylvania rural divided highways, Hu and Donnell (2011) estimated the severity of cross-median and median rollover crashes using a binary logit model and a multinomial logit model, respectively. The results confirmed that the severity outcomes of median rollover crashes were related to the presence of horizontal curves, unbelted drivers, steeper median cross-slopes, and narrower medians. In fact, these factors significantly increased rollover severity outcomes. Funk *et al.* (2012) investigated the risk factors of cervical spine, serious, head, and fatal injury on rollover crash data extracted from 1995 to 2008 in the NASS–CDS. Nonparametric univariate analyses, univariate logistic regression, and multivariate logistic regression were developed. The results indicated that complete or partial ejection, a greater number of roof inversions, the lack of seatbelt use, and older occupant age were significantly associated with the risk of all types of injuries considered in rollovers. On the other hand, occupant height, vehicle type, and occupant gender did not have a significant effect on injury.

Dell'Acqua *et al.* (2013b) developed Safety Performance Functions (SPFs) to predict crash injury rates for different crash types on horizontal homogeneous segments of two-lane rural roads in Southern Italy. The generalized estimating equations (GEEs) with a negative binomial function were developed to fit the SPFs. For single-vehicle run-off-road crashes, which contained crashes involving vehicles that exit the roadway and strikes a fixed object or rollover, the results showed that the factors road surface (dry/wet), light condition (day/night), crash location (tangent/circular curve element), lane width, horizontal curvature indicator, and mean speed had the greatest effect on the crash incidence. Bambach *et al.* (2013a) analysed the severity of thoracic injury sustained by contained and restrained occupants involved in single vehicle pure rollover crashes that occurred in the US between 2000 and 2009. The study data was collected from the US NASS–CDS. A logistic regression model was developed to investigate the effects of risk factors related to human, vehicle and environmental characteristics on the incidence of serious thoracic injury of such crashes. The results indicated that number of quarter turns, occupant age, and body mass index were directly associated with the probability of serious thoracic injury of pure rollover crashes. Occupants of SUVs, pickups, van or trucks were more likely to be involved in serious thoracic injury than car occupants. In addition, it was found that the probability of serious thoracic injury increases where rollovers occur on curves, on dry roadway surfaces not on a level roadway and undivided roadways. Using the same data set (CDS data), Bambach *et al.* (2013b) also investigated the nature and causes of spine injury for contained and restrained occupants who were injured from contact with the vehicle roof structure during a pure rollover. An ordinal logistic regression

model was applied to determine the key factors associated with spine injury. The empirical results revealed that number of roof inversions, vehicle type, number of quarter turns, roof intrusion, far side seating position, occupant age, occupant gender (male vs. female), and body mass index were found to be significantly associated with the incidence of spine injury in rollover crashes. In addition, crashes on dry roadway surfaces contributed to increased odds of spine injury.

2. Methodology

2.1. Study Area and Data Collection

This research is part of a large research project undertaken to assess the safety effects of roadway characteristics on crash outcomes by collision type (e.g., head-on, rear-end, pedestrian, rollovers). For this purpose, Malaysian federal road system was selected as the case study. This road system is among the most important transport systems throughout the country. However, it experiences the highest rate of traffic crashes compared to other road types (e.g., state highways, expressways) where the network comprises about 20% of the total road length, yet it accounts for over 40% of all accident fatalities nationwide (ITF 2012).

Among the candidate federal roads, those finally selected were on the condition that detailed information on roadway characteristics, traffic flow, and crash data had been available and complete. Based on these conditions, the study area finally consisted of 543 km sections from five federal roads, including Malaysia Federal Route 2 (F2), Malaysia Federal Route 3 (F3), Malaysia Federal Route 4 (F4), Malaysia Federal Route 67 (F67), and Malaysia Federal Route 76 (F76) located in the states of Perak, Kedah, Kelantan, Pahang, and Terengganu in Peninsular Malaysia. The road segments under study are representative sample to Malaysian federal road system in which the vast majority of study roadways pass through rural and semi-urban areas. To investigate the relationship between the rollover occurrence and roadway geometry, the environment, and traffic characteristics, detailed information were collected from three different sources: Malaysian Institute of Road Safety Research (MIROS), Highway Planning Unit (HPU), and Royal Malaysian Police (RMP). The first database obtained from the MIROS includes a list of road geometric and environmental characteristics, such as horizontal curvature, land use, shoulder width, number of lanes, etc. The second database, which was collected from the HPU, contains traffic data for a 4-year period from 2007 to 2010, including average daily Light-Vehicle Traffic (LVT) and average daily Heavy-Vehicle Traffic (HVT). The third database consists of crash data, including location and time of crashes occurred on the considered segments between 2007 and 2010. The data were collected from the MIROS database. More than one year of crash records were used to reduce the variability of the crash frequency from year to year.

With these data at hand, the next step is to divide the study area into homogeneous segments. To do

this, the study sections were split into homogeneous segments in terms of traffic flow, land use, and cross-sectional characteristics, including shoulder width, the number of lanes, and median. After the segmentation process, the 543 km sections were segregated into 448 homogeneous segments with the length ranged between 1 km and 7 km, and an average of 1.2 km.

For a specific variable of each segment, the characteristic with the largest proportion was determined as the representative characteristic of that variable for that segment. Descriptive statistics of the data are presented in Tables 1 and 2 for continuous and categorical variables, respectively. In the context of rollover crashes, a total of 136 rollovers occurred on the study roadways during the 4-year period. Among the 448 segments,

there were 350 segments (approximately 78%) for which no rollover crash were reported. This indicates the potential presence of excess zeros in the crash data. Moreover, the ratio between variance and mean was found to be about 1.5, which implies some over-dispersion exist in the crash data.

2.2. Applied Count Models

Count-data models are generally used to model traffic crashes that are discrete, random, and non-negative integers. Because rollovers occur very rarely compared to other collision types, it is expected to have a large number of road segments for which no rollover occurred during the study period; this may result in a mass of zero counts in the crash data. In such a condition, zero-

Table 1. Descriptive statistics of continuous variables

Variable	Description	Minimum	Maximum	Mean	Standard deviation
Rollover crash frequency	The number of rollover crashes occurring on the study segments during the study period (2007–2010)	0	4	0.304	0.453
Segment length	The length of segment [km]	1	7	1.21	0.690
LVT	Average daily traffic including light vehicles (e.g., motorcycles, passenger cars, light vans, SUVs, etc.)	2790	27470	9294	6563
HVT	Average daily heavy-vehicle traffic including bus, tractor, lorry, large van, truck	456	4789	1335	659
Speed limit	Actual posted speed limit (ranging from 50 to 90 km/h)	50	90	82	11.9
Paved shoulder width	Paved shoulder width (ranging from 0 to 2.4 m)	0	2.4	1.26	0.59
UPSW	Unpaved shoulder width (UPSW) (ranging from 0 to 2.4 m)	0	2.4	1.15	0.71
Curvature	Horizontal curvature [1/km]	0.091	13.386	3.090	2.682
Access point	Number of intersection and minor access points per km	0	9	1.08	1.51

