Kufa Journal of Engineering
Vol. 15, No. 1, January 2024, P.P. 30-46
Article history: Received 22 September 2023, last revised 9 October 2023, accepted 12 October 2023

INVESTIGATING THE RISK FACTORS AFFECTING THE OCCURRENCE, FREQUENCY, AND SEVERITY OF LARGE TRUCK ACCIDENTS IN AL-NAJAF GOVERNORATE, IRAQ

Firas H. Asad[1] and Maysoon Z. Saeed[2]

1 Civil Engineering Dep., Faculty of Engineering, University of Kufa, Najaf, Iraq.
Email: firas.alwan@uokufa.edu.iq

2 Civil Engineering Department, Faculty of Engineering, University of Kufa, Najaf, Iraq

https://doi.org/10.30572/2018/KJE/150103

ABSTRACT

In spite of the established literature-based evidence regarding the consequences of large truck accidents, limited body of research has been done on the characteristics and risk factors of such road accidents in Iraqi cities and governorates. According to national statistics, there has been a steady increase in the number of trucks and truck-related crashes over the past ten years. This paper aims to investigate the characteristics and risk factors associated with accidents involving large trucks in Al-Najaf governorate. A sample of 400 truck drivers were randomly selected and interviewed to collect the needed accident data. Four generalized linear models have been built; ordinal regression model for total injuries, binary logistic model for fatal accident occurrence, multinomial logit model for accident frequency, and ordinal regression model for accident cost. The analysis results revealed several influential predictors including truck driver age, education level, type of collision, truck speed, truck type, and street lighting condition. The obtained findings should be enlightening and helpful for government organizations looking to promote safety measures for sustainable freight truck transport.

KEYWORDS: Truck crashes; Freight transport; Truck drivers; Multinomial logistic regression; Ordinal regression; Road safety.

Article information: This manuscript was handled by Abbas Al-Hdabi, Managing Editor.

1. INTRODUCTION

Highway accidents are a road transport challenge that have recognized and wide-ranging effects in jeopardizing community economy and public health. Such accidents decrease the nation's population or manpower that can support its social and economic development (Kumaresh et al., 2021; Taniform et al., 2023). Nowadays, road accidents take lives of around 1.35 million people annually (WHO, 2018). The situation is worse and more intense when it comes to large trucks, the corner stone of road-based city logistics, because of their significant size, weight and operational characteristics. The need for a deeper understanding and detailed insights is driven by the fact that truck-related accidents are more catastrophic as they usually result in more serious injuries than passenger car accidents (Chen et al., 2020; Evgenikos et al., 2016). In USA, 5,700 heavy trucks were involved in fatal crashes in 2021, an increase of 18% from 2020 (NCS, 2023). Recently, it was stated that commercial vehicles were involved in about 14% of fatal road incidents and 4.5% of police-reported traffic accidents in Europe (Schindler et al., 2022). It is central to emphasize that depending on the related risk factors, the features and pattern of truck crashes might vary significantly. Driver traits, vehicle characteristics, environmental circumstances, and traffic and geometry aspects of the crash location are the four elements that are frequently covered in recent literature (Yuan et al., 2021; Behnood and Al-Bdairi, 2020; Rahimi et al., 2020; Hyun et al., 2021). Truck accident investigations should be site-specific and routine to ensure the efficacy of the suggested intervening safety measures due to the variable nature of these components, both temporally and spatially.

However, despite the substantial research endeavors dedicated to the analysis of accidents involving trucks, there is a notable scarcity of comprehensive knowledge pertaining to the safety of truck transport in developing nations (Rahimi et al., 2020). Regarding Iraq, this assertion is true. The study of road truck transport in Iraq should be driven by two distinct factors; first, the Iraqi 2010-2014 national development plan (Ministry of Planning, 2010) has already emphasized the significance of supporting the transport sector through three key tasks: promoting freight transport, reducing accidents, and shortening travel times. Second, both the numbers of trucks and truck-related crashes have recently clearly increased. Tucks numbers increased from 806,000 at the end of 2016 (CSO, 2017) to around 1 million at the end of 2021 (CSO, 2022). In contrast, the number of trucks involved in accidents rose from 667 in 2019 (CSO, 2020) to 1004 in 2022 (CSO, 2023).

The overall aim of the current paper, therefore, is to investigate the risk factors that contribute to the occurrence, frequency, severity, and economic cost of large truck accidents in Al-Najaf governorate.

2. DATA AND METHODS

2.1. Study Area

Al-Najaf governorate is located in Iraq's central regions about 160 km to the south-west of the capital Baghdad. It is with spatial coordinates of 32° 01' 33.38" N (latitude) and 44° 20' 46.50" E (longitude). Fig. 1 depicts a GIS-based map for the study area. Al-Najaf has several truck trip generators that add to the economy of the governorate including major dry food wholesale trade center, major fresh food wholesale trade center, three Portland cement factories, an oil refinery and depot, several stone and sand quarries, and several grain storage silos. The number of trucks in the private sector in Al-Najaf increased from about 29450 by the end of 2016 (CSO, 2017) to about 34,050 by the end of 2021 (CSO, 2022). In contrast, the number of trucks involved in crashes increased from 34 in 2019 (CSO, 2020) to 40 in 2022 (CSO, 2023). The prior statistics boosts the importance of inspecting the freight transport safety in the governorate.

