

Article

A Comprehensive Analysis of Multi-Vehicle Crashes on Expressways: A Double Hurdle Approach

Jungyeol Hong , Reuben Tamakloe  and Dongjoo Park *

Department of Transportation Engineering, University of Seoul, Seoul 02504, Korea; jyhong9868@uos.ac.kr (J.H.); uosgrad02022@uos.ac.kr (R.T.)

* Correspondence: djpark@uos.ac.kr; Tel.: +82-2-6490-2825

Received: 4 March 2019; Accepted: 7 May 2019; Published: 15 May 2019



Abstract: To maintain safe expressways, it is necessary to investigate the causes of severe traffic accidents and establish a strategy. This study aims to analyze crashes and identify the influence of crash-risk factors on multi-vehicle (MV) crashes. Crashes involving three types of vehicles namely passenger cars, buses, and freight trucks were analyzed using a seven-year data spanning 2011 to 2017 which consists of crashes that occurred on expressways in South Korea. We applied a double hurdle approach in which a model consists of two estimators: The first estimation, which is a binary logit model selects MV crashes from the dataset; and the second estimation which is a truncated regression model estimates the number of vehicles involved in the MV crash. We found that driver traffic violations such as the improper distance between vehicles, reversing and passing increases the probability of MV crashes occurring. MV crashes in tunnels and mainlines were found to be positively correlated with the number of vehicles involved in the crash, whereas fewer vehicles were involved in MV crashes at ramps and toll-booths. Further, we found that the hurdle model with an exponential form of conditional mean of the latent variable provides better estimation parameters.

Keywords: multi-vehicle crashes; double hurdle; risk factor; exponential double hurdle; expressway

1. Introduction

The annual death rate (per 100,000 population) caused by road crashes declined from 25.3 in 2000 to 10.1 in 2016 [1]. Nevertheless, the impacts of these crashes cannot be overlooked since it is still one of the leading causes of socio-economic and logistics cost loss. A total of 66,592 traffic accidents occurred on the Korean expressway between 2011 and 2017, of which 17,873 were multi-vehicle (MV) crashes, accounting for 27% of all crashes. Traffic accidents involving three or more vehicles accounted for about 8% of all crashes, with 5224 accidents, and 92 crashes involving more than 10 vehicles [2]. Even if the ratio of multi-vehicle MV involved crashes is less than single-vehicle (SV) crashes, the MV crashes occurring on the expressway have more victims compared to SV crashes. In addition, the damage to vehicles and road structures is more severe in MV crashes. As a result, the interest in finding the relationship between crash risk factors and the number of vehicles involved in a crash has increased due to the critical role this knowledge plays in reducing vehicular crashes.

In this study, we examine the characteristics of MV crashes involving buses, freight trucks, and passenger cars after developing models for each vehicle type by applying Cragg's double hurdle regression model. This model first determines the probability that a crash will occur given that it is an MV crash and then models the relationship between crash risk factors and the number of vehicles involved in a crash based on the MV crashes. In the context of this study, a crash that involves only one car is referred to as an SV crash whereas a crash involving two or more cars is an MV crash. From the first stage of the modeling process, we can identify the variables that have high probabilities in causing MV crashes. The results of the second stage can help us know the effect each variable has in

increasing the number of vehicles involved in a crash. We can also identify the variables that cause more vehicles to pile up in a crash.

This study continues by literature reviews and data descriptions used in the research. The methodology used in developing the models in this research would be discussed in that order. Then, the results would be presented and discussed accordingly. Lastly, conclusions and recommendations would be made.

2. Literature Reviews

In terms of methodological analysis for the MV and SV involved crashes, many researchers have applied various statistical approaches with the aim of reducing the bias in model outputs. Research has established that different crash risk factors have different impacts on SV and MV crashes.

Islam, Jones & Dye [3] conducted comprehensive research on both SV and MV large truck-at-fault crashes in Alabama by developing four separate random parameter logit models. Data of 8328 crashes, which occurred in rural and urban environments, was obtained from a police-reported crash database for use in their study. It contained information such as driver, vehicle, temporal, roadway, crash and land use characteristics. The models were area-specific; SV-rural, MV-rural, SV-urban, and MV-urban. The study results showed that different risk factors had different influences on both SV and MV crashes.

Addressing the differences between the effect of crash impact factors on SV and MV crashes, Hassan et al. [4] studied crash data in the Emirate of Abu Dhabi and noted that factors contributing to SV crashes were quite different from those that affect MV crashes. They also concluded that Emirati drivers had a higher chance of being involved in SV crashes whereas Asian drivers were more often engaged in MV crashes. Wu, Chen, Zhang, Liu, Wang & Bogus [5] also iterated the existence of significant differences in crash risk factors when determining the injury severities of SV and MV crashes. Other researchers provided a brief understanding of this subject matter [6–9].

Ivan, Pasupathy, & Ossenbruggen [10] modeled SV and MV crashes using site-specific variables together with other explanatory variables by employing a Poisson regression model such that each resulting model had its own explanatory variables. They related the decrease in SV crashes to the increase in factors such as shoulder width and sight distance. The rise in MV crashes was attributed to the increase in shoulder width. Ivan, Wang, & Bernardo [11] also used Poisson regression models to investigate crash rates for SV, and MV expressway crashes. Driveway variables and time of day were used to explain crash rates, and it was evident that the effects of variables used differed for SV and MV crashes. SV crashes occurred mostly in the evening and at nighttime due to drowsy driving, whereas MV crashes occur mostly under daylight conditions. The authors concluded that the types of trips made, and the level of alertness are both correlated with the time of day.

It has been established that the mechanisms of both SV and MV vary, hence, some risk factors affect the probability of having either type of crashes differently [12,13]. For example, SV crashes are likely to occur when the driver loses control of his vehicle. On the other hand, research shows that MV crashes are also likely to occur when drivers interact improperly with other vehicles on the roadway. Bowen dong [14] showed that most roadway specific variables were significant in both SV and MV models, meanwhile some causes of crashes were random parameters. They identified that the random parameters affected both types of crashes differently. The probability of having SV crashes was high when the roadway surfaces are wet. On the other hand, MV crash probability increases when there is a chemical wet road. In another similar study [15], young drivers were identified to be associated with high chance of having SV crashes. In addition, it was established that a positive correlation existed between vehicle density and MV crashes.

In a study of SV and MV vehicle crashes in Iowa involving heavy trucks, models for SV and MV crash severity were developed using a binary probit and a nested logit model, respectively. More variables were found to be significant in the MV crash severity model as compared to the SV crash severity model [16]. Comparing both models, older drivers had a higher probability of getting more severe injuries in both SV and MV crash models. Meanwhile, the likelihood of increased injury severity

is more elevated in SV crashes when the crash involves a single unit truck, whereas in the case of an MV crash, the probability of a severe crash increases when a combined truck is involved.

SV and MV crashes were studied using a multinomial logistic regression model to determine the different impacts of crash risk factors on the type of crash. It was identified that the risk of MV crashes was higher on weekends whereas the risk of SV crashes on divided and undivided roads was higher [17]. Martensen & Dupont [18] found that the probability of trucks getting involved in SV crashes were low as compared to passenger cars. Moreover, the probability of impaired drivers with more passengers getting involved in SV crashes is higher compared to other drivers.

A further exploratory analysis was conducted by Lord et al. [19] using five-year data from urban and rural areas. This data contained information such as crash severity, the day of the week, location, time of day of the crash, and crash type among others. Three Poisson-gamma related safety performance functions (SPF) for urban and rural freeway segments were developed for different crash types and severities, and models for SV and MV crashes were presented for both rural and urban sections. More analysis was conducted to show that single predictive models which combine all crash types were not adequate for predicting crashes on freeway segments as it does not clearly show the effect of impact factors on each type of crash. Hence, they recommend that two distinct SPF's be developed for both SV and MV crashes where appropriate in order to provide better performing models which would produce more accurate results to describe a roadway facility.

Geedipally & Lord [20] sought to investigate the claim made by Lord et al. [19]. They developed Poisson-gamma models for SV, MV and all crashes (SV + MV) using crash data from a four-lane undivided expressway segment in Texas. The data consisted of information such as the number of horizontal curves, the severity of a crash, the number of vehicles that are involved in a crash, and others that affect the crash rate. They compared the models with the aim of finding an appropriate way of identifying hot spots on expressway facilities. A place that has a crash frequency slightly higher than expected is referred to as a hot spot. They determined that slightly fewer false positives and negatives were predicted when SV and MV crashes are modeled separately as compared to modeling them together. Since modeling these crash types separately improved efficiency and the prediction capabilities of models, they recommended that separate models be made for both SV and MV crashes.

Advancement in research has led to an increase in ideas and methodological strategies. Bayesian hierarchical modeling approaches were used to develop SV and MV SPF models. The authors showed that the method allows for the proper estimation of results when dealing with multilevel data structures [21,22]. In addition to the methodological approaches used so far, few researchers considered using hurdle models. The use of these models is essential when an individual is faced with a sequential decision-making process and to handle excess zeros present in data, which is very particular of crash data. This model has two parts; a first part that is modeled using a binary logit or probit distribution and, conditional on the positive counts, a second part is modeled with a left truncated count distribution if the hurdle is crossed. This model combines all crashes in the first stage of analysis. Ma, Yan & Weng [23] examined crash rates using lognormal hurdle models and compared their results with a Tobit model. From the AIC and BIC values, the hurdle model was found to be more superior since it fits the crash data more accurately. They also developed and compared the performance of the proposed hurdle model to Poisson and negative binomial models. The hurdle models performed better than the count models since they only provided the expected number of vehicles involved in crashes. Due to its superiority, researchers in both non-transportation [24,25] and transportation fields applied hurdle models in their studies. In the field of transportation, Boucher & Santolino [26] used a Negative Binomial, Zero-Inflated Negative Binomial and Hurdle-Negative Binomial regression to model the disability severity score of crash victims. The AIC and BIC estimates gave the hurdle-negative binomial a more significant advantage over the other models. It also produced the best statistically fit results upon comparing the results of the Vuong tests. They concluded that hurdle models were more appropriate in explaining a data generation process.