Table 2. Descriptive statistics of categorical variables

Variable	Description	Observation (proportion) in sample					
		1	2	3			
No of lanes	Number of lanes for each travel direction (1 for one-lane, 2 for two and more lanes)	397 (89%)	51 (11%)	–			
Terrain type	Indicator of vertical gradient along roadway (1 for flat terrain, 2 for rolling/undulating terrain)	320 (71%)	128 (29%)	–			
Side friction	Level of interaction between roadside activities (e.g., parking, bus stopping, trading) and through traffic (1 for low or non-interaction, 2 for high interaction)	377 (84%)	71 (16%)	–			
Median type	Indicator of two opposing traffic flows are separated or not (1 for unseparated, 2 for separated)	417 (93%)	31 (7%)	–			
Area type	Level of roadside development (1 for rural, 2 for semi-urban, 3 for urban)	388 (87%)	46 (10%)	14 (3%)			
Land use	Level of activity along roadway (1 for no activity level, 2 for low activity level (e.g., educational, industrial), 3 for high activity level (e.g., residential or commercial))	279 (62%)	103 (23%)	66 (15%)			
Roadside condition	1 for continuous safety barrier (e.g., guardrail), 2 for cut > 2 m depth, 3 for deep drainage ditches, 4 for embankment, 5 if distance to the nearest aggressive objects (e.g., rocks, tree, utility pole) is 0–5 m, 6 if distance to the nearest aggressive objects is greater than 5 m						
	Code	1	2	3	4	5	6
	Observation (proportion)	16 (4%)	55 (12%)	63 (14%)	87 (19%)	188 (42%)	39 (9%)

altered models, including zero-inflated Poisson (ZIP), zero-inflated NB (ZINB), Poisson hurdle (PH), and NB hurdle (NBH) models may be more plausible to handle the issue. Therefore, this study has developed these models together with standard count models to identify the factors associated with the occurrence of rollover crashes. A logit model was used for the zero-inflated and hurdle parts of all zero-altered models. The descriptions of the models are presented in the following.

2.2.1. Standard Count Models

The Poisson regression model is taken as the starting point for modelling count data, assuming that the mean is equal to the variance (that is, equal-dispersion) (Khan *et al.* 2011). However, in most crash data, the variance is greater than the mean, which is known as over-dispersion. The over-dispersion is a result of extra variation in crash means across road segments and could result from various factors, such as the omission of important covariates, model misspecification, and excess zero counts (Mitra, Washington 2007). In such a case, applying a Poisson regression model would lead to an underestimation of the standard error of the parameters, causing a biased selection of parameters (Khan *et al.* 2011). A more flexible approach to handle the extra-Poisson variation is to apply a negative binomial regression model. The NB model accommodates the over-dispersion by including an error term in the Poisson model and allows the variance to differ from the mean such that:

$$Y_i | \mu_i \sim \text{Poisson}(\mu_i); \tag{1}$$

$$\mu_i = \exp(\beta \cdot X_i) \cdot \exp(\varepsilon_i); \tag{2}$$

$$\begin{aligned} \text{Var}(y_i) &= E(y_i) \cdot (1 + \alpha \cdot E(y_i)) = \\ E(y_i) &+ \alpha \cdot E(y_i)^2, \end{aligned} \tag{3}$$

where: μ_i is the expected number of rollovers on segment i ; X_i is a vector of covariates (e.g., road geometry or traffic volume); β_i is a vector of estimable regression coefficients; $\exp(\varepsilon_i)$ is gamma distributed with mean one and variance α (dispersion parameter).

The Probability Density Function (PDF) of the NB model is given by Eq. (4):

$$\Pr(Y_i = y_i) = \frac{\Gamma(y_i + \phi)}{\Gamma(\phi) \cdot y_i!} \cdot \left(\frac{\mu_i}{\mu_i + \phi}\right)^{y_i} \cdot \left(\frac{\phi}{\mu_i + \phi}\right)^\phi, \tag{4}$$

where: ϕ is inverse dispersion parameter $1/\alpha$; $\Gamma(\cdot)$ is a value of the gamma distribution.

The superiority of the negative binomial model over the Poisson regression model depends on the value of dispersion parameter α . If α is not statistically different from zero, the NB model reduces to the Poisson model. Otherwise, the NB model is selected as the appropriate choice.

An important limitation of the above NB model, known as the traditional NB model, is that it assumes that the dispersion parameter α to be fixed across all road segments. However, this assumption may be violated since α may vary from segment to segment. Recently,

the structure of the NB model has been widely investigated by many researchers to increase its flexibility and accuracy of parameter estimates. A prominent extension of the NB model is the heterogeneous negative binomial (HTNB) model, which allows the dispersion parameter to vary across road segments as a function of roadway characteristics (in some documents, this approach is referred to as generalized negative binomial (GNB) model). The superiority of the HTNB model using a varying dispersion parameter has been confirmed by previous studies, e.g., see Miranda-Moreno *et al.* (2005), Miranda-Moreno and Fu (2006), Geedipally and Lord (2008), Usman *et al.* (2010). Similar to the traditional NB model, the HTNB model uses the same PDF as that given in Eq. (3). However, in the HTNB model, dispersion parameter is a function of site-specific attributes, as follows (Geedipally, Lord 2008):

$$\alpha_i = \exp(\gamma_0 + \gamma_1 \cdot Z_{i1} + \gamma_2 \cdot Z_{i2} + \dots + \gamma_m \cdot Z_{im}), \tag{5}$$

where: $Z_i = (Z_{i1}, \dots, Z_{im})$ is a vector of site-specific variables, which are not necessarily the same as those used for estimating μ_i and $\gamma_i = (\gamma_{i1}, \dots, \gamma_{im})$ is a vector of parameters to be estimated.

Thanks to Eq. (5), one can associate dispersion parameters α_i to the road segments' characteristics. If no variables were found to contribute to the dispersion parameters, the latter will only take a constant value, reducing to a traditional NB model (Abdelwahab, Abdel-Aty 2004).

2.2.2. Zero Altered Models

In addition to unobserved heterogeneity, excess zeros in crash data could be another source of over-dispersion. In such cases, the parent count models, including the Poisson and NB models are not appropriate because these models may under predict excess zeros. To better fit the data, zero-inflated and hurdle models should be used to handle mass zero counts.