2.2. Survey Design

The research design, generally, entails performing quantitative analysis using primary accident data that was gathered from a cross-sectional questionnaire driving survey during the time period between November 2022 and April 2023. The survey design process has been conducted in line with the best practices and design guidelines mentioned in Fink (2017).

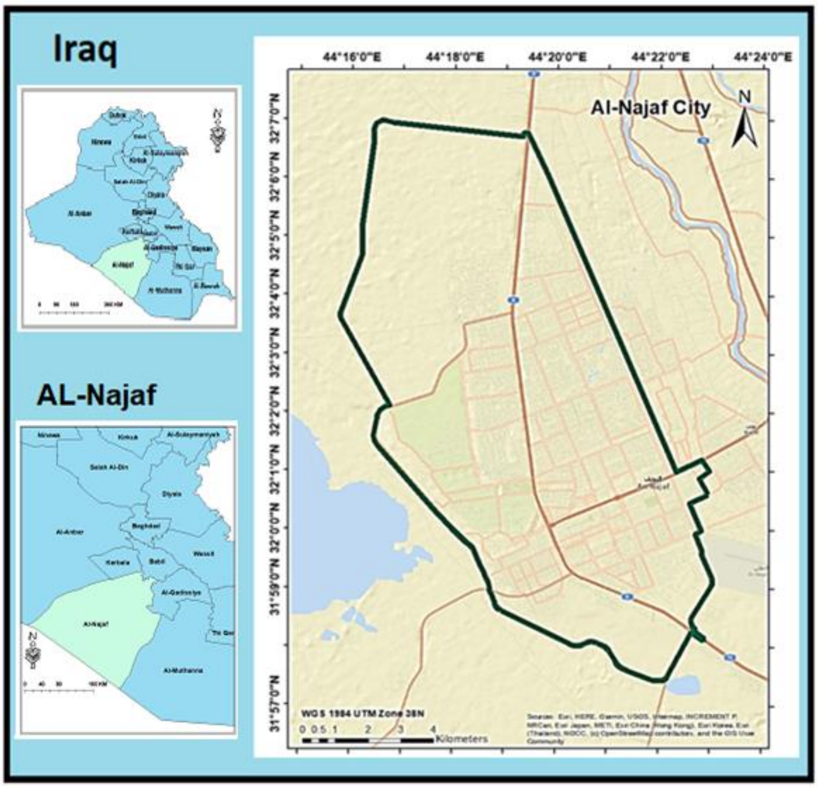

Fig. 1. A GIS-based map for the study area (by the researchers).

The structured interviewer-administered survey questionnaire consisted of four key parts. The first includes questions about truck drivers such as age, driving experience, level of education, and driving license. The second part is devoted to collect data about the accidents, including time, location, type, roadway characteristics, and speed. The third part is designed to measure the consequences of crashes, such as frequency, severity and monetary losses. Finally, the fourth part comprises the possible contributing factors for the accident, 23 possible causes are listed and the truck drivers were asked to choose among them. The truck drivers have been approached using two methods, telephone calls and in-person interviews. Two interviewers were well trained to conduct the interviews in addition to the researchers. All truck drivers involved in at least one accident in the previous five years are eligible to be interviewed. In this paper, and in line with the relevant guidelines adopted by (AASHTO, 2018), trucks are defined as any commercial vehicle that carries goods and with gross vehicle weight of at least 4000 kg or with more than four tires. Interviews were conducted with a random sample of truck drivers in order to ensure representative sample and hence generalizable research findings (Saunders et al., 2012). Numerous trips by interviewers were made to the truck trip generators in the governorate, petrol stations, and to truck dealers in order to recruit truck drivers. Drivers were also asked to willingly provide mobile phone numbers of one or more colleagues. A sample of 400 valid questionnaires has been obtained, which is based on the suggestions made in (Rea and Parker, 2014) is a statistically-approved sample size at 5% margin of error and 95% level of confidence. Finally, to ensure the internal validity of the questionnaire; that is, the questionnaire items can accurately represent the characteristics they are intended to measure, a pilot test was initially conducted. Twenty truck drivers were interviewed and their responses, viewpoints, and comments were considered to amend some questions to make them clearer. Pilot testing strengthens the reliability and validity of questionnaire surveys (Fink, 2017).

2.3. Data Processing and Analysis

In data processing stage, the collected questionnaires were extracted, coded, and digitized using the IBM SPSS (v. 26) software. The processing also involved data screening and cleaning to eliminates imprecisions in raw data that could compromise the output accuracy (Tabachnic et al., 2019). The data analysis, in contrast, involved two main stages; first, conducting descriptive frequency analysis for the variables that will be used as outcome or predictors in the later developed statistical accident models. Second, specifying, building and evaluation these accident models. This modelling stage involves developing several multiple regression models with discrete (categorical) dependent variables (DVs). Such models are members of a group of regression models typically designated as generalized linear models (GZLM). Two main essential features distinguish these models from the traditional ordinary least square (OLS) linear regression models; these are, (1) the dependent variable is not necessary to be continuous, and (2) they are free from the strict assumptions of errors in being normally distributed, independent, and with constant variance (Orme and Combs-Orme, 2009). In contrast, these models are more sophisticated than OLS linear regression models and hence they require more efforts and expertise in developing, evaluating their goodness and significance, and in interpreting the results. In order to sustain the linearity between the dependent (outcome) variable and the set of independent variables (IVs) (predictors), a specific link function should be used. For the case of binomial, multinomial, and ordinal regressions, this function is called "logit". Equation 1 shows the estimated logistic regression model based on a sample of population data (Orme and Combs-Orme, 2009):

$$\text{Logit (Y)} = a + B_1X_1 + B_2X_2 + \ldots + B_kX_k \tag{1}$$

Where Y is the estimated mean value of the DV; a is the estimated intercept; X_1 through X_k represents the included predictors (IVs); and B_1 through B_k are the regression slopes.