Hosseinpour et al. [27] and Hosseinpour et al. [28] also used this model framework to investigate crashes along 543 km and 448 km of Malaysian federal roads, respectively. The superiority of the hurdle model's modeling performance surpassed others. In addition to the many advantages in using this model formulation, Ma et al. [29] identified that regarding accommodating mixed skew data, the hurdle model is very flexible due to the nature of its framework. In their study, the log-normal hurdle model produced better estimates compared to the Tobit model and the random-parameters Tobit model.

Many studies have been conducted for studying the effects of crash risk factors on the type of crashes, that is, either SV or MV crashes. In a quest to find appropriate models for such investigations, researchers applied hurdle models in the field of transportation safety analysis due to its model structure flexibility. In their studies, it was evident that the hurdle model is more superior as compared with the standard model distributions used over the years. One main advantage of this model is its ability to separate one crash type from another, and then using the information obtained from the previous stage to develop a count model in the next stage provided the hurdle is crossed. Again, this analysis can be done using the same or different variables.

This study seeks to develop models to analyze risk factors and the characteristics of MV crashes on expressways. Since vehicles have different features, which influences the resulting type and number of vehicles involved in crashes [30,31], we developed separate models for bus, passenger car, and freight truck-involved crashes. The advantage of this is that it helps us distinguish the effects of MV crash features for each type of vehicles in detail.

3. Data Description

To analyze multi-vehicle crashes, this study used crash records from the entire expressway network in South Korea shown in Figure 1. As of 2017, a total of 38 routes had been constructed, with a total length of approximately 4746 km. The speed limit on most routes on the expressway ranges from 100 km/h~110 km/h, and it operates as a toll expressway. The route with the highest annual average daily traffic volume (AADT) is the Gyeongbu Expressway, which connects Seoul, the capital city of South Korea, to Busan in the southern part of the country. Table 1 summarizes 38 routes' characteristics of Korean highways.



Figure 1. Expressway routes in South Korea. Source: <http://map.ngii.go.kr/ms/map/NlipMap.do#>.

Table 1. Design features of the Korean expressway.

Expressway Route	Length (km)	Speed Limit (km/h)	Number of Tunnels	AADT (vehs/day)	Number of Tolls
Gyeongbu Route	416.05	100–110	26	1,335,770	37
Namhae Route	273.1	100	89	443,825	32
Muan-Gwangju Route	223.2	100	75	84,362	17
Seohaean Route	340.8	100–110	36	442,949	27
Ulsan Route	14.3	100	-	67,600	2
Iksan-Pohang Route	130.3	100	42	116,738	9
Nonsan Cheonan Route	82	110	2	307,683	9
Honam Route	194.2	100	14	366,838	19
Suncheon-Wanju Route	117.8	100	76	47,417	9
Dangjin-Yeongdeok Route	278.6	100–110	115	241,732	22
Tongyeong-Daejeon Route	332.5	100	52	436,399	29
Jungbu Route	117.2	110	5	384,402	13
Pyeongtaek-Jecheon Route	109.4	100	41	237,380	10
Jungbunaeryuk Route	301.7	100–110	70	295,640	23
Yeongdong Route	234.4	100	40	594,575	23
Gwangju-Wonju Route	59.65	100	12	35,634	7
Jungang Route	370.8	110	59	456,886	30
Seoul-Yangyang Route	150.21	100	125	37,906	12
Donghae Route	223.06	100	96	117,177	25
Seouloegwaksunhwan Route	128	100	13	900,959	13
Namhae 1jiseon Route	17.9	100	-	74,268	2
Namhae 2jiseon Route	20.6	90	-	132,535	3
Namhae 3jiseon Route	15.26	100	5	5957	3
2nd Gyeongin Route	70.02	100	2	78,783	2
Gyeongin Route	23.9	100	-	152,192	1
Incheon International Airport Route	36.6	80–100	1	66,200	3
Seocheon-Gongju Route	61.4	100–110	10	40,029	5
Pyeongtaek-Siheung Route	42.6	100	-	54,201	5
Yongin-Seoul Route	22.9	100	-	1953	4
Pyeongtaek-Hwaseong Route	54	100	-	245,646	4
Honam jiseon Route	42.5	100	-	64,235	3
Gochang-Damyang Route	13.3	100	14	63,009	2
Sangju-Youncheon Route	93.96	100	8	17,946	12
Bongdam-Dongtan Route	46.6	100	-	52,101	9
Jungbunaeryuk Jiseon Route	30	100	-	117,860	5
Jungang Jiseon Route	8.2	100	5	218,821	2
Busanoegwaksunhwan Route	48.8	100	6	239	2

The raw crash data were obtained from the Korea Expressway Corporation (KEC). Crash risk factors such as roadway geometric design features, weather condition, collision information such as crash severity level, vehicle malfunction, and driver's violations, spatiotemporal characteristics, seasonal and traffic volume information among others covering 38 expressway routes in South Korea were extracted from the raw crash data.

A total of 3481 bus crashes, 16,093 freight truck crashes, and 39,837 passenger car crashes were observed between 2011 and 2017. Information such as temporal, seasonal, roadway, vehicle, weather, and driver characteristics were all presented as dummy variables taking on only zeros if they satisfy the condition that a crash did not occur and ones if a crash occurs. Specific information about the nature of traffic and the crash were presented as count data. Table 2 provides the summary statistics of the data used to develop the models. The average number of vehicles involved in bus crashes is 1.523; that of truck and passenger car crashes are 1.479 vehicles and 1.409 vehicles, respectively.

In total, the number of zeros in this observed crash data is very high since SV crashes take a value of 0, and 1 for MV involved crashes. As per our definition of SV and MV crashes, the number of zeros reflected by the positively skewed distribution in Figure 2 tells us that many SV crashes exist in the data.

Table 2. Descriptive Statistics for bus, freight, and passenger car involved crash data.

Variables	Bus-Involved Crash Data ²		Truck-Involved Crash Data ³		Passenger Car-Involved Crash Data	
	Mean	St.d	Mean	St.d	Mean	St.d
Month of year						
January	0.088	0.283	0.069	0.254	0.094	0.292
February	0.064	0.244	0.062	0.242	0.078	0.268
March	0.065	0.246	0.072	0.259	0.073	0.260
April	0.090	0.286	0.083	0.276	0.081	0.273
May	0.077	0.267	0.086	0.280	0.083	0.275
June	0.082	0.274	0.091	0.287	0.082	0.275
July	0.099	0.298	0.104	0.305	0.105	0.306
August	0.102	0.303	0.096	0.295	0.100	0.300
September	0.079	0.270	0.090	0.286	0.081	0.273
October	0.097	0.295	0.087	0.282	0.075	0.263
November	0.082	0.274	0.080	0.272	0.074	0.261
December	0.076	0.265	0.079	0.270	0.076	0.264
Day of week						
Monday	0.144	0.351	0.163	0.369	0.143	0.350
Tuesday	0.134	0.341	0.174	0.379	0.121	0.326
Wednesday	0.138	0.345	0.156	0.363	0.126	0.331
Thursday	0.135	0.341	0.157	0.364	0.121	0.326
Friday	0.138	0.345	0.165	0.371	0.142	0.349
Saturday	0.164	0.370	0.120	0.325	0.175	0.380
Sunday	0.147	0.355	0.064	0.245	0.173	0.379
Weekday	0.689	0.463	0.816	0.388	0.652	0.476
Weekend	0.311	0.463	0.184	0.388	0.348	0.476
Time of day						
0 a.m.–3 a.m.	0.059	0.235	0.073	0.261	0.097	0.296
3 a.m.–6 a.m.	0.058	0.234	0.098	0.298	0.079	0.270
6 a.m.–9 a.m.	0.138	0.345	0.135	0.341	0.136	0.343
9 a.m.–12 p.m.	0.173	0.378	0.170	0.375	0.144	0.351
12 p.m.–3 p.m.	0.161	0.367	0.180	0.385	0.148	0.355
3 p.m.–6 p.m.	0.194	0.396	0.170	0.376	0.161	0.368
6 p.m.–9 p.m.	0.129	0.335	0.098	0.297	0.123	0.329
9 p.m.–12 a.m.	0.088	0.284	0.075	0.264	0.112	0.315
Location						
Mainline	0.715	0.452	0.630	0.483	0.735	0.442
Ramp	0.117	0.321	0.143	0.350	0.151	0.358
Toll booth	0.055	0.228	0.127	0.333	0.033	0.179
Rest area	0.011	0.105	0.015	0.122	0.013	0.113
Electronic toll	0.057	0.231	0.040	0.196	0.022	0.148
Tunnel	0.043	0.202	0.041	0.199	0.042	0.200
Others	0.003	0.054	0.005	0.068	0.004	0.062
Severity level ¹						
A	0.004	0.063	0.001	0.036	0.000	0.017
B	0.047	0.211	0.040	0.197	0.019	0.136
C	0.254	0.435	0.280	0.449	0.201	0.401
D	0.695	0.461	0.679	0.467	0.780	0.414
Reason for crash occurrence						
Over speeding	0.225	0.418	0.193	0.395	0.275	0.446
Improper safety gap	0.032	0.175	0.026	0.159	0.022	0.146
Improper passing	0.022	0.147	0.013	0.114	0.023	0.149
Improper reversing	0.002	0.045	0.001	0.036	0.002	0.044
Drowsy driving	0.104	0.305	0.170	0.376	0.118	0.322
Negligence	0.246	0.431	0.243	0.429	0.202	0.402
Other driving violence	0.081	0.273	0.081	0.273	0.105	0.307
Loads from other vehicles	0.001	0.024	0.032	0.177	0.001	0.023
Fire	0.036	0.186	0.040	0.196	0.014	0.118

Table 2. Cont.