Zero-inflated models

Zero inflated regression models are used for modelling data characterised by a significant amount of zeros or more zeros than expected in the standard Poisson and negative binomial models. While the zero-inflated Poisson (ZIP) model can handle over-dispersion caused by excess zeros, it does not accommodate over-dispersion arising from both unobserved heterogeneity and excess zeros. To deal with this problem, a zero-inflated negative binomial model (ZINB) is applied (Miranda-Moreno, Fu 2006).

Both ZIP and ZINB are interpreted as a mix of structural and sampling zeros that come from two different processes:

- the process that generates structural zeros estimated from a binary distribution (logit or probit distribution);
- the process that generates sampling zeros that are derived from the Poisson/NB distribution.

Structural zeros correspond to the outcomes that are never experienced (i.e., always zero), while sampling

zeros correspond to the outcomes that are experienced but not during a pre-defined short-term period (Martínez-Espiñeira 2007; Moineddin et al. 2011).

As the description of ZIP model, let P_i be the probability of segment i being an excess zero and $(1-P_i)$ be the probability of crash counts derived from the Poisson distribution. In general, the PDF for the ZIP model is:

$$P(Y = y_i) = \begin{cases} P_i + (1 - P_i) \cdot \exp(\mu_i), & y_i = 0; \\ (1 - P_i) \cdot \frac{\exp(-\mu_i) \cdot \mu_i^{y_i}}{y_i!}, & y_i > 0, \end{cases} \quad (6)$$

where: y_i is the number of rollover crashes for segment i ; μ_i is the expected outcome for segment i as a function of its covariates, $\mu_i = \exp(\beta \cdot X_i)$.

The probability of being in the zero-crash-state, P_i , is often fitted using a logistic regression model, as follows:

$$\text{logit}(P_i) = \ln\left(\frac{P_i}{1 - P_i}\right) = \gamma_o + \gamma_1 \cdot Z_1 + \dots + \gamma_N \cdot Z_N, \quad (7)$$

where: $Z = (Z_1, Z_2, \dots, Z_N)$ is a function of the explanatory variables and $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N)$ is the estimable coefficients.

Similar to the ZIP model, the PDF for the ZINB model is given by Eq. (8):

$$P(Y = y_i) = \begin{cases} a_1, & y_i = 0; \\ a_2, & y_i > 0; \end{cases} \quad (8)$$

$$a_1 = P_i + (1 - P_i) \cdot \frac{1}{(1 + \alpha \mu_i)^{\frac{1}{\alpha}}};$$

$$a_2 = (1 - P_i) \cdot \frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i + 1) \cdot \Gamma\left(\frac{1}{\alpha}\right)} \cdot \frac{(\alpha \cdot \mu_i)^{y_i}}{(1 + \alpha \cdot \mu_i)^{y_i + \frac{1}{\alpha}}},$$

where: α and $\Gamma(\cdot)$ are the dispersion parameter and the gamma function for the ZINB model, respectively.

Hurdle models

Hurdle models were first introduced by Cragg (1971) and subsequently reviewed by Mullahy (1986). As with ZI models, the hurdle models can handle data characterised by a mass of zeros, and fit the response variable as a mixture of binary and count distributions; however, they assume that all zeros in the crash data are sampling zeros. Hurdle models are interpreted as two state models; a zero state with no crashes and second state in which at least one crash occurs. The first part of the model can be modelled using a binary regression framework, such as a logit or probit model. Given that a crash occurs, the number of crashes can then be modelled by a left truncated Poisson or negative binomial distribution. In general, the hurdle models are typical count models in which the zeroes and positive counts are separately generated. First, to describe the Poisson hurdle model, let the probability of zero count be given by P . Furthermore, the probability of a non-zero count is given by $(1 - P)$. Therefore, a crash can be obtained from

a truncated Poisson with a probability of $(1 - P)$. The general hurdle Poisson (HP) density is given as follows:

$$P(Y = y_i) = \begin{cases} P_i, & y_i = 0; \\ a, & y_i > 0; \end{cases} \quad (9)$$

$$a = (1 - P_i) \cdot \frac{\exp(-\mu_i) \cdot \mu_i^{y_i}}{(1 - \exp(-\mu_i)) y_i!};$$

$$\text{logit}(P_i) = \ln\left(\frac{P_i}{1 - P_i}\right) = \beta_o + \beta_1 \cdot Z_1 + \dots + \beta_N \cdot Z_N, \quad (10)$$

where: P and μ are fitted by logit and count models, respectively, and their corresponding covariates may be fitted separately.

However, the HP model will only account for excess zeroes, and it will not account for over-dispersion caused by unobserved heterogeneity. To cope with this problem, the hurdle negative binomial (HNB) model can be used to handle the over-dispersion arising from both excess zeroes and unobserved heterogeneity. Similar to the HP model, if we use a logit approach to model the probability P of a zero versus a non-zero $(1-P)$ count and a left truncated negative binomial density for the count process, then our overall HNB density is:

$$P(Y = y_i) = \begin{cases} P_i, & y_i = 0; \\ a, & y_i > 0; \end{cases} \quad (11)$$

$$a = (1 - P_i) \cdot \left(1 - \frac{1}{(1 + \alpha \cdot \mu_i)^{\frac{1}{\alpha}}}\right) \times$$

$$\left(\frac{\Gamma\left(y_i + \frac{1}{\alpha}\right)}{\Gamma(y_i + 1) \cdot \Gamma\left(\frac{1}{\alpha}\right)}\right) \cdot \left(\frac{(\alpha \cdot \mu_i)^{y_i}}{(1 + \alpha \cdot \mu_i)^{y_i + \frac{1}{\alpha}}}\right),$$

where: α and $\Gamma(\cdot)$ are the dispersion parameter and the gamma function, respectively; μ_{it} is the predicted crash counts derived from left truncated negative binomial model.

Hurdle models have been applied in a variety of fields, such as economics, medical science, environment, industry. However, they have rarely been adopted in road safety literature (Boucher, Santolino 2010; Hosseinpour et al. 2013; Son et al. 2011).

2.3. Model Selection Criteria

The comparison and selection among the candidate models is based on the presence and the source of over-dispersion in the crash data (Son, 2011). To check if over-dispersion exists in the rollover crashes, a Wald t -statistical test on the dispersion parameter and a likelihood ratio test (LRT) were performed where the Poisson, NB, HP, and ZIP models were nested within the NB, HTNB, HNB, and ZINB models, respectively (Isgin et al. 2008). The LRT is based on differences in the log-likelihoods of two nested models, as given in Eq. (12):

$$LR = 2 \cdot \left(LL_{NB \text{ or } HNB} - LL_{PM \text{ or } HPM}\right) \cong \chi^2_{(d.f.=1)}. \quad (12)$$

The test follows a χ^2 distribution with one degree of freedom. A significant value for both the Wald t -statistical and LR tests indicates that the over-dispersion in the crash data is present, and that is thought to arise from unobserved heterogeneity. In this case, the NB-based models would be preferred to the Poisson counterparts. Otherwise, the Poisson-based models are used.