The logit (Y) is natural logarithm of the odds of occurring Y; in turn, odds of Y is the probability of occurring of Y, p(Y) to the probability of not occurring, 1 - p(Y). Hence, logit (Y) can be mathematically expressed as in Eq. 2 (Orme and Combs-Orme, 2009):

$$\text{Logit (Y)} = \ln(\text{Odds (Y)}) = \ln[p(Y) / 1 - P(Y)] \tag{2}$$

The odd ratio (OR) is the ratio of the odds of the event for one value of the IV divided by the odds for a different value of the IV, usually a value one unit lower. The OR specifies the amount of change in the odds and the direction of the relationship between an IV and the DV.

In this paper, four multivariate regression models were built; these are, an ordinal model for total injuries, a binary logistic model for fatal accident occurrence, a multinomial logit model for accident frequency, and an ordinal model for accident repairing cost. The PLUM procedure (Polytomous Universal Model) embedded in the IBM SPSS software has been used to run the ordinal regression analysis. The test of parallel lines assumption has been examined in order to certify the legitimacy of using ordinal regression. Further technical details can be found in (Orme and Combs-Orme, 2009; Tabachnic et al., 2019; George, and Mallery, 2019). Finally, the assumption of logistic regression regarding the linearity between the set of continuous IVs and the logit of the DV has been found valid after testing using the Box-Tidwell approach (Tabachnick and Fidell, 2019).

3. RESULTS AND DISCUSSION

3.1. Descriptive analysis

Table 1 depicts the personal, driving, and accident characteristics for the 400 interviewed large truck drivers that will be considered in building the four accident regression models in Section 3.2. It is worth noting that the number of accidents equals the number of truck drivers; that is because only drivers who involved in a road accident have been interviewed. According to the table, the majority of truck drivers (88.5%) are with less high school academic qualification. About 45% of drivers are within 10-20 years driving experience group. Most of the accidents were vehicle-vehicle collisions that occurred in ordinary weather, along major urban streets, and during day or good street lighting. Risky accidents such as head on and right-angle crashes represent just over 30% of total crashes.

Table 1 also shows that the standard deviation (sd) for some characteristics, such as monthly trips, speed, and repairing cost, are relatively high. The high sd in monthly trips could reflect the high variation in the frequency between internal and external truck trips; the latter are with less daily and thus monthly frequency as they usually take longer travel times. Regarding the high sd in speeds, this is most probably due to the absence of law enforcement techniques related to speed limit such as cameras and speed radars along most of the roads and streets; as a result, encouraging some drivers to exceed speed limits. Hence, the range between higher and lower speeds increases and, in turn, the standard deviation increases. Finally, the high standard deviation of truck repairing cost comes from the fact that some truck drivers reported that the repairing cost would almost equal the cost of a new truck; this is the case when the truck burnt during the crash.

3.2. Truck accidents models

This section aims to develop four truck accidents models to identify the predictors that are influential in predicting truck accident occurrence, frequency, severity and monitory cost.

Table 1. Personal and driving characteristics of truck drivers.

Characteristics	Statistics	Categories	Frequency	Share %
Highest education level	n = 393	High school or above	45	11.5
		Intermediate school	104	26.5
		Primary school	160	40.7
		Illiterate	84	21.4
Truck type	n = 400	Lorry (2 and 3 axles)	130	32.5
		Trailer/Semi-trailer	123	30.8
		Tanker	79	19.8
		Other SU trucks	68	17.0
Weather condition	n = 400	Non-ordinary weather	94	23.5
		Ordinary weather	306	76.5
Street location	n = 399	Urban	290	72.7
		Rural	109	27.3
Street lighting	n = 400	Day / Good lighting	331	82.8
		Bad or no lighting	69	17.3
Street class	n = 399	Major	369	92.5
		Minor	26	6.5
		Local	4	1.0
Accident location	n = 400	Along street	290	72.5
		Intersection/U-turn	110	27.5
Accident type	n = 400	Hitting a car	246	61.5
		Hitting a road user	56	14.0
		Others	98	24.5
Type of collision	n = 260	Rear end	75	28.8
		Head on	41	15.8
		Right angle	39	15.0
		Other angles	62	23.8
		Sideswipe	43	16.5
Truck driver age (years)			n = 400 ; x = 38.7 ; sd = 8.76	
Driving experience (years)			n = 396 ; x = 16 ; sd = 8.26	
Monthly truck trips			n = 393 ; x = 28.9 ; sd = 27.21	
No. of vehicles involved in crash			n = 400 ; x = 1.8 ; sd = 0.80	
Total injured per crash			n = 400 ; x = 1.1 ; sd = 1.49	
Total fatalities per crash			n = 400 ; x = 0.23 ; sd = 0.712	
Accident history in last 5 years			n = 400 ; x = 1.19 ; sd = 0.776	
Truck speed (kph)			n = 400 ; x = 55.6 ; sd = 26.17	
Other vehicle speed (kph)			n = 275 ; x = 59.3 ; sd = 31.26	
Truck repairing cost (x 1000 ID)			n = 393 ; x = 6230 ; sd = 13112.7	

n = number of valid cases; x = arithmetic mean; sd = standard deviation.