Variables	Bus-Involved Crash Data ²		Truck-Involved Crash Data ³		Passenger Car-Involved Crash Data	
	Mean	St.d	Mean	St.d	Mean	St.d
Brake malfunction	0.010	0.098	0.019	0.138	0.003	0.055
Tire burst	0.073	0.261	0.066	0.249	0.028	0.164
Vehicle defects	0.003	0.059	0.009	0.095	0.001	0.029
Roadway problem	0.011	0.104	0.006	0.077	0.019	0.137
Obstacle on the road	0.097	0.297	0.047	0.211	0.124	0.330
Pedestrian	0.005	0.072	0.002	0.041	0.002	0.042
Animal	0.008	0.088	0.003	0.053	0.017	0.129
Pothole	0.005	0.074	0.002	0.040	0.013	0.112
Others	0.039	0.182	0.046	0.162	0.032	0.189
Weather						
Sunny	0.600	0.490	0.644	0.479	0.598	0.490
Snowy	0.045	0.208	0.028	0.166	0.035	0.184
Rainy	0.216	0.412	0.187	0.390	0.216	0.412
Foggy	0.005	0.070	0.004	0.066	0.004	0.061
Cloudy	0.133	0.340	0.136	0.343	0.146	0.353
Windy	0.001	0.024	0.001	0.024	0.001	0.035
Number of vehicles involved crashes *	1.523	0.994	1.479	1.000	1.409	1.034
Horizontal alignment						
Straight	0.780	0.414	0.778	0.416	0.721	0.449
Curve length > 1000 m	0.212	0.409	0.215	0.411	0.269	0.443
Curve length > 500 m & ≤ 1000 m	0.003	0.056	0.004	0.059	0.005	0.072
Curve length > 100 m & ≤ 500 m	0.005	0.068	0.004	0.061	0.005	0.072
Vertical alignment						
No slope	0.619	0.486	0.650	0.477	0.586	0.493
Downward <1%	0.062	0.242	0.050	0.219	0.063	0.243
Downward 1–3%	0.094	0.292	0.093	0.291	0.119	0.324
Downward over 3%	0.033	0.180	0.034	0.182	0.043	0.203
Upward <1%	0.046	0.210	0.048	0.213	0.054	0.225
Upward 1–3%	0.097	0.297	0.084	0.277	0.095	0.293
Upward over 3%	0.047	0.196	0.041	0.174	0.040	0.176
Shoulder type						
Rock	0.016	0.125	0.019	0.137	0.015	0.122
Guardrail	0.434	0.496	0.398	0.489	0.490	0.500
Cable	0.000	0.017	0.000	0.021	0.001	0.030
Fence	0.006	0.079	0.008	0.090	0.008	0.091
Pipeline	0.001	0.038	0.002	0.047	0.003	0.054
Concrete	0.096	0.294	0.094	0.292	0.089	0.285
Others	0.093	0.290	0.102	0.302	0.090	0.286
No shoulder	0.354	0.478	0.377	0.485	0.304	0.460
Gender						
Male	0.920	0.272	0.981	0.137	0.797	0.403
Female	0.080	0.272	0.019	0.137	0.203	0.403
Age group						
≤20 years old	0.256	0.437	0.221	0.415	0.247	0.431
21~30 years old	0.055	0.228	0.038	0.190	0.150	0.357
31~40 years old	0.131	0.337	0.125	0.331	0.203	0.402
41~50 years old	0.208	0.406	0.240	0.427	0.191	0.393
51~60 years old	0.235	0.424	0.260	0.438	0.147	0.354
>60 years old	0.115	0.320	0.118	0.322	0.062	0.241
AADT (vehs/day)	60,668.09	49,118.88	60,930.58	49,940.37	60,028.19	48,721.1

¹ Abbreviations (Severity levels): A (deaths > 3, injured persons > 20 or damage cost > 1 bil. KR. Won), B (1 < deaths ≤ 3, 5 < injured persons ≤ 20 or 2.5 mil. KR. Won < damage cost ≤ 1 bil. KR. Won), C (1 < injured persons ≤ 5 or 300,000 Won < damage cost ≤ 2.5 mil. KR. Won), D (damage cost ≤ 300,000 KR. Won). ² Bus classifications by the South Korean Ministry of Land, Infrastructure, MoLIT, 2008. (Van—9 and 12 seater vehicles; minibus—15 seats; midibus—25 seats; full-size bus—35 to over 45). ³ Truck statistics, MoLIT, 2017. (Non-commercial trucks—3,072,915, commercial trucks—389,424; 1-ton non-commercial trucks—1,673,328 (48.3%), 1 ton commercial trucks—70,264 (2.0%)).

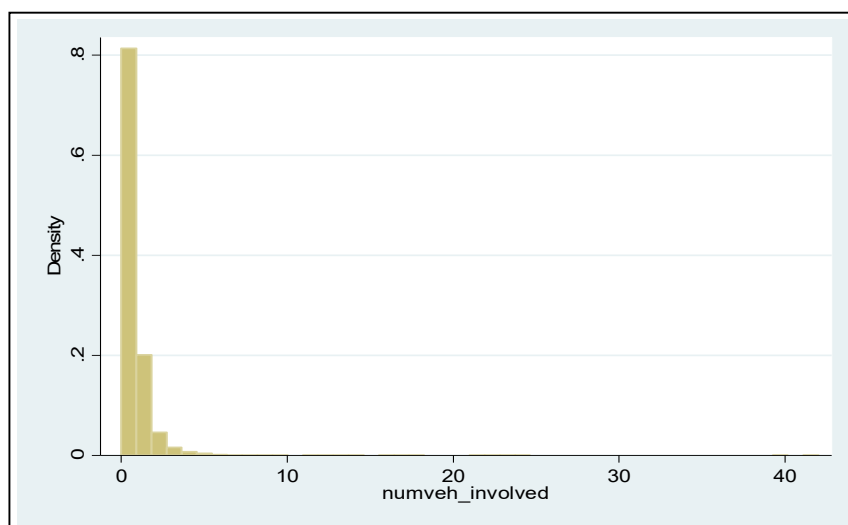


Figure 2. Histogram of number of partner-vehicles involved in a crash ($n = 0$ representing SV crash, and $n \geq 1$ represents MV crash).

4. Methodology

The double hurdle model was initially developed by Cragg [32] to explain the demand for durable goods. It has since been used in diverse fields such as healthcare delivery, agriculture, and transportation engineering. This discrete mixture model is very good at handling excess zeros. It is characterized by a two-process distribution, where the first process is a dichotomous distribution and the second is a count model. The dichotomous distribution models the probability of the crash type and separates SV crashes from MV crashes. If the number of vehicles involved in the crash is more than one, the hurdle is crossed. The second process then continues by determining the number of vehicles involved in the MV crash. In this model, the same or different explanatory variables may be used in both stages, but the explanations of the variables would be based on the distributions used at each stage.

In the present study, we follow the works of Boucher & Santolino [26], Hosseinpour et al. [27] and Hosseinpour et al. [28], and Ma et al. [29] which have demonstrated the effectiveness of this approach to analyze risk factors and the characteristics of MV crashes on expressways. Double hurdle models are generally of the form

$$y_i = s_i h_i^* \quad (1)$$

where y_i represents the dependent variable that is observed and s_i is the selection variable of the form

$$s_i = \begin{cases} 1 & \text{if } \mathbf{z}_i \gamma + \varepsilon_i > 0 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where ε_i is the standard normal error component, \mathbf{z}_i represents the independent variables, and γ is the vector of coefficients. When $s_i = 1$, the dependent variable is unbounded and the continuous latent variable h_i^* is observed. On the other hand, when $s_i = 0$, the dependent variable is bounded, and the continuous latent variable cannot be observed. Consider the case where the selection variable is one, then the continuous latent variable (h_i^*) is modeled as either a linear model or an exponential model of the form such that

$$\begin{aligned} h_i^* &= \mathbf{x}_i \boldsymbol{\beta} + \delta_i && \text{(linear)} \\ h_i^* &= \exp(\mathbf{x}_i \boldsymbol{\beta} + \delta_i) && \text{(exponential)} \end{aligned} \quad (3)$$

where $\boldsymbol{\beta}$ a vector of coefficients, \mathbf{x}_i is a vector of independent variables, and δ_i is the error component. The distribution of the error component differs for both the linear model and the exponential model. In the linear model, the error component follows a truncated normal distribution with $-\mathbf{x}_i \boldsymbol{\beta}$ its lower

truncation point whereas the error component in the exponential case follows a normal distribution. Regarding the exponential model, it follows that the conditional mean of the interior part of the hurdle model has an exponential mean.