Based on the existence of over-dispersion, the contribution of zeros to extra-dispersion is then examined by a Vuong (1989) test because zero-altered models are not nested within parent models. For the NB-based models (e.g., NBM vs. ZINB and HNB), a significant value for the Vuong test indicates that both excess zeros and unobserved heterogeneity account for over-dispersion, and thus zero-altered NB models are preferred to the parent NB model. Similarly, the Vuong test is also conducted for Poisson-based models. A significant value for the test indicates that only zero counts contribute to over-dispersion, and thus two-state models (either HP or ZIP) are preferred to the PM. In addition, the Vuong test is applied for comparing between hurdle and zero-inflated models (ZIP vs. HP or/and ZINB vs. HNB).

Given that $P_1(y_i | x_i)$ and $P_2(y_i | x_i)$ are the predicted probability of the standard models (Poisson or NB models) and the two-state model (zero-inflated and hurdle models), respectively, the Vuong test can be expressed as:

$$m_i = \ln \left(\frac{\sum_i P_1(y_i | x_i)}{\sum_i P_2(y_i | x_i)} \right); \tag{13}$$

$$m_i = \ln \left(\frac{\sum_i P_1(y_i | x_i)}{\sum_i P_2(y_i | x_i)} \right), \tag{14}$$

The Vuong test V follows a standard normal distribution. If V is greater than 1.96, then the test favours HP/ZIP or ZIP/ZINB over Poisson/NB, and if V is lower than -1.96 , the parent Poisson or NB model is favoured. A value of $-1.96 < V < 1.96$ indicates neither model is preferred over the other. In addition, two information criteria were used to compare both the nested and non-nested models: the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The AIC and BIC are defined as follows:

$$AIC = -2 \cdot LL + 2 \cdot P; \tag{15}$$

$$BIC = -2 \cdot LL + P(\ln(n)), \tag{16}$$

where: LL is the logarithm of the maximum likelihood estimation for each model; P is the number of model parameters, and n is the number of observations ($n = 448$).

A model with the lowest AIC and BIC values is preferred. To decide whether there is a statistically significant difference between two models, Hilbe's AIC and Raftery's BIC rule-of-thumb criteria were adopted in this study (Raftery 1995; Hilbe 2011). Table 3 shows the significance levels for both criteria. In this case study ($n = 448$), if the difference in the AIC value is greater than 2.5, then the model with lower AIC is favoured over another.

3. Results and Discussion

3.1. Model Development and Selection

Prior to the modelling procedure, a correlation analysis was conducted for the study variables to check the presence of multicollinearity. No evidence of high collinearity was found between different variables. Therefore, there is no concern regarding multicollinearity in the data. Next, all seven models, including single- and dual-state models were developed and compared. In all the considered models, segment length was modelled as an offset variable. With respect to the HNB model, the algorithm for model estimation was not converged during the model calibration. Therefore, the HNB model was excluded from further analysis. For the ZINB model, the dispersion parameter estimate was not significant at the 5% level. Consequently, the ZINB model was also excluded from further consideration. This may indicate that over-dispersion is likely to be due either to excess zeros or to unobserved heterogeneity rather than their combination. The results of parameter estimates as well as goodness-of-fit measures for the remaining models are presented in Table 4. The reason for presenting the results of all the models is to demonstrate how these models associate the various risk factors with rollover crashes. Note that original outcome of the logit model in the ZIP model is to estimate the probability of being in the zero rollover crash state. Nevertheless, to easily compare the results of the logit model to those of the Poisson model estimating the crash frequency, we changed the sign of the coefficients so that the zero state of the ZIP model reflects the probability of being in the non-zero rollover group.

To statistically confirm the presence of over-dispersion in the crash data, the standard NB model was compared to its Poisson counterpart. The likelihood ratio test for NB vs. Poisson was estimated $\chi^2 = 134.2$ (p -value < 0.0001) that is highly significant. In addition, the p -value for the dispersion parameter was found to be significant at the 1% level. These statistics

Table 3. Significance levels for AIC and BIC (Raftery 1995; Hilbe 2011)

ΔAIC for models A and B	Result if A < B	ΔBIC for models A and B	Result if A < B
< 0.0 and ≤ 2.5	No difference	< 0.0 and ≤ 2.0	Weak difference
< 2.5 and ≤ 6.0	Prefer A if $n > 256$	< 2.0 and ≤ 6.0	Positive difference
< 6.0 and ≤ 9.0	Prefer A if $n > 64$	< 6.0 and ≤ 10.0	Strong difference
10+	Prefer A	10+	Very strong difference

confirm the presence of over-dispersion, as well as the appropriateness of NB model over the Poisson model to fit the over-dispersed data. To compare between NB and HTNB, a LRT was applied since the former is nested in the latter. The test with a value of $\chi^2 = 10.18$

Table 4. Parameter estimates of the fitted models

Covariates	NB	HTNB	HP	ZIP
	Mean part as NB model		Count part as Poisson model	
Intercept	-8.096 (2.182)***	-8.512 (2.054)***	-4.715 (3.015)	-7.425 (1.369)***
ln(LVT)	0.411 (0.207)**	0.394 (0.200)**	1.659 (0.362)***	0.451 (0.148)***
ln(HVT)	-	-	-2.016 (0.444)***	-
Access point	0.086 (0.084)	0.128 (0.068)*	-	-
Horizontal curvature	0.151 (0.044)***	0.158 (0.046)***	0.376 (0.059)***	0.214 (0.031)***
Speed limit	0.015 (0.011)	0.021 (0.010)**	-	-
Median (1 if present, 0 otherwise)	1.012 (0.406)**	1.021 (0.349)**	1.641 (0.417)***	1.004 (0.252)***
	<i>Over-dispersion part</i>		<i>Zero part as logit model</i>	
Intercept	0.681 (0.223)***	1.218 (0.274)***	-6.524 (2.059)***	7.306 (1.952)***
ln(LVT)	-	-	0.437 (0.217)**	-
Access point	-	-0.652 (0.294)**	-	5.939 (1.660)***
Horizontal curvature	-	-	0.101 (0.049)**	-
UPSW	-	-	-0.319 (0.182)*	-2.926 (1.088)***
Median (1 if present, 0 otherwise)	-	-	1.078 (0.422)**	-
	<i>Summary statistics</i>			
No of observations	448	448	448	448
No of parameters	7	8	10	7
Log-likelihood at converge	-312.1	-307.0	-313.1	-321.0
Vuong test vs. Poisson (<i>p</i> -value) vs. NB (<i>p</i> -value)	-	-	3.63 (0.00014)*** -0.137 (0.445)	2.53 (0.0057)*** -0.876 (0.190)
AIC	638.2	630	646.1	655.9
BIC	669.5	662.8	687.2	684.6

Notes: Numbers in () denotes the S.E. of the parameter estimates; * indicates significance at $\alpha = 0.10$; ** indicates significance at $\alpha = 0.05$; *** indicates significance at 0.01.

