Australasian Transport Research Forum 2021 Proceedings
8-10 December, Brisbane, Australia
Publication website: http://www.atrf.info

Analysis of Head-on Heavy Vehicle Crashes in Queensland using Correlated Random Parameters with Heterogeneity in Means and Lindley

Krishna N.S. Behara[1], Yasir Ali[1], Alexander Paz[1], Owen Arndt[2], Douglas Baker[3]

[1] School of Civil & Environmental Engineering, Faculty of Engineering, Queensland University of Technology, Brisbane, Australia

[2]Queensland Department of Transport and Main Roads, Brisbane, Australia

[3] School of Architecture & Built Environment, Faculty of Engineering, Queensland University of Technology, Brisbane, Australia
Email for correspondence: alexander.paz@qut.edu.au

1.Introduction

The number of fatal crashes involving heavy vehicles (HVs) has dramatically increased world-wide over the last decade. In the United States, fatalities due to collisions involving large trucks increased by 41% in 2017 compared to 2009 (National Center for Statistics and Analysis 2019). The percentage increase in deaths from 2015 to 2016 was around 9.6% in the European Union (European Commission 2018). In Australia, 188 people died from 173 fatal crashes involving HVs in 2019; the increase was approximately 27.2% compared to 2018 (BITRE 2019). Due to Queensland's vast geographical size, drivers of HVs travel large distances, thereby increasing their chances of fatigue and isolation (Raftery et al. 2011). This becomes more critical as the number of HVs on Queensland's road network increase (TMR 2019). For instance, the number of HVs registered in Queensland increased by 0.5% in 2020 compared to 2019 (TMR 2020a). Statistics from 2011 revealed that the percentage of fatal and serious injury crashes involving HVs in Australia were nearly 18% while they were only 3% of the total registered vehicles (Assemi and Hickman 2018).

The collisions involving HVs, especially large trucks travelling on undivided highways, often result in injury severities relatively higher than Rear-end and collisions at signalised Intersections (Zhu and Srinivasan, 2011). The Queensland Department of Transport and Main Roads (TMR) has established a five-stage heavy vehicle study to analyze HV crashes on its state-controlled road network between 2012 and 2016 (Paz et al., 2020). Their study reported that the fatality rate for Head-on collisions was 35% compared to other types. Very few studies (e.g., Sassi et al. (2018)) in the literature have exclusively focused on Head-on HV crashes.

Recent literature on road safety research has focussed on variants of random parameters models to capture unobserved heterogeneity that may influence the occurrence of crashes. For instance, our previous study using random parameters with heterogeneity in means and a Lindley outperformed multiple alternative state-of-the-art specifications in terms of fit as well as prediction ability (Behara et al., 2021). However, that study did not investigate the presence of correlation across the random parameters. Therefore, the main objective of the current study is to investigate the correlation across random parameters with heterogeneity in means as well as address the issue related to excessive zeroes in the crash data using Lindley distribution.

2. Data

The data used in this study includes information regarding Head-on Fatal and Serious Injury (FSI) crashes over 1,848 homogeneous segments on the Queensland state road network. The crash data involved Multi-Combination Vehicles (MCVs) which were heavy freight trucks including two or more trailers with a minimum length of 25 meters (Haldane and Bunker, 2010). The percentage share of annual average daily traffic (AADT) by vehicle category averaged over all homogeneous road segments was as follows: 87% were cars, 8% were trucks and buses, 3% were articulated vehicles, and 2% were HVs or MCVs. The crash data for the road segments were from an analysis period of 10 years (97%), 5 to 9 years (2%), and 1 to 4 years (1%). The number of road segments was 88 (5%) and 1,760 (95%) with and without FSI crashes, respectively. Most road segments were rural roads (76%) and level terrain (97%). Among the 88 road segments with FSI crashes, the majority were rural (90%), have a high-speed limit (89%), level (98%) and rolling terrains (76%). Single carriageway rural roads and single carriageway urban roads witnessed a greater number of crashes compared to other types as shown in Figure 1. The data used in this study was provided by TMR, Australia.