3.2.1. Total injured accident model

The ordinal logistic regression has been chosen because of the ordinal nature of the dependent variable, the total injured individuals in a crash, which is coded as 0, 1, and 2 or more injury cases. The explanatory variables (predictors) included in the model are: driving experience, truck speed, weather condition, street location, street lighting, accident type, and the number of vehicles involved in accident. Table 2 depicts the key details of the model parameter estimates in addition to key statistics regarding the goodness of model fit. Based on Table 2, only the type of accident and number of vehicles involved in crash are with statistically significant influence in predicting the outcome variable (the total injured) at 5% level of significance (LOS). Regarding the accident type, hitting a car is the type of accident that more probable to result in larger number of injuries. Statistically speaking, the odds of "hitting a car" crash to have injuries is 2.745 higher than those types of crashes in "others" category which include roll-over, run off road, and hitting fixed objects or animals. The table also shows that as the number of vehicles involved in accident increases the probability of high-injuries accidents. The odds of accidents (with one- or two-vehicles involved) to be in a higher injuries category is less than the corresponding odds of accidents (with 2 or more vehicles category) by 0.023 and 0.355, respectively. Furthermore, the probability of high injuries increases in non-normal weather (fog, heavy rain, dust storms), over urban streets, and at inadequate street lighting conditions; however, these influences are not statistically significant. The effects of the characteristics of drivers, vehicles, and roadway on crash severity have been confirmed in literature (Jin et al., 2019; Decker et al., 2016).

With respect to the modelling goodness and model evaluation statistics, the footnotes of Table 2 illustrate that the final model Chi-Square is significant (p-value = < 0.001) which implies the significant contribution of the added variables in comparison with the reduced model which involves only the constant. The model's goodness of fit based on Pearson's Chi-Square test statistic is not significant (sig. = 0.587) which indicates the non-existence of significant difference between the predicted and actual category probabilities. The pseudo Nagelkerke R-square is 0.42 which comparatively indicate medium to good prediction goodness. Finally, the "test of parallel lines" is insignificant and hence the using of ordinal regression is legitimate.

Table 2. The parameter estimates for the total injured ordinal logit regression model

Variable	Estimate (B)	Std. Error	Sig.	Exp (B) (OR)	95% C.I for Exp(B) Lower	Upper
Total injured (0 = 0 injured)	-1.582	.724	.029	0.205	-3.002	-.162
Total injured (1 = 1 injured)	-.001	.718	.999	0.999	-1.409	1.407
Total injured (ref. cat. = ≥2)	-	-	-	-	-	-
Driving experience (years)	-.002	.013	.879	0.998	-.028	.024
Truck Speed (kph)	.002	.004	.665	1.002	-.007	.011
Weather (1= normal)	-.272	.269	.311	0.762	-.798	.254
Weather (ref. cat. = not normal)	0^a	-	-	-	-	-
Street location (1= urban)	.105	.245	.668	1.111	-.375	.586
Street location (ref. cat. = rural)	0^a	-	-	-	-	-
St. lighting (1= day/good lighting)	-.321	.291	.269	0.725	-.891	.249
St. lighting (ref cat.= bad/no light)	0^a	.	.	-	.	.
Acc. type (1= hitting a car)	1.010	.409	.013	2.745	.209	1.811
Acc. type (2=hitting road users)	.742	.447	.097	2.100	-.134	1.619
Acc. type (ref. cat. = others)	0^a	-	-	-	-	-
Veh. involved (1 = one)	-3.757	.514	.000	0.023	-4.764	-2.749
Veh. involved (2= two)	-1.036	.359	.004	0.355	-1.739	-.334
Veh. involved (ref. cat. = ≥3)	0^a	-	-	-	-	-

a. This parameter is set to zero because it is redundant; ref. cat. = Reference category.

- Predicted variable is the total injured (0, 1, ≥2); Reference category (≥2).

- The reference category for all the predictors has been set as the last category.
- Model fitting information: Final model Chi-Square is significant ($\chi 2$ (df=9) = 182.649, p = <0.001).

- Goodness of fit based on Pearson's Chi-Square ($\chi 2$ (df=673) = 664.271, p = 0.587).

- Pseudo Nagelkerke R^2 = 0.42

- Test of parallel lines: General Chi-Square ($\chi 2$ (df=9) = 10.849, p = 0.286).