If ll and ul are the lower and upper limits, then the probabilities of being at the limits are

$$\begin{aligned}\Pr(y_i = ll | \mathbf{z}_i) &= \Phi(ll - \mathbf{z}'_i \gamma_{ll}) \\ \Pr(y_i = ul | \mathbf{z}_i) &= \Phi(\mathbf{z}'_i \gamma_{ul} - ul)\end{aligned}\quad (4)$$

where Φ is the standard normal cumulative distribution function and γ_{ll} and γ_{ul} are the parameter vectors of the lower and upper limits of the selection model, respectively. If we assume that the error component δ_i in the underlining distribution for an individual i (linear or exponential) is normally truncated with lower and upper truncation points $ll - \mathbf{x}'_i \beta$ and $ul - \mathbf{x}'_i \beta$ respectively and with a homoscedastic variance, the log-likelihood function is given as

$$\begin{aligned}\ln L = \sum_{i=1}^n & (y_i \leq ll) \log \Phi(ll - \mathbf{z}'_i \gamma_{ll}) + (y_i \geq ul) \log \{1 - \Phi(ul - \mathbf{z}'_i \gamma_{ul})\} \\ & + (ul > y_i > ll) \left[\log \{ \Phi(ul - \mathbf{z}'_i \gamma_{ul}) - \Phi(ll - \mathbf{z}'_i \gamma_{ll}) \} \right] \\ & - (ul > y_i > ll) \left[\log \left\{ \Phi \left(\frac{ul - \mathbf{x}'_i \beta}{\sigma} \right) - \left(\frac{ll - \mathbf{x}'_i \beta}{\sigma} \right) \right\} \right] \\ & + (ul > y_i > ll) \left[\log \left\{ \phi \left(\frac{y_i - \mathbf{x}'_i \beta}{\sigma} \right) \right\} - \log(\sigma) \right].\end{aligned}\quad (5)$$

The corresponding log-likelihood when the exponential model is used is given as

$$\begin{aligned}\ln L = \sum_{i=1}^n & (y_i \leq ll) \log \Phi(ll - \mathbf{z}'_i \gamma) + (y_i > ll) \left[\log \{1 - \Phi(ll - \mathbf{z}'_i \gamma)\} \right] \\ & + (y_i > ll) \left\{ \log \left\{ \phi \left[\log(y_i - ll) - \mathbf{x}'_i \beta \right] / \sigma \right\} - \log(\sigma) - \log(y_i - ll) \right\}.\end{aligned}\quad (6)$$

5. Results and Discussions

We developed double hurdle models to estimate the probability of MV involved crash occurrence and the number of vehicles involved the crash when it is the MV involved a crash. Since the results show different relations with crash risk factors depending on vehicle type, we analyzed the impacts on vehicle types independently. In all, six different models were developed for each vehicle type, that is, trucks, buses and passenger cars. We considered two different distribution of error terms in this study and compared the results from the linear hurdle model and the exponential hurdle model.

5.1. Impact of Crash-Risk Factors on the Probability of Having MV Crashes on Expressways

The first stage of the double hurdle model which involved a logit model provided us with an output which helps us determine the variables that are likely to cause a particular type of crash (either an MV or SV crash). Tables 3 and 4 show the results of the first stage estimation of the probability of MV crash occurrence using the linear double hurdle model and the exponential double hurdle model, respectively.

Table 3. Results of linear double hurdle models for bus, truck and passenger car crashes: The first-stage estimation for the probability of multi-vehicles involved crashes.

Variables	Bus-Involved Crash		Truck-Involved Crash		Passenger-Involved Car Crash	
	Coef.	Z	Coef.	Z	Coef.	Z
Time, day and month of crash						
Weekday-Time of day (12 p.m.–3 a.m.)	-	-	-0.123	-2.59	-	-
Weekday-Time of day (3 a.m.–6 a.m.)	-	-	-0.185	-4.18	-0.059	-2.16
Weekday-Time of day (6 a.m.–9 a.m.)	-	-	-0.130	-3.27	-0.093	-3.45
Weekday-Time of day (12 p.m.–3 p.m.)	-	-	-0.157	-4.10	-	-
Weekday-Time of day (3 p.m.–6 p.m.)	-	-	-0.210	-5.25	-0.090	-3.53
Weekday-Time of day (6 p.m.–9 p.m.)	-	-	-	-	0.067	2.51
Weekend -Time of day (3 a.m.–6 a.m.)	-	-	-0.322	-3.93	-0.184	-4.20
Weekend -Time of day (6 a.m.–9 a.m.)	-	-	-0.391	-4.95	-0.139	-3.59
Weekend -Time of day (9 a.m.–12 p.m.)	-	-	-0.173	-2.40	-0.068	-1.92
Weekend -Time of day (12 p.m.–3 p.m.)	-	-	-0.169	-2.19	-	-
Weekend -Time of day (3 p.m.–6 p.m.)	-	-	-0.338	-4.10	-	-
Month of year: June	-	-	-	-	-0.071	-2.61
Month of year: July	-	-	-0.116	-2.89	-0.114	-4.47
Month of year: November	0.155	1.70	-	-	0.110	3.90
Month of year: December	-	-	0.184	4.16	0.189	6.84
Location of crash						
Crash occurs on main road	1.007	11.02	0.986	25.27	0.577	9.28
Crash occurs on ramp	-	-	-	-	-0.340	-5.04
Crash occurs at toll booth	-0.586	-3.53	-0.899	-14.22	-0.279	-3.65
Crash occurs at an electronic toll	-0.422	-2.71	-0.806	-8.45	-0.400	-4.63
Crash occurs in a tunnel	1.144	7.80	0.988	15.16	0.770	10.99
Drivers violations, vehicle malfunctions						
Over speeding	-	-	-	-	-0.215	-9.23
Improper safety gap	2.330	11.05	1.906	22.97	1.667	31.32
Improper passing	1.816	9.95	1.581	15.82	1.623	32.98
Improper reversing	2.079	3.72	1.525	4.70	2.083	11.16
Drowsy driving	0.616	7.60	0.680	19.78	0.213	8.58
Negligence by driver	1.085	17.57	1.157	35.44	0.665	33.03
Loads dropping from other vehicles on roadway	-	-	-	-	0.770	2.74
Brake malfunction	0.982	3.91	0.941	10.58	0.377	2.85
Tire malfunction	-0.621	-5.27	-0.520	-9.06	-0.401	-7.44
Other vehicle defects	-	-	-	-	0.404	1.78
Roadway surface condition and geometry						
Obstacle in roadway	-	-	0.457	8.72	-	-
Roadway problem	-	-	-	-	0.171	3.26
Pothole in roadway	-	-	-	-	0.151	2.40
Horizontal alignment: straight	0.902	2.17	-	-	0.317	3.65
Horizontal alignment: curve > 1000 m	0.774	1.85	-	-	0.230	2.62
Weather condition						
Snowy	-	-	-0.311	-4.22	-	-
Rainy	-0.215	-3.25	-0.251	-7.19	-	-
Cloudy	-	-	-0.085	-2.38	-	-
Drivers characteristics						
Age group (21–30 years old)	-0.480	-3.97	-0.236	-3.60	-0.340	-13.89
Age group (31–40 years old)	-0.263	-3.25	-0.166	-4.20	-0.338	-15.26
Age group (41–50 years old)	-0.264	-3.81	-0.132	-4.15	-0.365	-16.12
Age group (51–60 years old)	-0.212	-3.20	-0.190	-6.04	-0.373	-15.01
Age group (over 60 years old)	-	-	-	-	-0.261	-7.93
Constant	-1.370	-3.24	-0.373	-7.57	-0.319	-2.95

Note: All variables are significant at 90% confidence level.

Table 4. Results of exponential double hurdle models for bus, truck and passenger car crashes: The first-stage estimation for the probability of multi-vehicles involved crashes.

Variables	Bus-Involved Crash		Truck-Involved Crash		Passenger-Involved Car Crash	
	Coef.	Z	Coef.	Z	Coef.	Z
Time, day and month of crash						
Weekday-Time of day (12 p.m.–3 a.m.)	-	-	-0.108	-2.30	-	-
Weekday-Time of day (3 a.m.–6 a.m.)	-	-	-0.169	-3.85	-0.058	-2.11
Weekday-Time of day (6 a.m.–9 a.m.)	-	-	-0.114	-2.89	-0.095	-3.55
Weekday-Time of day (12 p.m.–3 p.m.)	-	-	-0.142	-3.74	-	-
Weekday-Time of day (3 p.m.–6 p.m.)	-	-	-0.194	-4.91	-0.094	-3.68
Weekday-Time of day (6 p.m.–9 p.m.)	-	-	-	-	0.069	2.60
Weekend-Time of day (12 p.m.–3 a.m.)	0.285	1.77	-	-	-	-
Weekend-Time of day (3 a.m.–6 a.m.)	-	-	-0.306	-3.75	-0.182	-4.14
Weekend-Time of day (6 a.m.–9 a.m.)	-	-	-0.376	-4.77	-0.139	-3.58
Weekend-Time of day (9 a.m.–12 p.m.)	-	-	-0.158	-2.19	-0.070	-1.95
Weekend-Time of day (3 p.m.–6 p.m.)	-	-	-0.319	-3.89	-	-
Month of year: July	-	-	-	-	-0.070	-2.57
Month of year: November	-	-	-0.116	-2.89	-0.113	-4.42
Month of year: December	0.156	1.70	-	-	0.099	3.51
Location of crash						
Crash occurs on main road	1.014	11.02	0.991	25.08	0.606	9.63
Crash occurs on ramp	-	-	-	-	-0.312	-4.60
Crash occurs at toll booth	-0.581	-3.26	-0.905	-13.72	-0.274	-3.57
Crash occurs at an electronic toll	-0.410	-2.63	-0.811	-8.52	-0.385	-4.45
Crash occurs in a tunnel	1.151	7.74	0.984	14.99	0.763	10.86
Drivers violations, vehicle malfunctions						
Over speeding	-	-	-	-	-0.218	-9.45
Improper safety gap	2.334	11.04	1.907	22.91	1.672	31.31
Improper passing	1.816	9.92	1.583	15.82	1.616	32.81
Improper reversing	2.086	3.73	1.526	4.70	2.070	11.08
Drowsy driving	0.608	7.46	0.682	19.53	0.197	7.97
Negligence by driver	1.084	17.83	1.157	36.13	0.658	33.02
Loads dropping from other vehicles on roadway	-	-	-	-	0.787	2.79
Brake malfunction	0.988	3.63	0.945	9.79	0.349	2.63
Tire malfunction	-0.619	-5.23	-0.517	-8.96	-0.409	-7.60
Roadway surface condition and geometry						
Obstacle in roadway	-	-	0.458	8.73	-	-
Other vehicle defects	-	-	-	-	0.376	1.65
Roadway problem	-	-	-	-	0.166	3.14
Pothole in roadway	-	-	-	-	0.149	2.35
Downward slope (1~3%)	-	-	-	-	-0.081	-3.05
Horizontal alignment: straight	-	-	-0.094	-2.36	-	-
Horizontal alignment: curve > 1000 m	0.894	2.15	-	-	0.354	4.06
Shoulder type: rock	0.762	1.82	-	-	0.265	3.01
Shoulder type: guardrail	-	-	-	-	0.125	2.18
Shoulder type: fence	-	-	-	-	-0.062	-3.60
Shoulder type: concrete	-	-	-	-	-0.136	-1.67
Weather condition						
Snowy	-	-	-0.303	-4.11	-	-
Rainy	-0.215	-3.25	-0.250	-7.10	-	-
Cloudy	-	-	-0.082	-2.30	-	-
Drivers characteristics						
Male driver	-	-	-	-	0.112	4.22
Age group (21–30 years old)	-	-	-	-	0.058	3.11
Age group (31–40 years old)	-0.486	-3.98	-0.232	-3.54	-0.283	-11.94
Age group (41–50 years old)	-0.267	-3.28	-0.163	-4.14	-0.281	-13.18
Age group (51–60 years old)	-0.270	-3.86	-0.131	-4.11	-0.305	-14.03
Age group (over 60 years old)	-0.212	-3.15	-0.189	-6.01	-0.312	-13.00
Constant	-	-	0.182	4.09	0.175	6.29