(*p*-value = 0.0014) gave an advantage to the HTNB model over the standard NB model, which indicates that the HTNB is a more flexible approach to handle the over-dispersion in the crash data. In the next step, the Vuong test was applied to compare the HP and ZIP models with the parent Poisson and NB models in order to check the contribution of excess zeros in over-dispersion. For pairs of Poisson vs. HP and ZIP, the test showed that the Poisson was rejected in favour of the HP and ZIP models. However, the test for HP and ZIP against NB revealed that neither model was favoured over the other. Under such circumstances, the two AIC and BIC criteria are used to determine the best fit model. These criteria are also used to compare the HTNB with HP and ZIP since no specific test exists to compare the HTNB model with zero-altered models. Both AIC and BIC favoured the HTNB model over the others. In terms of AIC, the HTNB model has the lowest value. Based on Hilbe's rule of thumb for this study ($n = 448$), if ΔAIC is greater than 2.5, then the model with the lowest AIC is preferred. For this study sample, the minimum difference in AIC was found for HTNB versus NB by 8, and thus the HTNB is preferred over the other models. The superiority of the HTNB model was also supported by the BIC. Based on Raftery's rule of thumb, the minimum ΔBIC was found to be strong for the HTNB versus NB by about 7, which favours highly the extended NB model. Overall, according to the LL, AIC, and BIC, the HTNB model was determined to be the best fit model for the current data. The appropriateness of HTNB model over the standard and zero-altered count models (e.g., NB, HP, and ZIP) indicates that the model is more flexible to accommodate over-dispersion due to unobserved heterogeneity in which excess zeros are not of concern. Overall, five variables were found to be statistically associated with rollover crash frequency in the HTNB model; these variables are the logarithm of LVT, access point, horizontal curvature, speed limit, and the presence of median. The variable access point was found to statistically contribute to dispersion parameter. To ease interpretation of the significant variables, the Incidence Rate Ratios (IRR), i.e. $\exp(\beta)$ was estimated and presented in Table 5. For a given variable with an IRR greater than 1.0, an increase in the value of the variable is correlated to an increase in rollover occurrence and vice versa.

As shown in the table, the logarithm of LVT was found to contribute positively to rollovers. This implies that as the amount of traffic flow increases, the exposure

Table 5. IRR for the HTNB model coefficients

Variables	Coefficient	IRR
ln(LVT)	0.394	1.483
Access point	0.128	1.137
Horizontal curvature	0.158	1.172
Speed limit	0.021	1.021
Median (1 if present, 0 otherwise)	1.021	2.775

to rollover crashes increases. LVT increases the risk of rollovers in two ways:

- a vehicle is struck by another vehicle, causing the vehicle to rollover;
- the vehicle is not struck by another vehicle but rolled over as a result of the driver's swerving manoeuvre to avoid a collision with the other vehicle (McKnight, Bahouth 2009).

For both circumstances, the risk of rollover crashes tends to rise as the number of LVT increases. The IRR for LVT is 1.48, which implies that one unit increase in the log of LVT corresponds to a 48% increase in the rollover crashes, with the remaining predictor values held constant. An important finding reached in this study is the impact of horizontal curvature on rollovers. Curve segments are more likely than straight segments to experience rollover crashes because sharp curves reduce drivers' visibility and their ability to control the vehicle, especially when travelling at high speeds, which increases the risk of rollover occurrence.

This finding is consistent with those of previous studies by Farmer and Lund (2002), Khattak and Rocha (2003), Khattak *et al.* (2003), Schneider *et al.* (2010), Hu, Donnell (2011), Dell'Acqua *et al.* (2013a). For example, McKnight and Bahouth (2009) implied that the effect of curvature on rollovers is mainly due to drivers' misjudgement of speed and their failure to adjust speed when they negotiate a curve. In such cases, straightening sharp curves is thought to be the best way to reduce the risk of rollovers. However, it seems to be very expensive, and may not be cost effective in most cases. Alternatively, other improvements, such as posting lower speed limits and warning signs, installing continuous guardrails and rumble strips, and widening clear zone could be more effective, especially for curves with hazardous roadside obstacles. According to the IRR value, a 1 km^{-1} sharper curve is associated with a 17% increase in the rollover occurrence.

The number of access points was positively correlated to the occurrence of rollover crashes. The reason for this finding is that as the number of minor driveways increases, there are more conflicts among vehicles approaching from different directions which increase the probability that a vehicle strikes or is struck by another vehicle in side, increasing the probability that the vehicle rolls over. The IRR for 'access point' is about 1.14, indicating that one unit increase in the number of access points will result in a 14% increase in the risk of rollover crashes. An interesting finding was reached for the effect of median on rollover crashes. As seen in the table, road segments with centreline physical median were more likely to be associated with rollover crashes. This may be because centreline medians (e.g., raised or curbed medians) increase the risk of rollover crashes when a driver loses control of the vehicle, and strikes the centre median, causing the vehicle to slide sideways and ultimately rollover. From the computed IRR value, a road segment with median is more than twice as likely to experience rollovers.

Speed limit was found to have a positive impact on rollover crashes. This finding is intuitive and is consistent with expectation. Road segments with higher speed limits are more prone to rollover risk than those with lower speed limits. This result is because higher speed limits are typically posted in rural areas with lower vehicular traffic. In such locations, drivers tend to go fast, and thus they are more likely to lose the control of vehicle at high speeds, especially when they negotiate a sharp curve or attempt to avoid an unexpected event, increasing the risk of rollover. This finding is consistent with the results of past studies (Dell'Acqua *et al.* 2013b; Donelson *et al.* 1999; Keall, Newstead 2009; Khattak, Rocha 2003; Krull *et al.* 2000; Viner 1995). In addition, excessive speeds are often related to reckless driving behaviours in most speeding-related crashes (McKnight, Bahouth 2009). As a result, speed limit could be as a surrogate for actual speeds because the latter is not available in most accident records.