3.2.2. Fatal accident occurrence model

The aim is to develop a regression model that best estimate the probability of truck driver to be in an accident based on a set of predictors. Since the predicted variable has only two categories (no death, one or more deaths), the binary logistic regression is the most suitable modelling technique. Table 3 lists the SPSS output for the parameter estimates of the developed logistic model along with the key statistics for modelling significance and goodness. Table 3 reveals two variables that are significant at 5% LOS (accident type and truck speed) and one at 10% LOS (truck type). Collisions that involved hitting a car or hitting a road user (pedestrians, bicyclists, or motor cyclists) have odds of being fatal accidents of 7.392 and 48.185 higher than those in "others" group of accident types, respectively. The "other" accident type group involves roll-over, run off road, and hitting fixed objects or animals. There is a very high probability that crashes involving hitting road users will include death cases; this is evidently because of vulnerability of road users. Regarding truck speed, the analysis reveals that a one unit (1 kph) increase in the speed of a truck would rise the odd of occurring a fatal accident by 1.054. Regarding truck type, accidents that involve tanker are more probable to be fatal; there odds are 2.667 higher the odds of accidents with “others” truck type group. The latter groups usually involve other single unit trucks. Tankers are comparatively heavier and harder to be controlled or stopped. Additionally, the probability of fatal accident increases as the number of involved vehicles increases; also, literate drivers are less probable than illiterate ones to be in fatal crash. However, the previous two factors are not statistically significant. The potential influences of crash type, truck type and combination, truck speed and driver education on the severity of fatal truck accidents have been confirmed in relevant studies (Yuan et al., 2021; Rahimi et al., 2020; Rezapour and Ksaibati, 2018; Wang et al., 2019).

Regarding the quality of model fitness, the table shows that the pseudo Nagelkerke R-square is 0.32 which comparatively indicate medium to good prediction goodness. The Hosmer and Lemeshow test statistic is not significant (p-values = 0.915) which is a good indicator as it implies the absence of significant variance between the observed and computed outcome results.

3.2.3. Accident frequency model

This model investigates the explanatory variables that best predict the accident frequency of partially or totally at-fault truck drivers (the dependent variable). The accident frequency has been coded into three categories that represent the number of truck driver accidents during the last five years (from 2017 to 2022). These groups are, no accident (61 cases), one accident (227 cases), and 2 or more accidents (112 cases). Whereas firstly the ordinal regression was chosen to construct the model, the parallel line test was found significant (p-value = 0.013); i.e, the null hypothesis is rejected and hence the assumption is violated. The most proper alternative is to run a multinominal logistic regression analysis. Table 4 presents a summary for the key modelling outputs for two scenarios; first, when the reference category is the (2 or more) group and the second when the reference category is the (no accident) group. According to Table 4, four predictors have been found with significant influence on the crash frequency; these are, driver age, education, street lighting, and monthly trips. Driver age has been only significant at 10% LOS. The odds of young drivers (18-25 years) to be in the “no accident” category comparing to the “2 or more accident” category is 23.528 higher than those elderly drivers (> 55 years). That is probably because senior drivers react more slowly and also their general health status may also have an impact on their driving safety (Zheng et al., 2018; Lee et al., 2020).

Table 3. Variables details for the fatality binary logistic regression model.

	B	S.E.	Sig.	Exp(B) (OR)	95% C.I for EXP(B) Lower	Upper
Acc. type (ref. cat. = others)	-	-	-	-	-	-
Acc. type (1= hitting a car)	2.000	.660	.002	7.392	2.029	26.933
Acc. type (2= hitting a road use)	3.875	.693	.000	48.185	12.384	187.484
Education (ref. cat. = illiterate)	-	-	-	-	-	-
Education (1=high school or above)	-.287	.582	.622	.751	.240	2.348
Education (2=intermediate school)	-.282	.479	.556	.754	.295	1.929
Education (3=primary school)	-.618	.457	.176	.539	.220	1.320
Street location (ref. cat. = rural)	-	-	-	-	-	-
Street location (1=urban)	.030	.367	.936	1.030	.502	2.115
St. lighting (ref, cat.= bad/no light.)	-	-	-	-	-	-
St. lighting (1= day/good lighting)	.182	.460	.692	1.200	.487	2.954
Weather (ref, cat, = non-ordinary)	-	-	-	-	-	-
Weather (1= ordinary weather)	.324	.440	.461	1.383	.584	3.279
Truck speed (kph)	.050	.009	.000	1.052	1.033	1.071
Vehicles involved	.237	.222	.286	1.268	.820	1.959
Truck type (ref. cat. = others)	-	-	-	-	-	-
Truck type (1=lorry 2-3 axles)	.390	.506	.442	1.476	.547	3.982
Truck type (2=trailer/semi-trailer)	.436	.524	.405	1.547	.554	4.322
Truck type (3= tanker)	.985	.560	.079	2.677	.893	8.026
Age	.001	.020	.976	1.001	.963	1.040
Constant	-8.091	1.503	.000	.000	-	-

- The predicted variable: the death status (no death=0, one or more = 1); Reference category = No death.
- The reference category for all the predictors has been set as the last category.
- Pseudo Nagelkerke R^2 = 0.32.
- Hosmer and Lemeshow Test: Chi-square details: = ($\chi 2$ (df=8) = 3.289, p = 0.915).

Regarding education, drivers with high school or those with intermediate school education have odds to be in the "no accidents" category that are 4.398 and 2.551 higher than illiterate drivers, respectively. This implies that truck drivers with low education have high tendency to get into accidents due to speeding or other risky driving behavior (Rahimi et al., 2020). For street lighting conditions, the occurrence of accidents in day light or in good lighted streets are less probable comparing those in bad or not lighted streets. In numbers, the odds of zero-accidents in day or good lighting conditions are 4.913 higher than those at bad or no lighting street conditions. Finally, for monthly trips, a one-unit increase in monthly trips can result in an odd to be in the "one accident" category that is 0.986 the odd of being in the reference category (2 or more accidents). That is, as the monthly trips increase, the probability of being in the "2 or more accidents" category increases with respect to the "one accident" category.