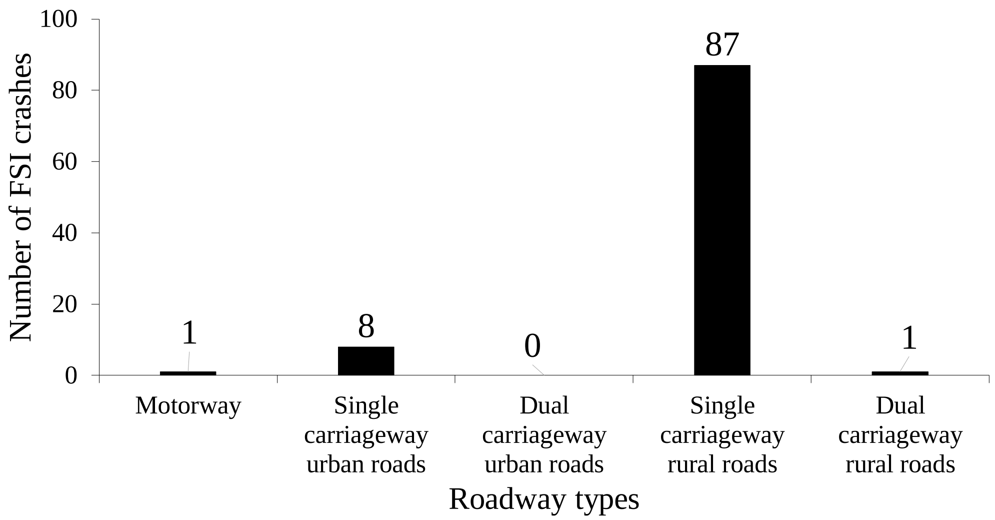

Figure 1 Statistics of FSI crashes on Queensland state road network

Table 1 and **Table 2**, respectively, list key continuous and categorical variables along with descriptive statistics for the crash data used in this study.

Table 1: Key Continuous Variables and Descriptive Statistics

Continuous variables	Minimum	Maximum	Mean	Std. dev.
AADT (vehicles per day) per road segment	16	133040	4591	10872
Average MCV traffic (vehicles per day) per road segment	1	1732	98	169
Length of road segment gazetted to MCV traffic (km)	0.19	10	6	4
Number of lanes per road segment	1	8	2.1	0.7
Total formation width (m)	5	38	9.7	3.3

Table 2: Key Categorical Variables and Descriptive Statistics

Categorical variables	Count	Percentage
Motorway	30	2%
Urban single carriageway	315	17%
Urban dual carriageway	100	5%
Rural single carriageway	1373	74%
Rural dual carriageway	30	2%
Presence of shoulder	1831	99%
Presence of median	238	13%
High-speed limit (>=100 kph)	1327	72%
Medium speed limit (>=50 and <100 km/h)	518	28%
Low speed limit (<50 km/h)	3	0.2%
Level terrain	1785	97%
Rolling terrain	835	45%
Mountainous terrain	156	8%
Curve (moderate/sharp/very sharp) longer than 50% of the corresponding segment length	222	12%
Wide centerline >50% of road length	79	4%
Presence of Audio Tactile Lane Marking (ATLM)	369	20%

3. Methodology

A step-by-step methodological approach presented by Arndt (2004) is adopted to analyse Head-on FSI and Fatal crashes on MCV gazetted roads as follows:

1. Initial selection of the variables to consider for analysis. In this study, 20 variables were initially considered. Length of road segment and period of analysis were included as offsets in the model.
2. Suitable relationships between response and explanatory variables were identified.
3. Suitable advanced regression models were selected.

4. Statistical check to obtain a statistically significant and logical result was performed.
5. A cross-validation technique was adopted to check consistency of significant variables as well as the predictive ability of crash models.