Note: All variables are significant at 90% confidence level.

5.1.1. Location of Crash

By examining the crash location variables, it was evident that crashes involving buses, passenger cars, and freight trucks are more likely to be MV crashes if they occur in a tunnel or on the main road section of the expressway. Tunnels are closed spaces. There is not enough space to evacuate a vehicle when a crash occurs. Drivers tend to be less aware of their speed and distance from the front vehicle because of the darker environment in the tunnel compared to outside the tunnel where daylight is abundant. Therefore, crashes occurring in tunnels are likely to be multiple collision accidents.

Furthermore, there are many vehicles conflicts among the vehicles during heavy traffic periods or near weaving sections of the main roadway; therefore, bus, freight trucks, and passenger car-involved crashes are likely to be involved in MV crashes on the main roadway. It was found that crashes involving passenger cars were less likely to be MV crashes if they occur at the ramp. Crashes involving all vehicle types also have a low probability of being an MV crash if they occur at the toll booth section of the expressway. Most of these crashes involve the vehicle running into stationary objects at toll booth sections or into the shoulders at the ramp section. In addition, at toll booth sections, vehicles generally slowdown in order to pay the toll. Hence, crashes at these sections are mostly SV crashes.

5.1.2. Drivers Violations, Vehicle Malfunctions, and AADT

Regarding the driver's traffic violations and faults, specific variables were seen to positively affect the likelihood of having MV crashes while others affected it negatively. The chance of buses, passenger cars or freight trucks being involved in an MV crash increases when the driver fails to keep a proper gap or safe distance between vehicles. In addition, this probability increases when the drivers of any of the three vehicle types wrongly pass other vehicles as they move on the expressway. Higher probability of the crashes caused by this fault result in MV crashes rather than SV crashes. Interestingly, it was discovered that drowsy driving caused more MV crashes in bus, truck, and passenger car crash incidents. This is because this type of driver's traffic violation is dynamically associated with other vehicles. The probability of a crash being an MV crash increases in bus, truck, and passenger cars when the drivers of such vehicles are negligent. The result shows that crashes caused by over speeding on the expressways by passenger car drivers are less likely to result in MV crashes. This observation was in line with literature [15,33]. This variable was insignificant in the models for freight trucks and buses.

Many vehicles develop faults such as tire and brake malfunctions when in use. The results of this study show that brake malfunctions of all types of vehicles mainly lead to MV crashes, but the influence on commercial vehicles such as trucks and buses are more significant than passenger cars.

The model results show that the probability of multi-vehicles involved crashes occurring is not related with AADT. The log of AADT was insignificant in both linear and exponential double hurdle models. This seems to deviate from literature which suggests an existence of a positive correlation between AADT and the probability of MV crash occurrence [21].

5.1.3. Drivers Characteristics

The relationship between the driver's features and the type of crash was also investigated. Considering their ages, the linear double hurdle model shows a high reduction in the probability of buses getting involved in MV crashes for drivers between the ages of 21 and 30 compared to the other vehicle types. This variable was found to be insignificant in the exponential double hurdle model for bus and freight trucks but showed an increase in odds for private car-involved crashes.

The other age groups had negative coefficients which connote a reduction in the probability of having MV crashes. In general, we noticed that even though young people are involved in many crashes as depicted by our data and supported by other research [34], they are more likely to be involved in SV crashes. Based on gender, the variable was not significant in the linear double hurdle model, but in the exponential double hurdle model, it was shown that male passenger car drivers are likely to be involved in MV crashes as compared to their female counterparts.

5.1.4. Roadway Surface Condition and Geometry

Sometimes, drivers encounter obstacles on the road. If they are not able to avoid these obstacles on time, they end up getting involved in a crash which may further result in an MV crash. Freight trucks are mostly involved in MV crashes when their drivers face obstacles. This can be explained by the fact that unsuspecting drivers hasten to take decisions quickly when they encounter an obstacle, and hence, oncoming vehicles can easily run into them in case they decelerate or come to a halt abruptly. The presence of potholes at sections on the expressway was also identified as a factor that significantly increases the chance of having a passenger car crash resulting in a higher probability of MV crashes. As a driver dodges potholes, he or she may lose control and hit other vehicles or be hit by unsuspecting drivers.

Factors about the effect of roadway geometry were also studied in this research. In the exponential double hurdle model, the chance of passenger cars being involved in an MV crash decreases when the crash occurred at a roadway section which is 1–3% downward elevated. In the models for other vehicle types, this variable was insignificant. However, we identified that multi-vehicle crashes involving buses and private cars are likely to occur in curve sections of radius greater than 1000 m.

5.1.5. Weather Condition

Based on the sign of the coefficient associated with the weather condition, we noted that slippery road surfaces caused by rain or snow reduce the probability of having bus and freight truck MV crash. Additionally, the results also suggest that the likelihood of having MV crashes involving trucks is reduced in cloudy weather conditions. Weather condition variables were insignificant in the passenger car-involved MV crash models.

5.1.6. Time, Day and Month of Crash

Time of day variables were predominantly insignificant in the bus-involved crash model, except for an increase in the probability of having MV crashes on weekends at 12 p.m. to 3 a.m. Generally, time of day variables were found to negatively influence the probability of trucks and passenger cars getting involved in MV crashes, except for evening peak hours (6 p.m. to 9 p.m.) where the probability of passenger car crashes resulting in MV crashes was found to be positive.

Considering the month of the year, we noticed from the exponential double hurdle model that the odds of having a private car and a bus involved in multi-vehicle crashes increased in December. In the case of the linear double hurdle model, the probability of having a truck involved in an MV crash increased in the same month. In November, the probability of buses and passenger cars being involved in an MV crash increased in the linear model case, and that of trucks and passenger cars decreased in the exponential model case. In July, both models showed a decrease in the probability of having trucks and passenger cars involved in MV crashes.

5.2. Number of Vehicles Involved in MV Crashes

Predicting number of vehicles on expressways is crucial since it provides as a decision-making tool to quantify the roadway risk and to set priorities for safety policies. The number of vehicles that are involved in an MV crash was modeled in the second stage of the double hurdle models. Tables 5 and 6 present results obtained from the second stages of the linear double hurdle model and exponential double hurdle model.

Table 5. Results of linear double hurdle models for bus, truck and passenger car crashes: The second-stage estimation for the number of vehicles-involved crashes.

Variables	Bus-Involved Crash		Truck-Involved Crash		Passenger-Involved Car Crash	
	Coef.	Z	Coef.	Z	Coef.	Z
Time, day and month of crash						
Weekday-Time of day (12 a.m.–3 a.m.)	-	-	-0.426	-1.77	-0.874	-1.66
Weekday-Time of day (3 a.m.–6 a.m.)	-	-	-0.499	-2.27	-	-
Weekday-Time of day (6 a.m.–9 a.m.)	-	-	-	-	-0.880	-1.94
Weekday-Time of day (6 p.m.–9 p.m.)	-	-	-	-	1.026	2.77
Weekend-Time of day (12 p.m.–3 p.m.)	0.610	2.12	-	-	-	-
Weekend-Time of day (3 p.m.–6 p.m.)	0.514	2.11	-	-	-	-
Month of year: March	-	-	-	-	1.060	2.51
Month of year: November	-	-	-	-	1.431	3.61
Location of crash						
Crash occurs on main road	0.861	3.09	0.798	3.05	1.899	2.02
Crash occurs on ramp	-	-	-	-	-2.155	-1.79
Crash occurs at toll booth	-	-	-1.229	-2.18	-3.564	-2.28
Crash occurs in a tunnel	1.327	3.82	1.486	4.50	3.530	3.40
Drivers violations, vehicle malfunctions, and AADT						
Over speeding	-	-	-	-	-2.302	-4.98
Improper safety gap	0.706	3.07	0.713	3.60	1.412	3.33
Improper passing	-0.770	-2.30	-1.540	-3.97	-4.331	-6.35
Drowsy driving	-0.582	-2.38	-0.691	-4.49	-4.732	-7.71
Negligence by driver	0.327	2.00	-	-	-1.247	-4.05
Tire malfunction	-	-	-	-	-4.471	-3.20
Brake malfunction	-	-	2.298	4.97	-	-
Log of AADT	-	-	0.136	1.96	0.238	1.65
Roadway surface condition and geometry						
Obstacle in roadway	0.437	1.80	1.175	5.76	-	-
Roadway problem	-	-	-	-	3.684	5.76
Pothole in roadway	1.753	2.33	2.931	2.94	2.756	3.77
No slope	0.235	1.77	-0.315	-2.56	-	-
Upward slope (1%)	-	-	-0.482	-1.87	-	-
Upward slope (3%)	-	-	-0.888	-2.69	-	-
Downward slope (3%)	-	-	-	-	1.187	2.10
Weather condition						
Snowy	-	-	-	-	5.513	7.43
Foggy	1.269	1.95	-	-	3.502	2.28
Drivers characteristics						
Age group (21–30 years)	-1.113	-2.91	-1.002	-2.90	-4.203	-8.05
Age group (31–40 years)	-0.675	-2.98	-0.756	-4.00	-3.291	-7.71
Age group (41–50 years)	-0.641	-3.43	-0.741	-4.85	-3.191	-7.47
Age group (51–60 years)	-0.280	-1.74	-0.786	-5.17	-3.227	-6.89
Constant	1.026	3.04	-0.993	-1.19	-8.500	-3.98
<i>Insignia constant</i>	0.405	9.41	0.778	24.52	1.406	30.58
<i>Number of observations</i>	3481		16,093		37,837	

Note: All variables are significant at 90% confidence level.