Table 4: Key modelling outputs for the accident frequency multinomial logit model

	No. of truck driver accidents [a]	B	Std. Error	Sig.	Exp(B)	95% C.I for Exp(B) Lower	Upper
No accident	Intercept	-4.850	2.110	.022	-	-	-
	Driving experience (years)	.013	.025	.610	1.013	.964	1.064
	Monthly trips	.000	.007	.954	1.000	.988	1.013
	Driver age (0 = 18-25)	3.158	1.673	.059	23.528	.887	624.175
	Driver age (1 = >25-55)	1.683	1.242	.175	5.380	.472	61.326
	Driver age (ref. cat. ≥ 55)	0^b	-	-	-	-	-
	Education (0=high school or above)	1.481	.662	.025	4.398	1.202	16.092
	Education (1= intermediate school)	.937	.501	.061	2.551	.957	6.805
	Education (2= primary school)	.374	.485	.441	1.453	.562	3.762
	Education (ref. cat.= illiterate)	0^b	-	-	-	-	-
	Truck type (1=lorry 2-3 axles)	-.147	.511	.773	.863	.317	2.350
	Truck type (2=trailer/semi-trailer)	.624	.588	.289	1.866	.589	5.911
	Truck type (3= tanker)	-.662	.635	.297	.516	.149	1.789
	Truck type (ref. cat. = others)	0^b	-	-	-	-	-
	Weather (1= ordinary)	-.287	.430	.505	.751	.323	1.744
	Weather (2= non-ordinary)	0^b	-	-	-	-	-
	Street class (1 = major)	.452	1.366	.741	1.571	.108	22.832
	Street class (2 = minor)	-.328	1.552	.833	.720	.034	15.091
	Street class (ref. cat. = local)	0^b	-	-	-	-	-
	Street location (1 = urban)	.194	.387	.617	1.214	.568	2.592
	Street location (ref. cat. = rural)	0^b	-	-	-	-	-
	St. lighting (day/good lighting)	1.592	.658	.016	4.913	1.352	17.852
	St. lighting (ref. cat. = bad/none)	0^b	-	-	-	-	-
One accident	Intercept	-.564	1.615	.727	-	-	-
	Driving experience (years)	-.006	.019	.740	.994	.958	1.031
	Monthly trips	-.014	.005	.008	.986	.976	.996
	Driver age (0 = 18-25)	2.168	1.141	.058	8.737	.933	81.841
	Driver age (1 = >25-55)	.306	.716	.669	1.359	.334	5.532
	Driver age (ref. cat. ≥ 55)	0^b	-	-	-	-	-
	Education (0=high school or above)	.665	.505	.188	1.945	.723	5.229
	Education (1= intermediate school)	.318	.362	.379	1.375	.677	2.792
	Education (2= primary school)	.421	.325	.195	1.524	.806	2.882
	Education (ref. cat.= illiterate)	0^b	-	-	-	-	-
	Truck type (1=lorry 2-3 axles)	-.254	.373	.496	.776	.374	1.610
	Truck type (2=trailer/semi-trailer)	.266	.439	.545	1.304	.552	3.082
	Truck type (3= tanker)	-.500	.446	.262	.606	.253	1.453
	Truck type (ref. cat. = others)	0^b	-	-	-	-	-
	Weather (1= ordinary)	-.217	.310	.484	.805	.438	1.479
	Weather (2= non-ordinary)	0^b	-	-	-	-	-
	Street class (1 = major)	1.052	1.289	.415	2.864	.229	35.847
	Street class (2 = minor)	1.029	1.352	.447	2.798	.198	39.576
	Street class (ref. cat. = local)	0^b	-	-	-	-	-
	Street location (1 = urban)	.324	.293	.268	1.383	.779	2.456
	Street location (ref. cat. = rural)	0^b	-	-	-	-	-
	St. lighting (day/good lighting)	.101	.320	.751	1.107	.592	2.071
	St. lighting (ref. cat. = bad/none)	0^b	-	-	-	-	-

- (a) The predicted variable (none, 1, 2 or more); Reference Category (ref. cat.) = 2 or more.
- (b) This parameter is set to zero because it is redundant.
- The reference category for all the predictors has been set as the last category.
- Model fitting information: Final model Chi-Square is significant ($\chi 2$ (df=30) = 53.813, p = 0.005).
- Goodness of fit based on Pearson's Chi-Square ($\chi 2$ (df=708) = 764.378, p = 0.070).

- Pseudo Nagelkerke R^2 = 0.15

Finally, for modelling goodness and evaluation statistics, the final model Chi-Square is significant (p-value =0.005) which implies the overall significance of model parameters. The model's goodness of fit based on Pearson's Chi-Square test statistic is not significant (p-value=0.07) which indicates the non-existence of significant difference between the predicted and actual category probabilities. Lastly, the pseudo Nagelkerke R-square is 0.15 which to some extent indicates a weak prediction capacity.