In both HP and ZIP models (Table 4), UPSW was found to have a significant effect on reducing the likelihood of rollover occurrence. Generally, wider shoulders decrease the risk of rollovers by giving more recovery room to errant or uncontrolled vehicles, so that they can avoid encountering roadside hazardous barriers (e.g., trees, guardrail, curbs) or side-slopes as potential causes of rollovers. In the HP model, a somewhat surprising finding was reached for the effect of HVT. The HVT was found to have a negative sign, indicating an increase in the number of HVT will correspond to a reduction in the rollover frequency. Such a result contradicts findings from earlier studies in the literature that found that heavy vehicles are more likely to rollover than other vehicle types, due to their higher centre of gravity and low roll stability (Khattak *et al.* 2003; McKnight, Bahouth 2009). However, a potential reason for such a contradictory finding is that heavy vehicles comprise a lower proportion of total traffic compared to passenger cars or other light vehicle types. Therefore, they are less exposed to rollover than other vehicle types. Another possible explanation is that heavy vehicles travel at lower speeds. In addition, they decrease the risk of speeding and passing actions made by other vehicles, and thus reduce the propensity of rollover crashes (Milton, Mannering 1998).

It is worthy to note that the factor 'roadside condition' did not contribute to rollovers in all the fitted models while this factor is believed to have a potential impact on some types of rollovers, such as fall over, flip over, and trip over. As a response, one may conclude that most of the rollovers occurred on-the-road rather than off-the-road. The contribution of two variables access points and centreline median to the rollover occurrence may be an evidence for reaching such a conclusion. However, this finding is somewhat questionable because no information was available in the dataset to address the type and location (i.e., on-road or off-road) of the rollovers occurred. Furthermore, since there are a few observations of rollovers in the present study, it is difficult to reach a reliable finding with respect to the effect of roadside conditions on rollover crashes.

3.2. Prevention Strategies

Based on the results presented in this study, rollover crashes are directly related to higher vehicular traffic, the larger number of access point, presence of centreline physical barrier, sharper horizontal curvature, narrower shoulder widths, and higher posted speed limits. As such, to reduce the risk of rollover crashes, preventive strategies should target those road sections that have such substandard safety conditions that are likely to pose rollover crashes. The findings of this study may assist road safety authorities to propose and develop countermeasures relevant to those roadway factors associated with the risk of rollover crashes. As instances of potential treatments, rollovers could be reduced by straightening sharp horizontal curves, proper management of minor access to the roadway, widening shoulder width, better design of centreline medians, posting lower speed limits and warning signs in areas with high rollover tendency, and installation of speed enforcement cameras at locations that are more prone to speeding manoeuvres. In addition, a low-cost countermeasure is to install rumble strips along the shoulder or centreline on curved sections. Rumble strips can assist in preventing rollover crashes by producing a rumbling sound and thus warns sleeping or inattentive drivers not to leave the road and possibly rollover (Spainhour, Mishra 2008). In addition, installing a continuous barrier (e.g., guardrails) along the outside edge of road curve sections could be implemented to reduce the risk of rollover crashes though preventing errant or uncontrolled vehicles from encountering off-road side-slopes.

Recently, stability-enhancing technologies, such as Electronic Stability Control (ESC) and Roll Stability Control (RSC), have been widely used worldwide to prevent rollovers. Among these technologies, ESC is regarded as the most prevalent system for promoting vehicle stability. This system has been applied since the late 1990s. ESC prevents rollovers by applying brakes individually to the wheels to correct for oversteering and understeering and thus help drivers maintain control of the vehicle (Keall, Newstead 2009). ESC is more effective when the road is wet or slippery. Numerous reports and studies around the world have recognised and confirmed the effectiveness of ESC in reducing deaths and serious injuries resulting from rollovers and Loss-Of-Control (LOC) collisions (Dang 2004; Farmer, Lund 2002; Ferguson 2007; MacLennan *et al.* 2008; Woodrooffe *et al.* 2011; Yim *et al.* 2012). As a consequence, most developed countries, such as the US, Canada, and Australia have recently required all new vehicles to be equipped with an ESC system. It seems that a similar fashion could be implemented in Malaysia so that all new vehicles (especially those more vulnerable to rollovers, such as trucks, buses, SUVs, pickups, and vans) are equipped with such stability-enhancing technologies. This could be achieved through public awareness, education, and legislation.

Conclusions and Recommendations

A number of previous studies have focused on rollover likelihood and severity mainly by fitting a binary model using a large context of vehicular, driver behavioural, and road factors. However, research on the effects of road geometry, environment, and traffic volumes on the rollover occurrence is still limited. In response, the current study has provided an empirical analysis to investigate the safety effects of road factors on rollover incidents occurred on 448 homogeneous segments from Malaysian federal roads. Because rollovers occur rarely compared to other collision types, a high number of zero counts are expected in the crash data. Thus, zero-altered models, including zero-inflated and hurdle models are typically applied to handle potential overdispersion due to excess zeros. To achieve the objective of this study, seven count models, including the Poisson, negative binomial (NB), heterogeneous negative binomial (HTNB), hurdle Poisson (HP), hurdle negative binomial regression (HNB), zero-inflated Poisson (ZIP), and zero-inflated negative binomial (ZINB) models, were developed and compared. The results showed that the Poisson model failed to fit the data due to the presence of over-dispersion. The HNB was excluded from further consideration due to convergence problem during the model calibration. For the ZINB models, the dispersion parameter was found to be non-significant. The overall conclusion from the comparative analyses indicated that the HTNB model was preferred to the other candidate models to fit the rollover data characterised by unobserved heterogeneity. The results showed that LVT, access points, HVT, curvature, shoulder width, speed limit, and presence of centreline median were significant factors that influence the occurrence of rollover crashes. The findings of this study could be useful for suggesting prevention efforts through developing a number of cost-effective countermeasures. To conclude, rollover crashes may be potentially prevented by implementing some remedial improvements, such as widening shoulder width, implementing rumble strips on road curved sections, the use of ESC, enforcing speed zone strategies, etc.

To the authors' best knowledge, this is one of the first efforts that provide new insight into the effects of roadway characteristics on rollover crashes. It is considered as a leading step forward in modelling such rarely-occurring but more-severe collision types in real-world traffic accidents. Further research in this domain should be conducted for larger samples and longer time period. In addition, a national-scale study is recommended to evaluate the potential benefits of ESC and other stability-enhancing technologies in reducing accident casualties resulting from rollovers and LOC crashes (as a major cause of rollovers), where these collision types overall comprise approximately 67% of single and 27% of total fatal crashes across the country.

Acknowledgments

The authors would like to thank the Malaysian Institute of Safety Research (MIROS) and the Highway Planning Unit (HPU) for the data used in this research.

This study is a part of a research project sponsored by the University of Science, Malaysia.

The authors acknowledge the support and assistance.