Table 6. Results of exponential double hurdle models for bus, truck and passenger car crashes: The second-stage estimation for the number of vehicles-involved crashes.

Variables	Bus-Involved Crash		Truck-Involved Crash		Passenger-Involved Car Crash	
	Coef.	Z	Coef.	Z	Coef.	Z
Time, day and month of crash						
Weekday-Time of day (12 a.m.–3 a.m.)	-	-	-	-	-0.042	-2.17
Weekday-Time of day (3 a.m.–6 a.m.)	-0.191	-2.74	-0.051	-2.12	-0.043	-1.92
Weekday-Time of day (6 a.m.–9 a.m.)	-	-	-	-	-0.047	-2.65
Weekday-Time of day (6 p.m.–9 p.m.)	-	-	0.069	2.89	-	-
Weekend-Time of day (6 a.m.–9 a.m.)	-	-	-	-	-0.064	-2.43
Weekend-Time of day (9 a.m.–12 p.m.)	-	-	-0.065	-1.67	0.039	1.77
Month of year: March	-	-	-	-	0.050	2.76
Month of year: October	-	-	0.046	1.95	0.070	4.03
Month of year: December	-	-	-	-	0.038	2.27
Location of crash						
Crash occurs on main road	0.197	3.59	0.113	3.94	0.098	2.76
Crash occurs on ramp	-	-	-	-	-0.079	-1.86
Crash occurs at toll booth	-	-	-0.136	-2.29	-0.138	-2.72
Crash occurs in a tunnel	0.339	4.49	0.219	5.60	0.209	5.27
Drivers violations, vehicle malfunctions, and AADT						
Over speeding	-	-	-	-	-0.111	-6.63
Improper safety gap	0.187	3.35	0.136	5.15	0.139	6.87
Improper passing	-0.157	-2.39	-0.197	-5.20	-0.170	-8.42
Improper reversing	-	-	-0.318	-2.44	-0.129	-2.21
Drowsy driving	-0.120	-2.40	-0.122	-7.23	-0.198	-11.72
Negligence by driver	0.105	2.87	-	-	-0.039	-3.11
Tire malfunction	-	-	-	-	-0.197	-4.42
Brake malfunction	-	-	0.310	4.85	-	-
Loads dropping from other vehicles on roadway	-	-	-	-	0.315	2.30
Log of AADT	-	-	0.017	2.05	0.016	2.75
Roadway surface condition and geometry						
Obstacle in roadway	0.130	2.25	0.140	4.96	-	-
Roadway problem	-	-	-	-	0.273	7.86
Pothole in roadway	0.386	1.80	0.510	2.87	0.217	5.28
Upward slope (3%)	-0.121	-1.76	-0.094	-2.70	-	-
Shoulder (fence)	-	-	-	-	-0.094	-1.78
Shoulder (concrete)	-	-	-	-	0.025	1.71
Weather condition						
Snowy	-	-	-	-	0.200	5.59
Foggy	0.465	2.55	-	-	0.182	2.39
Rainy	-	-	-0.066	-3.08	-	-
Drivers characteristics						
Male driver	-	-	0.085	1.92	0.022	1.86
Age group (21–30 years)	-0.224	-2.97	-0.142	-3.71	-0.235	-15.29
Age group (31–40 years)	-0.140	-2.89	-0.115	-5.21	-0.203	-14.57
Age group (41–50 years)	-0.163	-4.02	-0.105	-5.88	-0.194	-13.57
Age group (51–60 years)	-0.074	-1.98	-0.113	-6.38	-0.198	-12.50
Age group (over 60 years old)	-	-	-	-	-0.169	-8.30
Constant	0.169	2.95	0.003	0.03	0.168	2.24
<i>Insgma constant</i>	-0.742	-35.20	-0.756	-75.24	-0.749	-107.27
<i>Number of observations</i>	3481		16,093		39,837	

Note: All variables are significant at 90% confidence level.

5.2.1. Location of Crash

Considering the crash location variables, an interesting trend was observed. In both double hurdle models, parameter estimates showed an increase in the number of vehicles involved in bus, freight truck and passenger car-involved crashes in both tunnels and on main road sections. MV crashes are likely to occur in these areas partly because of the low visibility and narrow shoulders which are a characteristic of South Korean tunnels. Bus-involved crashes are more likely to result in MV crashes in darker places. However, the exponential double hurdle model results predict that the number of vehicles involved increases at these locations would be higher for bus-involved crashes and lower in passenger car-involved crashes; the linear double hurdle model estimates the opposite trend. Meanwhile, again, both model results show that the number of vehicles involved in a freight truck and passenger car-involved crash will reduce if the crash occurs in a toll booth section. In addition, the models show that passenger car-involved crashes that occur in ramps will have a few vehicles involved. This variable was found insignificant in the bus and freight truck-involved crash models.

5.2.2. Drivers Violations, Vehicle Malfunctions, and AADT

With regards to the driver's traffic violations and faults that were analyzed, it was seen that failing to keep a proper or safe distance could lead to an increase in the number of vehicles involved in a bus, freight truck and passenger car-involved crashes. Intuitively, this is possible because vehicles can quickly pile up in a crash if the preceding vehicle makes an abrupt halt because of a crash due to the high speeds on the expressways. However, our models reveal that over speeding in itself is likely to result in passenger car-involved MV crashes involving few vehicles.

Crashes involving all vehicle types that occur as a result of improper passing and drowsy driving are very likely to result in a reduction of the number of vehicles involved in the crashes. The models also revealed that negligence on the part of drivers could lead to a reduction in the number of vehicles involved in passenger car crashes, but an increase in the number of vehicles involved in a bus crash. It was also found that the number of vehicles involved in an MV crash involving trucks and passenger cars is positively correlated with the logarithm of AADT, which signifies that the increment in AADT levels results in more vehicles involved in a freight truck crash.

Variables for brake malfunctions in bus and passenger car-involved crashes were insignificant. However, it was positively correlated with the number of vehicles involved in truck crashes. Similarly, the variable for tire malfunction was only significant in the passenger car-involved crash double hurdle model. The variable showed a negative relationship with the number of vehicles involved in passenger car crashes.

5.2.3. Drivers Characteristics

The drivers' characteristics also provide some vital information about the number of vehicles involved in MV passenger car-involved crashes. The number of vehicles involved in these crashes decreases throughout all age groups (from 21 to 60 years and over). However, it decreases largely when the driver is between the age of 21 and 30. Hence, young passenger car drivers are more likely to be involved in MV crashes where fewer vehicles are involved compared to the case of older drivers. Compared to females, the results from the exponential double hurdle model in Table 6 show that male drivers are more likely to have MV passenger car-involved crashes in which many vehicles are involved.

5.2.4. Roadway Surface Condition and Geometry

The presence of obstacles and potholes in the roadway are associated with an increase in the number of vehicles involved in MV crashes of all vehicle types. Additionally, from the exponential double hurdle model, roadway geometry factors such as 3% upward elevated slopes were associated with a reduction in the number of vehicles involved in MV freight truck crashes.

5.2.5. Weather Condition

Weather conditions such as snow and fog are related to an increment in the number of vehicles involved in passenger car-involved crashes, and fog was found to increase the number of vehicles involved in MV bus-involved crashes. Meanwhile, the relationship between the number of vehicles involved in truck crashes and snow or foggy weather conditions were found to be insignificant. Rainy weather was estimated to cause an increase in the number of vehicles involved in truck crashes.

5.2.6. Time, Day and Month of Crash

With regards to the time variables, both linear and exponential double hurdle models showed very similar trends. It was estimated that the number of vehicles involved in passenger car and freight truck crashes decreased in all time periods. However, this trend changed from 9 a.m. to 12 p.m. on weekends, and from 6 p.m. to 9 p.m. on weekdays, respectively. In the case of buses, the dependent variable was insignificant at all time periods except for 3 a.m. to 6 a.m. (weekday) which shows a decreasing relationship. Considering the month of the year, the number of vehicles involved in passenger car-involved crashes is observed to correlate positively with variables for March, October, November, and December. Moreover, in October, the number of vehicles involved in truck-involved crashes is likely to increase.