3.2.4. Truck accident cost model

The ordinal outcome variable in this model is the total repairing cost for truck vehicles (in thousands ID) which has been coded into three levels. The model predictors and the modelling statistics are listed in Table 5. The table reveals that only three variables are significant at 5% LOS, these are truck speed, other-vehicle speed, and accident type. For truck and other-vehicle speeds, a one-unit increase in these speeds will increase the odds of being in higher repairing cost category by 1.026 and 1.011, respectively. That is a reasonable finding, as high speeds increase the damage intensity, and hence, the cost of repairment. For accident type, the odds of head on and right-angle collisions to be in higher repairing cost categories are 17.672 and 6.679 with respect to sideswipe crashes, respectively. That is most probably due to the high damage expected for these two risky types of crashed. Furthermore, the table finds non-significant but interesting influence of some variables. For example, the cost of truck repairing increases when the truck driver is illiterate. The crashes occurred along rural street, in ordinary weather, and during inadequate street lighting have higher repairing cost than others; this is most likely because of expected high speed associated with such conditions.

Finally, all the modelling performance indicators listed in the lower part of Table 5 are in their adequate ranges.

Table 5: Parameter estimates and statistics for the accident ordinal regression model

Model variables	B	Std. Error	Sig.	Exp (B)	95% C.I for Exp(B)	
					Lower	Upper
Repairing cost (1= <1000k)	.637	.809	.431	1.891	-.948	2.222
Repairing cost (2=1000-5000k)	4.064	.859	.000	58.207	2.381	5.747
Repairing cost (3= ≥5000k)	-	-	-	-	-	-

Model variables	**B**	**Std. Error**	**Sig.**	**Exp (B)**	**95% C.I for Exp(B)**	
					Lower	**Upper**
Truck speed, kph	.026	.007	.000	1.026	.013	.039
Other vehicle speed, kph	.011	.005	.038	1.011	.001	.020
Education (0=high school or above)	-.091	.518	.860	0.913	-1.107	.925
Education (1= intermediate school)	-.551	.402	.170	0.576	-1.338	.236
Education (2= primary school)	-.184	.374	.622	0.832	-.916	.548
Education (ref. cat.= illiterate)	0[a]	-	-	-	-	-
Acc. type (1= head on)	2.872	.513	.000	17.672	1.866	3.879
Acc. type (2= rear end)	.427	.426	.316	1.533	-.408	1.262
Acc. type (3= right angle)	1.899	.500	.000	6.679	.919	2.880
Acc. type (4= other angles)	.341	.431	.429	1.406	-.504	1.186
Acc. type (ref. cat.= sideswipe)	0[a]	-	-	-	-	-
Acc. location (1=along street)	-.021	.322	.948	0.979	-.652	.609
Acc. loc. (ref. cat.=intersection/u-turn)	0[a]	-	-	-	-	-
Weather (1= ordinary)	.458	.341	.179	1.581	-.210	1.127
Weather (2= non-ordinary)	0[a]	-	-	-	-	-
Street location (1 = urban)	-.248	.305	.416	0.780	-.847	.350
Street location (ref. cat. = rural)	0[a]	-	-	-	-	-
St. lighting (day/good lighting)	-.301	.375	.422	0.740	-1.035	.433
St. lighting (ref. cat. = bad/none)	0[a]	-	-	-	-	-
Truck type (1=single-unit truck)	-.307	.376	.414	0.736	-1.043	.430
Truck type (2=trailer/semi-trailer)	-.295	.391	.451	0.745	-1.061	.471
Truck type (ref. cat. = tankers)	0[a]	-	-	-	-	-

Model variables	B	Std. Error	Sig.	Exp (B)	95% C.I for Exp(B)	
					Lower	Upper

a. This parameter is set to zero because it is redundant; ref. cat. = reference category.
- Predicted variable is the truck repairing cost; ref cat. (≥ 5000k ID).

- The reference category for all the predictors has been set as the last category.
- Model fitting information: Final model Chi-Square is significant ($\chi2$ (df=15) = 113.031, p = <0.001).

- Goodness of fit based on Pearson's Chi-Square ($\chi2$ (df=485) = 436.93, p = 0.942).

- Pseudo Nagelkerke R^2 = 0.415.

- Test of parallel lines: General Chi-Square ($\chi2$ (df=15) = 11.856, p = 0.690.

4. CONCLUSIONS

Based on the preceding statistical analysis and modelling for the truck accident data, the following key findings are vital to be addressed:

1. The majority of truck drivers (88.5%) that have been interviewed are with low academic qualification. About 45% of drivers are within 10-20 years driving experience group. Most of the accidents were vehicle-vehicle collisions that occurred in ordinary weather, along major urban streets, and during day or good street lighting. Risky accidents such as head on and right-angle crashes represent over 30% of total crashes.

1. For the total injured prediction model, the ordinal regression analysis revealed that "hitting a car" accident type and the number of vehicles involved in accident can significantly increase the probability of high-injuries accidents.

2. For the total deaths model, the binary logistic regression revealed that crash types such as hitting a car/road user, truck speed, and tankers truck type are the more influential risk factors that rise the odds of truck-involved accidents to be fatal.

3. For the 5-years crash frequency, the multinominal logit model implies that elderly drivers and illiterate ones are more probable to get involved in more crashes. In addition, not adequate street lighting and high rate of monthly trips could rise the likelihood of being involved in a crash.