References

- Abdel-Aty, M.; Haleem, K. 2011. Analyzing angle crashes at unsignalized intersections using machine learning techniques, *Accident Analysis & Prevention* 43(1): 461–470. <http://doi.org/10.1016/j.aap.2010.10.002>
- Abdelwahab, H.; Abdel-Aty, M. 2004. Investigating the effect of light truck vehicle percentages on head-on fatal traffic crashes, *Journal of Transportation Engineering* 130(4): 429–437. [http://doi.org/10.1061/\(ASCE\)0733-947X\(2004\)130:4\(429\)](http://doi.org/10.1061/(ASCE)0733-947X(2004)130:4(429))
- Albertsson, P.; Falkmer, T.; Kirk, A.; Mayrhofer, E.; Björnstig, U. 2006. Case study: 128 injured in rollover coach crashes in Sweden – injury outcome, mechanisms and possible effects of seat belts, *Safety Science* 44(2): 87–109. <http://doi.org/10.1016/j.ssci.2005.07.005>
- Bambach, M. R.; Grzebieta, R. H.; McIntosh, A. S. 2013a. Thoracic injuries to contained and restrained occupants in single-vehicle pure rollover crashes, *Accident Analysis & Prevention* 50: 115–121. <http://doi.org/10.1016/j.aap.2012.03.033>
- Bambach, M. R.; Grzebieta, R. H.; McIntosh, A. S.; Mattos, G. 2013b. Cervical and thoracic spine injury from interactions with vehicle roofs in pure rollover crashes, *Accident Analysis & Prevention* 50: 34–43. <http://doi.org/10.1016/j.aap.2012.10.006>
- Boucher, J.-P.; Santolino, M. 2010. Discrete distributions when modeling the disability severity score of motor victims, *Accident Analysis & Prevention* 42(6): 2041–2049. <http://doi.org/10.1016/j.aap.2010.06.015>
- Chin, H. C.; Quddus, M. A. 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections, *Accident Analysis & Prevention* 35(2): 253–259. [http://doi.org/10.1016/S0001-4575\(02\)00003-9](http://doi.org/10.1016/S0001-4575(02)00003-9)
- Conroy, C.; Hoyt, D. B.; Eastman, A. B.; Erwin, S.; Pacyna, S.; Holbrook, T. L.; Vaughan, T.; Sise, M.; Kennedy, F.; Velky, T. 2006. Rollover crashes: predicting serious injury based on occupant, vehicle, and crash characteristics, *Accident Analysis & Prevention* 38(5): 835–842. <http://doi.org/10.1016/j.aap.2006.02.002>
- Cragg, J. G. 1971. Some statistical models for limited dependent variables with application to the demand for durable goods, *Econometrica* 39(5): 829–844. <http://doi.org/10.2307/1909582>
- Dang, J. N. 2004. *Preliminary Results Analyzing the Effectiveness of Electronic Stability Control (ESC) Systems*. Report No DOT HS 809 790. National Highway Traffic Safety Administration, US Department of Transportation. Washington, DC. 5 p.
- Dell'Acqua, G.; Busiello, M.; Russo, F. 2013a. Safety data analysis to evaluate highway alignment consistency, *Transportation Research Record: Journal of the Transportation Research Board* 2349: 121–128. <http://doi.org/10.3141/2349-14>
- Dell'Acqua, G.; Russo, F.; Biancardo, S. A. 2013b. Risk-type density diagrams by crash type on two-lane rural roads, *Journal of Risk Research* 16(10): 1297–1314. <http://doi.org/10.1080/13669877.2013.788547>
- Dinu, R. R.; Veeraragavan, A. 2011. Random parameter models for accident prediction on two-lane undivided highways in India, *Journal of Safety Research* 42(1): 39–42. <http://doi.org/10.1016/j.jsr.2010.11.007>
- Donelson, A.; Ramachandran, K.; Zhao, K.; Kalinowski, A. 1999. Rates of occupant deaths in vehicle rollover: importance of fatality-risk factors, *Transportation Research Record: Journal of the Transportation Research Board* 1665: 109–117. <http://doi.org/10.3141/1665-15>
- Farmer, C. M.; Lund, A. K. 2002. Rollover risk of cars and light trucks after accounting for driver and environmental factors, *Accident Analysis & Prevention* 34(2): 163–173. [http://doi.org/10.1016/S0001-4575\(01\)00010-0](http://doi.org/10.1016/S0001-4575(01)00010-0)
- Ferguson, S. A. 2007. The Effectiveness of electronic stability control in reducing real-world crashes: a literature review, *Traffic Injury Prevention* 8(4): 329–338. <http://doi.org/10.1080/15389580701588949>
- Funk, J. R.; Cormier, J. M.; Manoogian, S. J. 2012. Comparison of risk factors for cervical spine, head, serious, and fatal injury in rollover crashes, *Accident Analysis & Prevention* 45: 67–74. <http://doi.org/10.1016/j.aap.2011.11.009>
- Geedipally, S.; Lord, D. 2008. Effects of varying dispersion parameter of poisson-gamma models on estimation of confidence intervals of crash prediction models, *Transportation Research Record: Journal of the Transportation Research Board* 2061: 46–54. <http://doi.org/10.3141/2061-06>
- Hilbe, J. M. 2011. *Negative Binomial Regression*. 2nd edition. Cambridge University Press. 576 p.
- Hosseinpour, M.; Prasetijo, J.; Yahaya, A. S.; Ghadiri, S. M. R. 2013. A comparative study of count models: application to pedestrian-vehicle crashes along Malaysia federal roads, *Traffic Injury Prevention* 14(6): 630–638. <http://doi.org/10.1080/15389588.2012.736649>
- Hu, W.; Donnell, E. T. 2011. Severity models of cross-median and rollover crashes on rural divided highways in Pennsylvania, *Journal of Safety Research* 42(5): 375–382. <http://doi.org/10.1016/j.jsr.2011.07.004>
- Isgin, T.; Bilgic, A.; Forster, D. L.; Batte, M. T. 2008. Using count data models to determine the factors affecting farmers' quantity decisions of precision farming technology adoption, *Computers and Electronics in Agriculture* 62(2): 231–242. <http://doi.org/10.1016/j.compag.2008.01.004>
- ITF. 2012. *Road Safety Annual Report 2011*. International Road Traffic and Accident Database (IRTAD) Report. International Transport Forum (ITF), Organisation for Economic Co-operation and Development (OECD). 342 p. Available from Internet: http://www.oecd-ilibrary.org/transport/road-safety-annual-report-2011_irtad-2011-en
- Keall, M.; Newstead, S. 2009. Induced exposure estimates of rollover risk for different types of passenger vehicles, *Traffic Injury Prevention* 10(1): 30–36. <http://doi.org/10.1080/15389580802383117>
- Khan, A.; Ullah, S.; Nitz, J. 2011. Statistical modelling of falls count data with excess zeros, *Injury Prevention* 17(4): 266–270. <http://doi.org/10.1136/ip.2011.031740>
- Khattak, A.; Rocha, M. 2003. Are SUVs “supremely unsafe vehicles”? analysis of rollovers and injuries with sport utility vehicles, *Transportation Research Record: Journal of the Transportation Research Board* 1840: 167–177. <http://doi.org/10.3141/1840-19>