Model specifications: This study compares random parameters, and random parameters with heterogeneity in means, and Lindley with their correlated extensions.

Model estimation: Given that Lindley extensions to random parameters and random parameters with heterogeneity in means models are hierarchical, Bayesian inference with MCMC simulation was adopted in this study to estimate all models using WINBUGS tool.

Model performance measures: Average log-likelihood (average deviance) and Deviance Information Criteria (DIC) are used as goodness of fit measures. The Mean Absolute Deviation (MAD) and Mean Squared Prediction Error (MSPE) are used as predictive performance measures.

Comparative analysis*:* Percentage improvement in the Deviance, MAD and MSPE for the best model is compared to other models.

4. Preliminary findings

Our preliminary results using random parameters with heterogeneity in means and a Lindley outperformed multiple alternative state-of-the-art specifications in terms of fit as well as prediction ability as shown in the Table 3 (Behara et al., 2021). The correlated random parameters with heterogeneity in means and its Lindley extension is the prime focus of this study. Results shown in Table 3 for correlated random parameters with heterogeneity in means showed higher improvement compared to random parameters with heterogeneity in means. The on-going study investigates if correlated random parameters with heterogeneity in means and Lindley outperforms random parameters with heterogeneity in means and Lindley.

The factors likely to affect the likelihood of Head-on Fatal and Serious Injury crashes involving HVs in Queensland including volume, segment length, period of analysis, rural single carriageway with high (>=100 kph) and medium (>=50 and <100 kph) speed limits, urban single carriageway, and curve longer than 50% of the corresponding road length. Unobserved heterogeneity regarding curve longer than 50% of the corresponding segment length was explained using additional geometric characteristic including rolling terrain. These factors have been consistent with our previous research (Behara et al., 2021).

5. Conclusion

The purpose of this study is to perform statistical analyses of Head-on collisions involving Multi-Combination Vehicles (MCVs) to identify likely crash contributing factors and investigate the presence of correlation across random parameters. This study intends to compare several advanced models for crash analysis using Head-on crash data from Queensland. Recent literature on crash frequency analysis has focussed on variants of random parameters models to capture unobserved heterogeneity. Our previous study also revealed that random parameters with heterogeneity in means and Lindley out-performed some state-of-art specifications. However, that study did not investigate the presence of correlation across random parameters. Therefore, this study addresses this potential limitation and intends to develop correlated random parameters with heterogeneity in means and a Lindley distribution. The purpose of this advanced model is to capture unobserved heterogeneity using additional variables, capture correlation across the random parameters, and account for site-specific variation from excessive zero crash observations. Findings from this study can be used by traffic safety practitioners to select appropriate methods for crash analysis and highway safety evaluation. Advanced models are recommended when reliable data is available. This is because these models offer better statistical fit, greater explanatory power, and better estimation.

References

ARNDT, O. K. 2004. *Relationship between unsignalised intersection geometry and accident rates.* Doctoral Dissertation, Queensland University of Technology, Brisbane, Australia.

BEHARA, K. N. S., PAZ, A., ARNDT, O. & BAKER, D. 2021. A random parameters with heterogeneity in means and Lindley approach to analyze crash data with excessive zeros: A case study of head-on heavy vehicle crashes in Queensland. *Accident Analysis and Prevention,* 106308.

HALDANE, M. & BUNKER, J. 2010. Assessing the Impacts of Multi-Combination Vehicles on Traffic Operations and Safety: A Literature Review. Queensland Department of Transport and Main Roads, and Queensalnd University of Technology.

PAZ, A., BEHARA, K. N. S. & HAQUE, M. 2020. Analysis of Heavy Vehicle Crashes on Queensland's State-controlled Road Network - Stage 4. Report submitted to Queensland Department of Transport and Main Roads (TMR), Queensland.