5.3. Comparing Model Parameters across Vehicle Types

The estimates of model parameters for bus, freight truck, and passenger car-involved crashes were mainly comparable in terms of change in signs. Given the probability of having an MV crash modeled using the exponential double hurdle models in all cases of vehicle types, variables such as crashes occurring in tunnels and on the main roadway, failure to leave safe distances between vehicles, wrong passing, drowsy driving, negligence and brake malfunction showed an increase in the odds of having MV crashes. The variables in the first stage double hurdle model relating to all vehicle types had the same signs. This shows that the probability of having an MV bus crash has either an increasing or decreasing effect on all vehicle types. Considering the models for estimating the number of vehicles involved in MV crashes, variables such as crashes occurring in tunnels and the main roadway, failure to leave a safe distance between vehicles, and potholes in the roadway all led to an increase in the number of vehicles in MV crashes. For bus, freight truck and passenger car-involved crash models, it was also identified that variables such as crashes occurring at toll booths and those involving persons of all age groups from 31 to over 60 years showed a reduction in the probability of crashes resulting in an MV crash. Variables such as improper passing, drowsy driving also increased the risk having MV crashes.

In terms of number of vehicles involved in MV crashes, the second stage model results present few varying observations. In both linear and exponential double hurdle models, negligence of bus drivers leads to an increase in the number of vehicles involved in a crash, and a reduction in the case of passenger car crashes. Since buses are bigger than passenger cars, the impact they have on the crash is likely to cause more vehicles involved in MV crashes. In roadway segments with no slope, the number of vehicles likely to be involved in bus crashes will increase compared to freight trucks. Additionally, results from the exponential double hurdle model also shows that passenger car-involved crashes tend to have many vehicles involved compared to truck-involved crashes on weekends between 9 a.m. to 12 p.m. As there are many trips made by passenger cars on the expressways in South Korea on weekends, they are likely to end up in MV crashes.

5.4. Model Fit Tests

The number of vehicles involved in a crash was modeled using the double hurdle approach with both linear and exponential distributions. We compared the accuracy results of both models to the Poisson and negative binomial models which are generally used for count data. We computed four

forecast accuracy metrics namely the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) as shown in Table 7 in order to select a model that best fits the data. For the accuracy test, out of sample raw crash data from January to August in 2018 were applied. The number of observations in the accuracy test data was 469 for bus-involved crashes, 2441 for truck-involved crashes, and 5521 for passenger car-involved crashes, respectively. These accuracy metric methods estimate the precision of a number of observations by expressing the average model prediction error. The model with the least accuracy metric value is selected as the best. The results displayed in Table 7 show that the exponential double hurdle model had lower error values in the bus, freight truck, and passenger car-involved crash models compared to the other three model frameworks. Even though the linear double hurdles had a bad fit than the exponential double hurdle model, they were found to have a better fit than the general count models, Poisson and negative binomial model. It indicates that accounting for the dependence between the SV or MV crash occurrence and the number of vehicles involved in crashes provide more efficient estimates.

Table 7. Out of sample tests for Cragg's double hurdle models.

	Measures	MAE	MAPE	MSE	RMSE
Bus-involved crash model	Linear double hurdle	0.530	0.334	0.870	0.933
	Exponential double hurdle	0.499	0.248	0.745	0.863
	Poisson model	0.541	0.524	0.756	0.869
	Negative binomial model	0.539	0.522	0.752	0.867
Truck-involved crash model	Linear double hurdle	0.484	0.310	0.919	0.959
	Exponential double hurdle	0.465	0.232	0.815	0.903
	Poisson model	0.545	0.503	0.798	0.893
	Negative binomial model	0.544	0.501	0.798	0.893
Passenger car-involved crash model	Linear double hurdle	0.501	0.325	1.014	1.007
	Exponential double hurdle	0.444	0.229	0.936	0.967
	Poisson model	0.464	0.586	0.729	0.854
	Negative binomial model	0.462	0.580	0.725	0.851

In most empirical applications, the specification of the variance equation including functional form and variables is likely to be random. Therefore, the exponential specification is also considered in the study. It imposes the property that the standard deviation be strictly positive, which is desirable [35]. As shown in Table 8, we compared the log-likelihood improvement values of each model pair for each vehicle type. The log-likelihood improvement value was defined as the degree of the improved log-likelihood from baseline log-likelihoods. The models involving the exponential formulations all had better log-likelihood values and more log-likelihood improvements compared to the linear double hurdle, Poisson, and Negative Binomial model cases; hence we concluded and selected it as the best model. The log-likelihood improvement for bus-involved crash models estimated using the linear double hurdle model was found to be 15.86% while that of the exponential double hurdle model was 25.91%. The log-likelihoods of the exponential double hurdle model for the truck and passenger car crash models were improved by 27.63% and 18.53%, which were higher compared to their corresponding log-likelihood improvements in the linear double hurdle models. We further compared the models based on their AIC and BIC values. The exponential double hurdle models showed superiority over the other three models by having lower values of AIC and BIC.

Table 8. Comparison of model fits.

Measures		L-L	L-L Improvement	AIC	BIC
Bus-involved crash model	Linear double hurdle	−3170.5	15.86%	6416.9	6650.8
	Exponential double hurdle	−1740.3	25.91%	3556.5	3790.4
	Poisson model	−3034.5	17.5%	6109.0	6232.1
	Negative binomial model	−2975.3	12.4%	5992.6	6121.8
Truck-involved crash model	Linear double hurdle	−14,150.8	16.43%	28,407.6	28,814.9
	Exponential double hurdle	−7437.3	27.63%	14,986.6	15,417.0
	Poisson model	−14,058.6	13.2%	28,157.3	28,311.0
	Negative binomial model	−13,457.2	9.9%	26,956.5	27,117.9
Passenger car-involved crash model	Linear double hurdle	−33,582.8	10.98%	67,293.5	67,843.5
	Exponential double hurdle	−19,370.9	18.53%	38,893.8	39,546.8
	Poisson model	−32,168.8	13.3%	64,397.6	64,655.4
	Negative binomial model	−30,046.1	9.7%	60,154.1	60,420.5

6. Conclusions

In this study, models for bus, freight truck, and passenger car crashes were developed using linear and exponential double hurdle approaches to investigate the causes of multi-vehicle (MV) crashes, and their characteristics focusing on the influence of crash-risk factors on the number of vehicles involved in MV crashes associated with passenger cars, buses, and freight trucks. Independent variables used in this study ranged from factors such as time of day, drivers' violations and characteristics, the location of the crash, roadway geometry and condition, weather characteristics, vehicle malfunction, and log of AADT.

Key findings regarding the probability of MV involved crash occurrence are as follows.

- It was found that bus, truck, and passenger car-involved crashes were likely to be involved in multiple collisions in tunnels.
- Driver traffic violations such as the improper distance between vehicles, reversing and passing increases the probability of MV crashes occurring.
- Vehicle defects such as brake malfunctions increase the probability of MV crashes occurring, while tire punctures are more likely to be linked to SV crashes.
- Potholes increase the probability that passenger cars will be involved in an MV crash. However, it was shown that the variable indicator for “pothole in roadway” did not correlate with the probability of MV involved crash occurrence in bus and truck-involved crash models.
- Vertical curves on segments of the expressway were found to be related to the probability of MV crash occurrence in passenger car-involved crashes; however, it was found to be insignificant in bus and freight truck-involved crash models.

In terms of the number of vehicles involved in MV crashes, we found that;

- MV crashes in tunnels and mainlines were positively correlated with the number of vehicles involved in the crash, whereas fewer vehicles were involved in MV crashes at ramps and toll-booths.
- Crashes caused by not maintaining safe distances involved more vehicles, while crashes caused by improper passing were likely to involve a smaller number of vehicles.
- For crashes involving a bus or a passenger car, foggy or snowy weather increased the number of vehicles involved in the resulting MV crash. In contrast, crashes involving trucks showed an insignificant association between weather conditions such as snow and fog and the number of vehicles involved in the MV crash.
- MV crashes involving passenger cars that occur on weekends between 9 a.m. and 12 p.m. was associated with an increment in the number of vehicles, while the truck-involved MV crash model

also showed an increment in the number of vehicles involved in a crash on weekdays between 6 p.m. and 9 p.m. An MV crash involving buses was likely to have few vehicles involved between 3 a.m. and 6 a.m.

- The impact of AADT on the number of vehicles involved in an MV crash was significant only in crashes involving trucks and passenger cars. The model results showed that the number of vehicles involved in the crash was likely to rise as the AADT increases.

From our analysis, we found that operational management for tunnels, pavement, and adverse weather should be thoroughly implemented in order to reduce MV involved crashes. Additionally, the results discovered by this study validates reasons why we should intensify the education of drivers, enforcement of traffic laws, and vehicle inspections in the quest for preventing MV involved crashes.

Although this research is exploratory, the modeling approach used in the study provides a more robust way to analyze MV crash characteristics. For our analysis, we created a dummy dependent variable for estimating the probability of MV crash occurrence which gives a value of zero to SV involved crashes, and one to MV involved crashes in the dataset. This response showed a strong positively skewed distribution with many zeros. The double hurdle can capture the issue and allow the errors of the probability of the crash occurrence and the number of vehicles involved in a crash to be correlated [36]. We developed both linear and exponential double hurdle models, and the AIC, BIC and log-likelihood improvement results in Table 8 presented enough evidence to show that the exponential double hurdle model performed better compared to the linear double hurdle model, Poisson, and Negative Binomial model frameworks. The statistical tests showed that the exponential double hurdle model which considers that the conditional mean of the interior part of the hurdle model has an exponential mean is more efficient in dealing with a proper error distribution and excess zeros in the crash data compared to the other models.