- Khattak, A. J.; Schneider, R. J.; Targa, F. 2003. Risk factors in large truck rollovers and injury severity: analysis of single-vehicle collisions, in *TRB 2003 Annual Meeting CD-ROM*, 12–16 January 2003, Washington, DC, 1–22.
- Krull, K.; Khattak, A.; Council, F. 2000. Injury effects of rollovers and events sequence in single-vehicle crashes, *Transportation Research Record: Journal of the Transportation Research Board* 1717: 46–54. <http://doi.org/10.3141/1717-07>
- MacLennan, P. A.; Marshall, T.; Griffin, R.; Purcell, M.; McGwin, G.; Rue, L. W. 2008. Vehicle rollover risk and electronic stability control systems, *Injury Prevention* 14(3): 154–158. <http://doi.org/10.1136/ip.2007.016576>
- Martínez-Espiñeira, R. 2007. 'Adopt a hypothetical pup': a count data approach to the valuation of wildlife, *Environmental and Resource Economics* 37(2): 335–360. <http://doi.org/10.1007/s10640-006-9051-x>
- McKnight, A. J.; Bahouth, G. T. 2009. Analysis of large truck rollover crashes, *Traffic Injury Prevention* 10(5): 421–426. <http://doi.org/10.1080/15389580903135291>
- Milton, J.; Mannering, F. 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies, *Transportation* 25(4): 395–413. <http://doi.org/10.1023/A:1005095725001>
- Miranda-Moreno, L. F.; Fu, L. 2006. A comparative study of alternative model structures and criteria for ranking locations for safety improvements, *Networks and Spatial Economics* 6(2): 97–110. <http://doi.org/10.1007/s11067-006-7695-2>
- Miranda-Moreno, L.; Fu, L.; Saccomanno, F.; Labbe, A. 2005. Alternative risk models for ranking locations for safety improvement, *Transportation Research Record: Journal of the Transportation Research Board* 1908: 1–8. <http://doi.org/10.3141/1908-01>
- Mitra, S.; Washington, S. 2007. On the nature of over-dispersion in motor vehicle crash prediction models, *Accident Analysis & Prevention* 39(3): 459–468. <http://doi.org/10.1016/j.aap.2006.08.002>
- Moineddin, R.; Meaney, C.; Agha, M.; Zagorski, B.; Glazier, R. H. 2011. Modeling factors influencing the demand for emergency department services in Ontario: a comparison of methods, *BMC Emergency Medicine* 11: 1–14. <http://doi.org/10.1186/1471-227x-11-13>
- Mullahy, J. 1986. Specification and testing of some modified count data models, *Journal of Econometrics* 33(3): 341–365. [http://doi.org/10.1016/0304-4076\(86\)90002-3](http://doi.org/10.1016/0304-4076(86)90002-3)
- Pape, D.; McMillan, N.; Greenberg, A.; Mayfield, H.; Chitwood, J.; Winkler, C.; Blower, D.; Gordon, T.; Barnes, M.; Brock, J.; Harback, K. 2008. Benefits and costs of four approaches to improving rollover stability of cargo tank motor vehicles, *Transportation Research Record: Journal of the Transportation Research Board* 2066: 114–121. <http://doi.org/10.3141/2066-13>
- Poch, M.; Mannering, F. 1996. Negative binomial analysis of intersection-accident frequencies, *Journal of Transportation Engineering* 122(2): 105–113. [http://doi.org/10.1061/\(ASCE\)0733-947X\(1996\)122:2\(105\)](http://doi.org/10.1061/(ASCE)0733-947X(1996)122:2(105))
- Raftery, A. E. 1995. Bayesian model selection in social research, *Sociological Methodology* 25: 111–163. <http://doi.org/10.2307/271063>
- RMP. 2011. *Statistical Report on Road Accidents in Malaysia 2010*. Royal Malaysia Police (RMP). Available from Internet: <http://www.rmp.gov.my>
- Schneider, W.; Savolainen, P.; Moore, D. 2010. Effects of horizontal curvature on single-vehicle motorcycle crashes along rural two-lane highways, *Transportation Research Record: Journal of the Transportation Research Board* 2194: 91–98. <http://doi.org/10.3141/2194-11>
- Son, H. D.; Kweon, Y.-J.; Park, B. B. 2011. Development of crash prediction models with individual vehicular data, *Transportation Research Part C: Emerging Technologies* 19(6): 1353–1363. <http://doi.org/10.1016/j.trc.2011.03.002>
- Spainhour, L.; Mishra, A. 2008. Analysis of fatal run-off-the-road crashes involving overcorrection, *Transportation Research Record: Journal of the Transportation Research Board* 2069: 1–8. <http://doi.org/10.3141/2069-01>
- Thomson, R.; Fagerlind, H.; Martínez, A. V.; Amengual, A.; Naing, C. L.; Hill, J.; Hoschopf, H.; Dupre, G.; Bisson, O.; Kelkka, M.; Van der Horst, R.; Garcia, J. M. 2006. *Roadside Infrastructure for Safer European Roads: D06 European Best Practice for Roadside Design: Guidelines for Roadside Infrastructure on New and Existing Roads*. Released by Chalmers University of Technology on behalf of the RISER Consortium. 148 p.
- Usman, T.; Fu, L.; Miranda-Moreno, L. F. 2010. Quantifying safety benefit of winter road maintenance: Accident frequency modeling, *Accident Analysis & Prevention* 42(6): 1878–1887. <http://doi.org/10.1016/j.aap.2010.05.008>
- Viner, J. G. 1995. Rollovers on sideslopes and ditches, *Accident Analysis & Prevention* 27(4): 483–491. [http://doi.org/10.1016/0001-4575\(95\)00004-J](http://doi.org/10.1016/0001-4575(95)00004-J)
- Vuong, Q. H. 1989. Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica* 57(2): 307–333. <http://doi.org/10.2307/1912557>
- WHO. 2009. *Global Status Report on Road Safety: Time for Action*. World Health Organization (WHO). 301 p.
- Woodrooffe, J.; Blower, D.; Green, P. 2011. Prevention of tractor-semitrailer rollovers: effectiveness of electronic stability control systems, *Transportation Research Record: Journal of the Transportation Research Board* 2265: 27–33. <http://doi.org/10.3141/2265-03>
- Yim, S.; Jeon, K.; Yi, K. 2012. An investigation into vehicle rollover prevention by coordinated control of active anti-roll bar and electronic stability program, *International Journal of Control, Automation and Systems* 10(2): 275–287. <http://doi.org/10.1007/s12555-012-0208-9>