SASSI, S., HAKKO, H., RÄTY, E. & RIIPINEN, P. 2018. Light motor vehicle collisions with heavy vehicles—Psychosocial and health related risk factors of drivers being at-fault for collisions. *Forensic Science International,* 291**,** 245-252.

ZHU, X. & SRINIVASAN, S. 2011. A comprehensive analysis of factors influencing the injury severity of large-truck crashes. *Accident Analysis and Prevention,* 43**,** 49-57.

Table 3 Analysis of FSI Crashes using Bayesian Inference and MCMC Simulation

Response variable: FSI Crashes	**Random Parameters Poisson**			**Random Parameters Poisson with Heterogeneity in Means**			**Random Parameters Poisson-Lindley**			**Random Parameters Poisson with Heterogeneity in Means - Lindley**			**Correlated Random Parameters Poisson with Heterogeneity in Means**		
Explanatory Variables	**Mean**	**Std. Dev**	**MC error**	**Mean**	**Std. Dev**	**MC error**	**Mean**	**Std. Dev**	**MC error**	**Mean**	**Std. Dev**	**MC error**	**Mean**	**Std. Dev**	**MC error**
Constant	-13.08	1.22	0.05	-13.26	1.37	0.05	-13.24	1.58	0.06	-13.20	1.23	0.05	-13.52	1.14	0.04
Logarithmic of AADT	0.16	0.13	0.05	0.21	0.09	0.04	0.07	0.11	0.04	0.31	0.13	0.05	0.27	0.12	0.00
Logarithmic of AADT (Std. Dev.)	0.03	0.03	0.01	0.03	0.02	0.01	0.04	0.03	0.01	0.02	0.02	0.01	0.53	0.33	0.01
Logarithmic of MCV volume	0.98	0.15	0.05	0.93	0.12	0.04	1.04	0.13	0.05	0.89	0.14	0.05	0.90	0.14	0.00
Urban single carriageway	2.32	0.98	0.04	2.28	0.92	0.03	1.76	0.76	0.03	1.78	0.63	0.02	2.15	0.81	0.02
Rural single carriageway with high-speed limit	3.58	0.91	0.04	3.62	0.90	0.04	2.99	0.71	0.03	3.20	0.51	0.02	3.54	0.76	0.02
Rural single carriageway with high-speed limit (Std. Dev.)	0.31	0.21	0.01	0.24	0.19	0.01	0.19	0.17	0.01	0.14	0.13	0.00	6.10	3.52	0.13
Rural single carriageway with medium-speed limit	2.37	1.24	0.04	2.36	1.18	0.03	1.80	1.07	0.03	1.81	1.00	0.02	2.22	1.10	0.02
Curve longer than 50% of the segment length	-0.65	0.53	0.02	-0.51	0.62	0.02	-0.64	0.48	0.02	-0.56	0.62	0.02	-0.58	0.62	0.01
Curve longer than 50% of the segment length (Std. Dev.)	0.70	0.69	0.02	0.71	0.70	0.02	0.66	0.70	0.02	0.67	0.68	0.02	5.13	3.56	0.13
	Heterogeneity in Means of Random Parameters														
Curve longer than 50% of the segment length: Roll Terrain	-	-	-	-0.49	0.89	0.02	-	-	-	-0.47	0.93	0.02	-0.47	0.90	0.01
Average Log Likelihood	-274.50			-276.00			-262.74			-253.56			-273		
DIC	594.73			594.75			586.70			550.03			545.99		
MAD	0.084	0.01	0.00	0.085	0.01	0.00	0.082	0.01	0.00	0.079	0.01	0.00	0.084	0.007	0.000
MSPE	0.106	0.01	0.00	0.107	0.01	0.00	0.104	0.01	0.00	0.103	0.01	0.00	0.106	0.012	0.000
theta	-	-	-	-	-	-	0.90	0.77	0.03	1.55	1.21	0.05	-	-	-