4. For the truck accident cost model, the outputs of the ordinal regression analysis revealed that the high speeds of the collided vehicles can rise the odds of being in higher truck repairing cost category. Alao, being in a head-on or in a right-angle collision can increase the repairing cost.

Finally, based on the analysis results several truck transport safety measures are important to be considered. Examples include strict laws (especially for speed limits), adequate street design, and smart traffic operation to assure safe and smooth driving for large truck freight transport.

5. REFERENCES

AASHTO (2018). A Policy on Geometric Design of Highways and Streets (The Green Book). Washington, DC: American Association of State Highway and Transportation Officials.

Behnood, A., and Al-Bdairi, N. S. S., (2020). *Determinant of injury severities in large truck crashes: A weekly instability analysis.* Safety science. Vol 131, pp.104911.

Central Statistical Organization (2022). *Statistics of Private Sector Motorcars Registered at General Directorate of Traffic until December 2021*. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq.

Central Statistical Organization. (2017). *Statistics of Private Sector Motorcars Registered at General Directorate of Traffic until December 2016*. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq.

Central Statistical Organization. (2020). *Traffic Accidents Statistics for 2019*. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq.

Central Statistical Organization. (2023). *Traffic Accidents Statistics for 2022*. Directorate of Transport and Communication Statistics, CSO, Baghdad, Iraq.

Chen, S., Zhang, S., Xing, Y., and Lu, J., (2020). *Identifying the factors contributing to the severity of truck-involved crashes in Shanghai River-Crossing Tunnel*. International journal of environmental research and public health. Vol. 17, Issue 9, pp. 3155.

Evgenikos, P., Yannis, G., Folla, K., Bauer, R., Machata, K., and Brandstaetter, C., (2016). *Characteristics and causes of heavy goods vehicles and buses accidents in Europe*. Transportation research procedia. Vol. 14, pp. 2158-2167.

Fink, A. (2017). *How to conduct surveys: A step-by-step guide*. 6th ed. Sage Publications.

George, D., and Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference*. 16th ed. Routledge.

Hyun, K. K., Mitra, S. K., Jeong, K., and Tok, A., (2021). *Understanding the effects of vehicle platoons on crash type and severity.* Accident Analysis & Prevention. Vol. 149, pp. 105858.

Kumaresh, G., Saldanha, F., Lich, T., and Moennich, J., (2021). *Overview of Truck Accidents in India and Its Economic Loss Estimation* (No. 2021-26-0007). SAE Technical Paper.

Lee, J., Yeo, J., Yun, I., & Kang, S. (2020). Factors affecting crash involvement of commercial vehicle drivers: evaluation of commercial vehicle drivers' characteristics in South Korea. Journal of advanced transportation, 2020, 1-8.

Ministry of Planning. (2010). *National Development Plan 2010-2014*. Baghdad, Iraq: Ministry of Planning.

National Safety Council NSC (2023). Large Trucks – Injury Facts. Retrieved March, 2023, from: https://injuryfacts.nsc.org/motor-vehicle/road-users/large-trucks/.

Orme, J.G. and Combs-Orme, T., 2009. Multiple regression with discrete dependent variables. Oxford University Press.

Rahimi, E., Shamshiripour, A., Samimi, A., and Mohammadian, A. K., (2020). *Investigating the injury severity of single-vehicle truck crashes in a developing country*. Accident Analysis & Prevention. Vol. 137, pp. 105444.

Rea L. M. and Parker R. A (2014). *Designing and conducting survey research: A Comprehensive Guide*. 4th ed., Jossey-Bass.

Rezapour, M., and Ksaibati, K., (2018). *Application of multinomial and ordinal logistic regression to model injury severity of truck crashes, using violation and crash data*. Journal of modern transportation. Vol. 26, Issue 4, pp. 268-277.

Saunders, M., Lewis, P. and Thornhill, A., 2012. Research methods for business students. 6th edition Pearson Education Limited, UK.

Schindler, R., Jänsch, M., Bálint, A., and Johannsen, H., (2022). *Exploring European heavy goods vehicle crashes using a three-level analysis of crash data.* International Journal of Environmental Research and Public Health. Vol. 19, Issue 2, pp. 663.

Tabachnick, B. G., Fidell, L. S., and Ullman, J. B. (2019). *Using multivariate statistics* (7th ed.). Pearson Education, Inc.

Taniform, P., Persia, L., Usami, D. S., Kunsoan, N. B., Karumba, M. M., and Wijnen, W., (2023). *An Assessment of the Social Costs of Road Traffic Crashes in Cameroon*. Sustainability. Vol. 15, Issue 2, pp. 1316.

Wang, Y., Luo, Y. and Chen, F., 2019. Interpreting risk factors for truck crash severity on mountainous freeways in Jiangxi and Shaanxi, China. European transport research review, 11, pp.1-11.

WHO (2018). *Global status report on road safety 2018. Geneva*. World Health Organization.

Yuan, Y., Yang, M., Guo, Y., Rasouli, S., Gan, Z., and Ren, Y., (2021). *Risk factors associated with truck-involved fatal crash severity: Analyzing their impact for different groups of truck drivers.* Journal of safety research, Vol. 67, pp. 154-165.

Zheng, Z., Lu, P., and Lantz, B., (2018). *Commercial truck crash injury severity analysis using gradient boosting data mining model*. Journal of safety research, 65, 115-124.