Previous research mainly focused on predicting crash severity and frequency. However, we contribute to literature by separating total crash data based on the vehicle types (bus, truck, and passenger car) and crash types (MV or SV) and estimating the number of vehicles involved in crashes affected by specific factors on the expressway. Predicting the number of vehicles involved in crashes on expressways is important because it serves as a step for quantifying the damage caused in terms of socio-economic losses. MV involved crashes are emerging as a national crisis depending on how many vehicles are involved and how much it affects the society. Therefore, we focused on estimating the probability of a vehicle being involved in an MV crash, and the number of vehicles involved in the crashes. The double hurdle methodology for analyzing MV-involved crashes and findings presented in this paper may provide an avenue for the establishment of future traffic management strategies, and consequence and performance-based expressway designs.

Author Contributions: Conceptualization, Methodology, and Software, J.H.; Validation, R.T.; Formal Analysis, J.H. and R.T.; Investigation, J.H.; Data Curation, J.H.; Writing—Original Draft Preparation, J.H. and R.T.; Writing—Review and Editing, J.H., R.T. and D.P.; Visualization, J.H.; Supervision, D.P.

Funding: This research was founded by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (NRF-2017R1A2A2A05001395).

Acknowledgments: The authors wish to thank the National Research Foundation of Korea for the financial support and the South Korean Expressway Corporation (KEC) for providing the crash data used in this paper.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

1. Statistics Korea, Birth and Death. Available online: <http://kostat.go.kr/portal/eng/pressReleases/8/10/index.board> (accessed on 19 September 2018).
2. Korea Expressway Corporation, Traffic Accident Statistics. Available online: <https://www.data.go.kr/dataset/3038489/fileData.do> (accessed on 19 September 2018).

3. Islam, S.; Jones, S.L.; Dye, D. Comprehensive analysis of single-and multi-vehicle large truck at-fault crashes on rural and urban roadways in Alabama. *Accid. Anal. Prev.* **2014**, *67*, 148–158. [[CrossRef](#)]
4. Hassan, H.M.; Ahmed, M.S.; Garib, A.M.; Al-Harthei, H. Examining the differences between contributing factors affecting the severity of single and multi-vehicle Crashes. In Proceedings of the Transportation Research Board 94th Annual Meeting, Washington, DC, USA, 11–15 January 2015. Report/Paper Numbers: 15-3625.
5. Wu, Q.; Chen, F.; Zhang, G.; Liu, X.C.; Wang, H.; Bogus, S.M. Mixed logit model-based driver injury severity investigations in single-and multi-vehicle crashes on rural two-lane expressways. *Accid. Anal. Prev.* **2014**, *72*, 105–115. [[CrossRef](#)] [[PubMed](#)]
6. Lee, C.; Li, X. Analysis of injury severity of drivers involved in single-and two-vehicle crashes on expressways in Ontario. *Accid. Anal. Prev.* **2014**, *71*, 286–295. [[CrossRef](#)]
7. Chen, F.; Chen, S. Injury severities of truck drivers in single-and multi-vehicle accidents on rural expressways. *Accid. Anal. Prev.* **2011**, *43*, 1677–1688. [[CrossRef](#)] [[PubMed](#)]
8. Savolainen, P.; Mannering, F. Probabilistic models of motorcyclists' injury severities in single-and multi-vehicle crashes. *Accid. Anal. Prev.* **2007**, *39*, 955–963. [[CrossRef](#)]
9. Kockelman, K.M.; Kweon, Y.J. Driver injury severity: An application of ordered probit models. *Accid. Anal. Prev.* **2002**, *34*, 313–321. [[CrossRef](#)]
10. Ivan, J.N.; Pasupathy, R.K.; Ossenbruggen, P.J. Differences in causality factors for single and multi-vehicle crashes on two-lane roads. *Accid. Anal. Prev.* **1999**, *31*, 695–704. [[CrossRef](#)]
11. Ivan, J.N.; Wang, C.; Bernardo, N.R. Explaining two-lane expressway crash rates using land use and hourly exposure. *Accid. Anal. Prev.* **2000**, *32*, 787–795. [[CrossRef](#)]
12. Jonsson, T.; Ivan, J.N.; Zhang, C. Crash prediction models for intersections on rural multilane highways: Differences by collision type. *Transp. Res. Rec.* **2007**, *2019*, 91–98. [[CrossRef](#)]
13. Knipling, R.R. Car-truck crashes in the national motor vehicle crash causation study. In Proceedings of the Transportation Research Board 92nd Annual Meeting, Washington, DC, USA, 13–17 January 2013.
14. Dong, B.; Ma, X.; Chen, F.; Chen, S. Investigating the Differences of Single-Vehicle and Multivehicle Accident Probability Using Mixed Logit Model. *J. Adv. Transp.* **2018**, *2018*, 2702360. [[CrossRef](#)] [[PubMed](#)]
15. Chen, H.Y.; Ivers, R.Q.; Martiniuk, A.L.C.; Boufous, S.; Senserrick, T.; Woodward, M.; Stevenson, M.; Williamson, A.; Norton, R. Risk and type of crash among young drivers by rurality of residence: Findings from the DRIVE Study. *Accid. Anal. Prev.* **2009**, *41*, 676–682. [[CrossRef](#)]
16. Cerwick, D.M. A Study of Single and Multiple Vehicle Crashes Involving Heavy Trucks in Iowa. Graduate Theses, Iowa State University, Ames, IA, USA, 2013.
17. Bham, G.H.; Javvadi, B.S.; Manepalli, U.R. Multinomial logistic regression model for single-vehicle and multivehicle collisions on urban US expressways in Arkansas. *J. Transp. Eng.* **2011**, *138*, 786–797. [[CrossRef](#)]
18. Martensen, H.; Dupont, E. Comparing single vehicle and multivehicle fatal road crashes: A joint analysis of road conditions, time variables and driver characteristics. *Accid. Anal. Prev.* **2013**, *60*, 466–471. [[CrossRef](#)]
19. Lord, D.; Manar, A.; Vizioli, A. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accid. Anal. Prev.* **2005**, *37*, 185–199. [[CrossRef](#)] [[PubMed](#)]
20. Geedipally, S.R.; Lord, D. Identifying hot spots by modeling single-vehicle and multivehicle crashes separately. *Transp. Res. Rec.* **2010**, *2147*, 97–104. [[CrossRef](#)]
21. Yu, R.; Abdel-Aty, M. Multi-level Bayesian analyses for single-and multi-vehicle freeway crashes. *Accid. Anal. Prev.* **2013**, *58*, 97–105. [[CrossRef](#)] [[PubMed](#)]
22. Yu, R.; Abdel-Aty, M.; Ahmed, M. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. *Accid. Anal. Prev.* **2013**, *50*, 371–376. [[CrossRef](#)]
23. Ma, L.; Yan, X.; Weng, J. Modeling traffic crash rates of road segments through a lognormal hurdle framework with flexible scale parameter. *J. Adv. Transp.* **2015**, *49*, 928–940. [[CrossRef](#)]
24. Al Mamun, M.A. Zero-Inflated Regression Models for Count Data: An Application Tounder-5 Death. Master's Thesis, Ball State University, Muncie, IN, USA, 2014.
25. Irianti, S.; Prasetyoputra, P. Environmental, spatial, and sociodemographic factors associated with nonfatal injuries in Indonesia. *J. Environ. Public Health* **2017**, *2017*, 5612378. [[CrossRef](#)]
26. Boucher, J.P.; Santolino, M. Discrete distributions when modeling the disability severity score of motor victims. *Accid. Anal. Prev.* **2010**, *42*, 2041–2049. [[CrossRef](#)] [[PubMed](#)]

27. Hosseinpour, M.; Prasertijo, J.; Yahaya, A.S.; Ghadiri, S.M.R. A comparative study of count models: Application to pedestrian-vehicle crashes along Malaysia federal roads. *Traff. Inj. Prev.* **2013**, *14*, 630–638. [[CrossRef](#)]
28. Hosseinpour, M.; Shukri Yahaya, A.; Farhan Sadullah, A.; Ismail, N.; Reza Ghadiri, S.M. Evaluating the effects of road geometry, environment, and traffic volume on rollover crashes. *Transport* **2016**, *31*, 221–232. [[CrossRef](#)]
29. Ma, L.; Yan, X.; Wei, C.; Wang, J. Modeling the equivalent property damage only crash rate for road segments using the hurdle regression framework. *Anal. Methods Accid. Res.* **2016**, *11*, 48–61. [[CrossRef](#)]
30. Rezapour, M.; Moomen, M.; Ksaibati, K. Ordered logistic models of influencing factors on crash injury severity of single and multiple-vehicle downgrade crashes: A case study in Wyoming. *J. Saf. Res.* **2019**, *68*, 107–118. [[CrossRef](#)] [[PubMed](#)]
31. Li, Z.; Ci, Y.; Chen, C.; Zhang, G.; Wu, Q.; Qian, Z.S.; Prevedouros, P.D.; Ma, D.T. Investigation of driver injury severities in rural single-vehicle crashes under rain conditions using mixed logit and latent class models. *Accid. Anal. Prev.* **2019**, *124*, 219–229. [[CrossRef](#)] [[PubMed](#)]
32. Cragg, J.G. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econom. J. Econ. Soc.* **1971**, *39*, 829–844. [[CrossRef](#)]
33. Khattak, A.J.; Schneider, R.J.; Targa, F. Risk factors in large truck rollovers and injury severity: Analysis of single-vehicle collisions. In Proceedings of the Transportation Research Board 82nd Annual Meeting, Washington, DC, USA, 12–16 January 2003.
34. Dong, C.; Dong, Q.; Huang, B.; Hu, W.; Nambisan, S.S. Estimating factors contributing to frequency and severity of large truck-involved crashes. *J. Transp. Eng. Part A* **2017**, *143*, 04017032. [[CrossRef](#)]
35. Yen, S.T.; Su, S.J. Modeling US butter consumption with zero observations. *Agric. Res. Econ. Rev.* **1995**, *24*, 47–55. [[CrossRef](#)]
36. García, B. Implementation of a double hurdle model. *Stata J.* **2013**, *13*, 776–794. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).